



**UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO
ESCUELA DE POSGRADO**

MAESTRÍA EN ESTADÍSTICA

TESIS

**FACTORES ASOCIADOS A LA VIOLENCIA DE GENERO EN
EL PERÚ: UN ANÁLISIS MEDIANTE ARBOLES DE DECISIÓN
CON DATOS DE LA ENDES 2024**

**PARA OPTAR AL GRADO ACADÉMICO DE MAESTRO EN
ESTADÍSTICA**

AUTORA:

Br. JULIETA FLORES VELARDE

ASESOR:

Mgt. EDGAR CENTENO HUAMANI

CODIGO ORCID: 0000-0002-7736-1299

**CUSCO – PERÚ
2026**



Universidad Nacional de San Antonio Abad del Cusco

INFORME DE SIMILITUD

(Aprobado por Resolución Nro.CU-321-2025-UNSAAC)

El que suscribe, el Asesor Edgar Centeno Huamani
..... quien aplica el software de detección de similitud al
trabajo de investigación/tesis titulada:
Factores asociados a la violencia de genero en el Perú:
Un análisis mediante árboles de decisión con datos del
ENDES 2024.

Presentado por: Julieta Flores Velarde DNI N° 41706918 ;
presentado por: DNI N°:
Para optar el título Profesional/Grado Académico de Maestro en Estadística

Informo que el trabajo de investigación ha sido sometido a revisión por 2 veces, mediante el
Software de Similitud, conforme al Art. 6° del **Reglamento para Uso del Sistema Detección de
Similitud en la UNSAAC** y de la evaluación de originalidad se tiene un porcentaje de 8.....%.

Evaluación y acciones del reporte de coincidencia para trabajos de investigación conducentes a grado académico o título profesional, tesis

Porcentaje	Evaluación y Acciones	Marque con una (X)
Del 1 al 10%	No sobrepasa el porcentaje aceptado de similitud.	☒
Del 11 al 30 %	Devolver al usuario para las subsanaciones.	☐
Mayor a 31%	El responsable de la revisión del documento emite un informe al inmediato jerárquico, conforme al reglamento, quien a su vez eleva el informe al Vicerrectorado de Investigación para que tome las acciones correspondientes; Sin perjuicio de las sanciones administrativas que correspondan de acuerdo a Ley.	☐

Por tanto, en mi condición de Asesor, firmo el presente informe en señal de conformidad y **adjunto**
las primeras páginas del reporte del Sistema de Detección de Similitud.

Cusco, 14 de Julio de 2026

Firma

Post firma Edgar Centeno Huamani

Nro. de DNI 31032950

ORCID del Asesor 0000-0002-7736-1299

Se adjunta:

1. Reporte generado por el Sistema Antiplagio.
2. Enlace del Reporte Generado por el Sistema de Detección de Similitud: oid: 27259:607849263

Julieta Flores Velarde

1. FACTORES ASOCIADOS A LA VIOLENCIA DE GÉNERO EN EL PERÚ_ UN ANÁLISIS MEDIANTE ÁRBOLES DE DECISI-2.pdf

 Universidad Nacional San Antonio Abad del Cusco

Detalles del documento

Identificador de la entrega

trn:oid:::27259:607849263

Fecha de entrega

14 jul 2026, 10:24 a.m. GMT-5

Fecha de descarga

14 jul 2026, 10:35 a.m. GMT-5

Nombre del archivo

1. FACTORES ASOCIADOS A LA VIOLENCIA DE GÉNERO EN EL PERÚ_ UN ANÁLISIS MEDIANTE ÁRB....pdf

Tamaño del archivo

7.4 MB

166 páginas

43.023 palabras

227.545 caracteres

8% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Texto citado
- Texto mencionado
- Coincidencias menores (menos de 10 palabras)

Exclusiones

- N.º de coincidencias excluidas

Fuentes principales

- 6%  Fuentes de Internet
- 2%  Publicaciones
- 6%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alerta de integridad para revisión

-  **Texto oculto**
12 caracteres sospechosos en N.º de página
El texto es alterado para mezclarse con el fondo blanco del documento.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.



UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL
CUSCO
ESCUELA DE POSGRADO

INFORME DE LEVANTAMIENTO DE OBSERVACIONES A TESIS

Dr. LEONCIO SOLIS QUISPE, Director de la Escuela de Posgrado, nos dirigimos a usted en condición de integrantes del jurado evaluador de la tesis intitulada **FACTORES ASOCIADOS A LA VIOLENCIA DE GENERO EN EL PERÚ: UN ANÁLISIS MEDIANTE ARBOLES DE DECISIÓN CON DATOS DE LA ENDES 2024** del / de la Br. JULIETA FLORES VELARDE. Hacemos de su conocimiento que el (la) sustentante ha cumplido con el levantamiento de las observaciones realizadas por el Jurado el día **VEINTISÉIS DE JUNIO DE 2026**.

Es todo cuanto informamos a usted fin de que se prosiga con los trámites para el otorgamiento del grado académico de MAESTRO EN ESTADÍSTICA.

Cusco, 01 de Julio del 2026

MTRO. YARIAN BARRETO SERRANO
Primer Replicante

MGT. JUDITH DURAN HUAMAN
Segundo Replicante

DR. GUILLERMO PAUCAR CARLOS
Primer Dictaminante

MTRO. JOEL GRIMALDO OLARTE ESTRADA
Segundo Dictaminante

AGRADECIMIENTOS

Expreso un profundo y sincero agradecimiento a Dios, que a través de la fe me brinda salud, fortaleza y perseverancia a lo largo de todo este proceso académico así como en la vida, permitiéndome culminar satisfactoriamente esta etapa importante de mi formación profesional.

Agradezco de forma muy especial a mi asesor(a) de tesis, por su apoyo constante, sus valiosos aportes académicos y su disposición para guiarme con responsabilidad y compromiso durante el desarrollo de esta investigación.

Mi reconocimiento y eterno agradecimiento a los profesores de la carrera y en la Escuela de Posgrado, quienes con mucho empeño, conocimientos y experiencia ayudaron significativamente a la formación académica y profesional con propósito personal.

Asimismo, expreso mi gratitud al Instituto Nacional de Estadística e Informática (INEI)(PERÚ Instituto Nacional de Estadística e Informática, n.d.), por proporcionar el acceso a la base de datos de la Encuesta Demográfica y Salud Familiar (ENDES 2024)(Perú: Encuesta Demográfica y de Salud Familiar, Endes 2024 - Informes y Publicaciones - Instituto Nacional de Estadística e Informática - Plataforma Del Estado Peruano, n.d.), insumo fundamental para la realización de este estudio.

A mi familia y amigos, que con su soporte constante, comprensión y aliento en cada etapa de este camino sinuoso, siendo un pilar fundamental para alcanzar este logro académico.

Finalmente, agradezco a todas las personas que, con su apoyo y motivación constante, de manera directa o indirecta, contribuyeron al desarrollo de la presente tesis, en especial para Zeze.

ÍNDICE GENERAL

AGRADECIMIENTOS.....	iii
ÍNDICE GENERAL.....	iv
ÍNDICE DE TABLAS.....	v
ÍNDICE DE FIGURAS.....	vi
RESUMEN.....	vii
ABSTRACT.....	viii
CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA.....	1
1.1. Situación problemática.....	1
1.2. Formulación del problema.....	4
a. Problema general.....	4
b. Problemas específicos.....	4
1.3. Justificación de la investigación.....	5
1.4. Objetivos de la investigación.....	7
a. Objetivo General.....	7
b. Objetivos Específicos.....	7
CAPITULO II:: MARCO TEORICO CONCEPTUAL.....	8
2.1. Bases teóricas.....	8
2.1.1. Violencia de género.....	8
2.1.2 Tipos de violencia e impacto en las mujeres.....	10
2.1.3. Método de clasificación estadística, Árboles de Decisión.....	17
2.2. Marco conceptual.....	35
2.3. Antecedentes de la investigación.....	40
2.3.1. Antecedentes internacionales.....	40
2.3.2. Antecedentes nacionales.....	44
2.3.3. Antecedentes locales.....	48
2.4. HIPOTESIS.....	53
a) Hipótesis general.....	53
b) Hipótesis específicas.....	53
2.5. Identificación de variables e indicadores.....	53
2.6. Operacionalización de variables.....	55
CAPITULO III: METODOLOGIA.....	57
3.1. Ámbito de estudio: localización política y geográfica.....	57
3.2. Tipo y nivel de investigación.....	57

3.3. Unidad de análisis.....	58
3.4. Población de estudio.....	58
3.5. Tamaño de muestra.....	58
3.6. Técnicas de selección de muestra.....	59
3.7. Técnicas de recolección de información	60
3.8. Técnicas de análisis e interpretación de información.....	62
3.9. Técnicas para demostrar la verdad o falsedad de la hipótesis planteada.....	63
CAPITULO IV: RESULTADOS	66
4.1 Descripción de la población de mujeres peruanas analizadas.	69
4.2. Descripción de violencia sufrida en la población de mujeres peruanas	75
4.3. Correlación de los factores que influyen en la presencia de violencia.....	79
4.3.1. Violencia física y factores asociados.....	80
4.3.2. Violencia psicológica y factores asociados	94
4.3.3. Violencia psicológica y factores asociados	107
4.3.3. Violencia general y factores asociados.....	118
DISCUSIÓN.....	135
CONCLUSIONES.....	138
RECOMENDACIONES	141
BIBLIOGRAFIA.....	1
ANEXOS.....	6

ÍNDICE DE TABLAS

Tabla 1. Operacionalización de variables.....	55
Tabla 2 Matriz de Construcción	65
Tabla 3 Fiabilidad del cuestionario de presencia de violencia hacia la mujer.....	66
Tabla 4. Prueba de normalidad Kolmogórov-Smirnov.....	67
Tabla 5. Descripción de las características demográficas de la pareja.....	69
Tabla 6 Descripción de las características reproductivas de la pareja.....	71
Tabla 7. Descripción de las características económicas y sociales de la pareja.....	73
Tabla 8. Indicadores de la violencia de la mujer.	75
Tabla 9. Dimensiones y variables de la violencia de la mujer.....	77
Tabla 10. Correlaciones entre las dimensiones o tipo de violencia.....	78
Tabla 11 Características demográficas de la pareja y su asociación a la violencia física	80
Tabla 12. Características reproductivas de la pareja y su asociación a la violencia física	83
Tabla 13. Características económicas y sociales de la pareja y su asociación a la violencia física	85

Tabla 14 Frecuencia según nodos y perfiles mas sensibles a violencia física según OR	89
Tabla 15. Matriz de confusión y principales indicadores de la violencia física	92
Tabla 16. Características demográficas de la pareja y su asociación a la violencia psicológica	94
Tabla 17. Características reproductivas de la pareja y su asociación a la violencia psicológica	97
Tabla 18 Características económicas y sociales de la pareja y su asociación a la violencia psicológica	100
Tabla 19. Frecuencia según nodos y perfiles más sensibles a violencia psicológica según OR	103
Tabla 20. Matriz de confusión y principales indicadores de la violencia psicológica	105
Tabla 21. Características demográficas de la pareja y su asociación a la violencia sexual	107
Tabla 22. Características reproductivas de la pareja y su asociación a la violencia sexual	109
Tabla 23. Características económicas y sociales de la pareja y su asociación a la violencia sexual	111
Tabla 24. Frecuencia según nodos y perfiles más sensibles a violencia sexual según OR	115
Tabla 25. Matriz de confusión y principales indicadores de la violencia sexual	116
Tabla 26. Características demográficas de la pareja y su asociación a la violencia general	118
Tabla 27. Características reproductivas de la pareja y su asociación a la violencia general	121
Tabla 28. Características económicas y sociales de la pareja y su asociación a la violencia general	124
Tabla 29. Frecuencia según nodos y perfiles más sensibles a violencia general según OR	129
Tabla 30. Matriz de confusión y principales indicadores de la violencia general	132
Tabla 31. Resumen de índices según tipo de violencia que tipo de violencia es más predecible	133

ÍNDICE DE FIGURAS

Figura 1. Distribución del nivel de violencia percibida por las mujeres	67
Figura 2. Frecuencia de tipo de violencia sufrida por las mujeres	77
Figura 3. Árbol clasificación de presencia de violencia física	88
Figura 4. Árbol clasificación de presencia de violencia psicológica	102
Figura 5. Árbol clasificación de presencia de violencia sexual	114
Figura 6. Árbol clasificación de presencia de violencia general	127

RESUMEN

La violencia de género constituye una de las principales problemáticas sociales y de salud pública a nivel mundial, debido a su impacto sobre la integridad física, psicológica, sexual y económica de las mujeres, como por la vulneración de sus derechos fundamentales. En el Perú, esta problemática persiste a pesar de la implementación de políticas y marcos normativos orientados a su prevención y sanción. Según la (ENDES) 2024, porcentajes elevados de mujeres ha experimentado algún tipo de violencia, evidenciando la influencia de diversos factores sociodemográficos, económicos y culturales.

La presente investigación identificó los factores asociados a la violencia de género en el Perú mediante la aplicación de modelos de árboles de decisión, utilizando la base de datos de la ENDES 2024. Se empleó un enfoque cuantitativo, de tipo aplicado, con diseño no experimental y corte transversal. La población estuvo conformada por 18 836 mujeres encuestadas, analizándose la totalidad de los casos para mejorar la capacidad predictiva y la estabilidad del modelo.

Aplicando algoritmos de árboles de decisión para determinar las variables con mayor capacidad predictiva, evaluando el desempeño de los modelos mediante métricas como exactitud, sensibilidad, especificidad y el área bajo la curva ROC (AUC-ROC).

Los resultados permitirán identificar los principales factores de riesgo asociados a la violencia contra la mujer y generar evidencia científica que contribuya al diseño de políticas públicas, estrategias de prevención e intervenciones focalizadas en los grupos de mayor vulnerabilidad, fortaleciendo la toma de decisiones para la reducción de la violencia de género en el Perú.

Palabras clave: Violencia de género, Factores asociados, Arboles de decisión, ENDES 2024, Salud pública.

ABSTRACT

Gender-based violence is one of the major social and public health challenges worldwide due to its impact on women's physical, psychological, sexual, and economic well-being, as well as the violation of their fundamental rights. In Peru, this issue persists despite the implementation of policies and legal frameworks aimed at its prevention and punishment. According to the 2024 Demographic and Family Health Survey (ENDES), a high percentage of women have experienced some form of violence, highlighting the influence of various sociodemographic, economic, and cultural factors.

This study identified the factors associated with gender-based violence in Peru through the application of decision tree models using data from the 2024 ENDES. A quantitative, applied research approach was adopted, with a non-experimental, cross-sectional design. The study population consisted of 18,836 surveyed women, and the entire dataset was analyzed to improve the predictive performance and stability of the models.

Decision tree algorithms were applied to identify the variables with the greatest predictive power. Model performance was evaluated using metrics such as accuracy, sensitivity, specificity, and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC).

The findings will help identify the main risk factors associated with violence against women and provide scientific evidence to support the design of public policies, prevention strategies, and targeted interventions for the most vulnerable groups, thereby strengthening decision-making processes aimed at reducing gender-based violence in Peru.

Keywords: Gender-based violence, Associated factors, Decision trees, ENDES 2024, Public health

CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA

1.1. Situación problemática

En salvajismo y agresión contra la mujer es reconocida mundialmente como las principales violaciones de derechos humanos, siendo persistentes y devastadoras. La Organización Mundial de la Salud (OMS) la define como cualquier acto de violencia de género que tenga como resultado un daño físico, sexual o psicológico para la mujer, incluyendo amenazas, coacción o privación arbitraria de la libertad, ya sea en el ámbito público o privado. (WHO, 2021)

Según datos de todo el mundo, aproximadamente una de cada tres mujeres o el 30%, ha sufrido alguna forma de violencia física, sexual o psicológica en su vida, ya sea por parte de su pareja o de otras personas en general. (WHO, 2021).

Este problema afecta a personas en todo el planeta, no importando su origen, clase social o cultura, los mismos actos poseen el impacto directo en el bienestar físico, mental y reproductiva de las féminas, en adición con la influencia de su participación en la economía, la sociedad y la política. La Asamblea General de las Naciones Unidas, mediante su Declaración sobre la Eliminación de la Violencia contra la Mujer, explica claramente que la violencia puede ser física, sexual o psicológica. (ONU, 1994)

En América Latina, el Caribe entre otras se están enfrentando un problema muy serio en cuanto a la violencia contra las mujeres. Según datos de 2023 proporcionados por organismos oficiales al Observatorio para la Igualdad de Género en América Latina y el Caribe de la CEPAL, por lo menos 3,897 féminas fueron afectadas por feminicidio en 27 países y territorios de cada una de las regiones. Esto significa que en promedio al

menos 11 mujeres mueren cada día en la región ya mencionada por razones relacionadas con su género. (CEPAL, 2024)

En 2023, Honduras tuvo el valor más alto de feminicidio, con 7,2 casos por 100.000 mujeres. A estos resultados siguieron República Dominicana, con veinte y cuatro; y el país de Brasil con uno coma cuatro, en términos de números absolutos, Brasil registró el mayor número de víctimas, con 1.463 mujeres fallecidas por feminicidio en ese año, seguido por México, que tuvo 852 casos.

En Perú se reportaron 146 feminicidios en 2023, equivalente a una tasa de 1.3 por 100.000 féminas, sin cambios significativos desde 2015, esto significa que, en promedio, unas 0.4 mujeres mueren por feminicidio cada día en ese país. (CEPAL, 2018).

La agresión de género en la región aparece con diferentes formas, tanto en lo cotidiano como en lo estructural. Esto incluye agresiones físicas y sexuales, pero también violencia simbólica, económica e institucional. Todo esto crea un entorno de desigualdad que limita las oportunidades de crecimiento y desarrollo para niñas y mujeres.

En Perú, las agresiones contra las mujeres siguen siendo uno de los controversias sociales más serias. Según resultados de la Encuesta Demográfica en Salud Familiar (ENDES) 2024, donde 52% de féminas entre quince a cuarenta y nueve años que hayan tenido una relación en algún momento sufrieron algún tipo de violencia por parte de sus acompañantes amorosos, las formas en agresión que enfrentan estas mujeres, siendo la más común la violencia psicológica y verbal, que afectó al 48,4%; la violencia física fue experimentada por el 25,5% y la sexual por el 5,6%. En el último año antes de la encuesta, el 7,5% de mujeres entre 15 a 49 años se observó que estas sufrieron violencia física y/o sexual donde los victimarios fueron los esposos o parejas. Esta cifra es bastante similar a la del año pasado, que fue del 8,3%; dentro de estas, la violencia física fue la más

reportada, con un 7%, mientras que la violencia sexual fue menor, con solo un 1,7%. (INEI, 2024)

En el año 2024, se registraron oficialmente 170 casos de feminicidio, aunque muchas organizaciones defensoras de derechos humanos creen que esa cifra podría ser mucho mayor debido a que muchas veces no se denuncian o no se procesan adecuadamente. (Defensoría del pueblo, 2025).

En la región Cusco, según el Centro de Emergencia Mujer (CEM Cusco), en 2024 se atendieron 846 denuncias de violencia donde la mayoría fueron psicológicas el 47.2% seguido de física en un 37% y sexual en un 15.5%, representando ser el tercer departamento con más denuncias solo superado por Lima y Arequipa con un 6,51% del total de casos (CEM, 2025)

Comprender la violencia de género va mucho más allá de identificar sus formas evidentes, también implica entender las raíces estructurales, los factores que la alimentan y los contextos sociales y culturales que la mantienen activa, entonces usar herramientas como los árboles de decisión resulta muy útil para recopilar datos con solidez, con estos métodos se puede descubrir patrones, determinar cuáles son los riesgos y predecir posibles situaciones, ayudando así a crear políticas públicas que se basen en hechos reales y confiables.

Los árboles de decisión son herramientas muy utilizadas en diferentes áreas, como la medicina con el objetivo de diagnosticar enfermedades y planear tratamientos, así también en las finanzas para evaluar riesgos de crédito, como en la ingeniería para mantener la calidad, y en la educación para predecir cómo les irá a los estudiantes. Lo que los hace especialmente útiles es que son fáciles de entender, muestran de manera gráfica y sencilla cómo diversas variables se relacionan entre sí, ayudando a entender datos que pueden ser complicados.

Con relación al tema de la violencia doméstica, los árboles para decisión pueden ayudar a entender mejor los casos. Analizan características sociales, económicas, educativas y del contexto tanto de las víctimas. Esto facilita identificar qué situaciones requieren atención prioritaria y dónde enfocarse en la prevención.

Analizar la violencia o abuso de género usando métodos estadísticos no sustituye la preponderancia de derechos humanos sino que la complementa, identificar factores de riesgo y áreas críticas ayuda a prevenir mejor; aprovechar recursos de forma más eficiente y diseñar programas de protección, atención y acceso a la justicia, esto permite evidenciar la magnitud del problema ayuda a desmontar creencias equivocadas, por ello se plantea las siguientes problemáticas de evaluación.

1.2. Formulación del problema

a. Problema general

- ¿Cuál es el grado de precisión del modelo de árboles de decisión para pronosticar la violencia de género en mujeres, utilizando las variables socioeconómicas, demográficas y culturales de la Encuesta Demográfica para la Salud Familiar (ENDES) 2024?

b. Problemas específicos

- ¿Cuáles son los factores más influyentes en la predicción de la violencia doméstica de género según el modelo de árboles de decisión aplicado a los datos de la ENDES 2024?
- ¿Cuál es la tasa de error del modelo de árboles de decisión al clasificar casos de violencia doméstica de género en mujeres en edad reproductiva?

- ¿Cuál es el tipo de violencia (física, sexual o psicológica) que presenta mayor eficacia predictiva en el modelo basado en árboles de decisión?

1.3. Justificación de la investigación

Oportunidad

En la investigación se presenta un momento clave para proceder con los derechos de las mujeres, ya que en el año 2024, se cumplen treinta años con la firma para la Convención de Belém do Pará “Convención Interamericana para Prevenir, Sancionar y Erradicar la Violencia contra la Mujer”, siendo el primer tratado de derechos humanos que declara la violencia de clase en género contra las mujeres como un tipo de transgresión en derechos humanos, convenio donde han impulsado acciones contra la violencia de género.

La oleada de COVID también mostró claramente el aumento en casos de violencia doméstica y de género, mostrando la evidencia la relevancia del tema como una prioridad de salud pública. Donde el propósito de Desarrollo Sostenible “Eliminar todas las formas de violencia contra todas las mujeres y las niñas en los ámbitos público y privado, incluida la trata y la explotación sexual y otros tipos de explotación” ONU 2015, donde se propone acabar con todas las formas de violencia hacia el género femenino, es fundamental contar con datos recientes, precisos y comparables, a lo cual el estudio permitirá medir los avances y plantear respuestas más efectivas.

Implicancias Prácticas

Los resultados de esta investigación serán de aplicación práctica directa en la forma en que se crean e implementan las políticas públicas, al entender las diferentes maneras y contextos en los que se presenta la violencia ya sea familiar lo que posibilita diseñar intervenciones más específicas y efectivas desde la prevención hasta la atención integral

a quienes lo sufren. Analizar casos de feminicidios que tuvieron denuncias previas o medidas cautelares ayuda a detectar fallas en los sistemas de protección del Estado, permitiendo actuar de manera más rápida en situaciones de riesgo, los datos recopilados permitirán recomendar la asignación eficiente de recursos públicos en programas que involucren a varios sectores.

Relevancia Social y Económica

La violencia hacia las féminas, a nivel personal afectan tanto al bienestar físico como mental causando problemas como depresión, trastornos de estrés postraumático, enfermedades crónicas y complicaciones en el área reproductiva, así como los menores que conviven con personas que ejercen violencia igualmente padecen, enfrentando dificultades en su desarrollo y muchas veces repitiendo estos patrones de violencia en sus propias vidas, lo cual genera costos importantes para los sistemas de salud, justicia y seguridad, la violencia revela y perpetúa las discriminaciones de género existentes en entornos sociales, siendo especialmente evidente para casos como los feminicidios.

Valor Teórico

Este trabajo propone un enfoque donde podemos ver mejor cómo las desigualdades en la sociedad se reflejan en actos de violencia, permitiendo medirla de manera uniforme, algo muy importante para poder comparar datos tanto a nivel nacional como internacional. analizar cuánta gente enfrenta este problema nos ayuda a entender mejor su historia. Esto es clave para crear modelos que expliquen no solo cómo pasa la violencia, sino también por qué sigue ocurriendo.

Utilidad Metodológica

Desde la posición metodológica, la presente investigación garantiza en un nivel alto, la comparabilidad así como la estandarización de los datos, dado uso de la encuesta ENDES 2024. También se siguen protocolos éticos y de seguridad, con lo que la

información recolectada es de tipo documental no involucra relación directa con los entrevistados.

1.4. Objetivos de la investigación

a. Objetivo General

- Analizar y estudiar la eficacia de árboles para decisión como herramienta predictiva para identificar factores asociados a la violencia doméstica, utilizando datos de la Encuesta Demográfica para la Salud Familiar (ENDES) 2024.

b. Objetivos Específicos

- Identificar los factores principales asociados a la violencia domestica mediante el uso de árboles de decisión aplicados a los datos de la ENDES 2024.
- Evaluar el nivel de precisión del modelo de árboles de decisión en la predicción en casos de violencia doméstica en mujeres.
- Comparar para desempeño predictivo de modelo en árboles de decisión según la forma de violencia (sexual, física o psicológica), usando datos en la ENDES 2024

CAPITULO II:: MARCO TEORICO CONCEPTUAL

2.1. Bases teóricas

2.1.1. Violencia de género

Es un fenómeno de alta complejidad que recae a muchos habitantes en todo el globo terráqueo. En OMS así como la ONU coinciden en que la violencia domestica incluye cualquier acción que cause lesión o sufrimiento sexual, físico o psicológico hacia las mujeres por motivos relacionados con su género. Esto abarca chantajes, abuso y privación para la libertad, así como en espacios públicos y en privados. Esta definición está alineada con la Declaración de Naciones Unidas de pacto para la Eliminación de la Violencia contra la Mujer de 1993.

El informe de OMS de 2021 detalla especialmente dos tipos comunes en violencia: la violencia de conyugue y la violencia genésica por parte de personas que no son conyugue. Violencia de pareja (VPI), significa cualquier comportamiento dañino de una pareja actual o pasada: ya sea física, sexual o emocional. Incluye agresiones como: golpes, bofetadas, patadas, estrangulamientos, amenazas con armas, coerción sexual, humillaciones y control emocional. A veces, las conductas de control, como apartar a alguien de sus familiares o amigos, controlar sus movimientos así como limitar admisión a recursos, también entran en esta categoría.

Violencia sexual por no pareja (VSNP): Se relaciona con actos de violencia sexual en contra de mujeres mayores entre 15 años que son cometidos por alguien que no es un esposo o pareja, como familiares, amigos, conocidos o extraños. Esto puede incluir ser forzada o amenazada para tener relaciones no deseadas, violación, intentos de violación o contactos sexuales no consentidos.

Desde 2015, Agenda 2030 en el Desarrollo Sostenible para la ONU pone énfasis eliminar la violencia de género. La meta 5.2 busca acabar con todo tipo de violencia que afecten a mujeres y niñas, en todos campos y formas, abarcando trata y explotación sexual. Este objetivo es clave para lograr igualdad y a las mujeres.

CEPAL ahora destaca que la violencia más extrema, como el femicidio, evidencia la persistente desigualdad y discriminación en América Latina y el Caribe. Organización que señala estos problemas son enraizados en normas patriarcales y relaciones de poder desiguales que sostienen la discriminación, para enfrentarlo propone cambiar las relaciones de poder y actuar sobre cuatro nudos estructurales: desigualdad socioeconómica, división rígida del trabajo, concentración del poder y patrones culturales patriarcales. El CEPAL reconoce la violencia domestica cruza y se agudiza con otras formas en segregación, como racismo, discriminación étnica, orientación sexual, identidad de género, discapacidad, migración y edad, por lo que requiere una mirada integral en las políticas públicas.

Convención de Belém do Pará, la cual fue adoptada en 1994 y vigente desde 1996, primer tratado por el cual se reconoció el derecho para las mujeres a vivir sin violencia y en definir la agresión de género como una transgresión de derechos humanos. Propuso crear mecanismos para proteger y defender los derechos para las mujeres y luchar para que la violencia disminuya en contra su integridad física, sexual y psicológica. Su influencia ha sido importante, inspirando otros instrumentos similares en África y Europa. En 2004, se creó un mecanismo de seguimiento para evaluar cómo se cumple esta convención. En 2024, la novena reunión de los países parte reafirmó su compromiso con acelerar esfuerzos para eliminar la violencia domestica de género.

Según lo adoptado por la OMS, la ONU y la CEPAL, es un estupro grave a la dignidad humana y un estancamiento para el desarrollo. Con el objetivo de acabar con

ello, se necesita un compromiso fuerte y acordado que aborde tanto sus formas visibles como sus raíces más profundas en las desigualdades de género.(OMS, 2021)

2.1.2 Tipos de violencia e impacto en las mujeres

Violencia Física

El constreñimiento físico de pareja íntima se define prácticamente como actos que pueden causar daño físico a la agredida. Esto incluye, ser abofeteada, empujada, golpeada con el puño o algún objeto, pateada, arrastrada, golpeada, asfixiada, quemada a propósito, y/o amenazada o agredida con un arma de fuego, cuchillo u otra arma.

En Perú, según la Encuesta Demográfica y de Salud Familiar (ENDES) 2024:

El 25,5% de las mujeres entre quince a cuarenta y nueve años que en un momento estuvieron unidas, sufrieron constreñimiento físico por parte de su esposo o compañero en su vida. Este porcentaje forma parte del 52,0% de mujeres que experimentaron un tipo de constreñimiento (psicológica / verbal, física o sexual) por parte de su pareja o acompañante.

Pasados los doce meses anteriores a la encuesta, la violencia física ejercida por el esposo o pareja afectó al 7,0% de las mujeres entre 15 a 49 años. Este dato se enmarca dentro del 7,5% de mujeres que fueron víctimas de constreñimiento físico así como sexual por parte de su pareja o compañero en el mismo periodo.

En un entorno mundial, según las consideraciones de la Organización Mundial para la Salud (OMS) de 2018 (basadas en datos de 2000-2018):

La incidencia de por vida de violencia física así como sexual del acompañante íntimo en contra de mujeres entre 15 a 49 años que en algún momento estuvieron casadas o en convivencia es del 27%. Significa que, en promedio, 641 millones y 753 millones

de mujeres entre 15 años o más han sufrido de este tipo de violencia al menos una vez en el periodo de vida desde los 15 años.

En los últimos doce meses, se infiere que alrededor del 13% de las mujeres entre quince a cuarenta y nueve años que han estado casadas o en unión han experimentado algún tipo de violencia física y/o sexual por parte de su pareja íntima, con un intervalo de incertidumbre entre el 10% y el 16%. Esto equivale a que aproximadamente 307 millones de mujeres mayores de 15 años han sido víctimas de esta forma de violencia en un periodo reciente. Cabe destacar que este fenómeno suele iniciarse en etapas tempranas de la vida, afectando con mayor frecuencia a adolescentes y mujeres jóvenes entre los quince y cuarenta y cuatro años, lo que refleja la vulnerabilidad de estas etapas de desarrollo y el profundo impacto que puede generar en la salud física, emocional y social de quienes lo sufren.

Violencia Sexual

La violencia sexual en el ámbito de la pareja se entiende como aquellas situaciones en las que la fémina es obligada físicamente en mantener relaciones sexuales no concesionadas, accede a ellas por intimidación o coacción, o es forzada a realizar actos sexuales que percibe como humillantes o degradantes, ya sea por parte de su pareja actual o de una anterior.

En el caso del Perú, el estudio de la Encuesta Demográfica y para Salud Familiar (ENDES) 2024 revelan una realidad preocupante: el 5,6% de las mujeres entre quince y cuarenta y nueve años declaró ser víctima de violencia sexual proveniente de su esposo o compañero en un momento de sus vidas. Esta cifra se enmarca dentro de un panorama aún más amplio, donde más de la mitad de las mujeres (52,0%) reportó haber sufrido alguna clase de violencia —psicológica / verbal, física o sexual— ejercida por su pareja

o conviviente. Estos datos reflejan no solo la medición del problema, sino también la emergencia de robustecer estrategias en prevención con atención integral a las víctimas.

En los 12 meses previos a la encuesta, se identificó que en el 1,7% de féminas peruanas entre quince y cuarenta y nueve años fue víctima de agresión sexual ejercida por su esposo o compañero. Aunque este porcentaje puede parecer reducido, refleja la persistencia de una problemática que afecta directamente la integridad y la salud de millones de féminas en el país.

Si ampliamos la mirada al contexto mundial, los datos de la Organización Mundial de la Salud (OMS, 2018) muestran una realidad aún más alarmante: alrededor del 27% de las mujeres de 15 a 49 años han sufrido violencia física y/o sexual por parte de su pareja íntima al menos una vez en su vida. Esta cifra deja en evidencia que la violencia por parte en acompañante íntima constituye un problema estructural y global de salud pública, cuya magnitud trasciende fronteras y requiere respuestas integrales y sostenidas.

Violencia Psicológica/Emocional

La violencia psicológica, también conocida como violencia o abuso emocional, constituye una de las maneras más constantes de maltrato en la acompañante íntima y, en muchos contextos, tiende a ser normalizada o minimizada. Sin embargo, su impacto en el bienestar mental y en el estilo de vida de las víctimas es profundo y sostenido en el tiempo.

En el plano mundial, el informe de la Organización Mundial de la Salud (OMS, 2021), que recoge estimaciones globales del año 2018, señala una limitación importante: la violencia psicológica no fue incluida dentro de las estimaciones de prevalencia específicas, ello se debe a la ausencia de un consenso internacional sobre definiciones y herramientas estandarizadas que permitan medirla y compararla de manera uniforme entre países. Esta falta de medición homogénea ha dificultado dimensionar la magnitud real del problema, a pesar de que constituye una de las formas más persistentes de

violencia de género. En Perú, según la Encuesta Demográfica y de Salud Familiar (ENDES) 2024, fue la forma más declarada de violencia sufrida por mujeres alguna vez unidas. Afectó al 48,4% de las mujeres de quince a cuarenta y nueve años.

Violencia Económica/Financiera

La violencia económica o financiera es una forma de violencia de género que se caracteriza por el control, limitación o restricción de los recursos económicos de una persona, especialmente de las mujeres, con el propósito de generar dependencia y mantener una relación de poder desigual. Esta violencia puede manifestarse mediante la prohibición de trabajar, el control del dinero, la negación de acceso a recursos básicos o la supervisión constante de las decisiones financieras. Aunque suele ser menos visible que la violencia física, sus efectos afectan gravemente la autonomía, la libertad y la calidad de vida de las víctimas, la Organización Mundial de la Salud (OMS) señala que muchas encuestas aún no incorporan preguntas relacionadas con el abuso económico, lo que limita la información disponible sobre esta forma de violencia. Estas prácticas terminan por aislar a las víctimas, generar dependencia y atraparlas en relaciones desiguales, afectando profundamente su autonomía y calidad de vida.

Naturaleza Interseccional de la Violencia

La violencia de clase en género no afecta a todas las mujeres del mismo tipo; su impacto está atravesado por factores como el origen de, la etnicidad, la preferencia sexual, la menos valía, la condición migratoria y la edad. Esto significa que ciertos grupos como las mujeres indígenas, migrantes, con discapacidad o transgénero enfrentan mayores niveles de vulnerabilidad y sufren formas más complejas de violencia. Por esto, resulta indispensable que las políticas públicas incorporen un enfoque interseccional, capaz de reconocer estas diferencias y responder a las necesidades específicas de cada grupo, garantizando una atención más justa e inclusiva.(MIMDES, 2019)

Principales desafíos en la medición:

Subestimación: El estigma y taboo, especialmente sobre violencia sexual, impiden que muchas mujeres reporten sus experiencias

Falta de estandarización: Variabilidad en definiciones, periodos de medición y criterios dificulta comparar datos entre estudios

Brechas de información: Escasez de datos sobre violencia psicológica/económica en ciertas regiones.

Factores influyen en la violencia

Factores individuales: Baja educación, alcoholismo, violencia infantil y trastornos de personalidad en agresores.

Factores relacionales: Problemas familiares, mala comunicación, dependencia económica y roles tradicionales en las parejas.

Factores comunitarios: Pobreza, exclusión social, falta de servicios básicos e impunidad.

Factores sociales: Normas patriarcales, privilegios masculinos y leyes discriminatorias que mantienen la violencia domestica como un constructo problemático.

Consecuencias en la salud física, mental y social de los agraviados

Las consecuencias de las violencias de clase en género son profundas, complejas y multidimensionales, afectando de manera significativa la salud corporal, mental y el bienestar integral de las mujeres, así como de sus hijos y familias, tanto a corto como a largo plazo.

En el ámbito físico, las víctimas suelen presentar lesiones, problemas de salud crónicos y complicaciones sexuales y reproductivas, como embarazos no planificados o infecciones de transmisión sexual, incluido el VIH. En su expresión más extrema, la violencia puede derivar en feminicidios, los cuales constituyen la forma más grave de

violencia de género; a nivel mundial, se calcula que entre el 38% y el 40% de los asesinatos de mujeres son perpetrados por sus propias parejas íntimas.

En el plano psicológico, la violencia en especial la sexual se relaciona con un mayor riesgo de depresión, intentos de suicidio y suicidios consumados, generando un impacto devastador en la salud mental de las víctimas.

Estas secuelas trascienden lo individual, ya que repercuten directamente en el entorno familiar. Los hijos que crecen en contextos de violencia contra sus madres tienen mayor probabilidad de reproducir o sufrir violencia de pareja en la adultez, lo que contribuye a perpetuar un ciclo intergeneracional de violencia.

En el ámbito social, la violencia contra las mujeres restringe su acceso a oportunidades educativas y laborales, limitando su autonomía y reforzando los círculos de pobreza y desigualdad. Esto evidencia que la violencia de género no solo es un problema de salud pública, sino también un obstáculo estructural para el desarrollo social y humano.

Perspectiva en derechos de la humanidad

Desde el sentido de los atributos humanos, toda manifestación de violencia contra la femina en general constituye una violación grave e inaceptable. Diversos organismos internacionales —como la Organización Mundial de la Salud (OMS), la Organización de los Estados Americanos (OEA), la Corte Interamericana de Derechos Humanos y ONU— han sido enfáticos en señalar que los Estados tienen la obligación de prevenir, investigar, sancionar y erradicar la violencia o agsio en género, garantizando la protección integral de las víctimas.

Si bien se han logrado avances normativos y en la producción de estadísticas, aún persisten limitaciones en la estandarización, periodicidad y calidad de los datos. En América Latina, muchos países han incorporado módulos específicos sobre violencia en

encuestas nacionales y han creado registros únicos, incluyendo los de feminicidios. Sin embargo, la información disponible presenta desafíos: en varios casos, las series no son estrictamente comparables debido a diferencias metodológicas o a la diversidad de definiciones legales del delito. A ello se suma la falta de desagregación por variables como el origen étnico-racial, lo que invisibiliza a grupos particularmente vulnerables.

La Agenda 2030 con el Desarrollo Sostenible, con su ODS 5, reconoce la eliminación de la violencia contra mujeres y niñas como un objetivo prioritario. Alcanzarlo implica no solo un compromiso político sostenido, sino también una inversión pública multisectorial que permita implementar estrategias de prevención y atención integrales.

Entre las acciones más relevantes se incluyen:

Cuestionar las normas sociales que promueven masculinidades asociadas al control y la violencia.

Reformar legislaciones discriminatorias en el ámbito familiar.

Fortalecer los derechos económicos de las mujeres, asegurando igualdad de oportunidades en empleo formal y educación secundaria.

Reducir la exposición temprana a la violencia en la infancia y atender factores de riesgo como el consumo de sustancias.

Promover cambios a nivel individual, trabajando en la transformación de actitudes que justifican la violencia y en la superación de roles de género rígidos dentro del hogar.

Con la información acotada, es fundamental que los esfuerzos de prevención se desarrollen en paralelo con el acceso garantizado a servicios integrales para las sobrevivientes, de modo que la respuesta frente a la violencia de género sea al mismo tiempo preventiva, protectora y reparadora.

Enfoque de salud pública

Afecta física, sexual, reproductiva y mentalmente a millones de mujeres, generando altos costos para sistemas de salud y economía. Las víctimas necesitan atención integral, apoyo psicológico, cuidados médicos inmediatos y redes de acompañamiento.

Se plantea un enfoque preventivo que requiere capacitación de personal sanitario para identificar y atender casos, e inclusión en campañas educativas sobre salud sexual, reproductiva y género.

2.1.3. Método de clasificación estadística, Árboles de Decisión

Definición y fundamentos matemáticos.

Para los árboles de decisión o clasificación en algoritmos para tipificar utilizando particiones sucesivas. Son apropiados cuando hay un número elevado de datos, siendo una de sus ventajas su carácter descriptivo que permite entender e interpretar fácilmente las decisiones tomadas por el modelo/(Parra, 2019)

Los árboles de decisión o de clasificación son un modelo surgido en el ámbito del aprendizaje automático (Machine Learning) y de la Inteligencia Artificial que, partiendo de una base de datos, crea diagramas de construcciones lógicas que nos ayudan a resolver problemas. La técnica de árboles de decisión, también conocida como segmentación jerárquica, consiste en organizar un conjunto de datos con el fin de clasificar y predecir resultados, para esto se parte de una muestra de entrenamiento en la que ya se conoce a qué grupo pertenece cada caso, lo que permite establecer los criterios de clasificación.

El proceso comienza con un nodo raíz, que representa la variable más relevante para separar la información, a partir de este nodo se van generando divisiones sucesivas de acuerdo con las variables independientes, buscando siempre formar grupos lo más

homogéneos posibles y reduciendo la “impureza” de los datos. De esta manera, se construye un esquema compuesto por nodos, ramas y hojas, los nodos corresponden a las variables ingreso, las ramas que simbolizan los distintos valores que estas variables pueden asumir y las hojas señalan los resultados o categorías finales de la variable dependiente.

El árbol de decisión funciona como un mapa visual de opciones y consecuencias que ayuda a comparar diferentes cursos de acción considerando probabilidades, costos y beneficios. Su utilidad abarca desde la toma de decisiones estratégicas en contextos simples hasta la aplicación en modelos matemáticos de predicción más complejos, pero una limitación importante es el sobreajuste, cuando los datos presentan inconsistencias, el modelo tiende a ajustarse demasiado a esas particularidades, lo que afecta su capacidad de generalización y disminuye su rendimiento predictivo en nuevos escenarios.

Estructura para un árbol de determinación

Un árbol de decisión empieza con nodo inicial único que se ramifica hacia distintos resultados posibles. Cada resultado da lugar a nuevos nodos que a su vez, generan más ramificaciones, adoptando así una forma parecida, como un árbol.

Tipos en nodos

Dentro del árbol de decisión se distinguen tres tipos principales de nodos:

Nodo de probabilidad (círculo): Indica la probabilidad de que ocurran determinados resultados.

Nodo de decisión (cuadrado): Representa un punto en el que es necesario tomar una decisión.

Nodo terminal: Corresponde al desenlace final de una ruta o trayectoria dentro del árbol.

Reglas de parada o “pre-poda”

Durante la construcción del árbol es necesario establecer criterios que indiquen si conviene seguir expandiendo una rama o detener su crecimiento. A estas condiciones se les denomina reglas de pre-poda, y su función es evitar que el modelo se vuelva excesivamente complejo, entre ellas destacan:

Límite de profundidad: Define un nivel máximo de ramificación del árbol.

Umbral de soporte: Establece un número mínimo de observaciones necesarias para justificar la creación de un nuevo nodo.

Definición y naturaleza:

Los árboles de decisión constituyen la familia de modelos predictivos ampliamente utilizados en tareas de clasificación y regresión dentro del aprendizaje supervisado. Su estructura se basa en una representación jerárquica compuesta por nodos de decisión, ramas y nodos terminales o hojas, mediante la cual se modela el desarrollo de toma de decisiones partiendo de la recolección de información. Cada nodo interno evalúa una condición lógica sobre una o más variables explicativas, dividiendo el conjunto de observaciones en subconjuntos cada vez más homogéneos respecto a la variable respuesta, mientras que cada hoja proporciona una predicción final, ya sea una categoría o un valor continuo (Breiman et al., 1984).

Es un modelo predictivo que representa decisiones y sus posibles consecuencias mediante una estructura jerárquica de nodos. En aprendizaje supervisado, un árbol particiona recursivamente el espacio de propiedades X en sectores homogéneos respecto a la variable objetivo Y . Cada nodo interno evalúa una condición sobre una o más características y conduce a ramas; cada hoja (nodo terminal) ofrece una predicción (una clase para clasificación o un valor numérico para regresión).

Formalmente, un árbol busca una partición $\{R_m\}_{m=1}^M$ del espacio de entrada tal que la predicción en la región R_m sea la más adecuada según un criterio de calidad (error de clasificación o error cuadrático medio).

Objetivo matemático del aprendizaje de un árbol

Dado un conjunto de entrenamiento $D = \{(x_i; y_i)\}_{i=1}^n$ la etiqueta (categórica o continua), el objetivo es encontrar una partición recursiva que minimice una función de pérdida empírica agregada sobre las regiones:

Para clasificación (minimizar tasa de error o maximizar pureza), se busca:

$$\min_{\{R_m\}} \sum_{m=1}^M \sum_{i: x_i \in R_m} L_{class}(y_i, \hat{y}_m) \quad (1)$$

donde $\hat{y}_m$ es la predicción de la hoja m (la clase mayoritaria) y L_{class} es la pérdida (0-1, entropía relativa, etc.).

Para regresión (minimizar error cuadrático):

$$\min_{\{R_m\}} \sum_{m=1}^M \sum_{i: x_i \in R_m} (y_i - \hat{y}_m)^2, \quad \hat{y}_m = \frac{1}{n_m} \sum_{i: x_i \in R_m} y_i \quad (2)$$

Donde n_m es el número de observaciones en R_m .

La búsqueda exacta de la partición óptima es combinatoriamente intratable, por ello los métodos prácticos emplean algoritmos voraces (greedy) que seleccionan en cada paso la mejor división local.

Criterios de partición (impurezas y ganancia)

La elección del mejor split (división) en un nodo se basa en una medida de impureza o pureza. Los criterios más usados:

Entropía e Information Gain (ID3 / C4.5)

La entropía de Shannon para una distribución de clases en un nodo es:

$$H(S) = - \sum_{c=1}^K p_c \log_2 p_c \quad (3)$$

donde p_c es la proporción de instancias de la clase c en el nodo S . La ganancia de información al partir S en S_L y S_R por una prueba t es:

$$\text{Gain}(S, t) = H(S) - \frac{|S_L|}{|S|} H(S_L) - \frac{|S_R|}{|S|} H(S_R) \quad (4)$$

Se elige el t que maximiza Gain .

Índice Gini (CART)

Mide la probabilidad esperada de error al clasificar al azar según la distribución de clases:

$$G(S) = 1 - \sum_{c=1}^K p_c^2 \quad (5)$$

La reducción de impureza por el split es:

$$\Delta G = G(S) - \frac{|S_L|}{|S|} G(S_L) - \frac{|S_R|}{|S|} G(S_R) \quad (6)$$

CART usa ΔG para elegir el mejor split.

Error de clasificación (misclassification error)

$$E(S) = 1 - \max_c p_c \quad (7)$$

Menos sensible que entropía o Gini; se usa menos en selección de splits, más como medida de rendimiento.

Reducción de varianza

Criterio de partición para regresión basado en la suma de cuadrados:

$$\Delta Var = Var(S) - \frac{|S_L|}{|S|} Var(S_L) - \frac{|S_R|}{|S|} Var(S_R) \quad (8)$$

donde $Var(S) = \frac{1}{|S|} \sum_{i \in S} (y_i - \bar{y}_S)^2$

Algoritmo general (esquema greedy recursivo)

Pseudo-código (versión simplificada – CART-like):

Función BUILD_TREE(S, depth):

Si criterio_detener(S, depth) entonces:

 return hoja con predicción = aggregate(y in S)

Mejor_gain = -inf

Mejor_split = None

Para cada variable j:

 Para cada umbral t posible en j:

 Particionar S en $S_L = \{x: x_j \leq t\}$ y $S_R = \{x: x_j > t\}$

 Calcular gain según criterio (Gini, Entropía, Varianza)

 Si gain > Mejor_gain:

 Mejor_gain = gain

 Mejor_split = (j, t)

Si Mejor_split es None:

 return hoja con predicción = aggregate(y in S)

Crear nodo con condición (j, t)

Nodo.izq = BUILD_TREE(S_L , depth+1)

Nodo.der = BUILD_TREE(S_R , depth+1)

return Nodo

Criterios de detención comunes:

- profundidad máxima alcanzada;
- número mínimo de ejemplos en una hoja;
- ganancia mínima absoluta;
- impureza ya suficientemente baja.

Complejidad computacional

Si se prueba todos los umbrales posibles para cada variable en cada nodo, la construcción exacta tiene costo en el peor caso cercano a $O(pn \log n)$ por nivel (considerando preordenamiento) y peor $O(pn^2)$ si se hace búsqueda ingenua. Implementaciones eficientes (p. ej. CART) ordenan las instancias por variable una vez por nodo y actualizan contadores, logrando complejidad aproximada $O(pn \log n)$ en la práctica.

La predicción para un nuevo ejemplo es $O(d)$ donde d es la profundidad del árbol.

Poda (pruning) y control de sobreajuste

Los árboles grandes tienden a sobre ajustar. Estrategias para controlar esto:

Poda previa (pre-pruning)

Detener crecimiento usando criterios: profundidad máxima, mínimo de ejemplos por hoja, mejora mínima de impureza. Simple y eficiente, pero puede dejar sin capturar interacciones complejas.

Poda posterior (post-pruning)

Se deja crecer el árbol y luego se podan ramas que no aportan mejora en alguna métrica de validación. Ejemplo clásico (CART): poda por costo-complejidad $C_\alpha(T)$:

$$C_\alpha(T) = R(T) + \alpha|T| \quad (9)$$

donde $R(T)$ es el error empírico del árbol T (p. ej. error en conjunto de entrenamiento o validación) y $|T|$ el número de hojas; α regula la penalización por complejidad. Se busca la subestructura T que minimice C_α .

Bias–Variance y estabilidad

Los árboles de decisión son modelos de baja bias (pueden aproximar relaciones complejas, no lineales) pero alta varianza (pequeñas variaciones en los datos pueden cambiar completamente la estructura del árbol).

Técnicas como bagging (Bootstrap Aggregating) y Random Forests reducen la varianza creando muchos árboles y promediando/votando sus asociaciones.

Manejo de variables: numéricas, categóricas y faltantes

Variables numéricas

Se consideran posibles umbrales t en valores observados y se evalúa la partición $x_j \leq t / x_j > t$. A menudo se prueban umbrales entre valores consecutivos ordenados.

Variables categóricas

Para una variable con k categorías, hay $2^k - 1$ particiones binarias posibles (exponencial). En la práctica, se:

ordenan categorías por target medio (para regresión) y se prueban cortes (heurística), o si k es pequeño, se prueban todas las particiones; si k grande, se agrupan.

Valores faltantes

Estrategias comunes:

Imputación antes del modelado (media, mediana, modelo),

En CART: en el momento del split, enviar ejemplos con faltantes por la rama con mayor probabilidad (surrogate splits) o calcular frecuencias condicionales para dividir ejemplos faltantes entre ramas.

Interpretación y explicación

Una fortaleza clave de los árboles de decisión es su interpretabilidad: la ruta desde la raíz a una hoja constituye una regla “if–then” fácil de comprender. Para ejemplo x , su predicción sigue condiciones lógicas:

$$\text{Si } x_{j_1} \leq t_1 \text{ y } x_{j_2} > t_2 \text{ y } \dots \Rightarrow \hat{y} = \hat{y}_{\text{hoja}} \quad (10)$$

Esto facilita su uso en contextos donde la aplicabilidad es crítica.

Importancia de variables y medidas derivadas

Se pueden cuantificar las variables más relevantes:

Importancia por reducción de impureza

La importancia de la variable j es la suma, sobre todos los nodos donde j fue usada, de la reducción de impureza ponderada por el número de instancias que llegan al nodo:

$$\text{Imp}(j) = \sum_{t: \text{split en } j} \frac{n_t}{n} \Delta \text{Impurity}(t), \quad (11)$$

donde n_t es el número de ejemplos en el nodo t .

Importancia por permutación

En evaluación empírica, permutar aleatoriamente los valores de j y medir la degradación en desempeño (aumento del error) da un estimador potente de importancia causal.

Evaluación de modelos de árboles

Métricas habituales:

Clasificación: accuracy, precisión, recall, F1, matriz de confusión, AUC-ROC.

Regresión: $RMSE$, MAE , R^2 .

Validación: validación cruzada $k - fold$, hold-out, curvas de aprendizaje para diagnosticar varianza/bias.

Para comparar splits y poda, se usan curvas de validación y selección basada en C_α (pruning por complejidad).

Ventajas y limitaciones

Ventajas

- Intuitivos y fáciles de interpretar.
- Capturan relaciones no lineales e interacciones sin requerir transformaciones elaboradas.
- Pueden manejar variables mixtas (numéricas y categóricas).
- No requieren escalado de variables.

Limitaciones

- Alta varianza $\rightarrow$ susceptibles al sobreajuste.
- Particiones axiales (basadas en umbrales en una sola variable) pueden ser subóptimas para fronteras complejas.
- Sensibles a pequeñas perturbaciones en los datos (inestables).
- Para variables categóricas con muchas clases, la búsqueda de particiones es costosa.
- Los árboles de decisión constituyen una técnica flexible y explicable para tareas de clasificación y regresión. Su base matemática (medidas de impureza,

particiones recursivas y criterios de poda) permite construir modelos interpretables y efectivos, especialmente cuando se combinan con técnicas de validación y ensamblado para mitigar la alta varianza. Para una tesis, los árboles ofrecen un buen equilibrio entre interpretabilidad y capacidad de modelado de relaciones complejas, y son una plataforma adecuada tanto para análisis exploratorio como para sistemas predictivo

Tipos de poda en árboles de decisión

Existen diferentes estrategias para controlar el crecimiento de un árbol de decisión y evitar que se vuelva demasiado complejo o se sobreajuste a los datos de entrenamiento. Entre las más utilizadas se encuentran:

Poda a través de coste-complejidad: Busca un equilibrio entre la exactitud del modelo y el tamaño del árbol. La complejidad se mide en función del número de hojas, de manera que se eliminan aquellas ramas cuya contribución a la precisión no justifica el aumento de complejidad.

Poda pesimista: Parte de los errores cometidos en la clasificación de los casos para calcular un “error de sustitución”. Con base en este criterio, se eliminan los subárboles que no representan una mejora sustancial en la capacidad predictiva del modelo, simplificando así la estructura.

Elementos principales de los algoritmos de clasificación

Los algoritmos de clasificación se estructuran a partir de diversos componentes fundamentales que permiten organizar y depurar el modelo:

Criterio de partición: Define la manera en que se dividen los nodos, seleccionando la variable y el punto de corte que generan la mayor homogeneidad posible en los grupos resultantes.

Regla de nodo terminal: Establece el momento en que un nodo deja de subdividirse, convirtiéndose en un resultado final.

Asignación de clase: Indica la categoría o clase que se asigna a cada nodo terminal, en función de las observaciones predominantes.

Fusión: Permite agrupar categorías poco relevantes o con escaso peso dentro de las variables predictoras, con el fin de simplificar el modelo.

Partición: Consiste en la selección del punto exacto de división, tras comparar todas las variables disponibles.

Poda: Se encarga de eliminar ramas que no aportan valor predictivo significativo, evitando el sobreajuste del árbol.

Evaluación: Mide el desempeño global del clasificador, estimando su precisión y calculando el nivel de riesgo asociado a sus asociaciones.

Algoritmos: CART, CHAID, C5.0.

Árbol CHAID (Chi-square Automatic Interaction Detection)

El árbol de decisión se construye a partir del análisis de tablas de contingencia y del estadístico chi-cuadrado χ^2 . Aunque inicialmente fue concebido para el manejo exclusivo de variables categóricas, con el tiempo el algoritmo fue ampliado para permitir el uso de variables categóricas nominales, ordinales y variables continuas. Esta ampliación hizo posible la generación tanto de árboles de decisión orientados a problemas de clasificación como de árboles de regresión.

En este tipo de algoritmos, los nodos no se limitan a dividirse únicamente en dos ramas, sino que pueden generar múltiples subdivisiones. La estructura del árbol se define a partir de la significancia estadística obtenida mediante contrastes de hipótesis, los cuales permiten jerarquizar las variables predictoras o de salida, así como agrupar valores que presentan comportamientos similares respecto a la variable dependiente. De este modo,

los valores que no muestran diferencias estadísticamente significativas se integran en una misma categoría y se asignan a una sola rama, mientras que aquellos valores distintos se mantienen separados.

El criterio estadístico utilizado para realizar las divisiones depende del tipo de variable dependiente. Cuando la variable de respuesta es continua, se emplea la prueba F; en cambio, cuando la variable es categórica, se utiliza la prueba chi-cuadrado. Asimismo, la identificación de relaciones significativas se apoya en diferentes pruebas estadísticas según la naturaleza de la variable dependiente, ya sea nominal, ordinal o de intervalo. En el caso específico de variables dependientes nominales, se emplean principalmente dos contrastes: el estadístico chi-cuadrado χ^2 y la razón de verosimilitud (G^2).

El árbol de decisión se construye a partir de tablas de contingencia y del uso del estadístico χ^2 como base. Si bien inicialmente fue diseñado únicamente para trabajar con variables categóricas, con el tiempo se ha ampliado su alcance para incluir variables nominales, ordinales e incluso continuas, lo que ha permitido desarrollar tanto árboles de clasificación como árboles de regresión.

A diferencia de otros enfoques más limitados, este algoritmo permite que cada nodo se divida en más de dos ramificaciones. La organización de las variables predictoras y la conformación de grupos con valores similares se basan en los resultados de las pruebas de significancia estadística aplicadas. De este modo, aquellos valores que no presentan diferencias estadísticamente relevantes se reúnen en una misma categoría y se asignan a una sola rama del árbol, mientras que los valores que sí muestran diferencias se mantienen de manera independiente, sin ser alterados.

En cuanto al criterio estadístico, cuando la variable predicha es continua se utiliza la prueba F, y cuando la variable es categórica se aplica la prueba Chi-cuadrado (χ^2). Además, para establecer si una relación es significativa, se emplean distintas

pruebas según el tipo de variable dependiente: nominal, ordinal o de intervalo. En el caso específico de variables nominales, se utilizan principalmente dos métodos: el estadístico χ^2 y la razón de verosimilitud (G^2), ambos ampliamente reconocidos en el análisis de clasificación.

Árbol CART (Classification and Regression Trees)

El algoritmo CART es el acrónimo de Clasificación And Regresión Trees (Árboles de Clasificación y de Regresión) fue diseñado por Breiman et al. (1984). Constituye una de las técnicas más utilizadas en la construcción de árboles de decisión. A diferencia de otros métodos, este enfoque se caracteriza por generar árboles binarios, lo que implica que cada nodo se divide siempre en dos ramas.

Una de sus principales ventajas es su flexibilidad, ya que permite trabajar con variables de entrada y salida de distintos tipos: nominales, ordinales o continuas. Gracias a esta versatilidad, el algoritmo puede aplicarse tanto en problemas de clasificación (cuando el objetivo es predecir categorías) como en problemas de regresión (cuando se busca predecir valores numéricos).

Árbol C5.0

El algoritmo C5 y, en especial, su versión libre C4.5, se han consolidado como una de las herramientas más empleadas para la construcción de árboles de clasificación. Su popularidad radica en la eficacia y la robustez que ofrece al momento de organizar y analizar grandes volúmenes de datos, lo que lo convierte en un referente dentro del campo de la minería de datos y el aprendizaje automático. La forma de inferir árboles de decisión a través de este algoritmo es el resultado de la evolución del algoritmo C4.5 (Quinlan, 1993)

- Crea modelos de árbol de clasificación con variables de salida categóricas
- Variables de entrada pueden ser continuas o categóricas

- Basado en el concepto de "ganancia de información" de la teoría de la información

Funcionamiento del algoritmo

- Construye el árbol de manera descendente
- Comienza seleccionando el mejor atributo para la raíz mediante tests estadísticos
- Crea ramas y nodos para cada valor posible del atributo
- Reparte ejemplos de entrenamiento según valores del atributo raíz
- Repite el proceso para cada nodo generado

Representa la diferencia entre la información necesaria para identificar la categoría destino y la información necesaria cuando se conoce el valor de una variable de entrada, indicando menor incertidumbre en la clasificación.

Prueba de Mantel–Haenszel

La prueba es un procedimiento estadístico inferencial diseñado para evaluar la asociación entre dos variables categóricas dicotómicas, controlando simultáneamente el efecto de una o más variables de estratificación o de confusión. Esta prueba resulta especialmente útil en estudios observacionales, donde no es posible la asignación aleatoria, y se requiere estimar una asociación ajustada entre variables.

Desde un punto de vista formal, la prueba MH permite contrastar la hipótesis nula de independencia condicional entre una variable respuesta Y y una variable explicativa X , dadas una o varias variables de estrato S :

$$H_0: X \perp Y \mid S \quad (12)$$

Estructura matemática de la prueba Mantel–Haenszel

Supóngase que los datos están estratificados en K estratos (por ejemplo, definidos por nodos terminales de un árbol de decisión), y que en cada estrato k se dispone de una tabla de contingencia 2×2 :

	$Y = 1$	$Y = 0$	Total
$X = 1$	a_k	b_k	n_{1k}
$X = 0$	c_k	d_k	n_{0k}
Total	m_{1k}	m_{0k}	n_k

(13)

Estimador de razón de odds común (Mantel–Haenszel)

La razón de odds ajustada de Mantel–Haenszel se define como:

$$\hat{\theta}_{MH} = \frac{\sum_{k=1}^K \frac{a_k d_k}{n_k}}{\sum_{k=1}^K \frac{b_k c_k}{n_k}} \quad (14)$$

Este estimador representa una medida de asociación promedio ajustada, que controla la variabilidad introducida por los estratos.

Estadístico de prueba Mantel–Haenszel

El estadístico de contraste MH se define como:

$$X_{MH}^2 = \frac{(\sum_{k=1}^K (a_k - E[a_k]))^2}{\sum_{k=1}^K Var(a_k)} \quad (15)$$

donde:

$$E[a_k] = \frac{n_{1k}m_{1k}}{n_k}, \quad \text{Var}(a_k) = \frac{n_{1k}n_{0k}m_{1k}m_{0k}}{n_k^2(n_k-1)}$$

Bajo la hipótesis nula, el estadístico $2X_{MH}^2$ sigue aproximadamente una distribución chi-cuadrado con 1 grado de libertad.

Relación conceptual con los árboles de decisión

Desde la perspectiva de los árboles de decisión, la prueba de Mantel–Haenszel adquiere un rol complementario y altamente relevante:

Estratificación natural del árbol

Cada nodo terminal del árbol de decisión define un estrato homogéneo, construido a partir de reglas lógicas del tipo:

$$S_k = \{X_1 \leq c_1, X_2 \in A_2, \dots\} \quad (16)$$

Estos estratos permiten controlar automáticamente variables de confusión sin necesidad de modelarlas explícitamente.

Validación inferencial del árbol

Mientras el árbol optimiza criterios predictivos (Gini, entropía), la prueba MH permite evaluar si las asociaciones observadas en las ramas del árbol son estadísticamente significativas, fortaleciendo la interpretación causal y analítica del modelo.

Contraste de hipótesis sobre reglas del árbol

Cada división principal del árbol puede interpretarse como una hipótesis local del tipo:

$$P(Y = 1|X = 1, S_k) \neq P(Y = 1|X = 0, S_k) \quad (17)$$

La prueba MH permite contrastar estas hipótesis de forma agregada sobre múltiples nodos.

Consistencia entre clasificación y asociación

La matriz de confusión generada por el árbol (verdaderos positivos, falsos positivos, etc.) puede reorganizarse en tablas 2×2 por estrato, facilitando la aplicación de la prueba MH como mecanismo de validación estadística del modelo de clasificación.

Ventajas metodológicas del enfoque combinado

Control explícito de variables de confusión.

Integración de aprendizaje automático interpretativo con estadística inferencial clásica.

Mayor robustez en la demostración de hipótesis en estudios observacionales.

Complementa métricas predictivas (accuracy, AUC) con evidencia estadística formal.

La prueba de Mantel–Haenszel, integrada al análisis mediante árboles de decisión, permite superar la limitación puramente predictiva del modelo, aportando un sustento estadístico sólido para la aceptación o rechazo de hipótesis. Este enfoque híbrido fortalece la validez interna del estudio y ofrece una interpretación más rigurosa de los factores asociados a la variable de interés.

2.2. Marco conceptual

- **Violencia de género como problema social y de salud pública**

La violencia de género constituye una de las expresiones más persistentes de desigualdad estructural entre hombres y mujeres y representa una grave vulneración de los derechos humanos. Este fenómeno comprende no solo agresiones físicas, sino también manifestaciones psicológicas, sexuales y económicas que afectan de manera directa la salud, el bienestar emocional y la autonomía de las mujeres. La Organización Mundial de la Salud reconoce la violencia contra la mujer como un problema de salud pública de alcance global, debido a sus consecuencias físicas, psicológicas y sociales, así como a su impacto en el desarrollo sostenible de los países (OMS, 2021).

En el contexto peruano, la violencia de género mantiene una alta prevalencia pese a la existencia de marcos normativos orientados a su prevención y sanción. La Encuesta Demográfica y de Salud Familiar (ENDES), elaborada por el Instituto Nacional de Estadística e Informática, evidencia que un porcentaje significativo de mujeres ha experimentado algún tipo de violencia a lo largo de su vida, lo que pone de manifiesto la magnitud y complejidad del problema (INEI, 2024). Estos datos reflejan que la violencia de género se encuentra asociada a múltiples factores sociodemográficos, económicos y culturales, tales como el nivel educativo, la situación laboral, el entorno familiar y las relaciones de poder desiguales.

- **Análisis estadístico multivariable en estudios sociales**

El estudio de fenómenos sociales complejos, como la violencia de género, requiere enfoques metodológicos capaces de analizar simultáneamente la interacción entre múltiples variables explicativas. Los métodos estadísticos tradicionales, basados en supuestos lineales y paramétricos, pueden resultar

limitados cuando las relaciones entre variables son no lineales o presentan interacciones jerárquicas.

En este contexto, las técnicas de minería de datos y aprendizaje automático han adquirido creciente relevancia en las ciencias sociales y la salud pública, al permitir identificar patrones complejos y relaciones no evidentes en grandes bases de datos poblacionales. Según Hastie, Tibshirani y Friedman (2009), estos métodos proporcionan herramientas flexibles para la predicción y explicación de fenómenos, especialmente cuando se dispone de un número elevado de variables explicativas, como ocurre en encuestas nacionales de gran escala como la ENDES.

- **Árboles de decisión: fundamentos teóricos**

Los árboles de decisión constituyen una técnica fundamental del aprendizaje supervisado, utilizada tanto en problemas de clasificación como de regresión. Su estructura jerárquica permite representar decisiones y sus posibles consecuencias mediante nodos internos y nodos terminales, facilitando la interpretación de los resultados. Desde una perspectiva formal, un árbol de decisión particiona recursivamente el espacio de variables predictoras en regiones homogéneas respecto a la variable dependiente (Breiman et al., 1984).

El objetivo del aprendizaje de un árbol de decisión es encontrar divisiones sucesivas del conjunto de datos que maximicen la homogeneidad interna de cada nodo, minimizando una función de pérdida asociada a la variable respuesta. Este proceso se realiza mediante algoritmos voraces (greedy), que seleccionan en cada etapa la mejor división local según un criterio de calidad estadística.

- **Criterios de partición e impureza**

La selección de la mejor partición en cada nodo se basa en medidas de impureza. Entre las más utilizadas se encuentran la entropía de Shannon y el índice Gini. La entropía mide el grado de incertidumbre en la distribución de clases, mientras que el índice Gini evalúa la probabilidad de error al clasificar una observación de manera aleatoria (Quinlan, 1986; Breiman et al., 1984).

En particular, el algoritmo CART (Classification and Regression Trees), propuesto por Breiman et al. (1984), emplea el índice Gini para problemas de clasificación y la reducción de varianza para regresión. Este enfoque ha sido ampliamente adoptado en investigaciones empíricas debido a su robustez y simplicidad computacional.

- **Poda y control del sobreajuste**

Uno de los principales desafíos de los árboles de decisión es su tendencia al sobreajuste, ya que modelos muy profundos pueden ajustarse excesivamente a los datos de entrenamiento. Para controlar este problema, se emplean estrategias de poda previa y poda posterior. La poda por costo-complejidad, introducida por Breiman et al. (1984), permite penalizar la complejidad del árbol mediante un parámetro que equilibra el error de clasificación y el número de nodos terminales. Hastie et al. (2009) destacan que el uso de validación cruzada y conjuntos de prueba independientes resulta fundamental para seleccionar el nivel óptimo de complejidad del árbol y garantizar su capacidad de generalización.

- **Interpretabilidad y relevancia en estudios de violencia de género**

Una de las principales fortalezas de los árboles de decisión es su alto nivel de interpretabilidad. Cada ruta desde la raíz hasta una hoja puede expresarse como

una regla lógica del tipo “si–entonces”, lo que facilita la comprensión de los factores que influyen en la ocurrencia de un fenómeno determinado. Esta característica resulta especialmente relevante en estudios sobre violencia de género, donde no solo interesa predecir la ocurrencia del evento, sino también identificar y explicar los factores asociados, con el fin de orientar políticas públicas y estrategias de intervención (Hastie et al., 2009).

En investigaciones basadas en datos de la ENDES 2024, los árboles de decisión permiten identificar variables clave —como características sociodemográficas, económicas y del entorno familiar— y establecer jerarquías de importancia predictiva, contribuyendo a una comprensión integral del fenómeno estudiado.

- **Pertinencia del uso de árboles de decisión con datos ENDES 2024**

La ENDES 2024 constituye una fuente de información estadística confiable y representativa a nivel nacional, que recoge datos detallados sobre condiciones de vida, salud y experiencias de violencia contra la mujer. La naturaleza multidimensional de esta base de datos hace pertinente el uso de modelos de árboles de decisión, ya que estos permiten manejar variables mixtas, valores faltantes y relaciones no lineales sin requerir transformaciones complejas (Quinlan, 1993; Breiman et al., 1984).

En consecuencia, la aplicación de árboles de decisión en la presente investigación se sustenta en una sólida base teórica y metodológica, y resulta coherente con el objetivo de identificar los factores asociados a la violencia de género en el Perú, utilizando información proveniente de la ENDES 2024.

- **Violencia de género:** se entiende como cualquier forma de agresión, ya sea física, psicológica o sexual, dirigida hacia una persona que se encuentra en una situación de vulnerabilidad en el momento del acto. Este concepto fue reconocido y

promovido por la ONU para diferenciar una forma específica de violencia que, en la mayoría de los casos, es ejercida por hombres hacia las mujeres u otras identidades, afectando su dignidad, identidad y bienestar integral.

- **Clasificador de árboles:** En el análisis de decisiones, los árboles de decisión funcionan como una herramienta visual que permite representar de manera clara y estructurada las distintas alternativas y los posibles cursos de acción. En cambio, dentro de la minería de datos, los árboles de decisión no reflejan directamente las decisiones, sino que sirven para describir y organizar la información. El árbol de clasificación obtenido a partir de este proceso puede luego emplearse como un insumo valioso para apoyar la toma de decisiones.
- **Factores en violencia domestica:** Las investigaciones revisadas muestran que la violencia contra la mujer es un fenómeno con raíces históricas, que persiste hasta hoy y que, en gran medida, permanece silenciado, ya que muchas víctimas no denuncian por miedo a sufrir represalias o nuevas formas de violencia. Los factores que incrementan el riesgo de que ocurra esta violencia son múltiples y se relacionan con las particularidades de la dinámica familiar y social de cada país. La evidencia señala que estos factores se distribuyen en distintos niveles: desde el macroestructural, hasta el exosistema y el microsistema, lo que refleja la complejidad del problema. Asimismo, existen diversas teorías y modelos que buscan explicar los factores asociados a la violencia de género, destacando su carácter multifactorial y la necesidad de abordarlo desde diferentes enfoques.

2.3. Antecedentes de la investigación

2.3.1. Antecedentes internacionales

La tesis “Estudio longitudinal del impacto de la violencia de pareja sobre la salud física y el sistema inmune de las mujeres” presentada por Segunda Sánchez Lorente (2009), como tesis doctoral en la Universidad de Valencia, tuvo como propósito principal analizar la evolución de la salud de las mujeres víctimas de violencia de pareja durante un periodo de tres años. El estudio evaluó no solo la violencia sufrida, sino también la salud física, el funcionamiento del sistema inmune, así como factores personales, sociales y de pareja experimentados en ese lapso. Para esto, se contó con la participación de 182 mujeres, de las cuales 130 habían sido víctimas de violencia y 52 conformaron el grupo de control al mantener una relación de pareja sin violencia. La recolección de datos se realizó mediante cuestionarios y análisis de saliva, con el fin de corroborar indicadores de salud física. Dado que las variables no seguían una distribución normal, se empleó análisis de varianza (ANOVA) utilizando el test de Welch, junto con pruebas post-hoc de Tukey o T3 de Dunnet. Entre los hallazgos, se observó que el nivel educativo de las mujeres no estaba asociado con la violencia de pareja; sin embargo, el estado civil y los ingresos económicos mensuales sí mostraron relación con la violencia de género. Además, se evidenció que la violencia ejercida por la pareja podía disminuir con el tiempo, y que las mujeres que decidieron separarse del agresor experimentaron una notable mejoría en su salud física y mental. En conclusión, el estudio señala que la salud física de las mujeres víctimas de violencia tiende a mejorar con el tiempo, pero enfatiza que la separación del agresor es el factor más determinante para favorecer tanto la recuperación física como el bienestar emocional de la mujer (Sánchez Lorente, 2009).

En la tesis doctoral “Violencia de género en adolescentes. análisis de las percepciones y de las acciones educativas propuestas por la junta de Andalucía” de Francisco Miguel Guzmán Sánchez (2015), presentada en la Universidad de Sevilla, tuvo como propósito explorar las percepciones de los adolescentes respecto a la violencia de género, identificar los estereotipos que mantienen, analizar sus conocimientos previos sobre el tema y comprender las actitudes de tolerancia frente a actos de violencia de género. El estudio siguió un enfoque descriptivo, aplicando encuestas a adolescentes y posteriormente desarrolló un diseño comparativo causal para examinar diferencias sociodemográficas en función del género. La muestra estuvo conformada por 875 adolescentes: un 51,8% entre 14 y 16 años y un 48,2% entre 17 y 19 años. Para la recolección de datos, se empleó un cuestionario “ad hoc” validado, cuya confiabilidad se reforzó con la opinión de ocho especialistas: cuatro en el ámbito social y cuatro en el universitario. Entre las conclusiones más relevantes, el autor destaca que, a pesar de encontrarnos en pleno siglo XXI, los estereotipos de desigualdad de género aún persisten, así como la violencia de género aparece asociada a una dinámica bidireccional, es decir, que puede manifestarse en agresiones mutuas. En cuanto a las percepciones de los adolescentes, se identificó que tanto hombres como mujeres tienden a minimizar la violencia: los varones restan importancia a sus conductas agresivas, mientras que las mujeres tienden a subestimar los abusos que sufren. Estos hallazgos coinciden con estudios previos (González y Santana, 2001; Lorente, 2006; Walker, 2009; Moral de la Rubia et al., 2011; Roca y Masip, 2011), que muestran cómo los prejuicios de género pueden distorsionar la percepción de la violencia. Asimismo, Ferrer y Bosch (2005) advierten que, en ocasiones, los análisis pueden verse afectados por definiciones demasiado amplias de violencia o por la falta de claridad en las escalas utilizadas, lo que

Guzmán Sánchez plantea que su investigación debe ser ampliada en futuras ediciones para obtener resultados más sólidos y confiables (Guzmán Sánchez, 2015).

En la tesis de licenciatura “Proyecto de investigación sobre violencia de género en adolescentes” realizado por Auxiliadora Martos Delgado, Departamento de Psicología, Universidad de Jaén, cuyos objetivos fueron: analiza la violencia doméstica, considerar en adolescentes con relación de roles sexistas, estudiar el tipo de violencia más frecuente, analizar la sostenibilidad que tienen los adolescentes para hacer frente a las escalas de violencia de genero. En el estudio se tomó una muestra de 480 estudiantes entre 13 y 19 años de la en el instituto de educación secundaria Virgen del Carmen e instituto de Educación Secundaria Las cuentezuelas, el estudio fue realizado en forma de cuestionario a través de entrevistas, a los estudiantes en cuestión, el autor demuestra que a través de la concientización oportuna a los adolescentes es posible prevenir los actos de violencia, para lo cual se deben hacer más campañas de concientización en los colegios, así como en las escuelas de padres ya que los padres deben entender que la principal enseñanza vienen de los hogares, enseñando y aprendiendo a respetarse mutuamente. (Martos Delgado, 2015)

En la tesis doctoral “Violencia de género y prácticas tradicionales: una contribución al debate metodológico para su medición en Europa” presentada por Amalia Gómez Casillas, de la Universidad Autónoma de Barcelona, el objetivo principal del estudio de tesis fue el de hacer un plan metodológico para tratar de erradicar la violencia de género en prácticas tradicionales, tales como la mutilación genital femenina y el matrimonio precoz, para este objetivo la metodología usada fue la triangulación metodológica que confiere la evaluación de la información y los indicadores, cuyas conclusiones asertivas fueron: la veracidad de los datos disponibles no son confiables, la comparación supranacionales tienen limitaciones para cualquier tipo de investigación, las limitaciones

de los resultados dependen del recojo de ellos en cada país ya que no hay un ente controlador que pueda recabar información fidedigna para hacer las conclusiones respectivas. (Gómez Castillas, 2018)

En la tesis, “Trabajo de investigación sobre la violencia de género” de Roca Monjo Antonia, presentada en la Universidad Internacional de La Rioja (España), tuvo como propósito principal analizar diferentes fuentes de información relacionadas con la violencia de género. Entre sus objetivos se incluyeron: conocer las cifras de mujeres fallecidas por esta causa, examinar los datos judiciales sobre víctimas de feminicidio, analizar las estadísticas de llamadas a servicios de emergencia (ATEMPRO) y recoger la percepción social en torno a la violencia de género, así como el nivel de interés que la población manifiesta frente a esta problemática. La investigación adoptó una metodología cuantitativa, complementada con un enfoque histórico y la triangulación de datos, lo que permitió contrastar diversas fuentes para obtener una visión más amplia del fenómeno. Entre sus conclusiones, la autora señala que los datos disponibles reflejan principalmente los casos más extremos que son los feminicidios, lo que impide dimensionar con precisión el número real de mujeres que son víctimas de violencia de género. Asimismo, se evidenció que entre 1999 y 2004 se registraron 364 muertes, mientras que entre 2005 y 2010 la cifra aumentó a 400 fallecimientos. Roca Monjo también resalta que, pese a las campañas de sensibilización y las medidas de protección implementadas, los asesinatos por violencia de género se han mantenido constantes, lo que pone en cuestión la eficacia de las estrategias actuales (Roca Monjo, 2011).

En la tesis doctoral “Violencia contra las mujeres y alguien más...” sustentada por María de Lujan Piatti, de la Universidad de Valencia, Valencia (2013), la violencia contra la mujer tiene diversos tipos como también factores, no distingue de la raza ni el nivel socioeconómico ni cultural, ya que cualquier mujer puede ser víctima potencial de la

violencia de género, los objetivos principales son el definir más explícitamente los que es la violencia de género, así como el demostrar que para evitar la violencia de género no es suficientes las leyes que los amparan, para la realización del presente trabajo se utilizó una amplia bibliografía, así como los congresos y ponencias en los que se puso a discusión la violencia de género, las conclusiones fueron que hay un aumento considerable de la violencia de género en los últimos años en los cuales se realizó el estudio.(De Lujan Piatti, 2013)

2.3.2. Antecedentes nacionales

Vargas y Walde (2017) realizaron el trabajo “Factores socioculturales que influyen en la violencia intrafamiliar a partir de la diferencia de género en el distrito huancayo en la actualidad”, presentada en la Universidad del Centro del Perú (Vargas Armas & Walde Soto, 2017) tuvo como propósito principal analizar los valores socioculturales que influyen en la violencia intrafamiliar, abordándola desde la perspectiva de la violencia de género. Para ello, se trabajó con una muestra de 324 mujeres cuyos casos fueron reportados en la DEMUNA de la localidad de Áncash. Los hallazgos del estudio revelan que desde la niñez se asignan roles diferenciados y rígidamente marcados entre hombres y mujeres. En este esquema, la mujer queda relegada a las labores domésticas, una tradición transmitida de generación en generación, muchas veces por las propias mujeres. Asimismo, se evidenció que la violencia de género e intrafamiliar es más frecuente en zonas rurales y marginales, donde prevalecen creencias machistas profundamente arraigadas. El estudio también subraya la influencia de los medios de comunicación y la televisión, los cuales refuerzan estereotipos al mostrar a la mujer como objeto sexual subordinado al hombre. Estas representaciones contribuyen a consolidar la idea de superioridad masculina, lo que legitima prácticas de desigualdad y violencia dentro de la

familia y la sociedad. En conclusión, los resultados permiten afirmar que existe una relación estrecha entre los factores socioculturales y la perpetuación de la violencia intrafamiliar (Vargas Armas & Walde Soto, 2017).

Azcarruz y Pérez (2021), en su tesis titulada “Feminicidio y violencia de género en Perú, 2021” elaborada en la Universidad Peruana de las Américas, tuvieron como propósito identificar los factores que originan y fortalecen la violencia de género vinculada al feminicidio en el país, su estudio empleó un enfoque jurídico con una metodología de carácter cualitativo, utilizando procesos inductivos, deductivos, analíticos, sintéticos y descriptivos dentro de un diseño no experimental. Los autores concluyen que entre las causas más relevantes del feminicidio se encuentran los celos, la influencia persistente del machismo y la denominada “violencia innata”, factores que según señalan, continúan apareciendo en nuevos hallazgos investigativos, sostienen que el feminicidio constituye un delito que afecta múltiples bienes jurídicos y que se sustenta en dinámicas estructurales como la discriminación, los estereotipos de género, el androcentrismo, la misoginia, el patriarcado y el sexismo, expresados en conductas de control y dominación hacia las mujeres. El análisis también revela que la violencia de género en la actualidad opera como un sistema que restringe la autonomía femenina y favorece comportamientos dolosos por parte del agresor, quien busca someter la voluntad, sexualidad y decisiones de la víctima. Finalmente, los autores identifican que los casos de feminicidio suelen estar asociados a motivaciones como los celos, la infidelidad o la negativa de la mujer a continuar la relación, ocurriendo la mayoría de estos hechos dentro del domicilio compartido con la pareja.(Azcarruz Lopez, 2022)

Perez-Yari, Loo Valverde y Chanduvi Puicón (2024) realizaron un estudio sobre “Los factores sociodemográficos vinculados a la violencia intrafamiliar contra mujeres en el Perú durante el año 2021” publicado en la Revista de la Facultad de Medicina

Humana. La investigación parte de la premisa de que la violencia intrafamiliar surge dentro de dinámicas de subordinación, donde el agresor ejerce fuerza física, psicológica o simbólica con la finalidad de imponer control sobre la víctima, el objetivo del estudio fue identificar qué características sociodemográficas se relacionan con mayores probabilidades de experimentar violencia psicológica, física o sexual, los autores desarrollaron un análisis observacional, retrospectivo, analítico y transversal, utilizando la base de datos de la ENDES 2021. Se trabajó con información de 21 557 mujeres, y se empleó regresión de Poisson con varianza robusta para estimar razones de prevalencia, procesadas mediante el software SPSS 27, los resultados mostraron que la violencia psicológica presentó una prevalencia del 47,7%, siendo este el tipo de violencia más frecuente. El modelo multivariado permitió identificar que las mujeres con educación primaria (RPa: 1.10) y aquellas residentes en Lima Metropolitana (RPa: 1.14) tenían mayor probabilidad de sufrir violencia psicológica. En el caso de la violencia sexual, el estudio determinó que pertenecer al nivel de riqueza pobre incrementaba sustancialmente el riesgo (RPa: 2.35). Respecto a la violencia física, las mujeres entre 20 y 29 años (RPa: 1.21), las pertenecientes al estrato económico pobre (RPa: 1.49) y las residentes en la región costa (RPa: 0.79) mostraron asociaciones significativas, en síntesis, el estudio concluye que el nivel educativo, la región natural, el nivel de riqueza, la edad y ciertas características vinculadas al estado civil influyen significativamente en la probabilidad de que las mujeres peruanas experimenten distintos tipos de violencia intrafamiliar. Este trabajo aporta evidencia sólida para comprender los patrones sociodemográficos asociados a la violencia en el Perú, y constituye una referencia relevante para investigaciones que busquen modelar o predecir factores de riesgo mediante técnicas estadísticas o computacionales.(Perez-Yari et al., 2024)

Rodas, Del Castillo, Sandoval, Alatrística y Vela-Ruiz (2023) realizaron un estudio orientado a examinar “la situación actual de la violencia contra la mujer en el Perú, así como su evolución en los últimos años y el impacto que genera en las víctimas”. El trabajo consistió en una revisión sistemática de literatura científica con el objetivo de sintetizar los hallazgos más relevantes sobre la problemática y su persistencia en el contexto peruano, para ello, los autores llevaron a cabo una búsqueda electrónica en diversas bases de datos académicas, incluyendo PubMed, SciELO y Google Académico, utilizando términos clave como violencia contra la mujer, impacto, autoestima y Perú. Inicialmente se identificaron 99 artículos, de los cuales 69 fueron descartados por no cumplir los criterios de inclusión establecidos, quedando 30 estudios seleccionados para el análisis final, los resultados de esta revisión muestran que la violencia contra la mujer sigue siendo una problemática ampliamente extendida en el país, manifestándose a través de diversas formas como agresiones físicas, violencia psicológica y verbal. Los autores subrayan que estas experiencias no solo afectan la integridad física y emocional de las víctimas, sino que también tienen consecuencias profundas en dimensiones como la autoestima, las relaciones sociales y el bienestar psicológico general, la revisión concluye remarcando que la violencia hacia las mujeres continúa presentándose a lo largo de distintas etapas de la vida y bajo múltiples modalidades, lo que evidencia la urgencia de abordar este fenómeno desde perspectivas integrales. Este estudio constituye un aporte importante para investigaciones orientadas a identificar factores asociados o a predecir la violencia de género, ya que ofrece una síntesis actualizada del estado del problema en el Perú.(Roda G. et al., 2022)

En la tesis titulada “Un modelo de análisis multivariante para la medición de la pobreza en el Perú”, realizado por Caycho(2023) desarrolló un modelo predictivo orientado a determinar la condición de pobreza de los hogares peruanos utilizando

información de la Encuesta Nacional de Hogares 2018 (ENAH0-2018) del INEI. La investigación, presentada en la Universidad Nacional Mayor de San Marcos, tuvo como propósito construir un modelo de análisis multivariante sustentado en el enfoque de pobreza multidimensional, aplicando la técnica de árboles de decisión para clasificar los hogares como “pobres” o “no pobres”. El modelo se estructuró a partir de 21 variables independientes, agrupadas en dimensiones como: características de los miembros del hogar, servicios y condiciones de la vivienda, nivel educativo, tipo de combustible usado para cocinar, activos del hogar, y variables de ingresos y gastos. La variable dependiente fue la condición de pobreza (pobreza_2). Los resultados evidenciaron que tres variables fueron determinantes en la clasificación: el gasto monetario per cápita mensual, el número de integrantes del hogar, y la existencia de piso de tierra en la vivienda. Estas variables resultaron ser los principales factores predictivos para distinguir entre hogares pobres y no pobres, aportando evidencia metodológica útil para el seguimiento del Objetivo de Desarrollo Sostenible 1: Fin de la pobreza.(Roda G. et al., 2022)

2.3.3. Antecedentes locales

En la tesis titulada “tratamiento de las noticias sobre violencia de género en los diarios el sol, el matutino y diario del cusco – 2019”, elaborada por Ccahuantico-Quispe(2022), se analizó la manera en que los principales diarios de la ciudad del Cusco abordaron informativamente los casos de violencia de género durante el año 2019. El estudio tuvo como propósito describir y evaluar el tratamiento periodístico aplicado a estas noticias, considerando tanto la perspectiva de los lectores como el contenido producido por los medios. La investigación adoptó un diseño no experimental y retrospectivo, con un enfoque mixto, combinando métodos cualitativos y cuantitativos. En el componente cuantitativo participaron 280 lectores, seleccionados de los tres diarios,

a quienes se aplicó una encuesta para conocer sus percepciones respecto al tratamiento informativo. En la fase cualitativa se examinaron un total de 17 noticias: seis provenientes de El Sol, seis de El Matutino y cinco del Diario del Cusco, consideradas representativas por su relevancia local. Los resultados permitieron concluir que los tres diarios analizados abordan la violencia de género con bajo nivel de espectacularización y sensacionalismo, aunque los lectores consideran que el tratamiento periodístico carece de criterios éticos adecuados. Asimismo, los propios periodistas reconocen la ausencia de protocolos claros para la cobertura de hechos relacionados con la violencia contra la mujer, lo que evidencia limitaciones en la responsabilidad comunicativa y en la calidad del contenido informativo. Este antecedente resulta pertinente para investigaciones relacionadas con el análisis mediático y el abordaje social de la violencia de género, pues muestra cómo los medios regionales influyen en la percepción pública a partir de su tratamiento informativo.

La tesis desarrollada por Br. María Alejandra Pilares Aguirre (2024) titulada "Manifestación de la violencia de género desde la perspectiva y experiencias de los estudiantes de 3ero de secundaria de la i.e. "Tomasa Ttito Condemayta", distrito de acomayo, cusco 2024" analizó cómo se manifiesta la violencia de género entre los estudiantes de tercer grado de secundaria de la I.E. "Tomasa Ttito Condemayta", ubicada en el distrito de Acomayo, Cusco. A través de un enfoque cualitativo, sustentado en un diseño fenomenológico y el método etnográfico, la autora recopiló información mediante entrevistas, grupos focales y observación participante para comprender las experiencias y percepciones del estudiantado. Los resultados evidencian que la violencia de género se presenta en diversas formas —simbólicas, verbales, psicológicas y físicas— estrechamente asociadas a prácticas culturales y roles tradicionales que refuerzan desigualdades, particularmente hacia las mujeres. Se identificaron experiencias de

exclusión, control y agresiones que son normalizadas dentro del entorno escolar. Estas dinámicas afectan el bienestar emocional y el desempeño académico de los estudiantes, generando emociones como miedo, tristeza, inseguridad y baja autoestima. La investigación subraya la necesidad de implementar una educación con enfoque de género e interculturalidad, con el fin de prevenir la violencia y promover relaciones más equitativas, especialmente en contextos rurales donde estas prácticas suelen perpetuarse.(Roda G. et al., 2022)

La tesis elaborada por Ramos Perca, Synthia Shanery (2022) titulada “violencia contra la mujer y la aplicación de la ley 30364 en el distrito fiscal de cusco 2022” constituye un aporte significativo para la comprensión de la violencia contra la mujer desde la perspectiva legal e institucional. Su investigación tuvo como propósito analizar los casos de violencia atendidos en el Distrito Fiscal del Cusco y evaluar el nivel de aplicación de la Ley N.º 30364, norma destinada a prevenir, sancionar y erradicar la violencia contra las mujeres. Bajo un enfoque cualitativo, de tipo básico y con un diseño fenomenológico, la autora recopiló información mediante entrevistas dirigidas a 10 especialistas, entre fiscales y abogados. Esta aproximación permitió examinar, desde la experiencia profesional de los operadores de justicia, las limitaciones existentes en la implementación efectiva de la ley. Los resultados evidencian que, a pesar de la importancia de la Ley N.º 30364, su aplicación en el ámbito fiscal aún presenta serias deficiencias, especialmente en lo referido a la emisión y ejecución efectiva de medidas de protección para las víctimas. Los especialistas entrevistados señalaron que la respuesta institucional es insuficiente debido a factores como la falta de capacitación continua del personal, la ausencia de mecanismos adecuados para la atención oportuna y la dificultad de las víctimas para asumir su condición y participar activamente en las diligencias correspondientes. En ese sentido, el estudio argumenta que la persistencia de la violencia

contra la mujer en Cusco está vinculada no solo a factores sociales, sino también a fallas estructurales en el sistema de justicia, lo que limita el cumplimiento real del objetivo de la Ley 30364. Por ello, se concluye que es indispensable fortalecer la formación del personal jurisdiccional y mejorar la operatividad de las medidas de protección para garantizar la integridad física y psicológica de las mujeres afectadas.(Ramos Perca, 2024)

En su tesis titulada “Representaciones de violencia basada en género de mujeres víctimas en la provincia del Cusco”, Molina Cárdenas, Mayra Andrea (2023) desarrolló un estudio cualitativo orientado a comprender cómo las mujeres agredidas interpretan y significan la violencia que han vivido. Para ello, trabajó con 12 participantes, mujeres entre 20 y 45 años, todas con al menos dos denuncias previas por violencia de género. La información fue recopilada mediante entrevistas semiestructuradas, permitiendo acceder a sus percepciones y experiencias personales. Los hallazgos identifican siete representaciones de violencia, organizadas en tres dimensiones clave. La primera, permisividad y continuidad de la violencia, revela patrones familiares que reproducen agresiones, la dependencia emocional hacia la pareja y la normalización del maltrato. La segunda dimensión, significaciones de dolor y soledad, muestra cómo las víctimas experimentan sufrimiento emocional profundo y sentimientos de aislamiento, reflejados en expresiones como que “las palabras hieren más” y la percepción de que “solo con él” pueden sostener su vida afectiva o económica. La tercera dimensión, reconocimiento y visibilización de la violencia, evidencia momentos en los que las mujeres comienzan a identificar la agresión, apoyarse en su entorno familiar y emprender acciones para enfrentar o denunciar los hechos violentos. Este antecedente aporta una comprensión profunda del significado subjetivo que las mujeres atribuyen a la violencia basada en género, resaltando la importancia de considerar sus representaciones para el diseño de intervenciones más efectivas.(Molina Cárdenas, 2021)

Calero Qquenaya (2019) desarrolló una investigación titulada “La violencia familiar y su relación con el feminicidio en los distritos de San Jerónimo y Saylla del departamento de Cusco, 2018–2019”, cuyo objetivo fue analizar la relación existente entre la violencia familiar y la ocurrencia de feminicidios en dichos distritos durante el periodo señalado. La hipótesis central planteó que la violencia familiar constituye el principal factor causal de los casos de feminicidio registrados en estas jurisdicciones. El estudio adoptó un enfoque cuantitativo de carácter jurídico, apoyado en el análisis estadístico de información secundaria proveniente de fuentes oficiales, como el Ministerio Público y el Poder Judicial. A través del procesamiento de estos datos, se buscó establecer la magnitud y significancia de la relación entre ambas variables. Los resultados evidenciaron una correlación estadísticamente significativa entre la violencia familiar y los feminicidios ocurridos en San Jerónimo y Saylla, identificándose que la mayoría de estos crímenes fueron precedidos por episodios reiterados de violencia intrafamiliar. En este sentido, la investigación concluyó que el feminicidio representa la etapa final de una cadena de agresiones previas que no fueron atendidas de manera oportuna ni eficaz por el Estado. Asimismo, el autor resaltó la responsabilidad de los órganos jurisdiccionales y de las autoridades locales en la implementación efectiva de medidas de protección para las víctimas, como estrategia fundamental para prevenir desenlaces fatales. Finalmente, se recomendó fortalecer la capacitación y sensibilización de los funcionarios públicos y operadores de justicia en materia de prevención, atención y sanción de la violencia familiar, con el fin de reducir la incidencia de feminicidios. (Calero Qquenaya, 2022)

2.4. HIPOTESIS

a) Hipótesis general.

- El modelo de árboles de decisión presenta un nivel significativo de precisión en la predicción de violencia de género según los factores socioeconómicos, demográficos y culturales obtenidos de la Encuesta Demográfica y de Salud Familiar (ENDES 2024).

b) Hipótesis específicas

- Los principales factores asociados a la violencia de género son identificables mediante el modelo de árboles de decisión aplicado a los datos de la ENDES 2024.
- La tasa de error del modelo de árboles de decisión es menor al 30% en la clasificación de casos de violencia de género en mujeres en edad reproductiva.
- El modelo basado en árboles de decisión tiene mayor eficacia predictiva en la violencia psicológica que en la física o sexual cuando se aplica a los datos de la ENDES 2024.

2.5. Identificación de variables e indicadores

Variables independientes: factores asociados

- Edad
- Estado civil
- Nivel educativo
- Tiempo de convivencia
- Número de hijos
- Nivel socioeconómico
- Dependencia económica

- Autoestima
- Abuso infantil
- Traumas
- Alcoholismo
- Mariguana
- Otras drogas
- Exceso de hijos
- Madres adolescentes
- abandono de hijos
- denuncias

Variable dependiente: violencia de género:

- Física
- Psicológica
- Sexual

Indicadores:

- Socioeconómica: grado de instrucción, ingreso familiar, dependencia económica
- Psicológico: autoestima, abuso infantil, traumas.
- Consumo de drogas: alcoholismo, mariguana, otras drogas.
- Planificación familiar: exceso de hijos, madres adolescentes, total abandono de hijos.
- Participación del estado: policía nacional, DEMUNA, ministerio de la mujer, CEM.

2.6. Operacionalización de variables

Tabla 1. Operacionalización de variables

Variables	Definición Conceptual	Definición Operacional	Dimensiones	Indicadores
X: Factores asociados a la violencia de género	Es un conjunto de factores en los que la víctima es abusada física, Psicológica y sexual	La violencia de la mujer está determinada mediante el nivel educativo, la economía del hogar, alimentación sana, planificación familiar, agresión en la niñez, violencia Psicológica, física, sexual.	socioeconómica	-Grado de instrucción -Ingreso familiar -Dependencia económica
			Psicológico	Autoestima Abuso infantil Traumas
			Consumo de drogas	Alcoholismo Mariguana Otras drogas
			Planificación familiar	Exceso de hijos Madres adolescentes Total abandono de hijos
			Participación del estado	Policía nacional DEMUNA Ministerio de la mujer CEM (centro de emergencia mujer)

Variables	Definición Conceptual	Definición Operacional	Dimensiones	Indicadores
Y: Violencia de genero	Es un conjunto de factores en los que la víctima es abusada física, Psicológica y sexual	La violencia de la mujer está determinada mediante el nivel educativo, la economía del hogar, alimentación sana, planificación familiar, agresión en la niñez, violencia Psicológica, física, sexual.	física	Golpes Intento de homicidio Cualquier tipo de agresión física ya sea cuerpo a cuerpo o con algún objeto
			Psicológica	Insultos Gritos Amenazas de cualquier tipo que deje en claro que hubo un intento de agresión
			sexual	Violaciones Prostitución forzada Total, falta de pudor por parte de la pareja hacia la mujer o los hijos

CAPITULO III: METODOLOGIA

3.1. Ámbito de estudio: localización política y geográfica

La encuesta fue desarrollada en toda la república del Perú que se encuentra situado al medio de América del Sur, frente al océano pacifico, los datos fueron extraídos del Instituto Nacional de Estadística e Informática (INEI) ubicado en la página web: <http://inei.inei.gob.pe/microdatos/>

3.2. Tipo y nivel de investigación.

La presente investigación se caracteriza por poseer propósitos prácticos e inmediatos claramente definidos, ya que busca generar conocimiento orientado a la acción con el fin de intervenir, transformar y producir cambios en un sector específico de la realidad estudiada, con esto se enmarca dentro de la investigación aplicada, dado que utiliza los aportes conceptuales y teóricos provenientes de la investigación básica y sustantiva para dar respuesta a una problemática concreta.

De acuerdo con su alcance temporal, el estudio es de tipo transversal, debido a que la recolección de datos se realizó en un único momento, así como el enfoque metodológico es cuantitativo, puesto que se emplearon procedimientos estadísticos para el análisis de la información y el nivel de investigación es correlacional, ya que se orienta a identificar y analizar la relación existente entre las variables de estudio, finalmente el método utilizado fue el analítico–deductivo, el cual permitió partir de principios generales para interpretar y explicar los resultados obtenidos en el contexto específico de la investigación.

3.3. Unidad de análisis

La unidad de análisis de la presente investigación está constituida por las mujeres que han sido víctimas de algún tipo de abuso, entendido este como cualquier forma de violencia ejercida en su contra, ya sea de carácter físico, psicológico, sexual, económico, simbólico, en el contexto familiar, social o institucional. Estas mujeres representan el eje central del estudio, ya que sobre ellas se recaba la información necesaria para comprender la magnitud, características y factores asociados a la problemática investigada.

El análisis se centra en las experiencias, condiciones y situaciones vividas por las mujeres que han sufrido abuso, considerando que dichas vivencias permiten identificar patrones de violencia, niveles de afectación y respuestas institucionales frente a estos hechos, de este modo la unidad de análisis resulta pertinente para el cumplimiento de los objetivos de la investigación, en tanto posibilita examinar de manera directa el fenómeno de estudio desde la perspectiva de las personas directamente afectadas.

3.4. Población de estudio

Está constituida por todas las mujeres que participaron en la Encuesta Demográfica y de Salud Familia - ENDES 2024.

3.5. Tamaño de muestra

El tamaño de la muestra de la presente investigación estuvo conformado por mujeres encuestadas, seleccionadas bajo el criterio de proporción 1 a 1, es decir, se consideró una cantidad equivalente de mujeres víctimas de violencia y mujeres que no han experimentado situaciones de violencia, esta distribución equilibrada permitió realizar comparaciones objetivas entre ambos grupos, garantizando una adecuada representatividad y control de posibles sesgos en el análisis de los resultados.

La inclusión de ambos grupos responde a la necesidad de identificar diferencias y asociaciones entre las variables de estudio, así como de evaluar los factores que influyen en la ocurrencia de la violencia. La proporción 1:1 facilita el análisis estadístico y fortalece la validez interna del estudio, al permitir contrastar de manera equitativa las características y condiciones presentes en mujeres que han sufrido violencia y aquellas que no.

La determinación del tamaño muestral consideró los objetivos de la investigación, el enfoque cuantitativo y el diseño correlacional, asegurando que el número de participantes sea suficiente para la aplicación de las pruebas estadísticas y la obtención de resultados confiables y significativos, el tamaño de la muestra se estimó utilizando la fórmula para poblaciones finitas, considerando un nivel de confianza del 95% ($Z=1,96$), un margen de error del 5% ($e=0,05$) y una proporción esperada de éxito del 50% ($p=0,50$), obteniéndose una muestra de 386 mujeres.

Los parámetros utilizados fueron los siguientes:

- Población (N): 18 836 mujeres.
- Nivel de confianza: 95% ($CON=1.96Z=1,96CON=1.96$).
- Margen de error: 5% ($y=0.05e=0,05y=0.05$).
- Probabilidad de éxito: 50% ($pag=0.50p=0,50pag=0.50$).
- Tamaño de muestra calculado: 386 mujeres.

3.6. Técnicas de selección de muestra

Inicialmente, se planteó emplear un muestreo probabilístico aleatorio simple, mediante la generación de números aleatorios, con el propósito de garantizar que todas las mujeres

incluidas en la base de datos de la Encuesta Demográfica y de Salud Familiar (ENDES) tuvieran la misma probabilidad de ser seleccionadas.

No obstante, para el desarrollo de la presente investigación se decidió trabajar con la totalidad de la población, conformada por las 18 836 mujeres registradas en la base de datos de la ENDES que cumplieran con los criterios de inclusión, esta decisión se fundamentó en que los algoritmos de árboles de decisión (como CART, C4.5 o C5.0) mejoran su desempeño cuando se entrenan con una mayor cantidad de observaciones, ya que disponen de más información para identificar patrones, relaciones y reglas de clasificación entre las variables, el uso de toda la población reduce la pérdida de información asociada al muestreo, incrementa la estabilidad de las divisiones del árbol y favorece una mayor precisión en las etapas de entrenamiento, validación y poda, disminuyendo la probabilidad de obtener resultados influenciados por la variabilidad muestral, al contar con una base de datos completa y accesible, el análisis se realizó mediante un censo de la población de estudio, utilizando las 18 836 observaciones disponibles, lo que permitió construir modelos predictivos más robustos y obtener resultados con mayor confiabilidad para la identificación de los factores asociados a la violencia contra la mujer.

3.7. Técnicas de recolección de información

Para el desarrollo de la presente investigación, la recolección de información se realizó a partir del uso de fuentes secundarias, específicamente mediante el acceso y análisis de la base de datos del Instituto Nacional de Estadística e Informática (INEI) correspondiente a la Encuesta Demográfica y de Salud Familiar (ENDES) del año 2024. Esta encuesta constituye una fuente oficial y confiable de información estadística a nivel

nacional, ampliamente utilizada en estudios académicos y científicos relacionados con salud, demografía y violencia contra la mujer.

La ENDES 2024 proporciona datos estructurados y validados que permiten analizar diversas variables sociodemográficas, económicas y sociales vinculadas al fenómeno de estudio. El uso de esta base de datos posibilita obtener información representativa de la población femenina, garantizando la consistencia y calidad de los datos, así como el respeto a los principios éticos, ya que se trata de información anónima y de acceso público, la técnica de recolección empleada consistió en la extracción, depuración y sistematización de las variables relevantes contenidas en la base de datos, de acuerdo con los objetivos de la investigación. Este procedimiento permitió optimizar el análisis estadístico posterior y asegurar la pertinencia de la información utilizada para la contratación de hipótesis y la interpretación de resultados.

Para el desarrollo de la presente investigación, la recolección de la información se efectuó a partir del uso de fuentes secundarias, específicamente mediante el acceso al sistema de microdatos del Instituto Nacional de Estadística e Informática (INEI), disponible en su plataforma oficial (<http://inei.inei.gob.pe/microdatos/>). En particular, se empleó la base de datos correspondiente a la Encuesta Demográfica y de Salud Familiar (ENDES) del año 2024, por tratarse de una fuente estadística de carácter oficial, actualizada y ampliamente reconocida por su confiabilidad y rigor metodológico.

La ENDES 2024 proporciona información detallada y representativa a nivel nacional sobre diversas características demográficas, sociales y de salud de la población, incluyendo variables relevantes relacionadas con la violencia contra la mujer, condiciones sociodemográficas y dinámicas familiares. El uso de esta base de datos permitió disponer de información amplia y estandarizada, garantizando la validez y consistencia de los datos analizados, así como los datos recolectados se tomaron en forma transversal ya que se

realizaron en el año a la entrevistada y no se le vuelve a entrevistar, con la finalidad de observar la realidad en el momento.

El proceso de recolección consistió en la descarga, selección, depuración y sistematización de las variables de interés, de acuerdo con los objetivos y las hipótesis planteadas en la investigación. Dicho procedimiento facilitó la organización adecuada de la información y su posterior análisis estadístico, optimizando el uso de los datos y asegurando la pertinencia de la información utilizada.

Cabe precisar que, al tratarse de una base de datos de acceso público y de carácter anónimo, el uso de la información respetó los principios éticos de confidencialidad y protección de datos personales, no requiriéndose la identificación directa de las personas encuestadas.

3.8. Técnicas de análisis e interpretación de información

El proceso de análisis de la información se desarrolló en varias etapas, combinando técnicas estadísticas descriptivas, inferenciales y métodos de minería de datos, con el propósito de identificar y evaluar los factores asociados a la violencia contra la mujer.

En una primera fase, se realizó un análisis descriptivo y correlacional de las variables independientes y la variable dependiente (presencia o ausencia de violencia), con el fin de caracterizar a la población de estudio e identificar posibles asociaciones preliminares, para esto se empleó la prueba estadística de chi cuadrado, la cual permitió evaluar la existencia de relaciones significativas entre los factores sociodemográficos y la ocurrencia de violencia.

Posteriormente, se procedió a la construcción y entrenamiento de un modelo de árboles de decisión, utilizando un conjunto de datos de entrenamiento obtenido de la base de datos de la ENDES 2024. Esta técnica permitió modelar relaciones complejas entre

múltiples variables y generar reglas de decisión interpretables para la predicción de la violencia.

El desempeño del modelo fue evaluado mediante diversos indicadores de rendimiento, tales como la exactitud (accuracy), la sensibilidad, la especificidad y el área bajo la curva ROC (AUC-ROC), lo que permitió medir de manera integral la capacidad predictiva del modelo y su utilidad para la identificación de casos de violencia.

Las variables identificadas como factores significativos dentro del modelo de árbol de decisión fueron analizadas de forma complementaria mediante tablas de contingencia, con el objetivo de describir de manera clara y precisa la relación entre dichas variables y la presencia de violencia. Asimismo, se realizó un análisis detallado de las ramas y nodos del árbol, identificando aquellos que aparecen en las primeras divisiones, dado que estos representan los factores con mayor relevancia predictiva dentro del modelo.

Se entrenaron y validaron modelos de árboles de decisión independientes para cada tipo de violencia analizada: violencia física, violencia sexual y violencia psicológica. Esta estrategia permitió comparar el comportamiento del modelo en cada categoría y evaluar diferencias en su capacidad de predicción.

Se llevó a cabo una comparación de las métricas de desempeño obtenidas en los tres modelos, particularmente en términos de exactitud y AUC-ROC, con la finalidad de identificar el tipo de violencia en el cual el modelo presenta un mejor rendimiento predictivo, contribuyendo así a una comprensión más precisa de la aplicabilidad de los árboles de decisión en el estudio de la violencia contra la mujer.

3.9. Técnicas para demostrar la verdad o falsedad de la hipótesis planteada

Para la contrastación de la hipótesis planteada en la presente investigación, se emplearon pruebas estadísticas inferenciales orientadas a determinar la existencia de

asociaciones significativas entre las variables de estudio y a evaluar el desempeño de los modelos de clasificación utilizados.

En particular, se aplicó la prueba de Mantel–Haenszel, la cual resulta adecuada para el análisis de tablas de contingencia estratificadas, permitiendo evaluar la asociación entre variables categóricas mientras se controla el efecto de variables de confusión. Esta prueba fue utilizada sobre las matrices de conjunción (tablas de clasificación o matrices de confusión) generadas por los modelos de árboles de decisión, los cuales constituyen métodos de clasificación categórica.

La prueba de Mantel–Haenszel permitió contrastar si la asociación observada entre la variable dependiente (presencia o ausencia de violencia) y las variables predictoras identificadas por el modelo de árboles de clasificación es estadísticamente significativa, estableciendo así evidencia empírica para aceptar o rechazar la hipótesis de investigación. Asimismo, esta técnica facilitó la evaluación de la consistencia de la asociación a través de distintos estratos definidos por variables relevantes, fortaleciendo la validez de los resultados obtenidos.

De este modo, el uso de la prueba de Mantel–Haenszel contribuyó a demostrar, con un nivel de confianza estadística adecuado, la veracidad o falsedad de la hipótesis planteada, complementando el análisis predictivo realizado mediante los modelos de árboles de decisión y reforzando la interpretación de los hallazgos desde un enfoque cuantitativo y correlacional.

Tabla 2 Matriz de Construcción

variable	Dimensiones/componentes/sub variables	Indicadores	Preguntas/ítems/reactivos	Alternativas
Factores asociados a la violencia de genero	socioeconómica	Grado de instrucción	Cuál es el grado de instrucción de la mujer que sufrió el maltrato	a) primaria b) secundaria c) sup. tecnico d) sup. universitario
		Ingreso familiar	Cuál es el ingreso familiar	Bajo Medio Alto
		Dependencia económica	En que nivel depende la familia del trabajo o ingreso económico del varón	Bajo Medio Alto
	Psicológico	Autoestima	Cuál es el nivel de autoestima de la mujer que sufrió maltrato	a) bajo b) medio c) alto
		Abuso infantil	La mujer que sufrió maltrato en su niñez tuvo algún tipo de abuso por parte de sus progenitores	Si no
		traumas	La mujer que sufrió maltrato tuvo algún trauma en la infancia	si no
	Consumo de drogas	Alcoholismo	En la familia existe el consumo de alcohol	si no
		Mariguana	En la familia existe el consumo de mariguana	si no
		Otras drogas	En la familia existen el consumo de otras sustancias alucinógenas	si no
	Planificación familiar	Excesos de hijos	Cuántos hijos vivos tiene la familia	uno dos más de dos
		Madres adolescentes	A que edad la madre tuvo su primer hijo	antes de los 18 de 18 a 30 de 30 a mas
		Total abandono de los hijos	En el hogar salen a trabajar padre y madre	si no

CAPITULO IV: RESULTADOS

Tabla 3 *Fiabilidad del cuestionario de presencia de violencia hacia la mujer*

	Media	DE	Alfa de Cronbach
Cuestionario de violencia	0,0677	0,1140	0,750
1. ¿Fue maltratada físicamente durante el embarazo por su actual pareja?	0,0258	0,1584	0,743
2. ¿Su pareja alguna vez la empujó, sacudió o le lanzó algo?	0,0723	0,2589	0,724
3. ¿Ha sufrido violencia severa (estrangulamiento, quemaduras, amenazas con armas)?	0,0844	0,2779	0,724
4. ¿Su pareja alguna vez intentó estrangularla o quemarla?	0,0076	0,0868	0,746
5. ¿Su pareja la amenazó con un arma (cuchillo, pistola, etc.)?	0,0036	0,0600	0,749
6. ¿Su pareja la amenazó con hacerle daño a usted o a alguien cercano?	0,0334	0,1797	0,729
7. ¿Su pareja se pone celoso si habla con otro hombre?	0,2935	0,4554	0,735
8. ¿Su pareja la acusa frecuentemente de infidelidad?	0,1275	0,3336	0,709
9. ¿Su pareja la ha humillado delante de otros?	0,0810	0,2729	0,724
10. ¿Su pareja le impide ver o recibir a amistades?	0,0959	0,2945	0,712
11. ¿Su pareja limita sus contactos con su familia?	0,0749	0,2633	0,721
12. ¿Quién decide cómo se gasta el dinero en el hogar?	0,0115	0,1067	0,756
13. ¿Quién tiene la última palabra sobre su cuidado de salud?	0,0755	0,2642	0,776
14. ¿Su pareja la ha forzado a tener relaciones sexuales contra su voluntad?	0,0172	0,1298	0,739
15. ¿Su pareja la ha obligado a realizar actos sexuales que no aprueba?	0,0118	0,1082	0,743

Nota: elaboración propia

La Tabla 3 presenta la fiabilidad del cuestionario utilizado para evaluar la violencia ejercida contra la mujer en el marco de la encuesta ENDES Perú, mediante el coeficiente Alfa de Cronbach. El instrumento en su conjunto alcanzó un valor de 0,750, lo que evidencia una consistencia interna aceptable, de acuerdo con los criterios psicométricos establecidos ($\geq 0,70$). Este resultado indica que los ítems del cuestionario miden de manera coherente las diversas manifestaciones de violencia que afectan a la mujer.

Los valores del Alfa de Cronbach oscilan entre 0,709 y 0,776, lo cual sugiere que cada uno contribuye de forma homogénea a la medición del constructo, permitiendo captar adecuadamente las distintas dimensiones de la violencia que puede experimentar la mujer. Asimismo, se reportan la media y la desviación estándar (DE) de cada ítem. Se observa que las formas más severas de violencia, como las amenazas con armas (media = 0,0036), presentan una baja frecuencia en la experiencia de la mujer; mientras que otras manifestaciones, como los celos (media = 0,2935), son relativamente más frecuentes, lo que pone en evidencia la diversidad de situaciones de violencia que pueden afectar a la mujer en su vida cotidiana.

Figura 1. Distribución del nivel de violencia percibida por las mujeres

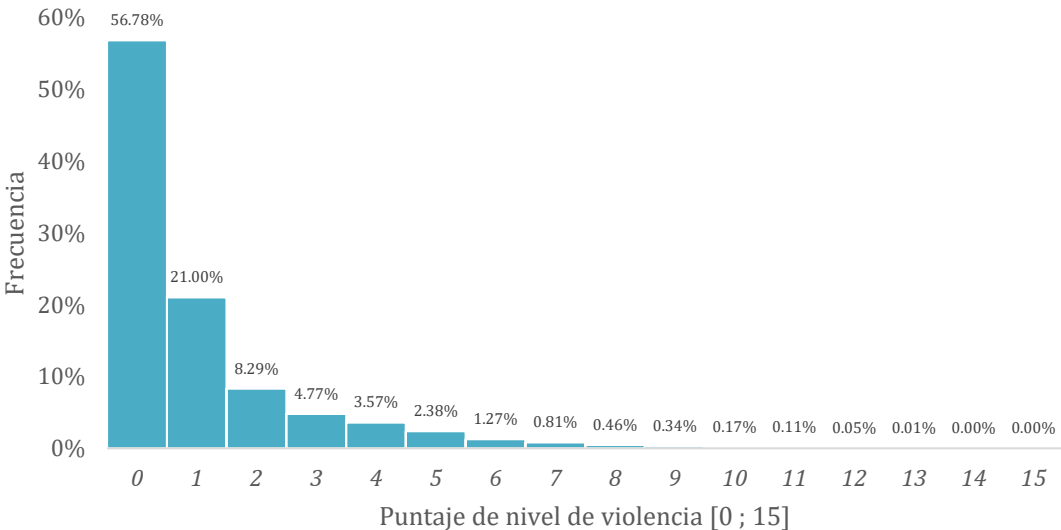


Tabla 4. Prueba de normalidad Kolmogórov-Smirnov

	n	Media	Desv.	Estadístico de prueba	Sig.
Violencia	18836	1,0158	1,7039	0,292	0,0000

La Figura 1 y la Tabla 4 presentan el análisis de la distribución del nivel de violencia percibida en las mujeres de la muestra. En la Figura 1, representada mediante un histograma, se observa que más de la mitad de las mujeres (56,78%) no reportaron haber experimentado ningún tipo de violencia, al obtener un puntaje igual a cero en los 15 ítems

del cuestionario. A medida que el puntaje de violencia aumenta, la frecuencia de mujeres que reportan estas experiencias disminuye considerablemente; por ejemplo, alrededor del 21% indicó un puntaje de 1, mientras que menos del 1% de las mujeres presentó puntajes superiores a 7, siendo prácticamente inexistentes los valores más altos (13, 14 y 15).

Esta distribución evidencia que la experiencia de violencia en la mujer se concentra mayormente en niveles nulos o bajos, mostrando una marcada asimetría con sesgo hacia la derecha. Este comportamiento se ve respaldado por los resultados de la prueba de normalidad de Kolmogórov-Smirnov (Tabla 2), la cual reporta un estadístico de 0,292 con un nivel de significancia de $p = 0,000$, permitiendo rechazar la hipótesis nula de normalidad, estos hallazgos indican que, si bien la mayoría de las mujeres no reporta haber vivido situaciones de violencia o las experimenta en baja intensidad, existe un grupo reducido de mujeres que sí presenta niveles más elevados. Asimismo, la naturaleza no normal de la distribución sugiere la necesidad de emplear métodos estadísticos no paramétricos para un análisis adecuado de la violencia que afecta a la mujer.

Los datos recolectados del ENDES 2024 presenta un total de 18836 registros validos considerando mujeres que pudieron realizar la encuesta de violencia de manera completa.

4.1 Descripción de la población de mujeres peruanas analizadas.

Tabla 5. Descripción de las características demográficas de la pareja

Variables	Categorías	Frecuencia	Porcentaje
Edad mujer 33,03 +/- 7,67	15 a 20	863	4,6%
	21 a 30	6601	35,0%
	31 a 40	7830	41,6%
	41 a 49	3542	18,8%
Edad pareja 36,9 +/- 8,35	15 a 20	256	1,4%
	21 a 30	4227	22,4%
	31 a 40	8114	43,1%
	41 a 50	5264	27,9%
	51 a 60	873	4,6%
	61 a 78	102	0,5%
Región	Sierra	6334	33,6%
	Resto costa	5386	28,6%
	Selva	5016	26,6%
	Lima metropolitana	2100	11,1%
Provincia	Lima	9024	47,9%
	Piura	1699	9,0%
	La libertad	1624	8,6%
	Arequipa	1182	6,3%
	Cajamarca	1036	5,5%
	Junín	992	5,3%
	Cusco	773	4,1%
	Lambayeque	670	3,6%
	Puno	636	3,4%
	Áncash	357	1,9%
	Loreto	285	1,5%
	San Martín	197	1,0%
	Ica	186	1,0%
	Otras	175	0,9%

Nota: promedio +/- desviación

La Tabla 5 presenta las características demográficas de las mujeres peruanas incluidas en el estudio, conformado por la población de 18 836 participantes. En relación con la edad, se observa que la mayor proporción de mujeres se concentra en los grupos de 21 a 30 años (35,0 %) y de 31 a 40 años (41,6 %), con una edad promedio de 33,03 años y una desviación estándar de 7,67.

Respecto a las parejas de las mujeres, se identifica una distribución etaria similar, aunque con una ligera tendencia a edades mayores, evidenciada por una media de 36,9 años ($DE = 8,35$), siendo el grupo predominante el de 31 a 40 años (43,1 %).

En cuanto a la procedencia geográfica de la mujer, cerca de un tercio pertenece a la región Sierra (33,6 %), seguida por la costa fuera de Lima (28,6 %) y la Selva (26,6 %), mientras que una menor proporción reside en Lima Metropolitana (11,1 %). A nivel provincial, casi la mitad de las mujeres (47,9 %) se ubica en la provincia de Lima, destacando también la participación de mujeres provenientes de regiones como Piura (9,0 %), La Libertad (8,6 %) y Arequipa (6,3 %), con los resultados evidencian que la muestra de mujeres es heterogénea tanto en términos de edad como de distribución territorial, con predominio de mujeres en edad adulta reproductiva y una representación que abarca diversas regiones del Perú.

Tabla 6 Descripción de las características reproductivas de la pareja

Variables	Categorías	Frecuencia	Porcentaje
Estado civil	Viviendo juntos	12758	67,7%
	Casado	4188	22,2%
	No viven juntos	1890	10,0%
Nivel de educación	Sin educación	1480	7,9%
	Inicial	1955	10,4%
	Primaria completa	1196	6,3%
	Primaria incompleta	3998	21,2%
	Secundaria completa	8314	44,1%
	Superior	1893	10,0%
Método anticonceptivo	Inyección	4552	24,2%
	No usando	3854	20,5%
	Ritmo	3013	16,0%
	Condón	2171	11,5%
	Implantes	1977	10,5%
	Esterilización femenina	1957	10,4%
	Píldora	983	5,2%
	Diu	172	0,9%
	Otro	113	0,6%
	Esterilización masculina	44	0,2%
Embarazada	No	18317	97,2%
	Sí	519	2,8%
Edad de primera relación 17,51 +/- 3,17	10 a 15	4934	26,2%
	16 a 20	11470	60,9%
	21 a 30	2332	12,4%
	31 a 44	100	0,5%
Total, de hijos nacidos 2,44 +/- 1,46	0	393	2,1%
	1	4733	25,1%
	2	6316	33,5%
	3	4056	21,5%
	4	1766	9,4%
	5	809	4,3%
	6	403	2,1%
	7 o mas	360	1,9%

Nota: promedio +/- desviación

La Tabla 6 presenta las características reproductivas de las mujeres peruanas incluidas en la encuesta ENDES 2024, con una muestra de 18 836 participantes. En relación con el estado civil, se observa que la mayoría de las mujeres mantiene una unión consensual con su pareja (67,7 %), seguida por aquellas que se encuentran casadas legalmente (22,2 %), mientras que un 10 % de las mujeres no convive actualmente con su pareja.

En cuanto al nivel educativo, predomina la educación secundaria completa (44,1 %) en la mujer, seguida por un porcentaje importante con primaria incompleta (21,2 %), y una menor proporción que ha accedido a estudios superiores (10 %), lo que evidencia brechas en el acceso a niveles educativos más altos.

Respecto a la planificación familiar, la mujer recurre principalmente a métodos anticonceptivos como la inyección (24,2 %) y el método del ritmo (16,0 %); sin embargo, una proporción relevante de mujeres (20,5 %) no utiliza ningún método, lo que podría implicar limitaciones en el acceso o uso de servicios de salud reproductiva.

En relación con la condición de embarazo, la gran mayoría de las mujeres no se encontraba gestando al momento de la encuesta (97,2 %). Asimismo, la edad promedio de inicio de la vida sexual en la mujer fue de 17,51 años ($DE = 3,17$), concentrándose principalmente entre los 16 y 20 años (60,9 %), aunque destaca que más de una cuarta parte de las mujeres (26,2 %) inició antes de los 16 años, el número promedio de hijos por mujer fue de 2,44 ($DE = 1,46$), siendo más frecuente que la mujer tenga dos (33,5 %) o tres hijos (21,5 %). En conjunto, estos resultados reflejan patrones reproductivos característicos de mujeres en un contexto de transición demográfica, donde coexisten distintos comportamientos en relación con la fecundidad y la planificación familiar.

Tabla 7. Descripción de las características económicas y sociales de la pareja

Variables	Categorías	Frecuencia	Porcentaje
Tipo ingresos	Efectivo y especie	1200	6,4%
	No le pagan	6237	33,1%
	Solo en efectivo	11114	59,0%
	Solo en especie	285	1,5%
Decisor de gastos de pareja	Entrevistada	y 9730	51,7%
	esposo/compañero		
	Esposo/compañero		
Seguro salud	Entrevistada	3832	20,3%
	No	15703	83,4%
	Sí	3133	16,6%
Frecuencia de consumo	Nunca	3220	17,1%
De alcohol por la pareja	Algunas veces	14647	77,8%
	Mucha frecuencia	969	5,1%
Entrevista interrumpida Por pareja	fue No	18675	99,1%
	Si	161	0,9%
Años del matrimonio 11,98 +/- 7,85	primer 0 a 4	3638	19,3%
	11 a 20	6549	34,8%
	21 a 30	2774	14,7%
	31 a mas	290	1,5%
	5 a 10	5585	29,7%

La Tabla 7 presenta las características económicas y sociales que influyen en la vida de las mujeres peruanas incluidas en la muestra de 18 836 participantes. En relación con los ingresos, se observa que la mayoría de las mujeres recibe remuneración únicamente en efectivo (59,0 %), mientras que una proporción considerable (33,1 %) no percibe ningún tipo de ingreso, lo que podría evidenciar situaciones de dependencia económica o su participación en actividades no remuneradas, como el trabajo doméstico.

Respecto a la toma de decisiones sobre los gastos del hogar, en más de la mitad de los casos (51,7 %) la mujer comparte esta responsabilidad con su pareja; sin embargo, en un porcentaje importante (28,0 %), es el hombre quien asume de manera exclusiva dicho control, lo que puede reflejar desigualdades en la distribución del poder dentro del hogar, en cuanto al acceso a servicios de salud, se identifica que la mayoría de las mujeres no

cuenta con ningún tipo de seguro (83,4 %), en contraste con una minoría (16,6 %) que sí dispone de cobertura, lo que evidencia una importante limitación en el acceso a protección en salud.

En relación con el consumo de alcohol por parte de la pareja, la mayoría de las mujeres reporta que este ocurre de manera ocasional (77,8 %); no obstante, un grupo reducido (5,1 %) señala un consumo frecuente, situación relevante debido a su posible vínculo con contextos de riesgo y violencia hacia la mujer, casi la totalidad de las mujeres (99,1 %) pudo responder la encuesta sin interrupciones por parte de su pareja, lo que sugiere condiciones adecuadas para la recolección de información. Finalmente, la duración promedio de la primera unión de la mujer fue de 11,98 años ($DE = 7,85$), concentrándose principalmente entre los 11 y 20 años (34,8 %), seguida del rango de 5 a 10 años (29,7 %), estos resultados muestran que muchas mujeres se encuentran en contextos socioeconómicos diversos, donde persisten condiciones de vulnerabilidad como la limitada cobertura de salud, la dependencia económica y, en algunos casos, una menor participación en la toma de decisiones dentro del hogar.

4.2. Descripción de violencia sufrida en la población de mujeres peruanas

Tabla 8. Indicadores de la violencia de la mujer.

Dimensiones	Indicador	Categorías	
		Si	No
Violencia física	1. ¿Fue maltratada físicamente durante el embarazo por su actual pareja?	485 (2,6%)	18351 (97,4%)
	2. ¿Su pareja alguna vez la empujó, sacudió o le lanzó algo?	1361 (7,2%)	17475 (92,8%)
	3. ¿Ha sufrido violencia severa (estrangulamiento, quemaduras, amenazas con armas)?	1589 (8,4%)	17247 (91,6%)
	4. ¿Su pareja alguna vez intentó estrangularla o quemarla?	143 (0,8%)	18693 (99,2%)
	5. ¿Su pareja la amenazó con un arma (cuchillo, pistola, etc.)?	68 (0,4%)	18768 (99,6%)
	6. ¿Su pareja la amenazó con hacerle daño a usted o a alguien cercano?	629 (3,3%)	18207 (96,7%)
Violencia psicológica	7. ¿Su pareja se pone celoso si habla con otro hombre?	5528 (29,3%)	13308 (70,7%)
	8. ¿Su pareja la acusa frecuentemente de infidelidad?	2402 (12,8%)	16434 (87,2%)
	9. ¿Su pareja la ha humillado delante de otros?	1526 (8,1%)	17310 (91,9%)
	10. ¿Su pareja le impide ver o recibir a amistades?	1807 (9,6%)	17029 (90,4%)
	11. ¿Su pareja limita sus contactos con su familia?	1411 (7,5%)	17425 (92,5%)
	12. ¿Quién decide cómo se gasta el dinero en el hogar?	217 (1,2%)	18619 (98,8%)
	13. ¿Quién tiene la última palabra sobre su cuidado de salud?	1422 (7,5%)	17414 (92,5%)
Violencia sexual	14. ¿Su pareja la ha forzado a tener relaciones sexuales contra su voluntad?	323 (1,7%)	18513 (98,3%)
	15. ¿Su pareja la ha obligado a realizar actos sexuales que no aprueba?	223 (1,2%)	18613 (98,8%)

La Tabla 8 muestra la prevalencia de las diferentes formas de violencia que afectan a las mujeres peruanas incluidas en la muestra de la ENDES 2024 ($n = 18\,836$), diferenciadas en violencia física, psicológica y sexual. En relación con la violencia física, se observa que una proporción menor de mujeres reporta haber experimentado este tipo de agresión; siendo más frecuentes acciones como empujones, sacudidas o lanzamiento de objetos (7,2 %), así como amenazas de daño (3,3 %). En contraste, las formas más severas, como el

intento de estrangulamiento o las amenazas con armas, afectan a menos del 1 % de las mujeres.

Por otro lado, la violencia psicológica se presenta como la forma más común en la experiencia de la mujer, destacando principalmente los celos por parte de la pareja frente a la interacción con otros hombres (29,3 %). Asimismo, un número importante de mujeres reporta restricciones en sus relaciones sociales (9,6 %) y situaciones de humillación pública (8,1 %), evidenciando dinámicas de control que impactan su bienestar emocional y social.

En cuanto a la violencia sexual, esta se reporta en menor proporción; sin embargo, no deja de ser relevante, ya que un 1,7 % de las mujeres señala haber sido forzada a mantener relaciones sexuales en contra de su voluntad, mientras que un 1,2 % indica haber sido obligada a realizar actos sexuales no deseados, estos hallazgos evidencian que, aunque la violencia física y sexual afectan a un menor número de mujeres en términos proporcionales, la violencia psicológica es la manifestación más extendida en la vida de la mujer, lo que resalta su papel central en las dinámicas de poder, control y dominación dentro de las relaciones de pareja.

Figura 2. Frecuencia de tipo de violencia sufrida por las mujeres

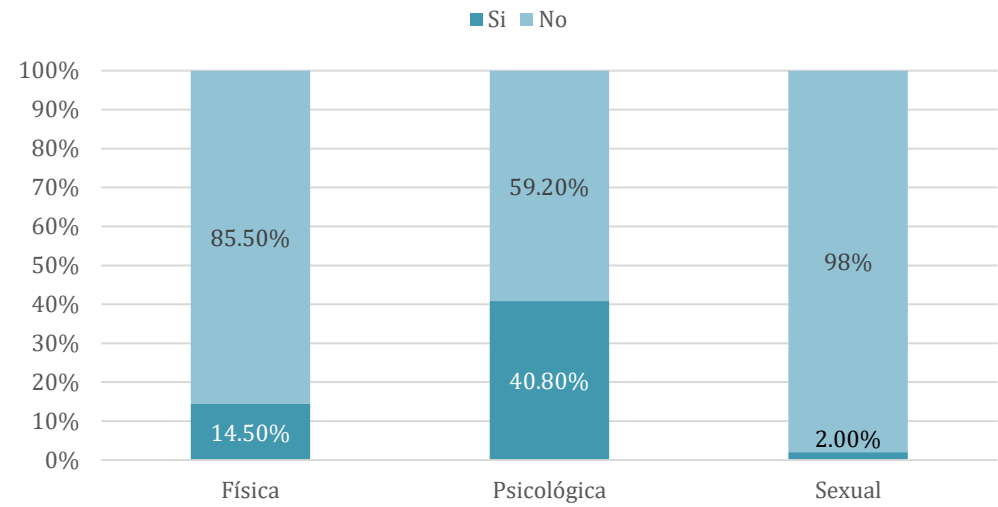


Tabla 9. Dimensiones y variables de la violencia de la mujer.

Dimensiones	Indicador	Categorías	
		Si	No
Dimensiones de la variable violencia	Física	2734 (14,5%)	16102 (85,5%)
	Psicológica	7685 (40,8%)	11151 (59,2%)
	Sexual	385 (2,0%)	18451 (98%)
Variable	Violencia	8140 (43,2%)	10696 (56,8%)

La figura2 y tabla 9 sintetiza la prevalencia de las distintas formas de violencia que afectan a las mujeres peruanas incluidas en la muestra de la ENDES 2024 (n = 18 836). Se evidencia que la violencia psicológica es la más frecuente en la experiencia de la mujer, alcanzando al 40,8 % (7 685 casos). En segundo lugar, se encuentra la violencia física, reportada por el 14,5 % de las mujeres (2 734 casos), mientras que la violencia sexual presenta la menor prevalencia, afectando al 2,0 % (385 casos), al analizar la variable global de violencia general, definida como la presencia de al menos una de estas formas, se observa que 8 140 mujeres (43,2 %) han experimentado algún tipo de violencia por parte de su pareja, la interpretación mencionada pone en evidencia que una proporción considerable de mujeres ha estado expuesta a situaciones de violencia en el ámbito de la relación de pareja. Asimismo, se confirma que la violencia psicológica constituye la

forma más extendida en la vida de la mujer, mientras que la violencia sexual, aunque menos frecuente en términos relativos, sigue representando una problemática relevante. Estos hallazgos resaltan la magnitud del problema y la diversidad de manifestaciones de violencia que afectan a la mujer.

Tabla 10. *Correlaciones entre las dimensiones o tipo de violencia*

		Violencia física		Total	Chi ²	Sig.
		Si	No			
Violencia	Si	2290 (12,2%)	5395 (28,6%)	7685 (40,8%)	2443	0,0001
Psicológica	No	444 (2,4%)	10707 (56,8%)	11151 (59,2%)		
Total		2734 (14,5%)	16102 (85,5%)	18836 (100%)		

		Violencia física		Total	Chi ²	Sig.
		Si	No			
Violencia	Si	300 (1,6%)	85 (0,5%)	385 (2%)	1273	0,0001
Sexual	No	2434 (12,9%)	16017 (85%)	18451 (98%)		
Total		2734 (14,5%)	16102 (85,5%)	18836 (100%)		

		Violencia psicológica		Total	Chi ²	Sig.
		Si	No			
Violencia	Si	363 (1,9%)	22 (0,1%)	385 (2%)	465,5	0,0001
Sexual	No	7322 (38,9%)	11129 (59,1%)	18451 (98%)		
Total		7685 (40,8%)	11151 (59,2%)	18836 (100%)		

La Tabla 10 presenta las asociaciones bivariados entre las dimensiones de violencia que afectan a la mujer como física, psicológica y sexual mediante el uso de tablas de contingencia y la prueba de chi-cuadrado. Los resultados evidencian la existencia de relaciones estadísticamente significativas entre todas las combinaciones analizadas ($p < 0,001$), lo que indica que las distintas formas de violencia tienden a coexistir en la experiencia de la mujer, se observa una alta concurrencia entre la violencia psicológica y la violencia física. Del total de mujeres que reportaron violencia física (14,5 %), una amplia mayoría (83,7 %) también experimentó violencia psicológica, lo que sugiere que estas formas de violencia suelen presentarse de manera conjunta en la vida de la mujer.

Asimismo, la violencia sexual muestra una fuerte asociación con las otras dimensiones. Entre las mujeres que reportaron haber sufrido este tipo de violencia, el 77,9 % también experimentó violencia física y el 94,3 % violencia psicológica, evidenciando que la violencia sexual rara vez ocurre de manera aislada y suele estar acompañada de otras formas de agresión, la significancia estadística observada en todos los casos ($\chi^2 > 465$; $p = 0,0001$) confirma que estas asociaciones no responden al azar, sino que reflejan un patrón consistente en el que las distintas manifestaciones de violencia se interrelacionan en la experiencia de la mujer, en conjunto, se evidencia que la violencia psicológica no solo es la más frecuente, sino también la que se presenta tanto de manera independiente como en combinación con otras formas de violencia. Por su parte, la violencia física en la mujer aparece, en gran medida, vinculada a la presencia previa o simultánea de violencia psicológica, mientras que la violencia sexual, aunque menos frecuente, se manifiesta casi exclusivamente en contextos donde la mujer ya está expuesta a múltiples formas de violencia. Estos hallazgos resaltan el carácter interrelacionado y acumulativo de la violencia que afecta a la mujer en el ámbito de la pareja.

4.3. Correlación de los factores que influyen en la presencia de violencia

Para la evaluación de factores se decidió el uso de Odd ratio para lo cual se requiere que el número de casos sea igual al número de controles razón por la cual el número de controles fue escogido al azar del total.

4.3.1. Violencia física y factores asociados

Tabla 11 Características demográficas de la pareja y su asociación a la violencia física

		Violencia física hacia la mujer				Odd Radio (OR)	IC 95% (OR)		Chi²	Sig.
		Sí		No			Li	Ls		
Variable	Categoría	F	%	F	%					
Edad mujer	15 a 20	125	48,1%	135	51,9%	0,922	0,719	1,184	43,12	0,000
	21 a 30	891	48,1%	961	51,9%	0,892	0,797	0,998		
	31 a 40	1038	47,3%	1156	52,7%	0,835	0,750	0,931		
	41 a 49	680	58,5%	482	41,5%	1,547	1,357	1,764		
Edad pareja	15 a 20	45	56,3%	35	43,8%	1,290	0,827	2,014	18,90	0,002
	21 a 30	611	49,8%	617	50,2%	0,987	0,870	1,121		
	31 a 40	1073	47,1%	1203	52,9%	0,822	0,738	0,916		
	41 a 50	858	54,0%	731	46,0%	1,253	1,115	1,409		
	51 a 60	131	49,6%	133	50,4%	0,984	0,769	1,260		
	61 a 78	16	51,6%	15	48,4%	1,067	0,526	2,163		
Región	Sierra	1021	53,2%	898	46,8%	1,219	1,090	1,362	18,94	0,000
	Resto Costa	716	47,5%	792	52,5%	0,870	0,773	0,980		
	Selva	747	50,5%	731	49,5%	1,030	0,914	1,161		
	Lima	250	44,4%	313	55,6%	0,778	0,653	0,928		
Provincia	Lima	1213	48,3%	1296	51,7%	0,885	0,796	0,984	33,40	0,001
	Piura	266	51,5%	251	48,5%	1,066	0,889	1,278		
	La Libertad	246	49,9%	247	50,1%	0,996	0,827	1,198		
	Arequipa	186	49,9%	187	50,1%	0,994	0,806	1,227		
	Cajamarca	178	53,3%	156	46,7%	1,151	0,922	1,437		
	Junín	168	56,6%	129	43,4%	1,322	1,044	1,674		
	Cusco	113	50,2%	112	49,8%	1,009	0,773	1,318		
	Lambayeque	118	56,7%	90	43,3%	1,325	1,002	1,753		
	Puno	74	43,5%	96	56,5%	0,764	0,562	1,040		
	Áncash	65	60,7%	42	39,3%	1,561	1,055	2,309		
	Loreto	39	46,4%	45	53,6%	0,865	0,561	1,332		
	San Martín	36	61,0%	23	39,0%	1,573	0,930	2,661		
	Ica	12	31,6%	26	68,4%	0,459	0,231	0,912		
	Otras	20	37,0%	34	63,0%	0,585	0,336	1,019		

La Tabla 11 presenta los resultados del análisis bivariado mediante razones de momios (odds ratios, OR), intervalos de confianza al 95 % (IC 95 %) y pruebas de chi-cuadrado (χ^2), con el propósito de identificar los factores demográficos asociados a la violencia física ejercida contra la mujer por parte de su pareja. La variable dependiente es dicotómica, diferenciando entre mujeres que han experimentado o no este tipo de violencia, en relación con la edad de la mujer, se evidencia una asociación estadísticamente significativa ($\chi^2 = 43,12$; $p < 0,001$), con un comportamiento no lineal.

Las mujeres más jóvenes (15–20 años) y adultas jóvenes (21–30 años) presentan una menor probabilidad de experimentar violencia física ($OR < 1$), en comparación con el grupo de referencia. En contraste, las mujeres de 41 a 49 años muestran un incremento significativo en el riesgo, con aproximadamente 1,5 veces más probabilidad de sufrir violencia física ($OR = 1,547$; IC 95 %: 1,357–1,764), lo que podría estar relacionado con procesos acumulativos de violencia a lo largo del tiempo, así como con cambios en las dinámicas de pareja y roles de género en etapas más avanzadas de la vida, respecto a la edad de la pareja, también se identifica una asociación significativa ($\chi^2 = 18,90$; $p = 0,002$). Las mujeres cuya pareja se encuentra en el rango de 41 a 50 años presentan una mayor probabilidad de sufrir violencia física ($OR = 1,253$; IC 95 %: 1,115–1,409), mientras que aquellas con parejas de 31 a 40 años muestran una menor probabilidad ($OR = 0,822$; IC 95 %: 0,738–0,916), lo que sugiere diferencias en el riesgo según la etapa del ciclo de vida de la pareja.

En cuanto a la ubicación geográfica, se observan diferencias significativas ($\chi^2 = 18,94$; $p < 0,001$). Las mujeres que residen en la Sierra presentan una mayor probabilidad de experimentar violencia física ($OR = 1,219$; IC 95 %: 1,090–1,362), lo que podría estar vinculado a desigualdades estructurales, barreras de acceso a servicios y persistencia de normas tradicionales. Por el contrario, las mujeres que viven en Lima Metropolitana presentan una menor probabilidad de sufrir este tipo de violencia ($OR = 0,778$; IC 95 %: 0,653–0,928), lo cual podría asociarse con mayores oportunidades de acceso a recursos, redes de apoyo y mecanismos de protección. La costa fuera de Lima también muestra una ligera reducción del riesgo ($OR = 0,870$), mientras que en la Selva no se evidencian diferencias significativas.

A nivel provincial, las asociaciones también resultan significativas ($\chi^2 = 33,40$; $p = 0,001$).

Las mujeres que residen en provincias como Áncash, San Martín y Lambayeque

presentan mayores probabilidades de experimentar violencia física, evidenciando contextos donde este problema puede ser más pronunciado. En contraste, en provincias como Ica y otras agrupadas en la categoría “Otras”, se observa una menor probabilidad, aunque estos resultados deben interpretarse con cautela debido al tamaño de muestra. Asimismo, las mujeres de Lima mantienen un ligero efecto protector, en concordancia con los resultados observados a nivel regional, se observa que la experiencia de violencia física en la mujer no es homogénea, sino que varía según características demográficas y contextuales, evidenciando la influencia de factores etarios y territoriales en el riesgo de sufrir este tipo de violencia dentro de la relación de pareja.

Tabla 12. Características reproductivas de la pareja y su asociación a la violencia física

		Violencia física hacia la mujer				IC_95% (OR)			Chi²	Sig.
Variable	Categoría	Si		No		Odd Radio (OR)	Li	Ls		
		F	%	F	%					
Estado civil	Viviendo juntos	1367	41,4%	1937	58,6%	0,411	0,368	0,460	753,06	0,000
	Casado	386	37,7%	638	62,3%	0,540	0,470	0,621		
	No viven juntos	981	86,1%	159	13,9%	9,063	7,584	10,831		
Educación Máxima	Sin educación	259	56,8%	197	43,2%	1,348	1,111	1,635	29,58	0,000
	Inicial	344	54,5%	287	45,5%	1,227	1,039	1,450		
	Primaria C.	212	55,9%	167	44,1%	1,292	1,047	1,594		
	Primaria I.	522	46,7%	596	53,3%	0,847	0,742	0,966		
	Secundaria C	1107	47,7%	1216	52,3%	0,849	0,763	0,946		
	Superior	290	51,7%	271	48,3%	1,078	0,905	1,284		
	Inyección	498	41,9%	691	58,1%	0,658	0,578	0,750		
Método Anticonceptivo	No usando	897	63,7%	511	36,3%	2,124	1,875	2,407	198,15	0,000
	Ritmo	318	40,6%	466	59,4%	0,641	0,549	0,747		
	Condón	236	41,7%	330	58,3%	0,688	0,577	0,821		
	Este. Femenina	321	53,7%	277	46,3%	1,180	0,995	1,399		
	Implantes	323	54,1%	274	45,9%	1,203	1,014	1,426		
	Píldora	97	40,1%	145	59,9%	0,657	0,505	0,854		
	DIU	25	55,6%	20	44,4%	1,252	0,694	2,260		
	Otro	15	50,0%	15	50,0%	1,000	0,488	2,050		
	Este. Masculina	4	44,4%	5	55,6%	0,800	0,215	2,981		
	Embarazada	No	2680	50,3%	2648	49,7%	1,612	1,142		
Sí		54	38,6%	86	61,4%	0,620	0,440	0,876		
Edad primera Relación	10 a 15	954	56,9%	724	43,1%	1,488	1,325	1,671	67,56	0,000
	16 a 20	1558	48,5%	1652	51,5%	0,868	0,779	0,966		
	21 a 30	214	38,8%	337	61,2%	0,604	0,504	0,723		
	31 a 44	8	27,6%	21	72,4%	0,379	0,168	0,857		
Total, hijos Nacidos	0	49	42,2%	67	57,8%	0,726	0,501	1,054	100,06	0,000
	1	565	45,0%	690	55,0%	0,772	0,680	0,876		
	2	769	44,9%	945	55,1%	0,741	0,661	0,831		
	3	651	52,0%	601	48,0%	1,109	0,978	1,258		
	4	359	62,4%	216	37,6%	1,762	1,475	2,105		
	5	141	59,5%	96	40,5%	1,494	1,146	1,948		
	6	109	63,4%	63	36,6%	1,760	1,285	2,412		
	7 o mas	91	61,9%	56	38,1%	1,647	1,175	2,307		

La Tabla 12 analiza la relación entre diversos factores reproductivos y la probabilidad de que la mujer experimente violencia física por parte de su pareja, utilizando razones de momios (OR), intervalos de confianza al 95 % y pruebas de significancia estadística. En general, los resultados evidencian asociaciones estadísticamente significativas, lo que resalta la influencia de estas características en la experiencia de violencia de la mujer.

En relación con el estado civil, se identifica que las mujeres que no conviven con su pareja presentan una probabilidad considerablemente mayor de sufrir violencia física, superando ampliamente a aquellas que sí viven en pareja. Este resultado sugiere que la inestabilidad o la interrupción de la convivencia puede situar a la mujer en contextos de mayor

vulnerabilidad. Por el contrario, las mujeres casadas o en unión consensual muestran una menor probabilidad de experimentar este tipo de violencia, lo que podría estar asociado a dinámicas de mayor estabilidad en la relación, con respecto al nivel educativo, se observa que la mujer con menor grado de instrucción presenta un mayor riesgo de sufrir violencia física, lo que evidencia el rol protector de la educación al favorecer mayores niveles de autonomía, capacidad de decisión y acceso a oportunidades. En cambio, alcanzar niveles como la educación secundaria completa se asocia con una disminución del riesgo, aunque el nivel superior no muestra una relación claramente protectora en este análisis, en cuanto al uso de métodos anticonceptivos, se evidencia que la mujer que no utiliza ningún método presenta una mayor probabilidad de experimentar violencia física, lo cual podría reflejar limitaciones en su autonomía reproductiva o conflictos dentro de la relación de pareja. Por el contrario, el uso de métodos anticonceptivos modernos se asocia con una menor probabilidad de violencia, lo que sugiere relaciones más equitativas en la toma de decisiones. Sin embargo, algunos métodos como la esterilización o los implantes muestran una ligera asociación con mayor riesgo, lo que podría estar vinculado a dinámicas de imposición o tensiones dentro de la relación, en relación con la condición de embarazo, se observa que la mujer que no se encuentra gestando presenta una mayor probabilidad de sufrir violencia física en comparación con aquella que sí lo está, lo que podría estar relacionado con factores contextuales como mayor supervisión institucional o normas sociales que tienden a proteger a la mujer durante el embarazo, el inicio de la vida sexual a edades tempranas constituye un factor de riesgo importante. Las mujeres que iniciaron relaciones sexuales antes de los 16 años presentan una mayor probabilidad de experimentar violencia física, mientras que este riesgo disminuye progresivamente conforme aumenta la edad de inicio, lo que sugiere una relación entre madurez, capacidad de decisión y menor exposición a relaciones desiguales, el número de hijos nacidos

muestra una asociación positiva con la violencia física. Las mujeres con mayor número de hijos presentan un incremento en la probabilidad de experimentar este tipo de violencia, lo que podría estar relacionado con la sobrecarga de responsabilidades, la dependencia económica y las dinámicas de control dentro del hogar.

La experiencia de violencia física en la mujer está estrechamente vinculada a factores reproductivos y sociales, los cuales pueden incrementar o disminuir su vulnerabilidad dentro de la relación de pareja.

Tabla 13. *Características económicas y sociales de la pareja y su asociación a la violencia física*

		Violencia física Hacia la mujer									
		Si		No		Odd Radio (OR)	IC_95% (OR)		Chi²	Sig.	
Variable	Categoría	F	%	F	%		Li	Ls			
Tipo ingresos	Solo en efectivo	1798	53,2%	1579	46,8%	1,405	1,259	1,568	78,62	0,000	
	No le pagan	643	40,7%	935	59,3%	0,592	0,526	0,666			
	Efectivo y especie	244	57,1%	183	42,9%	1,366	1,119	1,668			
	Solo en especie	49	57,0%	37	43,0%	1,330	0,865	2,045			
Decisor gasto De la pareja	Entrevistada y Esposo/compañero	795	34,9%	1484	65,1%	0,345	0,309	0,386	629,32	0,000	
	Esposo/compañero	1530	70,9%	628	29,1%	4,262	3,792	4,790			
	Entrevistada	409	39,7%	622	60,3%	0,597	0,520	0,686			
Tiene seguro salud	No	2397	51,3%	2276	48,7%	1,431	1,230	1,666	21,55	0,000	
	Si	337	42,4%	458	57,6%	0,699	0,600	0,813			
Frecuencia de Consumo de alcohol	Nunca	272	34,1%	526	65,9%	0,464	0,396	0,543	534,13	0,000	
	Algunas veces	1857	46,4%	2142	53,6%	0,585	0,518	0,661			
De la pareja	Mucha frecuencia	605	90,2%	66	9,8%	11,487	8,854	14,904			
Entrevista	No	2710	50,0%	2707	50,0%	1,126	0,648	1,957	0,18	0,673	
Interrumpida \por pareja	Si	24	47,1%	27	52,9%	0,888	0,511	1,543			
Años del primer	0 a 4	404	41,8%	562	58,2%	0,670	0,582	0,771	86,78	0,000	
Matrimonio	5 a 10	699	46,3%	810	53,7%	0,816	0,724	0,919			
	11 a 20	970	50,9%	936	49,1%	1,056	0,945	1,181			
	21 a 30	582	60,2%	385	39,8%	1,650	1,432	1,901			
	31 a mas	79	65,8%	41	34,2%	1,954	1,335	2,861			

La Tabla 13 presenta el análisis de la relación entre diversos factores económicos, de toma de decisiones, acceso a salud, conductuales y relacionales, y la probabilidad de que la mujer experimente violencia física por parte de su pareja. Los resultados, estimados mediante razones de momios (OR) e intervalos de confianza al 95 %, evidencian asociaciones estadísticamente significativas en la mayoría de los casos, lo que pone de manifiesto el carácter complejo y multidimensional de la violencia que afecta a la mujer.

En relación con el tipo de ingreso de la pareja, se observa que las mujeres cuyas parejas perciben ingresos en efectivo, ya sea exclusivamente o combinados con pagos en especie, presentan una mayor probabilidad de sufrir violencia física en comparación con aquellas cuyas parejas no reciben ingresos. Este resultado sugiere que ciertas condiciones económicas, especialmente vinculadas a la inestabilidad laboral, pueden incrementar la vulnerabilidad de la mujer frente a situaciones de violencia.

Uno de los hallazgos más relevantes se vincula con la toma de decisiones sobre los gastos del hogar. Cuando esta responsabilidad recae exclusivamente en el hombre, la mujer presenta un riesgo considerablemente mayor de experimentar violencia física. En contraste, cuando las decisiones son compartidas, la probabilidad de violencia disminuye notablemente, evidenciando que la participación equitativa de la mujer en la gestión del hogar actúa como un factor protector frente a dinámicas de control y dominación.

Respecto al acceso a seguro de salud, se identifica que la mujer que no cuenta con cobertura presenta una mayor probabilidad de sufrir violencia física, lo que refleja cómo la exclusión de sistemas de protección social puede incrementar su situación de vulnerabilidad y limitar sus posibilidades de acceso a apoyo institucional, el consumo de alcohol por parte de la pareja, este se constituye como uno de los factores más determinantes. Las mujeres cuyas parejas consumen alcohol con alta frecuencia presentan un riesgo significativamente mayor de sufrir violencia física, lo que evidencia el papel de este comportamiento en la intensificación de conflictos y conductas agresivas dentro de la relación, la interrupción de la entrevista por parte de la pareja no muestra una asociación estadísticamente significativa con la violencia física, lo que sugiere que este factor no permite explicar de manera consistente la experiencia de violencia en la mujer dentro de este contexto específico, la duración de la unión evidencia una relación directa con la violencia física. A medida que aumentan los años de convivencia, la mujer presenta

una mayor probabilidad de experimentar este tipo de violencia, lo que podría estar asociado a la acumulación de tensiones, la persistencia de relaciones desiguales o la dificultad de romper vínculos en contextos prolongados, con estos hallazgos se evidencian que la violencia física hacia la mujer no responde a un único factor, sino que está influenciada por condiciones económicas, relaciones de poder dentro del hogar, acceso a recursos y características propias de la dinámica de pareja, lo que refuerza la necesidad de abordajes integrales para su comprensión y prevención.

Figura 3. Árbol clasificación de presencia de violencia física

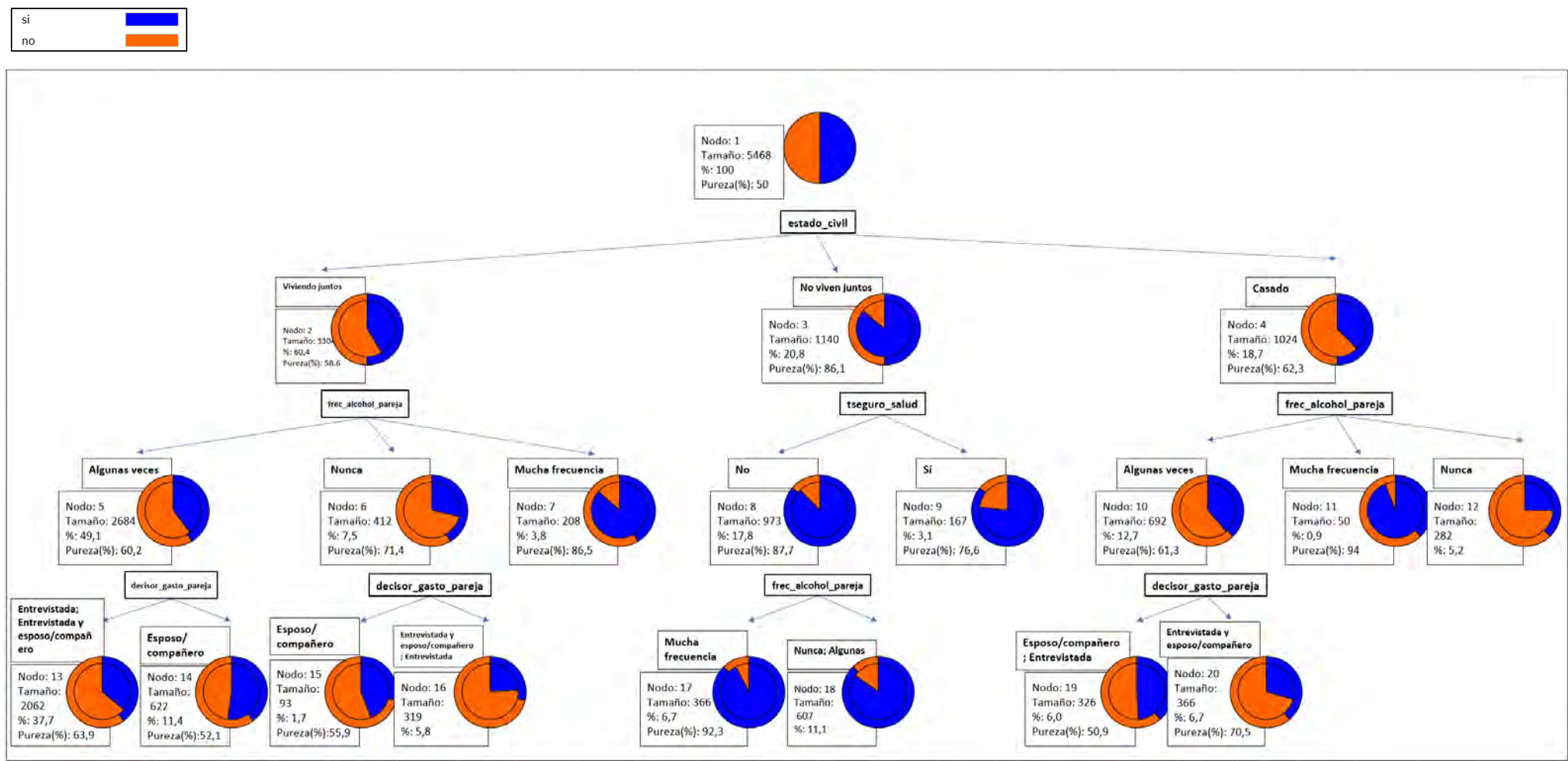


Tabla 14 Frecuencia según nodos y perfiles mas sensibles a violencia física según OR

Nodos	Características del nodo	Tamaño del nodo		Presencia de violencia				Odd Ratio (OR)	IC_95% (OR)	
		F	%	Si (pureza)		No			Li	Ls
Nodo 1	Inicial	5468	100,0%	2734	50,0%	2734	50,0%	1,000	0,928	1,078
Nodo 2	Vive con pareja	3304	60,4%	1367	41,4%	1937	58,6%	0,608	0,559	0,660
Nodo 3	No viven juntos	1140	20,8%	981	86,1%	159	13,9%	7,300	6,145	8,672
Nodo 4	Casados	1024	18,7%	386	37,7%	638	62,3%	0,575	0,504	0,656
Nodo 5	Vive con pareja – su pareja consume alcohol algunas veces	2684	49,1%	1069	39,8%	1615	60,2%	0,580	0,531	0,633
Nodo 6	Vive con pareja – su pareja consume nunca consume alcohol	412	7,5%	118	28,6%	294	71,4%	0,388	0,312	0,482
Nodo 7	Vive con pareja – su pareja consume alcohol muy frecuente	208	3,8%	180	86,5%	28	13,5%	6,613	4,433	9,866
Nodo 8	No viven juntos – no tiene seguro	973	17,8%	853	87,7%	120	12,3%	8,237	6,777	10,012
Nodo 9	No viven juntos – si tiene seguro	167	3,1%	128	76,6%	39	23,4%	3,337	2,327	4,785
Nodo 10	Casados - su pareja consume alcohol algunas veces	692	12,7%	268	38,7%	424	61,3%	0,613	0,524	0,718
Nodo 11	Casados – su pareja consume alcohol muy frecuente	50	0,9%	47	94,0%	3	6,0%	15,794	4,913	50,774
Nodo 12	Casados – su pareja consume nunca consume alcohol	282	5,2%	71	25,2%	211	74,8%	0,328	0,250	0,430
Nodo 13	Vive con pareja – su pareja consume alcohol algunas veces – ella y pareja decide gastos	2062	37,7%	745	36,1%	1317	63,9%	0,497	0,450	0,549
Nodo 14	Vive con pareja – su pareja consume alcohol algunas veces – su pareja decide gastos	622	11,4%	324	52,1%	298	47,9%	1,093	0,929	1,285
Nodo 15	Vive con pareja – su pareja consume nunca consume alcohol – su pareja decide gastos	93	1,7%	41	44,1%	52	55,9%	0,787	0,522	1,187
Nodo 16	Vive con pareja – su pareja consume nunca consume alcohol - ella y pareja decide gastos	319	5,8%	77	24,1%	242	75,9%	0,308	0,238	0,400
Nodo 17	No viven juntos – no tiene seguro - su pareja consume alcohol muy frecuente	366	6,7%	338	92,3%	28	7,7%	12,801	8,690	18,856
Nodo 18	No viven juntos – no tiene seguro - consume alcohol algunas veces o algunas veces	607	11,1%	515	84,8%	92	15,2%	6,076	4,851	7,610
Nodo 19	Casados - su pareja consume alcohol algunas veces – deciden gastos por separado	326	6,0%	160	49,1%	166	50,9%	0,963	0,772	1,200
Nodo 20	Casados - su pareja consume alcohol algunas veces – deciden gastos juntos	366	6,7%	108	29,5%	258	70,5%	0,407	0,324	0,511

La Figura 3 y la Tabla 14 del árbol de clasificación de la violencia física permiten identificar, de manera estructurada, los perfiles de mayor vulnerabilidad en las mujeres peruanas, a partir de un modelo de árbol de decisión (CART) aplicado a los datos de la ENDES 2024. En el nodo inicial se incluye a la totalidad de mujeres analizadas (5 468), donde aproximadamente la mitad reporta haber experimentado violencia física, estableciendo así el nivel base del fenómeno.

El primer criterio de segmentación corresponde al estado civil de la mujer, el cual se posiciona como el principal factor discriminante. A partir de este, se distinguen tres grupos: mujeres que conviven con su pareja, mujeres que no conviven y mujeres casadas. Esta primera división evidencia que la situación de no convivencia se asocia con mayor riesgo, mientras que la convivencia estable o el matrimonio se vinculan con menores probabilidades relativas de violencia física.

El perfil de mayor vulnerabilidad o factor de riesgo ($OR > 1$) se observa en las mujeres que no viven con su pareja, quienes presentan una alta proporción de violencia física. Dentro de este grupo, la ausencia de seguro de salud incrementa aún más el riesgo, y este se intensifica considerablemente cuando la pareja presenta un consumo frecuente de alcohol, configurando un contexto de alta exposición a situaciones de violencia. Esto sugiere que la combinación de inestabilidad en la relación, limitaciones en el acceso a servicios y conductas de riesgo por parte de la pareja agrava la situación de la mujer, en el caso de las mujeres que conviven con su pareja, el factor que marca mayores diferencias es el consumo de alcohol. Cuando la pareja consume alcohol con frecuencia, la mujer presenta una probabilidad significativamente mayor de sufrir violencia física; en cambio, cuando no existe consumo, el riesgo disminuye de manera importante, evidenciando el rol del alcohol como un factor que potencia la agresión, por su parte, en las mujeres casadas, el consumo de alcohol continúa siendo un elemento determinante, aquellas cuya pareja

consume alcohol de manera frecuente presentan niveles muy elevados de violencia física. No obstante, este riesgo se ve modulado por la toma de decisiones en el hogar: cuando la mujer participa de forma conjunta en las decisiones económicas, la probabilidad de violencia disminuye, mientras que cuando el hombre concentra este poder, el riesgo se incrementa, el modelo permite identificar que las mujeres en mayor situación de vulnerabilidad son aquellas que no conviven con su pareja, no cuentan con seguro de salud y tienen una pareja con consumo frecuente de alcohol. Esto evidencia que la violencia física en la mujer está influida por la interacción de factores relacionales, económicos y conductuales, lo que resalta la importancia de diseñar estrategias integrales que fortalezcan la autonomía de la mujer, mejoren su acceso a servicios de salud y aborden conductas de riesgo dentro de la pareja. Cabe resaltar los factores de protección ($OR < 1$) que se observan en la tabla de la mujer encuestada son las que vive con su pareja, las parejas casadas, vive con pareja – su pareja consume alcohol algunas veces, vive con pareja – su pareja consume nunca consume alcohol, vive con pareja – su pareja consume alcohol muy frecuente, casados - su pareja consume alcohol algunas veces, casados – su pareja consume nunca consume alcohol, vive con pareja – su pareja consume alcohol algunas veces – ella y pareja decide gastos, vive con pareja – su pareja consume nunca consume alcohol - ella y pareja decide gastos, casados - su pareja consume alcohol algunas veces – deciden gastos juntos; todos los nodos mencionados son los que tienen límites inferiores a la unidad.

Tabla 15. *Matriz de confusión y principales indicadores de la violencia física*

			Predicción de presencia de violencia física Hacia la mujer		
			Si	No	Total
Valor observado de violencia Física hacia la mujer	de Si		2246 (41,1%)	488 (8,9%)	2734 (50%)
	No		1202 (22%)	1532 (28%)	2734 (50%)
	Total		3448 (63,1%)	2020 (36,9%)	5468 (100%)

Indicadores	Valor	Descripción
Exactitud	0,691	Proporción total de asociaciones correctas (aciertos / total)
Sensibilidad	0,560	Proporción de casos positivos reales se identificaron correctamente (TP / TP+FN)
Especificidad	0,822	Proporción de casos negativos reales se identificaron correctamente (TN / TN+FP)
Valor predictivo positivo	0,758	Probabilidad de que sea realmente positivo (TP / TP+FP)
Valor predictivo negativo	0,651	Probabilidad de que sea realmente negativo (TN / TN+FN)
Exactitud balanceada	0,691	Promedio de sensibilidad y Especificidad

La Tabla 15 presenta la matriz de confusión y los principales indicadores de desempeño del modelo de árbol de clasificación utilizado para predecir la presencia de violencia física en la mujer dentro de la muestra analizada (n = 5 468). De manera general, el modelo alcanza una exactitud global del 69,1 %, lo que indica que aproximadamente siete de cada diez mujeres fueron clasificadas correctamente según hayan experimentado o no violencia física, al examinar el desempeño del modelo en función de la condición de la mujer, se evidencian diferencias importantes. La sensibilidad es moderada (56,0 %), lo que refleja que el modelo logra identificar solo a poco más de la mitad de las mujeres que efectivamente han sufrido violencia física, dejando sin detectar a un porcentaje considerable de ellas. Esto implica que existe un número relevante de mujeres que han vivido situaciones de violencia y que no son reconocidas adecuadamente por el modelo. En contraste, la especificidad es elevada (82,2 %), lo que indica que el modelo tiene una buena capacidad para identificar correctamente a las mujeres que no han experimentado violencia física. Este comportamiento sugiere que el modelo tiende a ser más preciso al

clasificar la ausencia de violencia que su presencia en la mujer, en cuanto a los valores predictivos, el valor predictivo positivo (75,8 %) evidencia que, cuando el modelo clasifica a una mujer como víctima de violencia física, existe una alta probabilidad de que esta clasificación sea correcta. No obstante, el valor predictivo negativo (65,1 %) es menor, lo que indica que existe mayor incertidumbre al afirmar que una mujer no ha sido víctima de violencia, la exactitud balanceada coincide con la exactitud global (69,1 %), lo que se explica por la distribución equilibrada de los casos en la muestra. En conjunto, estos resultados muestran que, si bien el modelo presenta un desempeño aceptable para identificar violencia física en la mujer, su limitada sensibilidad constituye una restricción importante, especialmente en contextos donde es prioritario detectar oportunamente a todas las mujeres que se encuentran en situación de violencia.

4.3.2. Violencia psicológica y factores asociados

Tabla 16. Características demográficas de la pareja y su asociación a la violencia psicológica

		Violencia psicológica Hacia la mujer									
Variable	Categoría	Si		No		Odd Radio (OR)	IC_95% (OR)		Chi2	Sig.	
		F	%	F	%		Li	Ls			
Edad Mujer	15 a 20	358	51,7%	334	48,3%	1,075	0,923	1,253	80,48	0,000	
	21 a 30	2560	48,0%	2778	52,0%	0,882	0,826	0,943			
	31 a 40	3080	48,2%	3314	51,8%	0,882	0,827	0,941			
	41 a 49	1687	57,3%	1259	42,7%	1,436	1,324	1,557			
Edad pareja	15 a 20	112	55,7%	89	44,3%	1,262	0,954	1,670	29,30	0,000	
	21 a 30	1658	48,8%	1743	51,2%	0,938	0,869	1,012			
	31 a 40	3208	48,3%	3433	51,7%	0,887	0,833	0,946			
	41 a 50	2267	52,6%	2045	47,4%	1,154	1,075	1,238			
	51 a 60	396	54,3%	333	45,7%	1,199	1,033	1,393			
	61 a 78	44	51,2%	42	48,8%	1,048	0,686	1,601			
Región	Sierra	2778	53,3%	2435	46,7%	1,221	1,142	1,305	39,19	0,000	
	Resto Costa	2087	47,7%	2289	52,3%	0,879	0,819	0,943			
	Selva	2032	49,6%	2064	50,4%	0,979	0,911	1,051			
	Lima	788	46,8%	897	53,2%	0,865	0,781	0,957			
Provincia	Lima	3475	47,8%	3789	52,2%	0,849	0,797	0,904	64,52	0,000	
	Piura	729	52,2%	667	47,8%	1,103	0,988	1,231			
	La Libertad	669	49,7%	678	50,3%	0,985	0,881	1,102			
	Arequipa	502	51,3%	476	48,7%	1,058	0,930	1,205			
	Cajamarca	453	52,9%	404	47,1%	1,129	0,983	1,296			
	Junín	460	56,2%	358	43,8%	1,303	1,131	1,502			
	Cusco	348	54,9%	286	45,1%	1,227	1,046	1,440			
	Lambayeque	297	53,3%	260	46,7%	1,148	0,969	1,360			
	Puno	233	45,5%	279	54,5%	0,830	0,695	0,991			
	Áncash	169	56,5%	130	43,5%	1,307	1,037	1,646			
	Loreto	126	52,3%	115	47,7%	1,097	0,850	1,416			
	San Martín	92	54,4%	77	45,6%	1,197	0,883	1,623			
	Ica	53	35,8%	95	64,2%	0,555	0,396	0,778			
	Otras	79	52,7%	71	47,3%	1,114	0,807	1,537			

La Tabla 16 analiza la relación entre variables demográficas —como la edad de la mujer y de su pareja, así como la región y provincia de residencia— y la experiencia de violencia psicológica en la mujer, entendida como conductas de control, celos, humillación, aislamiento o restricción de su autonomía. Los resultados, estimados mediante razones de momios (OR), intervalos de confianza al 95 % y pruebas de chi-cuadrado, evidencian

asociaciones estadísticamente significativas que permiten identificar patrones de mayor vulnerabilidad en la mujer.

En relación con la edad de la mujer, se observa un comportamiento diferenciado. Las mujeres de 41 a 49 años presentan una mayor probabilidad de experimentar violencia psicológica, lo que sugiere que, en etapas más avanzadas del ciclo de vida, pueden intensificarse dinámicas de control dentro de la relación de pareja. Por el contrario, las mujeres en edades más jóvenes y adultas tempranas muestran una menor probabilidad relativa, lo que podría estar asociado a relaciones más recientes, cambios generacionales o mayores recursos de apoyo. En el caso de las adolescentes, si bien se observa la presencia de violencia psicológica, esta no presenta una diferencia significativa en comparación con otros grupos etarios, respecto a la edad de la pareja, también se identifican asociaciones relevantes. Las mujeres cuyas parejas se encuentran en rangos de edad intermedios y avanzados presentan una mayor probabilidad de sufrir violencia psicológica, lo que podría estar relacionado con dinámicas de control, tensiones personales o cambios en los roles dentro de la relación. En el caso de parejas más jóvenes, aunque se observa una tendencia al incremento del riesgo, esta no presenta la misma consistencia estadística.

En cuanto a la ubicación geográfica, se evidencian diferencias importantes en la experiencia de la mujer. Aquellas que residen en la Sierra presentan una mayor probabilidad de experimentar violencia psicológica, lo que podría estar vinculado a factores estructurales, normas sociales tradicionales y menores oportunidades de autonomía. En contraste, las mujeres que viven en Lima Metropolitana y en la costa fuera de Lima muestran una menor probabilidad, lo que sugiere un posible efecto protector asociado a contextos urbanos, mayor acceso a educación y servicios de apoyo. Por su parte, en la Selva no se observan diferencias significativas, lo que podría reflejar una

mayor diversidad de contextos dentro de esta región, a nivel provincial, se identifican también variaciones relevantes. Las mujeres que residen en provincias como Junín, Áncash y Cusco presentan mayores probabilidades de experimentar violencia psicológica, lo que evidencia contextos donde este tipo de violencia puede estar más arraigado. En contraste, provincias como Ica muestran una menor probabilidad, lo que podría estar asociado a condiciones socioeconómicas que favorecen una mayor autonomía de la mujer. Otras provincias presentan efectos protectores más moderados. En conjunto, estos hallazgos evidencian que la violencia psicológica en la mujer no se distribuye de manera uniforme, sino que está influida por factores etarios y territoriales que condicionan su vulnerabilidad. Esto resalta la importancia de considerar el contexto demográfico y geográfico en el análisis y abordaje de la violencia que afecta a la mujer.

Tabla 17. Características reproductivas de la pareja y su asociación a la violencia psicológica

Variable	Categoría	Violencia psicológica Hacia la mujer				Odd Ratio (OR)	IC_95% (OR)		Chi2	Sig.
		Si F	%	No F	%		Li	Ls		
Estado civil	Viviendo Juntos	4745	46,1%	5545	53,9%	0,623	0,582	0,667	938,23	0,000
	Casado	1457	43,9%	1863	56,1%	0,731	0,677	0,790		
	No viven Juntos	1483	84,3%	277	15,7%	6,395	5,600	7,302		
Educación Máxima	Sin Educación	633	51,6%	594	48,4%	1,072	0,954	1,204	66,16	0,000
	Inicial	853	53,2%	749	46,8%	1,156	1,042	1,282		
	Primaria C	567	56,4%	438	43,6%	1,318	1,159	1,499		
	Primaria I	1550	47,6%	1705	52,4%	0,886	0,820	0,957		
	Secundaria completa	3212	47,8%	3510	52,2%	0,854	0,801	0,910		
	Superior	870	55,8%	689	44,2%	1,296	1,167	1,440		
Método Anticonceptivo	Inyección	1624	44,7%	2007	55,3%	0,758	0,703	0,817	208,11	0,000
	No usando	1920	59,1%	1331	40,9%	1,590	1,470	1,720		
	Ritmo	1231	49,4%	1261	50,6%	0,972	0,892	1,059		
	Condón	767	44,0%	975	56,0%	0,763	0,690	0,844		
	Est. Femenina	875	54,1%	741	45,9%	1,204	1,086	1,335		
	Implantes	819	51,3%	778	48,7%	1,059	0,955	1,175		
	Píldora	331	41,9%	459	58,1%	0,709	0,613	0,819		
	DIU	59	45,7%	70	54,3%	0,842	0,594	1,192		
	Otro	45	52,3%	41	47,7%	1,098	0,718	1,679		
	Est. Masculina	14	38,9%	22	61,1%	0,636	0,325	1,243		
Embarazada	No	7500	50,2%	7449	49,8%	1,284	1,057	1,561	6,35	0,012
	Sí	185	43,9%	236	56,1%	0,779	0,641	0,946		
Edad primer Relación	10 a 15	2296	56,2%	1790	43,8%	1,403	1,306	1,508	114,54	0,000
	16 a 20	4571	48,9%	4782	51,1%	0,891	0,835	0,951		
	21 a 30	790	42,7%	1060	57,3%	0,716	0,649	0,790		
	31 a 44	28	34,6%	53	65,4%	0,527	0,333	0,833		
Total, hijos Nacidos	0	150	46,4%	173	53,6%	0,864	0,693	1,078	193,62	0,000
	1	1658	44,1%	2098	55,9%	0,733	0,680	0,789		
	2	2406	46,9%	2724	53,1%	0,830	0,776	0,888		
	3	1799	53,8%	1547	46,2%	1,213	1,123	1,309		
	4	868	58,8%	609	41,2%	1,479	1,327	1,650		
	5	398	58,3%	285	41,7%	1,418	1,214	1,657		
	6	214	61,8%	132	38,2%	1,639	1,316	2,041		
	7 o mas	192	62,1%	117	37,9%	1,657	1,314	2,090		

La Tabla 17 examina la relación entre diversas características reproductivas de la mujer y la probabilidad de experimentar violencia psicológica por parte de su pareja, utilizando razones de momios (OR) e intervalos de confianza al 95 %. Los resultados evidencian asociaciones estadísticamente significativas que reflejan cómo las condiciones reproductivas de la mujer se vinculan con dinámicas de control dentro de la relación.

En relación con el estado civil, se observa que las mujeres que no conviven con su pareja presentan una probabilidad considerablemente mayor de sufrir violencia psicológica, lo

que sugiere que la inestabilidad en la relación puede aumentar su vulnerabilidad frente a conductas como los celos, el control o la manipulación emocional. En contraste, las mujeres casadas o en unión consensual muestran una menor probabilidad, lo que podría asociarse a contextos de mayor estabilidad relacional.

Respecto al nivel educativo, se identifican diferencias importantes. La mujer con menor nivel de instrucción presenta mayor riesgo de experimentar violencia psicológica, mientras que alcanzar la educación secundaria completa se asocia con una disminución del riesgo, evidenciando el rol protector de la educación. Sin embargo, en el caso de la educación superior, se observa una tendencia a mayor riesgo, lo que podría estar relacionado con tensiones en la relación cuando la mujer alcanza mayores niveles educativos.

En relación al uso de métodos anticonceptivos, se evidencia que la mujer que no utiliza ningún método presenta una mayor probabilidad de experimentar violencia psicológica, lo que podría reflejar limitaciones en su autonomía reproductiva. Por el contrario, el uso de métodos anticonceptivos modernos se asocia con una menor probabilidad de violencia, sugiriendo relaciones más equitativas en la toma de decisiones, algunos métodos como la esterilización femenina muestran una ligera asociación con mayor riesgo, lo que podría estar vinculado a dinámicas de imposición o conflicto dentro de la pareja, la edad de inicio de la vida sexual constituye un factor relevante. Las mujeres que iniciaron su vida sexual a edades tempranas presentan una mayor probabilidad de experimentar violencia psicológica, mientras que este riesgo disminuye conforme aumenta la edad de inicio, lo que sugiere una relación entre mayor madurez y menor exposición a relaciones desiguales, el número de hijos nacidos muestra una asociación positiva con la violencia psicológica. A medida que aumenta el número de hijos, la mujer presenta una mayor probabilidad de experimentar este tipo de violencia, lo que podría estar relacionado con

mayores responsabilidades, dependencia económica y dinámicas de control en el hogar. Por otro lado, las mujeres embarazadas presentan una menor probabilidad de violencia psicológica, lo que podría vincularse a factores de protección temporales durante esta etapa.

La violencia psicológica en la mujer está estrechamente relacionada con sus condiciones reproductivas, las cuales pueden influir en su nivel de autonomía, en las dinámicas de poder dentro de la relación y en su grado de vulnerabilidad frente a conductas de control y dominación.

Tabla 18 Características económicas y sociales de la pareja y su asociación a la violencia psicológica

Variable	Categoría	Violencia psicológica Hacia la mujer								Chi2	Sig.
		Si		No		Odd Radio (OR)	IC_95% (OR)				
		F	%	F	%		Li	Ls			
Tipo ingresos	Solo en efectivo	4654	51,1%	4448	48,9%	1,117	1,048	1,192	93,72	0,000	
	No le pagan	2280	45,5%	2732	54,5%	0,765	0,715	0,818			
	Efectivo y especie	610	59,6%	414	40,4%	1,514	1,331	1,723			
	Solo en especie	141	60,8%	91	39,2%	1,560	1,196	2,034			
Decisor gasto de la pareja	Entrevistada y esposo/compañero	3130	41,1%	4483	58,9%	0,491	0,460	0,523	812,39	0,000	
	Esposo/compañero	3130	67,2%	1527	32,8%	2,771	2,578	2,979			
	Entrevistada	1425	46,0%	1675	54,0%	0,817	0,755	0,884			
Tiene Seguro salud	No	6627	51,5%	6230	48,5%	1,463	1,342	1,595	74,98	0,000	
	Si	1058	42,1%	1455	57,9%	0,684	0,627	0,745			
Frecuencia de consumo de alcohol de la pareja	Nunca	1133	43,8%	1452	56,2%	0,742	0,682	0,808	549,98	0,000	
	Algunas veces	5755	48,5%	6113	51,5%	0,767	0,711	0,827			
	Mucha frecuencia	797	86,9%	120	13,1%	7,294	6,004	8,862			
Entrevista interrumpida por pareja	No	7624	50,0%	7627	50,0%	0,950	0,663	1,363	0,08	0,782	
	Si	61	51,3%	58	48,7%	1,052	0,733	1,509			
Años del Primer matrimonio	0 a 4	1260	43,7%	1624	56,3%	0,732	0,675	0,794	176,86	0,000	
	5 a 10	2087	46,5%	2398	53,5%	0,822	0,767	0,881			
	11 a 20	2800	51,5%	2632	48,5%	1,100	1,030	1,176			
	21 a 30	1375	59,2%	947	40,8%	1,550	1,418	1,696			
	31 a mas	163	66,0%	84	34,0%	1,961	1,504	2,556			

La Tabla 18 analiza la relación entre las características económicas y sociales de la pareja y la probabilidad de que la mujer experimente violencia psicológica, evidenciando asociaciones estadísticamente significativas en la mayoría de los casos. Estos resultados permiten comprender cómo distintos factores estructurales influyen en la situación de la mujer dentro de la relación de pareja, en relación con el tipo de ingreso de la pareja, se observa que las mujeres cuyas parejas perciben ingresos en efectivo, ya sea de forma exclusiva o combinada con pagos en especie, presentan una mayor probabilidad de experimentar violencia psicológica. Esto sugiere que condiciones de inestabilidad o precariedad económica pueden incrementar la vulnerabilidad de la mujer frente a dinámicas de control. Por el contrario, cuando la pareja no percibe ingresos, el riesgo es menor, aunque este resultado debe interpretarse con cautela debido a posibles contextos

de dependencia o apoyo externo, esto se vincula con la toma de decisiones sobre los gastos del hogar. Cuando el hombre asume esta responsabilidad de manera exclusiva, la mujer presenta una probabilidad considerablemente mayor de experimentar violencia psicológica. En contraste, cuando las decisiones son compartidas, el riesgo disminuye significativamente, lo que evidencia que la participación de la mujer en la toma de decisiones actúa como un importante factor protector frente a relaciones desiguales, respecto al acceso a seguro de salud, se identifica que la mujer que no cuenta con cobertura presenta una mayor probabilidad de sufrir violencia psicológica, lo que refleja cómo la falta de acceso a sistemas de protección social incrementa su situación de vulnerabilidad y limita sus opciones de apoyo. En cuanto al consumo de alcohol por parte de la pareja, se observa una asociación clara con el incremento del riesgo. Las mujeres cuyas parejas consumen alcohol con alta frecuencia presentan una probabilidad significativamente mayor de experimentar violencia psicológica, lo que evidencia el papel de este factor en la intensificación de conductas de control, humillación y conflicto dentro de la relación, la interrupción de la entrevista por parte de la pareja no muestra una asociación significativa con la violencia psicológica, lo que sugiere que este indicador no permite explicar de manera consistente la experiencia de violencia en la mujer en este contexto, la duración de la unión presenta una relación directa con la violencia psicológica. A medida que aumentan los años de convivencia, la mujer muestra una mayor probabilidad de experimentar este tipo de violencia, lo que podría estar asociado a la acumulación de tensiones, la normalización de conductas de control o la dificultad de modificar dinámicas establecidas a lo largo del tiempo.

Figura 4. Árbol clasificación de presencia de violencia psicológica



Tabla 19. Frecuencia según nodos y perfiles más sensibles a violencia psicológica según OR

Nodos	Características del nodo	Tamaño del nodo		Presencia de violencia				Odd Radio (OR)	IC_95% (OR)	
		F	%	Si (pureza)		No			Li	Ls
Nodo 1	Inicial	15370	100,0%	7685	50,0%	7685	50,0%	1,000	0,956	1,046
Nodo 2	Viven juntos	10290	66,9%	4745	46,1%	5545	53,9%	0,791	0,755	0,830
Nodo 3	No viven juntos	1760	11,5%	1483	84,3%	277	15,7%	5,819	5,108	6,629
Nodo 4	Casados	3320	21,6%	1457	43,9%	1863	56,1%	0,759	0,706	0,816
Nodo 5	Viven juntos – ella decide el gasto	2370	15,4%	1088	45,9%	1282	54,1%	0,837	0,770	0,911
Nodo 6	Viven juntos – ambos deciden gastos	5681	37,0%	2369	41,7%	3312	58,3%	0,663	0,626	0,703
Nodo 7	Viven juntos – el decide el gasto	2239	14,6%	1288	57,5%	951	42,5%	1,387	1,271	1,513
Nodo 8	No viven juntos - el consume alcohol algunas veces o nunca	1236	8,0%	1002	81,1%	234	18,9%	4,511	3,905	5,210
Nodo 9	No viven juntos – el consume alcohol frecuentemente	524	3,4%	481	91,8%	43	8,2%	11,515	8,422	15,744
Nodo 10	Casados – con 5 a más hijos	817	5,3%	473	57,9%	344	42,1%	1,387	1,205	1,596
Nodo 11	Casados – con 3 hijos	818	5,3%	397	48,5%	421	51,5%	0,941	0,819	1,082
Nodo 12	Casados – sin hijos o 1 hijo	552	3,6%	162	29,3%	390	70,7%	0,409	0,340	0,492
Nodo 13	Casados – con 2 hijos	1133	7,4%	425	37,5%	708	62,5%	0,589	0,521	0,666
Nodo 14	Viven juntos – ella decide el gasto – el consume alcohol algunas veces	1979	12,9%	911	46,0%	1068	54,0%	0,844	0,770	0,924
Nodo 15	Viven juntos – ella decide el gasto – el nunca consume alcohol	310	2,0%	110	35,5%	200	64,5%	0,547	0,433	0,691
Nodo 16	Viven juntos – ella decide el gasto – el consume alcohol frecuentemente	81	0,5%	67	82,7%	14	17,3%	4,802	2,698	8,546
Nodo 17	Viven juntos – ambos deciden gastos– el consume alcohol algunas veces	4776	31,1%	1991	41,7%	2785	58,3%	0,672	0,632	0,716
Nodo 18	Viven juntos – ambos deciden gastos – el nunca consume alcohol	779	5,1%	286	36,7%	493	63,3%	0,572	0,494	0,663
Nodo 19	Viven juntos – ambos deciden gastos – el consume alcohol frecuentemente	126	0,8%	92	73,0%	34	27,0%	2,716	1,832	4,028
Nodo 20	Viven juntos – el decide el gasto – ella tiene seguro	313	2,0%	130	41,5%	183	58,5%	0,708	0,565	0,887
Nodo 21	Viven juntos – el decide el gasto – ella no tiene seguro	1926	12,5%	1158	60,1%	768	39,9%	1,549	1,410	1,702
Nodo 22	Casados – con 5 a más hijos – el consume alcohol algunas veces	510	3,3%	301	59,0%	209	41,0%	1,449	1,213	1,731
Nodo 23	Casados – con 5 a más hijos – el nunca consume alcohol	274	1,8%	141	51,5%	133	48,5%	1,061	0,836	1,346
Nodo 24	Casados – con 5 a más hijos – el consume alcohol frecuentemente	33	0,2%	31	93,9%	2	6,1%	15,529	3,716	64,900
Nodo 25	Casados – con 3 hijos – vive en lima, costa o selva	477	3,1%	202	42,3%	275	57,7%	0,731	0,609	0,878
Nodo 26	Casados – con 3 hijos – vive en sierra	341	2,2%	195	57,2%	146	42,8%	1,340	1,080	1,662
Nodo 27	Casados – sin hijos o 1 hijo – ella inicio de vida sexual de 10 a 15 años	53	0,3%	26	49,1%	27	50,9%	0,963	0,562	1,651
Nodo 28	Casados – sin hijos o 1 hijo – ella inicio de vida sexual de 16 a 44 años	499	3,2%	136	27,3%	363	72,7%	0,369	0,303	0,450
Nodo 29	Casados – con 2 hijos – ella o ambos deciden de los gastos	929	6,0%	324	34,9%	605	65,1%	0,526	0,458	0,603
Nodo 30	Casados – con 2 hijos – solo el decide los gatos	204	1,3%	101	49,5%	103	50,5%	0,980	0,744	1,291

La Figura 4 y la Tabla 19 del árbol de clasificación de la violencia psicológica permiten identificar de manera estructurada los perfiles de mayor vulnerabilidad en las mujeres peruanas, a partir de un modelo de árbol de decisión (CART) aplicado a los datos de la ENDES 2024. En el nodo inicial se incluye a la totalidad de mujeres analizadas (15 370), donde la mitad reporta haber experimentado violencia psicológica, estableciendo así el nivel base del fenómeno, el primer criterio de segmentación corresponde al estado civil de la mujer, el cual se posiciona como el principal factor explicativo. A partir de este, se distinguen tres grupos: mujeres que conviven con su pareja, mujeres que no conviven y mujeres casadas. Esta primera división evidencia que la falta de convivencia o la inestabilidad en la relación incrementa el riesgo de violencia psicológica en la mujer, mientras que la convivencia estable o el matrimonio se asocian con menores probabilidades relativas.

El perfil de mayor vulnerabilidad se observa en las mujeres que no viven con su pareja, quienes presentan una alta proporción de violencia psicológica. Dentro de este grupo, el consumo de alcohol por parte de la pareja intensifica significativamente el riesgo, especialmente cuando este es frecuente, lo que refleja un contexto de mayor exposición a conductas de control, celos y manipulación emocional, en el caso de las mujeres que conviven con su pareja, la toma de decisiones en el hogar se constituye como un factor clave. Cuando el hombre asume de manera exclusiva el control de los gastos, la mujer presenta una mayor probabilidad de experimentar violencia psicológica. En cambio, cuando las decisiones son compartidas o asumidas por la mujer, el riesgo disminuye. No obstante, este efecto protector puede verse reducido si la pareja consume alcohol con frecuencia, lo que incrementa nuevamente la probabilidad de violencia.

Por su parte, en las mujeres casadas, el número de hijos emerge como un factor relevante. Aquellas con mayor número de hijos presentan una mayor probabilidad de experimentar

violencia psicológica, especialmente cuando este factor se combina con el consumo frecuente de alcohol por parte de la pareja. En contraste, las mujeres casadas con pocos o ningún hijo presentan menores niveles de riesgo, particularmente cuando han iniciado su vida sexual a edades más avanzadas, el modelo permite identificar que las mujeres en mayor situación de vulnerabilidad son aquellas que no conviven con su pareja y cuya pareja presenta consumo frecuente de alcohol; así como aquellas que, aun conviviendo, enfrentan dinámicas de desigualdad en la toma de decisiones y consumo de alcohol por parte de la pareja. Asimismo, las mujeres casadas con un mayor número de hijos también constituyen un grupo de riesgo cuando se combinan factores estructurales adversos, la violencia psicológica en la mujer está determinada por la interacción de factores relacionales, económicos y conductuales, lo que resalta la necesidad de implementar estrategias integrales que fortalezcan su autonomía, promuevan relaciones más equitativas y aborden factores de riesgo como el consumo de alcohol dentro de la pareja.

Tabla 20. *Matriz de confusión y principales indicadores de la violencia psicológica*

		Predicción de presencia de p psicológica violencia hacia la mujer		
		Si	No	Total
Valor observado de violencia	Si	6102 (39,7%)	1583 (10,3%)	7685 (50%)
Psicológica hacia la mujer	No	4217 (27,4%)	3468 (22,6%)	7685 (50%)
Total		10319 (67,1%)	5051 (32,9%)	15370 (100%)

Indicadores	Valor	Descripción
Exactitud	0,623	Proporción total de asociaciones correctas (aciertos / total)
Sensibilidad	0,451	Proporción de casos positivos reales se identificaron correctamente (TP / TP+FN)
Especificidad	0,794	Proporción de casos negativos reales se identificaron correctamente (TN / TN+FP)
Valor predictivo positivo	0,687	Probabilidad de que sea realmente positivo (TP / TP+FP)
Valor predictivo negativo	0,591	Probabilidad de que sea realmente negativo (TN / TN+FN)
Exactitud balanceada	0,623	Promedio de sensibilidad y Especificidad

La Tabla 20 presenta la matriz de confusión y los principales indicadores de desempeño del modelo de árbol de clasificación empleado para predecir la presencia de violencia

psicológica en la mujer dentro de una muestra de 15 370 casos. De manera general, el modelo alcanza una exactitud global del 62,3 %, lo que indica que poco más de seis de cada diez mujeres fueron correctamente clasificadas según hayan experimentado o no este tipo de violencia, al analizar el desempeño del modelo en función de la condición de la mujer, se evidencian limitaciones importantes. La sensibilidad es baja (45,1 %), lo que refleja que el modelo identifica correctamente a menos de la mitad de las mujeres que realmente han sufrido violencia psicológica, dejando sin detectar a una proporción considerable de ellas. Este aspecto resulta especialmente crítico, ya que implica que muchas mujeres que experimentan este tipo de violencia no son reconocidas en la predicción, en contraste, la especificidad es alta (79,4 %), lo que indica que el modelo tiene una buena capacidad para identificar correctamente a las mujeres que no han sido víctimas de violencia psicológica. Este comportamiento sugiere que el modelo tiende a clasificar con mayor precisión la ausencia de violencia que su presencia en la mujer.

En cuanto a los valores predictivos, el valor predictivo positivo (68,7 %) muestra que, cuando el modelo identifica a una mujer como víctima de violencia psicológica, existe una probabilidad moderada de que esta clasificación sea correcta. Sin embargo, el valor predictivo negativo (59,1 %) es menor, lo que evidencia una mayor incertidumbre al descartar que la mujer haya experimentado violencia, la exactitud balanceada coincide con la exactitud global (62,3 %), debido a la distribución equilibrada de los casos en la muestra. En conjunto, estos resultados indican que, si bien el modelo presenta cierta utilidad para identificar perfiles de riesgo en la mujer, su limitada capacidad para detectar la mayoría de los casos reales constituye una restricción importante, especialmente considerando que la violencia psicológica es una de las formas más frecuentes y muchas veces antecede a manifestaciones más severas de violencia.

4.3.3. Violencia sexual y factores asociados

Tabla 21. Características demográficas de la pareja y su asociación a la violencia sexual

		Violencia sexual Hacia la mujer				Odd Radio (OR)	IC_95% (OR)		Chi2	Sig.
Variable	Categoría	Si		No			Li	Ls		
Edad mujer	15 a 20	14	46,7%	16	53,3%	0,870	0,419	1,809	0,70	0,872
	21 a 30	129	48,5%	137	51,5%	0,912	0,678	1,228		
	31 a 40	157	50,5%	154	49,5%	1,033	0,774	1,378		
	41 a 49	85	52,1%	78	47,9%	1,115	0,789	1,576		
Edad pareja	15 a 20	5	55,6%	4	44,4%	1,253	0,334	4,703	7,33	0,197
	21 a 30	82	46,1%	96	53,9%	0,815	0,582	1,140		
	31 a 40	168	50,8%	163	49,2%	1,054	0,793	1,403		
	41 a 50	110	55,0%	90	45,0%	1,311	0,949	1,812		
	51 a 60	15	34,9%	28	65,1%	0,517	0,272	0,984		
	61 a 78	5	55,6%	4	44,4%	1,253	0,334	4,703		
Región	Sierra	149	53,6%	129	46,4%	1,253	0,933	1,682	7,29	0,063
	Resto Costa	85	43,4%	111	56,6%	0,699	0,505	0,969		
	Selva	114	53,8%	98	46,2%	1,232	0,897	1,691		
	Lima	37	44,0%	47	56,0%	0,765	0,485	1,206		
Provincia	Lima	155	46,7%	177	53,3%	0,792	0,595	1,054	12,58	0,481
	Piura	37	54,4%	31	45,6%	1,214	0,737	2,001		
	La Libertad	37	52,9%	33	47,1%	1,134	0,693	1,855		
	Arequipa	35	53,8%	30	46,2%	1,183	0,711	1,970		
	Cajamarca	25	48,1%	27	51,9%	0,921	0,524	1,617		
	Junín	26	56,5%	20	43,5%	1,322	0,725	2,410		
	Cusco	18	54,5%	15	45,5%	1,210	0,601	2,437		
	Lambayeque	19	55,9%	15	44,1%	1,281	0,641	2,559		
	Puno	8	32,0%	17	68,0%	0,459	0,196	1,077		
	Áncash	4	36,4%	7	63,6%	0,567	0,165	1,953		
	Loreto	7	53,8%	6	46,2%	1,170	0,389	3,513		
	San Martín	6	85,7%	1	14,3%	6,079	0,728	50,737		
	Ica	4	66,7%	2	33,3%	2,010	0,366	11,042		
	Otras	4	50,0%	4	50,0%	1,000	0,248	4,028		

La Tabla 21 analiza la posible asociación entre variables demográficas —como la edad de la mujer y de su pareja, así como la región y provincia de residencia— y la experiencia de violencia sexual en la mujer, entendida como la imposición de relaciones sexuales o actos sexuales no deseados. Sin embargo, los resultados deben interpretarse con cautela debido a dos aspectos fundamentales: la baja frecuencia de este tipo de violencia en la muestra y la ausencia de significancia estadística en las variables analizadas, lo que limita

la identificación de asociaciones concluyentes, en relación con la edad de la mujer, no se evidencian diferencias significativas en la probabilidad de experimentar violencia sexual. Aunque algunos grupos etarios, como las mujeres de 41 a 49 años, presentan ligeras variaciones en el riesgo, estas no son estadísticamente relevantes, lo que indica que la violencia sexual no muestra un patrón claro según la edad de la mujer, de manera similar, la edad de la pareja tampoco presenta una asociación significativa con la violencia sexual en la mujer. Si bien se observan algunas fluctuaciones en los valores de riesgo entre distintos grupos de edad, estas no permiten establecer una relación consistente ni concluyente.

En cuanto a la ubicación geográfica, se identifican diferencias leves entre regiones; por ejemplo, las mujeres que residen en la Sierra o la Selva muestran una tendencia a mayor riesgo en comparación con otras zonas. No obstante, estas diferencias no alcanzan significancia estadística, por lo que no es posible afirmar que el lugar de residencia influya de manera determinante en la experiencia de violencia sexual en la mujer, a nivel provincial, se observa una alta variabilidad en los resultados; sin embargo, esta se encuentra influenciada por el reducido número de casos en algunas regiones, lo que genera estimaciones poco precisas y limita su interpretación. En este sentido, las aparentes diferencias entre provincias deben ser consideradas con cautela, ya que no reflejan necesariamente patrones reales en la población, estos hallazgos sugieren que la violencia sexual en la mujer no sigue una distribución claramente asociada a las variables demográficas tradicionales analizadas. A diferencia de otras formas de violencia, como la física o psicológica, este tipo de violencia parece depender de factores más específicos del contexto relacional o situacional, que no son captados por variables como la edad o el lugar de residencia, en la que resaltan la complejidad de la violencia sexual en la mujer y la necesidad de profundizar su análisis mediante variables más específicas que permitan

comprender mejor las condiciones en las que ocurre y, de esta manera, fortalecer las estrategias de prevención y atención.

Tabla 22. *Características reproductivas de la pareja y su asociación a la violencia sexual*

		Violencia sexual Hacia la mujer								
Variable	Categoría	Si		No		Odd Radio (OR)	IC_95% (OR)		Chi2	Sig.
		F	%	F	%		Li	Ls		
Estado civil	Viviendo juntos	202	43,3%	265	56,7%	0,500	0,372	0,671	48,32	0,000
	Casado	58	43,6%	75	56,4%	0,733	0,503	1,068		
Educación máxima	No viven juntos	125	73,5%	45	26,5%	3,632	2,491	5,296	3,69	0,595
	Sin educación	36	51,4%	34	48,6%	1,065	0,651	1,741		
	Inicial	49	55,1%	40	44,9%	1,258	0,807	1,961		
	Primaria C	24	54,5%	20	45,5%	1,213	0,659	2,235		
	Primaria I	72	46,2%	84	53,8%	0,824	0,580	1,172		
	Secundaria	159	48,2%	171	51,8%	0,880	0,662	1,172		
Método anticonceptivo	Superior	45	55,6%	36	44,4%	1,283	0,808	2,039	14,82	0,096
	Inyección	74	43,5%	96	56,5%	0,716	0,508	1,009		
	No usando	103	53,4%	90	46,6%	1,197	0,864	1,660		
	Ritmo	48	44,9%	59	55,1%	0,787	0,522	1,186		
	Condón	36	43,9%	46	56,1%	0,760	0,479	1,205		
	Est. femenina	55	61,8%	34	38,2%	1,721	1,094	2,707		
	Implantes	45	53,6%	39	46,4%	1,174	0,746	1,849		
	Píldora	15	48,4%	16	51,6%	0,935	0,456	1,919		
	DIU	8	66,7%	4	33,3%	2,021	0,604	6,769		
	Otro	0	0,0%	1	100,0%					
Embarazada	Est. masculina	1	100,0%	0	0,0%				0,61	0,434
	No	379	50,2%	376	49,8%	1,512	0,533	4,290		
Edad primera relación	Sí	6	40,0%	9	60,0%	0,661	0,233	1,876	7,36	0,061
	10 a 15	134	55,8%	106	44,2%	1,405	1,034	1,909		
	16 a 20	216	46,7%	247	53,3%	0,714	0,535	0,954		
	21 a 30	33	50,8%	32	49,2%	1,034	0,622	1,719		
	31 a 44	2	100,0%	0	0,0%					
Total, hijos nacidos	0	5	35,7%	9	64,3%	0,550	0,183	1,656	18,01	0,012
	1	75	43,6%	97	56,4%	0,718	0,511	1,011		
	2	98	45,4%	118	54,6%	0,773	0,564	1,059		
	3	105	53,3%	92	46,7%	1,194	0,864	1,652		
	4	52	65,8%	27	34,2%	2,071	1,271	3,374		
	5	27	48,2%	29	51,8%	0,926	0,537	1,596		
	6	11	57,9%	8	42,1%	1,386	0,551	3,485		
	7 o mas	12	70,6%	5	29,4%	2,445	0,853	7,008		

La Tabla 22 analiza la relación entre diversas características reproductivas de las mujeres y la presencia de violencia sexual, aunque estos resultados deben interpretarse con cautela debido a la baja frecuencia de este tipo de violencia en la muestra y a la limitada significancia estadística en algunas variables. A pesar de ello, se identifican ciertos

patrones relevantes que permiten comprender mejor las condiciones de vulnerabilidad de la mujer.

En primer lugar, el estado civil de la mujer se presenta como un factor importante. Las mujeres que no conviven con su pareja muestran una probabilidad considerablemente mayor de experimentar violencia sexual, lo que sugiere que contextos de separación o relaciones inestables pueden incrementar su exposición a situaciones de coerción. Por el contrario, las mujeres que viven con su pareja o están casadas presentan menor riesgo, lo que indicaría que la estabilidad en la convivencia puede tener un efecto protector relativo, aunque no elimina completamente la posibilidad de violencia, el uso de métodos anticonceptivos, se observa que la esterilización femenina se asocia con una mayor probabilidad de violencia sexual en la mujer, lo que podría estar relacionado con dinámicas de control o conflictos dentro de la relación. Asimismo, el uso del DIU también muestra una tendencia hacia mayor riesgo, aunque esta asociación debe interpretarse con prudencia debido al reducido número de casos. En contraste, el uso de métodos temporales como la inyección, el condón o el método del ritmo se relaciona con una menor probabilidad de violencia, aunque sin evidencia estadística concluyente.

En relación con el número de hijos, se identifica que las mujeres con mayor cantidad de hijos presentan una mayor probabilidad de experimentar violencia sexual. Este patrón sugiere que en contextos de alta carga familiar, donde pueden coexistir tensiones económicas y sobrecarga de responsabilidades, la mujer podría estar más expuesta a dinámicas de coerción dentro de la pareja, la edad de inicio de la vida sexual muestra una tendencia relevante, donde las mujeres que iniciaron relaciones sexuales a edades tempranas presentan una mayor probabilidad de experimentar violencia sexual, lo que podría estar vinculado a relaciones desiguales o a experiencias tempranas de vulnerabilidad, la condición de embarazo no muestra una asociación significativa con la

violencia sexual en la mujer, lo que sugiere que este factor no influye de manera determinante en este tipo de violencia dentro de la muestra analizada, estos resultados indican que, aunque la violencia sexual en la mujer es poco frecuente en términos absolutos y difícil de explicar únicamente a partir de variables reproductivas, existen ciertos perfiles de mayor vulnerabilidad, especialmente en mujeres que no conviven con su pareja, presentan alta paridad o han atravesado procesos reproductivos como la esterilización. Estos hallazgos resaltan la importancia de abordar la violencia sexual desde una perspectiva integral, incorporando la autonomía de la mujer, el consentimiento y el acompañamiento en los servicios de salud reproductiva.

Tabla 23. *Características económicas y sociales de la pareja y su asociación a la violencia sexual*

Variable	Categoría	Violencia sexual hacia la mujer				Odd Ratio (OR)	IC_95% (OR)		Chi2	Sig.
		Si		No			Li	Ls		
		F	%	F	%					
Tipo ingresos	Solo en efectivo	256	50,6%	250	49,4%	1,072	0,796	1,443	3,43	0,329
	No le pagan	92	46,0%	108	54,0%	0,805	0,583	1,112		
	Efectivo y especie	29	60,4%	19	39,6%	1,569	0,864	2,850		
	Solo en especie	8	50,0%	8	50,0%	1,000	0,371	2,692		
Decisor gasto de la pareja	Entrevistada y esposo/compañero	96	34,7%	181	65,3%	0,374	0,276	0,508	63,39	0,000
	Esposo/compañero	227	65,6%	119	34,4%	3,211	2,387	4,320		
	Entrevistada	62	42,2%	85	57,8%	0,677	0,471	0,974		
Tiene seguro salud	No	344	50,7%	334	49,3%	1,281	0,827	1,985	1,23	0,267
	Sí	41	44,6%	51	55,4%	0,781	0,504	1,209		
Frecuencia de consumo de alcohol de la pareja	Nunca	41	38,0%	67	62,0%	0,566	0,373	0,859	65,29	0,000
	Algunas veces	235	44,5%	293	55,5%	0,492	0,360	0,671		
	Mucha frecuencia	109	81,3%	25	18,7%	5,687	3,583	9,027		
Entrevista interrumpida por pareja	No	382	50,2%	379	49,8%	2,016	0,500	8,119	1,01	0,314
	Si	3	33,3%	6	66,7%	0,496	0,123	1,998		
Años del primer matrimonio	0 a 4	65	47,4%	72	52,6%	0,883	0,610	1,278	6,98	0,137
	5 a 10	99	45,8%	117	54,2%	0,793	0,578	1,087		
	11 a 20	139	51,3%	132	48,7%	1,083	0,806	1,456		
	21 a 30	78	58,2%	56	41,8%	1,493	1,024	2,176		
	31 a mas	4	33,3%	8	66,7%	0,495	0,148	1,657		

La Tabla 23 analiza cómo diversos factores estructurales —relacionados con el poder económico, las dinámicas del hogar, los hábitos de la pareja y las condiciones sociales—

se asocian con la presencia de violencia sexual en la mujer, entendida como la imposición de actos sexuales sin su consentimiento. Aunque este tipo de violencia presenta una baja frecuencia en la muestra, se identifican algunos factores clave que permiten reconocer situaciones de mayor vulnerabilidad en la mujer, uno de los hallazgos más relevantes se relaciona con la toma de decisiones sobre los gastos del hogar. Cuando el hombre ejerce control exclusivo sobre los recursos económicos, la mujer presenta una probabilidad significativamente mayor de experimentar violencia sexual. En cambio, cuando las decisiones económicas son compartidas o asumidas por la mujer, el riesgo disminuye de manera importante. Esto evidencia que la participación de la mujer en la gestión económica del hogar constituye un elemento fundamental para su protección, ya que reduce las condiciones de desigualdad y subordinación dentro de la relación.

Otro factor determinante es el consumo de alcohol por parte de la pareja. Se observa que, a medida que aumenta la frecuencia de consumo, también se incrementa la probabilidad de que la mujer experimente violencia sexual. En particular, cuando el consumo es frecuente, el riesgo se eleva de manera considerable, lo que pone en evidencia el papel del alcohol como un factor que favorece conductas de coerción, disminuye el autocontrol y agrava los conflictos dentro de la relación, en contraste, otras variables como el tipo de ingreso de la pareja no muestran una relación clara con la violencia sexual en la mujer, lo que sugiere que no es tanto la forma de ingreso lo que influye, sino el control que se ejerce sobre los recursos. De manera similar, la posesión de seguro de salud no presenta una asociación significativa, aunque la ausencia de este podría reflejar condiciones generales de mayor vulnerabilidad, asimismo, el hecho de que la entrevista haya sido interrumpida por la pareja no evidencia una relación estadísticamente significativa con la violencia sexual, lo que limita su utilidad como indicador de control en este contexto específico.

Por último, la duración de la unión muestra una tendencia a un mayor riesgo conforme aumentan los años de convivencia, aunque esta relación no es completamente consistente. Esto podría sugerir que, con el tiempo, la mujer puede estar expuesta a dinámicas acumulativas de poder y control dentro de la relación, en conjunto, estos resultados evidencian que la violencia sexual en la mujer, aunque menos frecuente, está estrechamente vinculada a situaciones de desigualdad en la toma de decisiones y a conductas de riesgo como el consumo de alcohol por parte de la pareja. Esto resalta la importancia de promover la autonomía económica de la mujer, así como de implementar estrategias preventivas que aborden las dinámicas de poder dentro del hogar y los factores conductuales asociados a la violencia.

Figura 5. *Árbol clasificación de presencia de violencia sexual*

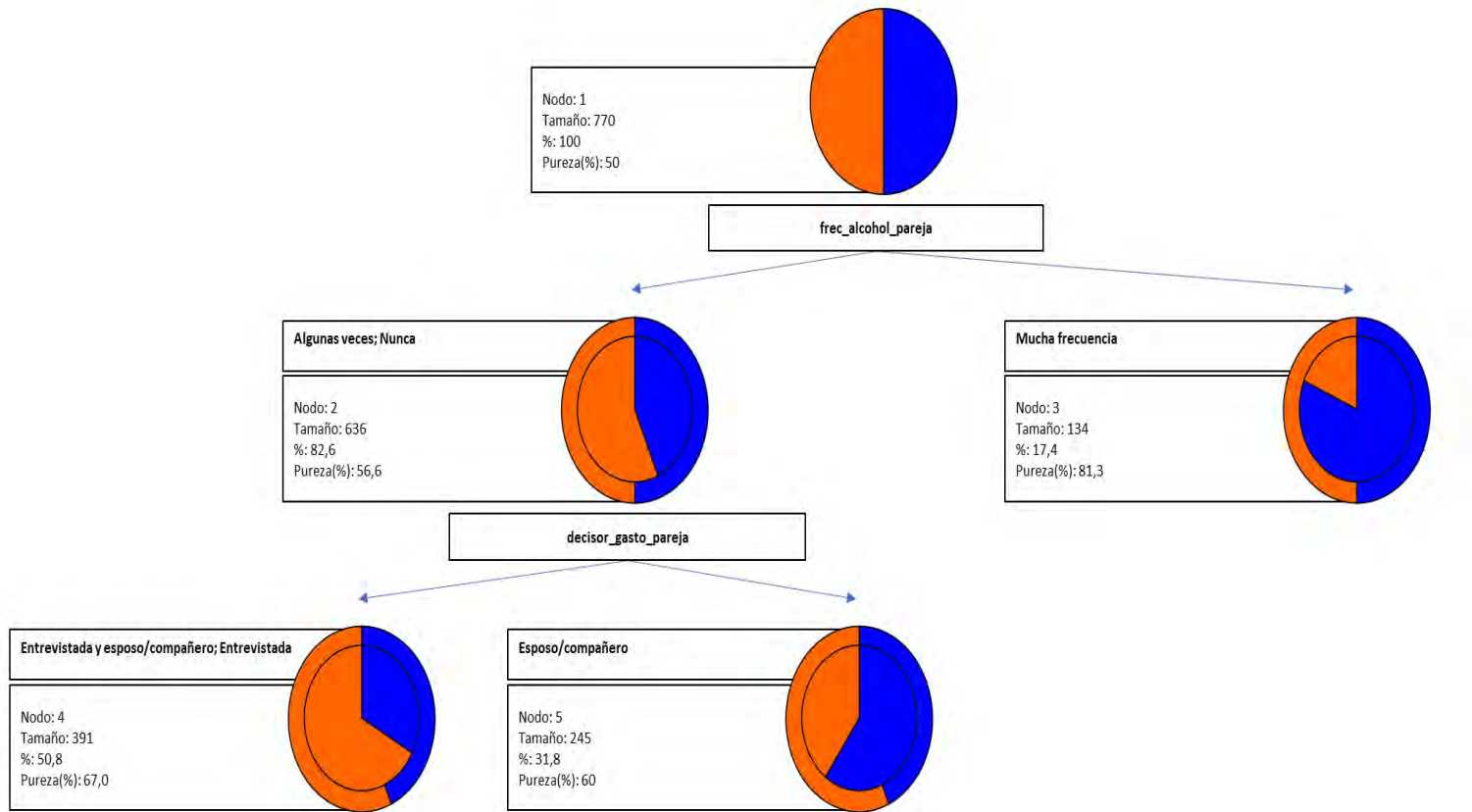
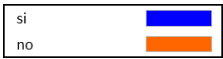


Tabla 24. Frecuencia según nodos y perfiles más sensibles a violencia sexual según OR

No dos	Características del nodo	Tamaño del nodo		Presencia de violencia sexual				Odd Radio (OR)	IC 95% (OR)	
		F	%	Si (pureza)		No			Li	Ls
No do 1	Inicial	77 0	14,1 %	38 5	50, 0%	38 5	50, 0%	1,000	0,8 19	1,2 21
No do 2	Consumo de alcohol por pareja algunas veces o nunca	63 6	11,6 %	27 6	43, 4%	36 0	56, 6%	0,636	0,5 19	0,7 81
No do 3	Consumo de alcohol por pareja mucha frecuencia	13 4	2,5%	10 9	81, 3%	25	18, 7%	4,914	3,1 43	7,6 84
No do 4	Consumo de alcohol por pareja algunas veces o nunca – la decisión del gasto es de la entrevistada o entrevistada y compañero	39 1	7,2%	12 9	33, 0%	26 2	67, 0%	0,390	0,3 07	0,4 96
No do 5	Consumo de alcohol por pareja algunas veces o nunca – la decisión del gasto es del esposo o compañero	24 5	4,5%	14 7	60, 0%	98	40, 0%	1,618	1,2 25	2,1 36

La Figura 5 y la Tabla 24 del árbol de clasificación de la violencia sexual permiten identificar de manera estructurada los perfiles de mayor riesgo en las mujeres peruanas, a partir de un modelo de árbol de decisión (CART) aplicado a una submuestra de 770 casos. En el nodo inicial se incluye a la totalidad de mujeres analizadas, donde la mitad reporta haber experimentado violencia sexual, estableciendo así el nivel base del fenómeno.

El principal factor que diferencia el riesgo en la mujer es el consumo de alcohol por parte de la pareja, el cual se posiciona como el predictor más influyente. Las mujeres cuyas parejas consumen alcohol con alta frecuencia presentan una probabilidad considerablemente mayor de experimentar violencia sexual, en comparación con aquellas cuyas parejas no consumen o lo hacen de forma ocasional. Este hallazgo evidencia que el consumo frecuente de alcohol incrementa significativamente la vulnerabilidad de la mujer frente a situaciones de coerción sexual, en el grupo de mujeres cuyas parejas no presentan

un consumo frecuente de alcohol, la toma de decisiones económicas dentro del hogar adquiere un papel determinante. Cuando la mujer participa en la toma de decisiones ya sea de manera compartida o individual, la probabilidad de experimentar violencia sexual disminuye de forma importante, lo que resalta el rol protector de la autonomía económica femenina. Por el contrario, cuando el hombre concentra el control de los recursos del hogar, la mujer presenta un mayor riesgo, lo que pone en evidencia la influencia de las desigualdades de poder dentro de la relación, el modelo identifica como situaciones de mayor vulnerabilidad aquellas en las que la mujer se encuentra en una relación donde la pareja consume alcohol con frecuencia, así como aquellas en las que, aun sin este factor, existe una distribución desigual del poder económico. Estos resultados muestran que la violencia sexual en la mujer está estrechamente vinculada a dinámicas de control y a conductas de riesgo dentro de la pareja.

Por ello se resalta la importancia de promover la participación de la mujer en la toma de decisiones del hogar y de abordar el consumo problemático de alcohol como estrategias clave para reducir la exposición de la mujer a este tipo de violencia.

Tabla 25. *Matriz de confusión y principales indicadores de la violencia sexual*

		Predicción de presencia de Violencia sexual hacia la mujer		
		Si	No	Total
Valor observado de violencia Sexual hacia la mujer	Si	262 (34%)	123 (16%)	385 (50%)
	No	129 (16,8%)	256 (33,2%)	385 (50%)
	Total	391 (50,8%)	379 (49,2%)	770 (100%)

Indicadores	Valor	Descripción
Exactitud	0,673	Proporción total de asociaciones correctas (aciertos / total)
Sensibilidad	0,665	Proporción de casos positivos reales se identificaron correctamente (TP / TP+FN)
Especificidad	0,681	Proporción de casos negativos reales se identificaron correctamente (TN / TN+FP)
Valor predictivo positivo	0,675	Probabilidad de que sea realmente positivo (TP / TP+FP)
Valor predictivo negativo	0,670	Probabilidad de que sea realmente negativo (TN / TN+FN)
Exactitud balanceada	0,673	Promedio de sensibilidad y Especificidad

La Tabla 25 presenta la matriz de confusión y los principales indicadores de desempeño del modelo de árbol de clasificación utilizado para predecir la presencia de violencia sexual en la mujer en una submuestra de 770 casos. En términos generales, el modelo muestra un rendimiento equilibrado, con una exactitud global del 67,3 %, lo que indica que aproximadamente siete de cada diez mujeres fueron correctamente clasificadas según su experiencia de violencia sexual.

A diferencia de otros modelos analizados, este presenta un comportamiento más homogéneo entre sus indicadores. La sensibilidad (66,5 %) evidencia que el modelo logra identificar correctamente a una proporción importante de mujeres que han sufrido violencia sexual, lo que representa una mejora en la capacidad de detección de casos reales. De igual manera, la especificidad (68,1 %) muestra que el modelo también es capaz de reconocer adecuadamente a las mujeres que no han experimentado este tipo de violencia. En cuanto a los valores predictivos, tanto el valor predictivo positivo (67,5 %) como el valor predictivo negativo (67,0 %) presentan valores similares, lo que sugiere que el modelo ofrece un desempeño equilibrado al momento de confirmar o descartar la presencia de violencia sexual en la mujer. Asimismo, la exactitud balanceada coincide con la exactitud global, lo cual se explica por la distribución equitativa de los casos dentro de la muestra analizada, en conjunto, indican que, a pesar de la menor frecuencia y la complejidad inherente de la violencia sexual, el modelo logra una capacidad predictiva aceptable en la identificación de mujeres en riesgo. Además, se refuerza la relevancia de factores como el consumo de alcohol por parte de la pareja y la toma de decisiones económicas en el hogar, los cuales permiten comprender mejor las condiciones que incrementan la vulnerabilidad de la mujer frente a este tipo de violencia, el modelo podría constituir una herramienta útil para la detección temprana de situaciones de riesgo en la

mujer, especialmente si se fortalece con mayor información y se integra en estrategias de prevención dentro de los servicios de salud y programas comunitarios orientados a la protección de la mujer.

4.3.3. Violencia general y factores asociados

Tabla 26. Características demográficas de la pareja y su asociación a la violencia general

Variable	Categoría	Violencia hacia la mujer				Odd Radio (OR)	IC_95% (OR)		Chi2	Sig.
		Si		No			Li	Ls		
		F	%	F	%					
Edad mujer	15 a 20	379	50,9%	366	49,1%	1,037	0,895	1,202	92,48	0,000
	21 a 30	2702	47,9%	2937	52,1%	0,880	0,825	0,939		
	31 a 40	3259	48,1%	3514	51,9%	0,879	0,826	0,936		
	41 a 49	1800	57,6%	1323	42,4%	1,463	1,352	1,583		
Edad pareja	15 a 20	118	55,1%	96	44,9%	1,233	0,940	1,616	30,83	0,000
	21 a 30	1760	48,8%	1848	51,2%	0,939	0,872	1,011		
	31 a 40	3391	48,3%	3624	51,7%	0,890	0,836	0,947		
	41 a 50	2402	52,3%	2189	47,7%	1,138	1,063	1,219		
	51 a 60	421	55,2%	341	44,8%	1,247	1,078	1,444		
	61 a 78	48	53,3%	42	46,7%	1,144	0,755	1,732		
Región	Sierra	2946	53,3%	2586	46,7%	1,218	1,142	1,300	45,25	0,000
	Resto Costa	2206	47,7%	2423	52,3%	0,877	0,819	0,939		
	Selva	2170	50,0%	2169	50,0%	1,001	0,933	1,073		
	Lima metropolitana	818	46,0%	962	54,0%	0,834	0,755	0,920		
Provincia	Lima	3663	47,4%	4065	52,6%	0,820	0,771	0,872	81,87	0,000
	Piura	765	52,2%	700	47,8%	1,102	0,990	1,228		
	La Libertad	725	51,0%	697	49,0%	1,044	0,936	1,164		
	Arequipa	525	51,4%	497	48,6%	1,060	0,934	1,203		
	Cajamarca	485	53,4%	423	46,6%	1,156	1,011	1,322		
	Junín	491	56,8%	374	43,2%	1,333	1,161	1,530		
	Cusco	365	54,5%	305	45,5%	1,206	1,033	1,409		
	Lambayeque	314	53,0%	279	47,0%	1,130	0,959	1,332		
	Puno	255	46,1%	298	53,9%	0,851	0,718	1,009		
	Áncash	185	59,1%	128	40,9%	1,456	1,160	1,827		
	Loreto	134	53,4%	117	46,6%	1,148	0,894	1,474		
	San Martín	97	53,3%	85	46,7%	1,143	0,853	1,532		
	Ica	55	35,5%	100	64,5%	0,547	0,393	0,761		
	Otras	81	52,9%	72	47,1%	1,126	0,819	1,549		

La Tabla 26 analiza cómo las características demográficas de la mujer y su pareja se relacionan con la ocurrencia de violencia general definida como la vivencia de al menos un tipo de violencia (física, psicológica o sexual) en la ENDES Perú 2024. Los resultados

evidencian asociaciones claras y estadísticamente significativas ($p < 0,001$), destacando especialmente el papel de las condiciones de vida de las mujeres.

En primer lugar, la edad de la mujer muestra una relación importante con la violencia. Se observa que las mujeres de 41 a 49 años presentan una mayor probabilidad de experimentar violencia general ($OR = 1,463$; IC 95 %: 1,352–1,583). Este hallazgo sugiere que, conforme avanza el ciclo de vida, las mujeres pueden encontrarse más expuestas a dinámicas de pareja marcadas por el desgaste emocional, patrones de control consolidados o mayores dificultades para abandonar relaciones conflictivas. En contraste, las mujeres más jóvenes, especialmente aquellas entre 21 y 40 años, presentan una menor probabilidad de sufrir violencia ($OR \approx 0,88$), lo que podría vincularse con relaciones más recientes, mayor acceso a información y redes de apoyo, así como cambios generacionales en torno a la equidad de género.

Respecto a la edad de la pareja, también se identifican efectos relevantes. Las mujeres con compañeros de 41 a 50 años ($OR = 1,138$) y de 51 a 60 años ($OR = 1,247$) tienen mayor riesgo de experimentar violencia, lo que podría asociarse a tensiones propias de la mediana edad o a la persistencia de roles tradicionales que refuerzan relaciones de poder desiguales. Aunque se observa un incremento del riesgo en parejas más jóvenes (15–20 años), este resultado no es concluyente debido a la amplitud del intervalo de confianza.

En el ámbito territorial, el lugar de residencia influye de manera significativa en la experiencia de violencia de las mujeres. Vivir en la Sierra se asocia con un mayor riesgo ($OR = 1,218$), lo que podría explicarse por la persistencia de normas socioculturales que limitan la autonomía femenina y por un menor acceso a servicios de protección. Por el contrario, residir en Lima Metropolitana ($OR = 0,834$) o en la costa fuera de Lima ($OR = 0,877$) se relaciona con una menor probabilidad de violencia, posiblemente debido a

mayores niveles de urbanización, educación y disponibilidad de recursos institucionales de apoyo para las mujeres.

A nivel regional, se evidencian diferencias importantes. Departamentos como Áncash (OR = 1,456), Junín (OR = 1,333) y Cusco (OR = 1,206) muestran mayores niveles de riesgo, lo que refuerza la idea de que las mujeres en contextos andinos enfrentan mayores desigualdades estructurales. En contraste, Ica presenta un efecto protector significativo (OR = 0,547), posiblemente relacionado con mayores oportunidades económicas y laborales para las mujeres, lo que favorece su autonomía y reduce su tolerancia frente a situaciones de violencia.

En conjunto, estos resultados evidencian que la violencia general contra las mujeres no es uniforme, sino que está condicionada por factores como la edad, la etapa del ciclo de vida y el contexto geográfico. Esto resalta la necesidad de implementar políticas públicas con enfoque diferenciado, priorizando a las mujeres en mayor situación de vulnerabilidad especialmente aquellas de mayor edad y residentes en zonas rurales de la sierra, así como fortalecer estrategias que promuevan su autonomía, acceso a servicios y equidad de género a lo largo de toda su vida.

Tabla 27. Características reproductivas de la pareja y su asociación a la violencia general

		Violencia hacia la mujer								
Variable	Categoría	Si		No		Odd Radio (OR)	IC_95% (OR)		Chi2	Sig.
		F	%	F	%		Li	Ls		
Estado civil	Viviendo juntos	5007	45,8%	5928	54,2%	0,596	0,558	0,637	1172,68	0,000
	Casado	1544	43,7%	1993	56,3%	0,722	0,670	0,778		
	No viven juntos	1589	87,9%	219	12,1%	8,773	7,589	10,142		
Educación Máxima	Sin educación	673	52,0%	621	48,0%	1,091	0,974	1,223	74,48	0,000
	Inicial	905	53,4%	791	46,6%	1,162	1,051	1,285		
	Primaria C	600	56,4%	463	43,6%	1,319	1,164	1,496		
	Primaria I	1630	47,3%	1814	52,7%	0,873	0,810	0,941		
	Secundaria	3411	47,8%	3726	52,2%	0,854	0,803	0,909		
Método Anticonceptivo	Superior	921	56,0%	725	44,0%	1,305	1,178	1,446	254,59	0,000
	Inyección	1726	44,8%	2124	55,2%	0,762	0,709	0,820		
	No usando	2048	60,0%	1367	40,0%	1,666	1,543	1,798		
	Ritmo	1289	48,9%	1345	51,1%	0,951	0,874	1,033		
	Condón	803	42,9%	1068	57,1%	0,725	0,658	0,799		
	Est femenina	923	54,3%	776	45,7%	1,214	1,097	1,342		
	Implantes	869	51,2%	828	48,8%	1,055	0,954	1,167		
	Pildora	355	41,8%	494	58,2%	0,706	0,614	0,812		
	DIU	63	46,3%	73	53,7%	0,862	0,614	1,209		
	Otro	49	53,3%	43	46,7%	1,140	0,756	1,720		
Est masculina	15	40,5%	22	59,5%	0,681	0,353	1,314			
Embarazada	No	7946	50,2%	7889	49,8%	1,303	1,078	1,576	7,51	0,006
	Sí	194	43,6%	251	56,4%	0,767	0,635	0,928		
Edad	10 a 15	2443	56,7%	1863	43,3%	1,445	1,347	1,550	143,21	0,000
Primera relación	16 a 20	4839	48,8%	5074	51,2%	0,886	0,832	0,943		
	21 a 30	829	41,9%	1148	58,1%	0,691	0,628	0,760		
	31 a 44	29	34,5%	55	65,5%	0,526	0,335	0,825		
Total	0	152	45,2%	184	54,8%	0,823	0,662	1,022	239,53	0,000
Hijos nacidos	1	1752	43,6%	2262	56,4%	0,713	0,663	0,766		
	2	2539	46,8%	2886	53,2%	0,825	0,773	0,881		
	3	1912	54,1%	1625	45,9%	1,231	1,142	1,326		
	4	924	59,1%	640	40,9%	1,501	1,350	1,668		
	5	425	59,4%	291	40,6%	1,486	1,276	1,730		
	6	230	62,3%	139	37,7%	1,674	1,353	2,071		
	7 o mas	206	64,6%	113	35,4%	1,844	1,463	2,325		

La Tabla 27 examina la relación entre las características reproductivas de las mujeres y sus parejas y la presencia de violencia general definida como la experiencia de al menos una forma de violencia (física, psicológica o sexual), evidenciando asociaciones sólidas y estadísticamente significativas que resaltan cómo las trayectorias reproductivas de las mujeres se vinculan estrechamente con dinámicas de poder dentro de la relación.

Uno de los resultados más relevantes se observa en el estado civil. Las mujeres que no conviven con su pareja presentan una probabilidad significativamente mayor de experimentar violencia general (OR = 8,773; IC 95 %: 7,589–10,142). Esto sugiere que

las situaciones de separación, inestabilidad o relaciones no convivientes pueden exponer a las mujeres a mayores niveles de control, conflictos y conductas abusivas. En contraste, las mujeres casadas ($OR = 0,722$) y aquellas en unión consensual ($OR = 0,596$) presentan menor riesgo, lo que indicaría que la convivencia estable puede actuar como un factor que atenúa aunque no elimina la violencia, en relación con el nivel educativo, se identifican asociaciones significativas que reflejan una dinámica compleja. Las mujeres con educación primaria completa ($OR = 1,319$) y educación superior ($OR = 1,305$) presentan mayor probabilidad de sufrir violencia, lo que podría estar vinculado a tensiones generadas cuando las mujeres alcanzan niveles educativos que cuestionan los roles tradicionales de género. Por otro lado, contar con educación secundaria completa se asocia con un menor riesgo ($OR = 0,854$), posiblemente porque favorece cierto grado de autonomía sin generar una ruptura abrupta con las normas sociales predominantes.

Respecto al uso de métodos anticonceptivos, se evidencia la importancia de la autonomía reproductiva. Las mujeres que no utilizan ningún método presentan un mayor riesgo de violencia ($OR = 1,666$), lo que podría reflejar limitaciones en la toma de decisiones sobre su propio cuerpo o conflictos en torno a la planificación familiar. En cambio, el uso de métodos temporales como la inyección, el condón o la píldora se asocia con menor probabilidad de violencia (OR entre $0,706$ y $0,762$), sugiriendo relaciones más equitativas. Sin embargo, la esterilización femenina muestra un incremento en el riesgo ($OR = 1,214$), lo que podría indicar contextos donde esta decisión no siempre es plenamente autónoma o genera tensiones en la pareja, la edad de inicio de la vida sexual también constituye un factor importante. Las mujeres que inician su vida sexual antes de los 16 años presentan mayor probabilidad de experimentar violencia ($OR = 1,445$), lo que podría estar relacionado con relaciones desiguales en términos de poder o menor

capacidad de negociación. A medida que la edad de inicio aumenta, el riesgo disminuye, siendo menor en aquellas que comienzan después de los 21 años ($OR = 0,691$).

En cuanto al número de hijos, se observa una relación directa entre mayor fecundidad y mayor riesgo de violencia. Las mujeres con cuatro o más hijos presentan un incremento progresivo del riesgo, destacando aquellas con siete o más hijos ($OR = 1,844$). Este patrón podría explicarse por la sobrecarga de responsabilidades, la dependencia económica y las tensiones en el hogar. Por otro lado, el hecho de estar embarazada se asocia con menor probabilidad de violencia ($OR = 0,767$), posiblemente debido a factores culturales o a una mayor protección social durante esta etapa.

En suma la violencia contra las mujeres está profundamente relacionada con sus experiencias reproductivas, sus decisiones sobre el cuerpo y las condiciones de su relación de pareja. Esto pone en evidencia la necesidad de fortalecer políticas públicas que integren la prevención de la violencia en los servicios de salud sexual y reproductiva, promuevan la autonomía de las mujeres en la toma de decisiones y brinden apoyo específico a aquellas en situaciones de mayor vulnerabilidad, como las que atraviesan procesos de separación o tienen alta carga reproductiva.

Tabla 28. Características económicas y sociales de la pareja y su asociación a la violencia general

		Violencia hacia la mujer								
Variable	Categoría	Si		No		Odd Radio (OR)	IC_95% (OR)		Chi2	Sig.
		F	%	F	%		Li	Ls		
Tipo ingresos	Solo en efectivo	4940	51,3%	4693	48,7%	1,134	1,065	1,207	115,62	0,000
	No le pagan	2405	45,1%	2928	54,9%	0,746	0,699	0,797		
	Efectivo y especie	647	60,4%	424	39,6%	1,571	1,385	1,783		
	Solo en especie	148	60,9%	95	39,1%	1,568	1,210	2,033		
Decisor gasto de la pareja	Entrevistada y esposo/compañero	3298	40,4%	4860	59,6%	0,460	0,432	0,489	985,80	0,000
	Esposo/compañero	3316	68,6%	1517	31,4%	3,001	2,795	3,222		
	Entrevistada	1526	46,4%	1763	53,6%	0,835	0,773	0,901		
Tiene seguro salud	No	7011	51,6%	6584	48,4%	1,468	1,350	1,596	81,32	0,000
	Si	1129	42,0%	1556	58,0%	0,681	0,627	0,741		
Frecuencia de consumo de alcohol de la pareja	Nunca	1185	43,0%	1569	57,0%	0,714	0,657	0,775	689,35	0,000
	Algunas veces	6102	48,5%	6484	51,5%	0,765	0,710	0,823		
	Mucha frecuencia	853	90,7%	87	9,3%	10,835	8,671	13,540		
Entrevista interrumpida por pareja	No	8075	50,0%	8077	50,0%	0,969	0,684	1,372	0,03	0,859
	Si	65	50,8%	63	49,2%	1,032	0,729	1,461		
Años del primer matrimonio	0 a 4	1328	43,3%	1741	56,7%	0,717	0,662	0,776	227,03	0,000
	5 a 10	2198	46,0%	2577	54,0%	0,799	0,746	0,854		
	11 a 20	2965	51,8%	2754	48,2%	1,121	1,051	1,195		
	21 a 30	1471	59,7%	991	40,3%	1,591	1,458	1,736		
	31 a mas	178	69,8%	77	30,2%	2,341	1,789	3,064		

La Tabla 28 analiza la relación entre las características económicas y sociales de la pareja y la presencia de violencia general entendida como la vivencia de al menos una forma de violencia (física, psicológica o sexual) en la ENDES Perú 2024, evidenciando asociaciones significativas que ponen en relieve cómo las condiciones estructurales influyen directamente en la situación de las mujeres.

En primer lugar, el tipo de ingreso de la pareja muestra una relación clara con el riesgo de violencia. Las mujeres cuyas parejas perciben ingresos solo en efectivo o combinan efectivo y especie presentan una mayor probabilidad de sufrir violencia (OR entre 1,134 y 1,571). Esto podría estar vinculado a contextos de inestabilidad laboral, informalidad o inseguridad económica, que generan tensiones en el hogar y favorecen dinámicas de control sobre las mujeres. En contraste, cuando la pareja no percibe ingresos, el riesgo disminuye (OR = 0,746), aunque este resultado debe interpretarse con cautela, ya que

podría reflejar situaciones de dependencia económica o arreglos familiares que no necesariamente implican ausencia de violencia.

Uno de los aspectos más relevantes es la toma de decisiones sobre los gastos del hogar. Cuando el hombre decide de manera unilateral, las mujeres tienen aproximadamente tres veces más probabilidad de experimentar violencia (OR = 3,001; IC 95 %: 2,795–3,222). Por el contrario, cuando las decisiones son compartidas, el riesgo disminuye considerablemente (OR = 0,460). Este hallazgo evidencia que la violencia contra las mujeres está estrechamente vinculada con relaciones desiguales de poder y con la falta de autonomía en el ámbito doméstico.

En relación con el acceso a seguro de salud, se observa que las mujeres que no cuentan con seguro presentan un mayor riesgo de violencia (OR = 1,468). Esto sugiere que la exclusión de los sistemas de protección social incrementa su vulnerabilidad, ya sea por limitaciones en el acceso a servicios de apoyo, menor independencia económica o por residir en entornos más desfavorecidos.

El consumo de alcohol por parte de la pareja se presenta como uno de los factores de mayor impacto. Cuando el consumo es frecuente, el riesgo de violencia hacia las mujeres aumenta de manera muy significativa (OR = 10,835; IC 95 %: 8,671–13,540). Incluso el consumo ocasional se asocia con un incremento del riesgo en comparación con la abstinencia. Esto pone en evidencia que el alcohol actúa como un factor que intensifica conductas agresivas y diversas formas de maltrato, la interrupción de la entrevista por parte de la pareja no mostró una asociación significativa con la violencia ($p = 0,859$), lo que indica que este indicador no refleja adecuadamente las dinámicas de control en este contexto, la duración de la relación también influye en la experiencia de violencia. A medida que aumentan los años de convivencia, el riesgo para las mujeres se incrementa de manera progresiva. Aquellas que tienen entre 21 y 30 años de unión presentan un

mayor riesgo (OR = 1,591), y este es aún más elevado en relaciones de 31 años o más (OR = 2,341). Este patrón podría explicarse por la normalización de conductas violentas, la acumulación de conflictos o las dificultades para romper vínculos prolongados, especialmente en contextos donde existen barreras sociales o económicas para la separación, los resultados muestran que la violencia contra las mujeres no es un fenómeno aislado, sino que está profundamente relacionado con las condiciones económicas, las relaciones de poder dentro del hogar y las trayectorias de vida en pareja, por esto resulta fundamental promover políticas públicas que fortalezcan la autonomía de las mujeres, impulsen la toma de decisiones equitativa en el hogar, amplíen el acceso a servicios de protección social y aborden factores de riesgo como el consumo de alcohol, especialmente en contextos de mayor vulnerabilidad.

Figura 6. Árbol clasificación de presencia de violencia general

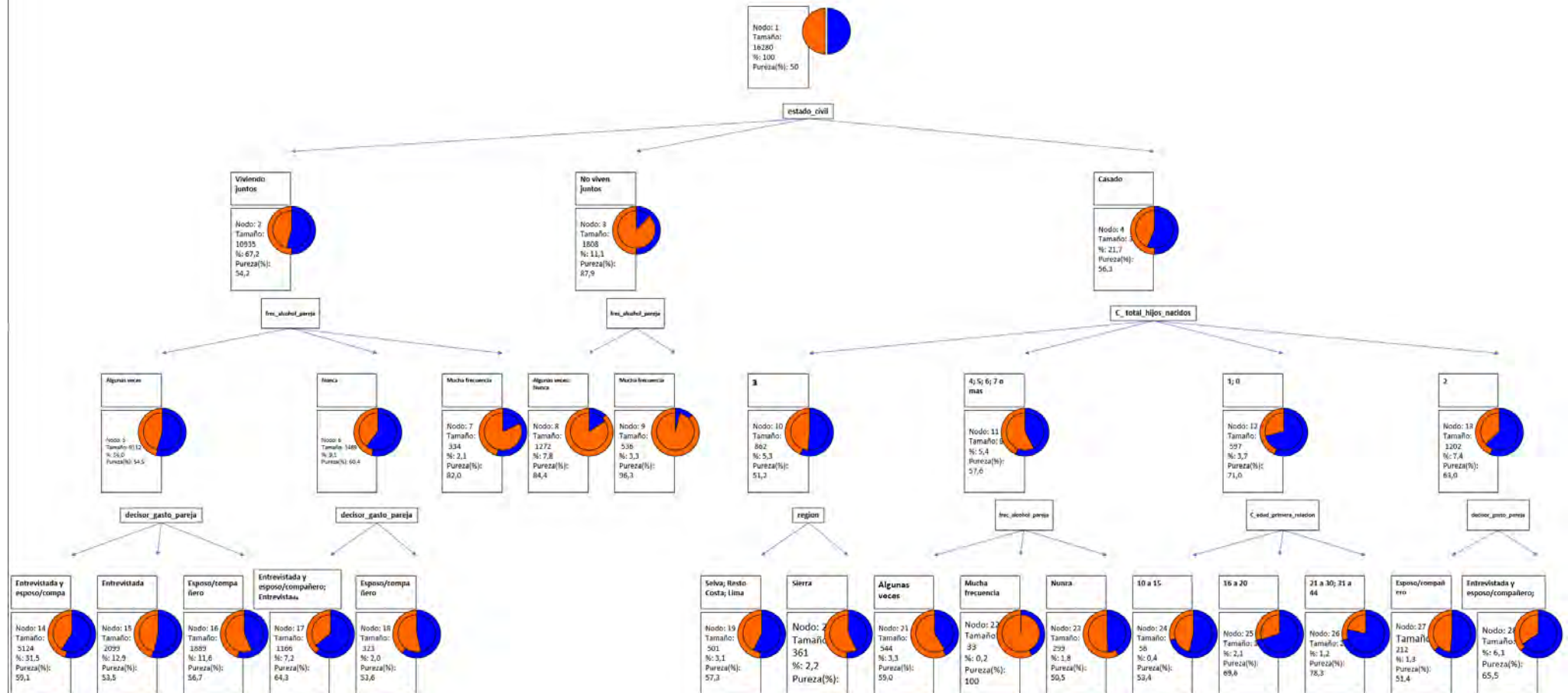
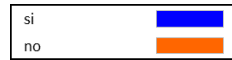


Tabla 29. Frecuencia según nodos y perfiles más sensibles a violencia general según OR

Nodos	Características del nodo	Tamaño del nodo		Presencia de violencia				Odd Radio (OR)	IC_95% (OR)	
		F	%	Si (pureza)		No			Li	Ls
Nodo 1	Inicial	16280	297,7%	8140	50,0%	8140	50,0%	1,000	0,957	1,044
Nodo 2	Viven juntos	10935	200,0%	5928	54,2%	5007	45,8%	1,289	1,231	1,350
Nodo 3	No viven juntos	1808	33,1%	219	12,1%	1589	87,9%	0,126	0,109	0,145
Nodo 4	Casados	3537	64,7%	1993	56,3%	1544	43,7%	1,331	1,241	1,428
Nodo 5	Viven juntos – el consume a veces alcohol	9112	166,6%	4968	54,5%	4144	45,5%	1,286	1,225	1,350
Nodo 6	Viven juntos – el nunca consume alcohol	1489	27,2%	900	60,4%	589	39,6%	1,559	1,402	1,734
Nodo 7	Viven juntos – el consume alcohol frecuentemente	334	6,1%	60	18,0%	274	82,0%	0,216	0,163	0,286
Nodo 8	No viven juntos – el consume alcohol nunca o algunas veces	1272	23,3%	199	15,6%	1073	84,4%	0,175	0,151	0,204
Nodo 9	No viven juntos – el consume alcohol frecuentemente	536	9,8%	20	3,7%	516	96,3%	0,038	0,024	0,059
Nodo 10	Casados – con 3 hijos	862	15,8%	441	51,2%	421	48,8%	1,049	0,916	1,201
Nodo 11	Casados – con 4 o más hijos	876	16,0%	371	42,4%	505	57,6%	0,728	0,636	0,834
Nodo 12	Casados – sin hijos o 1 hijo	597	10,9%	424	71,0%	173	29,0%	2,490	2,083	2,975
Nodo 13	Casados – con 2 hijos	1202	22,0%	757	63,0%	445	37,0%	1,735	1,540	1,955
Nodo 14	Viven juntos – el consume a veces alcohol – ambos deciden el gasto	5124	93,7%	3027	59,1%	2097	40,9%	1,545	1,454	1,641
Nodo 15	Viven juntos – el consume a veces alcohol – ella decide el gasto	2099	38,4%	1123	53,5%	976	46,5%	1,162	1,063	1,269
Nodo 16	Viven juntos – el consume a veces alcohol – el decide el gasto	1889	34,5%	818	43,3%	1071	56,7%	0,751	0,684	0,825
Nodo 17	Viven juntos – el nunca consume alcohol – ella o ambos deciden el gasto	1166	21,3%	750	64,3%	416	35,7%	1,842	1,630	2,080
Nodo 18	Viven juntos – el nunca consume alcohol - el decide el gasto	323	5,9%	150	46,4%	173	53,6%	0,866	0,695	1,079
Nodo 19	Casados – con 3 hijos – viven en lima, selva o costa	501	9,2%	287	57,3%	214	42,7%	1,347	1,127	1,610
Nodo 20	Casados – con 3 hijos – viven en sierra	361	6,6%	154	42,7%	207	57,3%	0,742	0,601	0,915
Nodo 21	Casados – con 4 o más hijos -el consume alcohol algunas veces	544	9,9%	223	41,0%	321	59,0%	0,690	0,581	0,820
Nodo 22	Casados – con 4 o más hijos -el consume alcohol frecuentemente	33	0,6%	0,5	1,5%	33	100,0%	0,015	0,001	0,247
Nodo 23	Casados – con 4 o más hijos -el nunca consume alcohol	299	5,5%	148	49,5%	151	50,5%	0,980	0,780	1,231
Nodo 24	Casados – sin hijos o 1 hijo – ella inicio su vida sexual a 10 a 15 años	58	1,1%	31	53,4%	27	46,6%	1,148	0,685	1,925
Nodo 25	Casados – sin hijos o 1 hijo – ella inicio su vida sexual a 16 a 20 años	336	6,1%	234	69,6%	102	30,4%	2,313	1,831	2,922
Nodo 26	Casados – sin hijos o 1 hijo – ella inicio su vida sexual de 21 a mas años	203	3,7%	159	78,3%	44	21,7%	3,639	2,605	5,086
Nodo 27	Casados – con 2 hijos - el decide el gasto	212	3,9%	109	51,4%	103	48,6%	1,059	0,808	1,387
Nodo 28	Casados – con 2 hijos – ella o ambos deciden el gasto	990	18,1%	648	65,5%	342	34,5%	1,932	1,692	2,206

La Figura 6 y la Tabla 29 presentan los resultados del árbol de clasificación (CART) sobre la presencia de violencia general definida como la experiencia de al menos una forma de violencia (física, psicológica o sexual) en una muestra de 16 280 mujeres de la ENDES Perú 2024. Este enfoque permite identificar de manera jerárquica los perfiles de mujeres con mayor o menor probabilidad de sufrir violencia, evidenciando patrones claros en función de sus características relacionales y contextuales.

En el nodo inicial, que incluye a toda la muestra, se observa que el 50 % de las mujeres reportó haber experimentado violencia general, lo que establece un punto de referencia importante. A partir de este nodo, el factor que primero divide a las mujeres es el estado civil, el cual se posiciona como el principal determinante del riesgo. En este sentido, se distinguen tres grupos: mujeres que viven con su pareja, aquellas que no conviven y las que están casadas, mostrando que la forma en que se estructura la relación influye significativamente en la experiencia de violencia.

El grupo de mayor vulnerabilidad corresponde a las mujeres que no viven con su pareja, donde la gran mayoría reporta haber sufrido violencia. Esta situación demuestra que la inestabilidad o discontinuidad en la relación expone a las mujeres a contextos más propensos al conflicto, control y abuso. Dentro de este grupo, el consumo de alcohol por parte de la pareja intensifica aún más el riesgo, especialmente cuando es frecuente, configurando un escenario de alta peligrosidad para las mujeres. Incluso en ausencia de consumo o cuando este es ocasional, el riesgo se mantiene elevado, lo que refuerza la idea de que la no convivencia por sí misma constituye un factor crítico.

En el caso de las mujeres que conviven con su pareja, el consumo de alcohol vuelve a ser un factor relevante en la diferenciación del riesgo. Sin embargo, los resultados muestran variaciones que deben interpretarse con cautela, ya que algunos subgrupos presentan comportamientos aparentemente contradictorios, posiblemente debido a tamaños muestrales

reducidos o a la influencia de variables no consideradas. En este grupo, la toma de decisiones en el hogar, particularmente en relación con los gastos, también influye en la probabilidad de violencia. Las mujeres que participan en decisiones compartidas presentan un riesgo intermedio, mientras que aquellas que deciden por sí mismas muestran una ligera reducción del riesgo. Estos hallazgos reflejan la importancia de la autonomía y la negociación dentro de la relación, entre las mujeres casadas, el número de hijos emerge como un factor clave. Las mujeres casadas sin hijos o con un solo hijo presentan un mayor riesgo de violencia, lo que podría estar asociado a tensiones dentro de la relación, expectativas no cumplidas o conflictos en torno a la maternidad. Este riesgo es aún más elevado en quienes iniciaron su vida sexual a una edad más tardía. Asimismo, las mujeres con dos hijos también muestran niveles importantes de riesgo, especialmente cuando las decisiones en el hogar son compartidas. En contraste, aquellas con tres o más hijos presentan una menor probabilidad de violencia, lo que podría sugerir una cierta estabilización en la relación o una mayor adaptación a las dinámicas familiar, en la que se pone en evidencia que la violencia contra las mujeres no responde a un único factor, sino que es el resultado de la interacción entre condiciones relacionales, comportamientos de la pareja y características del entorno familiar. El árbol de clasificación permite identificar perfiles específicos de mayor riesgo, lo que resulta clave para el diseño de intervenciones focalizadas, lo que se hace necesario priorizar acciones dirigidas a mujeres en situaciones de no convivencia, fortalecer estrategias de prevención del consumo problemático de alcohol en las parejas y promover relaciones más equitativas que garanticen la autonomía y el bienestar de las mujeres.

Tabla 30. Matriz de confusión y principales indicadores de la violencia general

		Predicción de presencia De violencia hacia la mujer		
		Si	No	Total
Valor observado de violencia	Si	3819 (23,5%)	4321 (26,5%)	8140 (50%)
Hacia la mujer	No	1772 (10,9%)	6368 (39,1%)	8140 (50%)
Total		5591 (34,3%)	10689 (65,7%)	16280 (100%)

Indicadores	Valor	Descripción
Exactitud	0,626	Proporción total de asociaciones correctas (aciertos / total)
Sensibilidad	0,782	Proporción de casos positivos reales se identificaron correctamente (TP / TP+FN)
Especificidad	0,469	Proporción de casos negativos reales se identificaron correctamente (TN / TN+FP)
Valor predictivo positivo	0,596	Probabilidad de que sea realmente positivo (TP / TP+FP)
Valor predictivo negativo	0,683	Probabilidad de que sea realmente negativo (TN / TN+FN)
Exactitud balanceada	0,626	Promedio de sensibilidad y especificidad

La Tabla 30 muestra la matriz de confusión y los principales indicadores de desempeño del modelo de árbol de clasificación empleado para predecir la presencia de violencia general — entendida como la experiencia de violencia física, psicológica o sexual— en mujeres de la ENDES Perú 2024 (n = 16 280). En términos generales, el modelo alcanza una exactitud global de 62,6 %, lo que indica que aproximadamente seis de cada diez clasificaciones coinciden con la situación real de las mujeres, el desempeño del modelo no es homogéneo entre quienes experimentan y quienes no experimentan violencia. Por un lado, presenta una alta sensibilidad (78,2 %), lo que significa que logra identificar correctamente a cerca de ocho de cada diez mujeres que han sido víctimas de violencia. Este resultado es especialmente relevante desde una perspectiva de salud pública y protección social, ya que prioriza la detección de mujeres en situación de riesgo o vulnerabilidad, permitiendo una intervención oportuna, esta capacidad de detección se acompaña de una baja especificidad (46,9 %), lo que implica que un número importante de mujeres que no han sufrido violencia son clasificadas erróneamente como víctimas (falsos positivos). En otras palabras, el modelo tiende a sobreestimar la presencia de violencia, lo cual responde a una lógica en la que se privilegia no dejar sin identificar a mujeres que realmente podrían necesitar apoyo.

Esta tendencia también se refleja en los valores predictivos, el valor predictivo positivo (59,6 %) indica que, cuando el modelo identifica a una mujer como víctima de violencia, esta clasificación es correcta en alrededor de seis de cada diez casos, este valor predictivo negativo (68,3 %) sugiere que, cuando el modelo indica ausencia de violencia, existe una probabilidad relativamente alta de que esta condición sea correcta, en conjunto el modelo tiene un mejor desempeño para detectar a mujeres que sí experimentan violencia que para descartar correctamente a quienes no la padecen, su utilidad principal radica en su aplicación como herramienta de tamizaje inicial, especialmente en contextos donde es fundamental identificar de manera temprana a mujeres en riesgo, debido a su limitada especificidad, es necesario complementar sus resultados con evaluaciones más detalladas, a fin de evitar clasificaciones erróneas y asegurar intervenciones adecuadas, integrales y centradas en las necesidades reales de las mujeres

Tabla 31. Resumen de indicadores según tipo de violencia que tipo de violencia es más predecible

	Indicadores		
	Exactitud	Sensibilidad	Especificidad
Violencia física	0,691	0,560	0,822
Violencia psicológica	0,623	0,451	0,794
Violencia sexual	0,673	0,665	0,681
Violencia en general	0,626	0,782	0,469

La Tabla 31 presenta los principales indicadores de desempeño de los modelos predictivos aplicados a los distintos tipos de violencia contra las mujeres, lo que permite comparar su capacidad para identificar correctamente cada forma de violencia en función de la exactitud, sensibilidad y especificidad.

En el caso de la violencia general que agrupa cualquier tipo de violencia, el modelo destaca por su alta sensibilidad (78,2 %), lo que indica que es eficaz para detectar a la mayoría de mujeres que han sido víctimas. Sin embargo, esta fortaleza se acompaña de una baja especificidad (46,9 %), lo que implica que también clasifica erróneamente a un número

considerable de mujeres que no han sufrido violencia. Esto sugiere que el modelo prioriza la identificación de posibles víctimas, incluso a costa de generar más falsos positivos, el modelo de violencia física presenta el mejor desempeño global. Alcanza la mayor exactitud (69,1 %) y una alta especificidad (82,2 %), lo que indica que es más preciso para diferenciar entre mujeres que han experimentado este tipo de violencia y aquellas que no, su sensibilidad es moderada (56,0 %), lo que significa que no logra identificar a todas las mujeres afectadas, dejando algunos casos sin detectar.

En cuanto a la violencia sexual, a pesar de su menor frecuencia, el modelo muestra un desempeño relativamente equilibrado, con niveles similares de sensibilidad (66,5 %) y especificidad (68,1 %), y una exactitud de 67,3 %. Esto sugiere que, cuando se incorporan variables relevantes como el consumo de alcohol de la pareja o la dinámica en la toma de decisiones económicas, es posible predecir este tipo de violencia con una precisión mayor a la esperada, considerando su baja prevalencia, la violencia psicológica, aunque es la más frecuente entre las mujeres, resulta ser la más difícil de predecir, resenta los valores más bajos de exactitud (62,3 %) y sensibilidad (45,1 %), lo que refleja su carácter menos visible, más subjetivo y, en muchos casos, socialmente normalizado, lo que dificulta su identificación mediante variables tradicionales, estos resultados evidencian que la capacidad predictiva de los modelos varía según el tipo de violencia que enfrentan las mujeres. Mientras que la violencia física es la más precisa de identificar en términos generales, la violencia general resulta útil como herramienta de detección amplia, y la violencia sexual puede modelarse de manera relativamente confiable cuando se consideran factores específicos. En cambio, la violencia psicológica continúa representando un desafío importante, lo que resalta la necesidad de desarrollar enfoques más sensibles e integrales que permitan visibilizar y abordar esta forma de violencia en la vida de las mujeres.

DISCUSIÓN

Los resultados obtenidos en la presente investigación permiten comprender con mayor claridad la complejidad de la violencia de género en el Perú, evidenciando que se trata de un fenómeno multifactorial que no puede explicarse a partir de una sola variable. En ese sentido, los hallazgos coinciden con los antecedentes teóricos y empíricos revisados, los cuales señalan que la violencia hacia la mujer está influenciada por factores estructurales, sociales y relacionales. En relación con el objetivo general, el modelo de árboles de decisión demostró ser una herramienta eficaz para la predicción de la violencia de género, lo cual respalda su utilidad dentro del análisis de fenómenos sociales complejos. Este resultado se encuentra en concordancia con lo señalado por diversos autores en el campo de la minería de datos, quienes destacan que los árboles de decisión permiten identificar patrones no lineales y relaciones ocultas entre variables, facilitando la interpretación de los resultados.

Respecto a los factores asociados, se encontró que el estado civil es uno de los predictores más relevantes en la ocurrencia de violencia. Este hallazgo es consistente con estudios previos que indican que las relaciones inestables o no convivientes presentan mayores niveles de conflicto y riesgo de violencia. En este sentido, los resultados refuerzan la idea de que la dinámica de pareja juega un papel determinante en la aparición de conductas violentas.

Asimismo, el consumo de alcohol por parte de la pareja se identificó como un factor altamente influyente, lo cual coincide con investigaciones previas que lo vinculan con el incremento de comportamientos agresivos y la disminución del control emocional. Este resultado evidencia que la violencia no solo responde a factores estructurales, sino también a conductas individuales que actúan como desencadenantes.

Por otro lado, el nivel educativo mostró una relación inversa con la violencia, donde menores niveles de educación se asocian con una mayor probabilidad de sufrir agresión. Este hallazgo

se alinea con lo planteado por organismos internacionales, que señalan que la educación constituye un factor protector al fortalecer la autonomía, el acceso a oportunidades y la capacidad de toma de decisiones de las mujeres.

En cuanto al tipo de violencia, los resultados indican que la violencia psicológica es la más frecuente y la que presenta mayor capacidad predictiva dentro del modelo. Este resultado coincide con los datos reportados por la ENDES 2024 y con estudios internacionales que señalan que este tipo de violencia suele ser más persistente y menos visibilizada, lo que dificulta su identificación y tratamiento oportuno.

En relación con el desempeño del modelo, se observó un nivel de precisión adecuado, aunque con una tasa de error esperada en este tipo de estudios. Esto se debe a que los fenómenos sociales, como la violencia de género, están influenciados por múltiples variables no observadas, lo que limita la capacidad predictiva absoluta de cualquier modelo estadístico. Sin embargo, el nivel de clasificación alcanzado resulta suficiente para identificar tendencias y perfiles de riesgo.

Un aspecto importante a destacar es que los resultados obtenidos mediante el árbol de decisión no solo permiten predecir la ocurrencia de violencia, sino también comprender la interacción entre diferentes factores. Esto representa una ventaja frente a otros métodos estadísticos más tradicionales, ya que facilita la interpretación de los resultados y su aplicación en el diseño de políticas públicas.

En comparación con los antecedentes revisados, los hallazgos de esta investigación refuerzan la idea de que la violencia de género está profundamente relacionada con desigualdades estructurales, normas sociales y dinámicas de poder dentro de las relaciones. Asimismo, se evidencia que factores como la dependencia económica, la educación y el comportamiento de la pareja continúan siendo determinantes en la persistencia de este problema.

Finalmente, los resultados permiten concluir que el uso de modelos de árboles de decisión constituye una alternativa metodológica válida y pertinente para el análisis de la violencia de género, ya que no solo aporta capacidad predictiva, sino también claridad en la interpretación de los factores asociados. Esto resulta fundamental para la formulación de estrategias de prevención más focalizadas y efectivas.

CONCLUSIONES

En relación con el objetivo general de la investigación, se concluye que los modelos de árboles de decisión permitieron analizar de manera adecuada la violencia de género en mujeres peruanas utilizando la información obtenida de la ENDES 2024. A través de estos modelos fue posible identificar patrones y factores relacionados con la presencia de violencia dentro de las relaciones de pareja. Los resultados mostraron que el 43,2% de las mujeres encuestadas manifestó haber sufrido algún tipo de violencia, evidenciando que esta problemática continúa siendo una situación frecuente en la sociedad peruana. Asimismo, la violencia psicológica fue la más predominante con un 40,8%, seguida de la violencia física con 14,5% y la violencia sexual con 2,0%. Estos hallazgos permiten comprender que muchas formas de violencia, especialmente la psicológica, suelen mantenerse invisibilizadas o normalizadas dentro del entorno familiar y social, afectando directamente el bienestar emocional y la calidad de vida de las mujeres.

Respecto al primer objetivo específico, orientado a identificar los factores sociodemográficos asociados a la violencia de género, se concluye que existen diferentes variables que influyen significativamente en la ocurrencia de esta problemática. Entre los factores más relevantes se encontraron el estado civil, el consumo frecuente de alcohol por parte de la pareja, la dependencia económica, el número de hijos y el inicio temprano de la vida sexual. En particular, se observó que las mujeres que no convivían con su pareja presentaban mayores probabilidades de sufrir violencia física, psicológica y sexual. Asimismo, el consumo de alcohol de la pareja se identificó como uno de los factores más influyentes, debido a que incrementa las conductas agresivas y los conflictos dentro del hogar. De igual manera, la limitada autonomía económica de las mujeres favorece relaciones de dependencia y

desigualdad, dificultando muchas veces que las víctimas puedan denunciar o alejarse de situaciones de violencia.

En relación con el segundo objetivo específico, se concluye que el modelo de árboles de decisión alcanzó resultados satisfactorios en la clasificación y predicción de casos de violencia de género. Los resultados evidenciaron que la violencia física presentó un mejor desempeño predictivo debido a que sus manifestaciones son más visibles y fácilmente identificables. Por otro lado, el modelo aplicado a la violencia psicológica obtuvo una exactitud de 62,3% y una sensibilidad de 45,1%, lo cual demuestra que este tipo de violencia es más difícil de reconocer y clasificar. Esto se debe a que muchas agresiones psicológicas suelen desarrollarse de forma progresiva mediante insultos, humillaciones, amenazas o control emocional, conductas que en algunos casos son consideradas normales dentro de las relaciones de pareja. A pesar de ello, el modelo logró identificar patrones importantes que permiten comprender mejor las condiciones que favorecen este tipo de violencia.

Respecto al tercer objetivo específico, se determinó que la capacidad predictiva de los árboles de decisión varió según el tipo de violencia analizada. La violencia física presentó la mayor precisión de clasificación debido a que sus indicadores y manifestaciones son más concretos y evidentes. Asimismo, la violencia sexual mostró una adecuada capacidad predictiva al relacionarse principalmente con factores como el consumo de alcohol, la desigualdad en la toma de decisiones económicas y las relaciones de poder dentro de la pareja. Sin embargo, la violencia psicológica fue la forma de violencia más compleja de predecir, pese a ser la más frecuente entre las mujeres evaluadas. Esto refleja que muchas veces las víctimas no identifican inmediatamente este tipo de agresión o la consideran parte de la convivencia cotidiana, lo que dificulta tanto su detección como su prevención.

Se concluye que la aplicación de técnicas de minería de datos y árboles de decisión permitió no solo predecir la ocurrencia de violencia de género, sino también identificar perfiles de mayor

vulnerabilidad en las mujeres peruanas. Los resultados obtenidos evidencian que la violencia contra la mujer no depende de un único factor, sino de la interacción de aspectos sociales, económicos, culturales y familiares que incrementan el riesgo de sufrir agresiones. En ese sentido, esta investigación aporta información relevante que puede ser utilizada para fortalecer las políticas públicas, diseñar programas de prevención más focalizados y mejorar las estrategias de atención integral dirigidas a mujeres víctimas de violencia. Además, el uso de herramientas estadísticas y modelos predictivos representa una alternativa importante para apoyar la toma de decisiones y contribuir a la reducción de la violencia de género en el Perú.

RECOMENDACIONES

Se recomienda que futuras investigaciones amplíen el análisis incorporando otros modelos de clasificación, tales como la regresión logística, Random Forest, Gradient Boosting o redes neuronales, además de los árboles de decisión. La comparación sistemática entre distintos enfoques permitiría identificar el modelo que ofrezca el mejor equilibrio entre interpretabilidad y capacidad predictiva, de acuerdo con la naturaleza de los datos y los objetivos de cada estudio.

Se sugiere al Instituto Nacional de Estadística e Informática (INEI) reforzar los procesos de recolección de datos del módulo de violencia de la ENDES, mediante programas de capacitación especializada dirigidos al personal encuestador. Una adecuada formación técnica y ética puede contribuir a generar un mayor nivel de confianza en las entrevistadas, reduciendo así la tasa de no respuesta y mejorando la calidad de la información recolectada, especialmente en temas sensibles como la violencia de género.

Se recomienda que los resultados de esta investigación sean considerados por las instituciones públicas y privadas responsables del diseño, implementación y evaluación de políticas públicas orientadas a la prevención y atención de la violencia de género. El uso de evidencia empírica basada en modelos predictivos puede facilitar la identificación de grupos de mayor riesgo y la focalización de estrategias de intervención más eficientes y sostenibles.

Se recomienda que investigaciones posteriores profundicen el análisis mediante el uso de enfoques longitudinales y técnicas avanzadas de validación, así como la desagregación de los distintos tipos de violencia (física, psicológica, sexual y económica). Este abordaje permitiría una comprensión más detallada de la dinámica temporal de la violencia de género y fortalecería la aplicación de técnicas de minería de datos en el estudio de problemáticas sociales complejas.

Se recomienda que futuras investigaciones amplíen el análisis de la violencia de género incorporando también a la población masculina, con el propósito de comprender de manera integral las distintas manifestaciones de violencia que pueden afectar tanto a hombres como a mujeres. Aunque la violencia contra la mujer constituye un problema prioritario de salud pública y de derechos humanos, es importante reconocer que los hombres también pueden ser víctimas de violencia en diferentes contextos, como el ámbito familiar, de pareja o social. La inclusión de esta población permitirá realizar estudios comparativos, identificar factores de riesgo y de protección, y generar evidencia científica que contribuya al diseño de políticas públicas, programas de prevención e intervenciones más inclusivas y equitativas, orientadas a la protección de todas las personas que experimentan situaciones de violencia.

BIBLIOGRAFIA

- CEM. (2025). *Instituciones - Ministerio de la Mujer y Poblaciones Vulnerables - Programa Nacional Aurora - Orientación - Ministerio de la Mujer y Poblaciones Vulnerables - Plataforma del Estado Peruano*. <https://www.gob.pe/75538-instituciones-ministerio-de-la-mujer-y-poblaciones-vulnerables-programa-nacional-aurora>
- CEPAL. (2024). *Violencia feminicida en cifras América Latina y el Caribe | Comisión Económica para América Latina y el Caribe*. <https://www.cepal.org/es/tipo-de-publicacion/violencia-feminicida-cifras-america-latina-caribe>
- De Lujan Piatti, M. (2013). *VIOLENCIA CONTRA LAS MUJERES Y ALGUIEN MÁS...*
- Defensoría del pueblo. (2025). *Defensoría del Pueblo Peru*. <https://www.defensoria.gob.pe/defensoria-del-pueblo-alerta-sobre-incremento-peligroso-de-casos-de-feminicidio-en-ultimos-tres-anos/>
- Gómez Castillas, A. (2018). *Violencia de género y prácticas tradicionales perjudiciales: una contribución al debate metodológico para su medición en Europa. TDX (Tesis Doctorals En Xarxa)*. <https://www.tdx.cat/handle/10803/665411>
- Guzmán Sánchez, F. M. (2016). *Violencia de género en adolescentes: análisis de las percepciones y de las acciones educativas propuestas por la Junta de Andalucía*. <https://dialnet.unirioja.es/servlet/tesis?codigo=47227&info=resumen&idioma=S>
- PA
- INEI. (2024). *Resumen-peru-encuesta-demografica-y-de-salud-familiar-endes-2024*.
- Martos Delgado, A. (2015). *PROYECTO DE INVESTIGACIÓN SOBRE VIOLENCIA DE GÉNERO EN ADOLESCENTES*.

- MIMDES. (2019). *PERÚ Ministerio de la Mujer y Desarrollo Social*.
- OMS. (2021). *Violencia contra la mujer*. <https://www.who.int/es/news-room/fact-sheets/detail/violence-against-women>
- ONU. (1994). *Declaración sobre la eliminación de la violencia contra la mujer Resolución de la Asamblea General 48/104 del 20 de diciembre de 1993*.
- Parra, F. (2019). *6 métodos de clasificación | Estadística y Machine Learning con R*. <https://bookdown.org/content/2274/metodos-de-clasificacion.html>
- Roca Monjo, A. (2011). *Trabajo de investigación sobre la violencia de genero*.
- Sánchez Lorente, S. (2009). *Estudio longitudinal del impacto de la violencia de pareja sobre la salud física y el sistema inmune de las mujeres*. <https://dialnet.unirioja.es/servlet/tesis?codigo=82259&info=resumen&idioma=SPA>
- VARGAS ARMAS, R., & WALDE SOTO, K. M. (2017). *FACTORES SOCIOCULTURALES QUE INFLUYEN EN LA VIOLENCIA INTRAFAMILIAR A PARTIR DE LA DIFERENCIA DE GÉNERO EN EL DISTRITO HUANCAYO EN LA ACTUALIDAD*.
- WHO. (2021). Violence against women, 2018 estimates. *WHO*, 1–87. <https://www.who.int/publications/i/item/9789240022256>
- Breiman, L., Friedman, JH, Olshen, RA y Stone, CJ (1984). *Árboles de clasificación y regresión*. Grupo Internacional Wadsworth.
- Hastie, T., Tibshirani, R. y Friedman, J. (2009). *Los elementos del aprendizaje estadístico: minería de datos, inferencia y predicción* (2ª ed.). Saltador. <https://doi.org/10.1007/978-0-387-84858-7>
- Breiman, L., Friedman, JH, Olshen, RA y Stone, CJ (1984). *Árboles de clasificación y regresión*. Grupo Internacional Wadsworth.

Quinlan, JR (1993). *C4.5: Programas para el aprendizaje automático*. Editores Morgan Kaufmann.

Sánchez-Lorente, S., Blasco-Ros, C., & Martínez, M. (2009). *Factores que contribuyen o impiden la recuperación de la salud física de mujeres expuestas a violencia de pareja: un estudio longitudinal*. **Problemas de salud de la mujer**, 19(6), 367–376. <https://doi.org/10.1016/j.whi.2009.07.003>

Guzmán Sánchez, F. M. (2016). *Violencia de género en adolescentes: Análisis de las percepciones y de las acciones educativas propuestas por la Junta de Andalucía* [Tesis doctoral, Universidad de Sevilla]. Repositorio Institucional de la Universidad de Sevilla.

Martos Delgado, A. (2015). *Proyecto de investigación sobre la violencia de género en adolescentes* [Trabajo de Fin de Grado, Universidad de Jaén]. Repositorio Institucional CREA, Universidad de Jaén. [Repositorio CREA – Universidad de Jaén](#)

Ascarruz López, M., & Pérez, J. (2022). *Feminicidio y violencia de género en Perú, 2021* [Tesis de licenciatura, Universidad Peruana de las Américas].

Calero Qquenaya, J. (2022). *La violencia familiar y su relación con el feminicidio en los distritos de San Jerónimo y Saylla del departamento del Cusco, 2018–2019* [Tesis de licenciatura].

Caycho, J. (2023). *Un modelo de análisis multivariante para la medición de la pobreza en el Perú* [Tesis de maestría, Universidad Nacional Mayor de San Marcos].

Ccahuantio Quispe, J. (2022). *Tratamiento de las noticias sobre violencia de género en los diarios El Sol, El Matutino y Diario del Cusco – 2019* [Tesis de licenciatura].

- Gómez Casillas, A. (2018). *Violencia de género y prácticas tradicionales: Una contribución al debate metodológico para su medición en Europa* [Tesis doctoral, Universidad Autónoma de Barcelona].
- Martos Delgado, A. (2015). *Proyecto de investigación sobre la violencia de género en adolescentes* [Trabajo de Fin de Grado, Universidad de Jaén].
- Molina Cárdenas, M. A. (2023). *Representaciones de violencia basada en género de mujeres víctimas en la provincia del Cusco* [Tesis de licenciatura].
- Pérez-Yari, A., Loo Valverde, J., & Chanduví Puicón, M. (2024). *Los factores sociodemográficos vinculados a la violencia intrafamiliar contra mujeres en el Perú durante el año 2021*. Revista de la Facultad de Medicina Humana.
- Piatti, M. de L. (2013). *Violencia contra las mujeres y alguien más...* [Tesis doctoral, Universidad de Valencia].
- Pilares Aguirre, M. A. (2024). *Manifestación de la violencia de género desde la perspectiva y experiencias de los estudiantes de 3.º de secundaria de la I.E. “Tomasa Tito Condemayta”, distrito de Acomayo, Cusco, 2024* [Tesis de licenciatura].
- Quinlan, JR (1993). C4.5: *Programas para el aprendizaje automático*. Editores Morgan Kaufmann.
- Ramos Perca, S. S. (2024). *Violencia contra la mujer y la aplicación de la Ley N.º 30364 en el distrito fiscal de Cusco, 2022* [Tesis de licenciatura].
- Roca Monjo, A. (2011). *Trabajo de investigación sobre la violencia de género* [Trabajo de investigación, Universidad Internacional de La Rioja].
- Rodas, G., Del Castillo, J., Sandoval, L., Alatrística, M., & Vela-Ruiz, R. (2023). *Situación actual de la violencia contra la mujer en el Perú: Revisión sistemática de la literatura científica*.

Sánchez-Lorente, S., Blasco-Ros, C., & Martínez, M. (2009). *Factores que contribuyen o impiden la recuperación de la salud física de mujeres expuestas a violencia de pareja: un estudio longitudinal*. Problemas de salud de la mujer, 19(6), 367–376.

<https://doi.org/10.1016/j.whi.2009.07.003>

Vargas Armas, M., & Walde Soto, R. (2017). *Factores socioculturales que influyen en la violencia intrafamiliar a partir de la diferencia de género en el distrito de Huancayo en la actualidad* [Tesis de licenciatura, Universidad Nacional del Centro del Perú].

ANEXOS

MATRIZ DE CONSISTENCIA

PROBLEMA GENERAL	OBJETIVO GENERAL	HIPOTESIS GENERAL	VARIABLE INDEPENDIENTE:	POBLACION:
¿Cuál es el nivel de precisión del modelo de árboles de decisión para predecir la violencia de género en mujeres, utilizando las variables socioeconómicas, demográficas y culturales de la Encuesta Demográfica y de Salud Familiar (ENDES) 2024?	Analizar y evaluar la eficacia de los árboles de decisión como herramienta predictiva para identificar factores asociados a la violencia de género, utilizando datos de la Encuesta Demográfica y de Salud Familiar (ENDES) 2024.	El modelo de árboles de decisión presenta un nivel significativo de precisión en la predicción de violencia de género según los factores socioeconómicos, demográficos y culturales obtenidos de la Encuesta Demográfica y de Salud Familiar (ENDES 2024)	Factores asociados Edad Estado civil Nivel educativo Tiempo de convivencia Número de hijos Nivel socioeconómico Dependencia económica Autoestima Abuso infantil Traumas Alcoholismo Marihuana Otras drogas Exceso de hijos Madres adolescentes abandono de hijos denuncias	las mujeres que fueron encuestadas En el año 2024 por la encuestadora ENDES
PROBLEMAS ESPECÍFICOS	OBJETIVOS ESPECIFICOS	HIPÓTESIS ESPECÍFICOS	VARIABLE DEPENDIENTE:	MUESTRA:
¿Cuáles son los factores más influyentes en la predicción de la violencia de género según el modelo de árboles de decisión aplicado a los datos de la ENDES 2024?	Identificar los principales factores asociados a la violencia de género mediante el uso de árboles de decisión aplicados a los datos de la ENDES 2024.	Los principales factores asociados a la violencia de género son identificables mediante el modelo de árboles de decisión aplicado a los datos de la ENDES 2024.	violencia de género, considerada como la agresión hacia la mujer (si, no) • Física • Psicológica • Sexual	las mujeres víctimas y no víctimas de violencia de genero
¿Cuál es la tasa de error del modelo de árboles de decisión al clasificar casos de violencia de género en mujeres en edad reproductiva?	Evaluar el nivel de precisión del modelo de árboles de decisión en la predicción de casos de violencia de género en mujeres.	La tasa de error del modelo de árboles de decisión es menor al 30% en la clasificación de casos de violencia de género en mujeres en edad reproductiva.		
¿En qué tipo de violencia (física, sexual o psicológica) presenta mayor eficacia predictiva el modelo basado en árboles de decisión?	Comparar el desempeño predictivo del modelo de árboles de decisión según el tipo de violencia (física, sexual y psicológica), usando datos de la ENDES 2024	El modelo basado en árboles de decisión tiene mayor eficacia predictiva en la violencia psicológica que en la física o sexual cuando se aplica a los datos de la ENDES 2024.		

variable	Dimensiones/componentes /sub variables	Indicadores	Preguntas/ítems/reactivos	Alternativas
Violencia de genero	Físico	golpes	Que tipo de agresión recibió por parte de su pareja	a) golpes b) patadas c) puñetes d) cabezazos
		Intento de homicidio	El agresor intento matarla al momento de la agresión	si no
		agresión física con algún objeto	El agresor uso algún objeto para golpearla	si no
	Psicológico	Insultos	El agresor insulto de manera despectiva a la pareja	a) si b) no
		Gritos	El agresor intento intimidar con gritos a la pareja	Si no
		Amenazas de cualquier tipo	El agresor amenaza de alguna otra manera	si no
	sexual	Violaciones	Hubo violación al momento del abuso	si no
		Prostitución forzada	El agresor obliga a la pareja a prostituirse	si no
		Falta de pudor	El agresor camina sin ropa por la casa	si no

ENCUESTA DEMOGRÁFICA Y DE SALUD FAMILIAR ENDES

DICCIONARIO DE VARIABLES CUESTIONARIO INDIVIDUAL (REC84DV)

N°	VARIABLE	DESCRIPCIÓN DE LAS VARIABLES	TIPO DE CARÁCTER	LONGITUD	RANGO DE VARIACIÓN	VALORES	DESCRIPCIÓN DE LAS ALTERNATIVAS	DB5
1	D11	Año	N	4				
2	CASEID	Identificación de caso	A	18				Variable (Llave de identificación)
3	D11A	Su esposo/compañero se pone (ponía) celoso o molesto si usted conversa (conversaba) con otro hombre	N	1	0,1,8	0 1 8	No Sí No sabe	
4	D11B	Su esposo/compañero la acusa (acusaba) frecuentemente de ser infiel	N	1	0,1,8	0 1 8	No Sí No sabe	
5	D11C	Su esposo/compañero le impide (impedía) que visite o le visiten sus amistades	N	1	0,1,8	0 1 8	No Sí No sabe	
6	D11D	Su esposo/compañero trata (trataba) de limitar las visitas/contactos a su familia	N	1	0,1,8	0 1 8	No Sí No sabe	
7	D11E	Su esposo/compañero insiste (insistía) siempre en saber todos los lugares donde usted va (iba)	N	1	0,1,8	0 1 8	No Sí No sabe	
8	D11F	Su esposo/compañero desconfía (desconfiaba) de Ud. con el dinero	N	1	0,1,8	0 1 8	No Sí No sabe	
9	D11G	CS control	N	1	0,1,8	0 1 8	No Sí No sabe	
10	D11H	CS control	N	1	0,1,8	0 1 8	No Sí No sabe	
11	D11I	CS control	N	1	0,1,8	0 1 8	No Sí No sabe	
12	D11J	CS control	N	1	0,1,8	0 1 8	No Sí No sabe	
13	D11K	Número de problemas de control	N	2		0 1 2 3	No Frecuentemente Algunas veces Nunca	
14	D11L	Su esposo/compañero alguna vez le ha dicho o le ha hecho cosas para humillarla delante de los demás	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
15	D11M	Su esposo/compañero la ha amenazado con hacerle daño a usted o a alguien cercano a usted	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
16	D11N	Su esposo/compañero siempre rasca o hace sentir mal	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
17	D11O	Su esposo/compañero la ha amenazado con hacerle daño a sus hijos e hijas o la ayuda económica	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
18	D11P	Abuso emocional CS: siempre	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	

N°	VARIABLE	DESCRIPCIÓN DE LAS VARIABLES	TIPO DE CARÁCTER	LONGITUD	RANGO DE VARIACIÓN	VALORES	DESCRIPCIÓN DE LAS ALTERNATIVAS	DB5
19	D11Q	Abuso emocional CS: siempre	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
20	D11R	Hubo violencia emocional	N	1	0,1	0 1	No Sí	
21	D11S	Su esposo/compañero alguna vez la empujó, le empujó o le tiró algo	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
22	D11T	Su esposo/compañero alguna vez la abofeteó o le retorcía el brazo	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
23	D11U	Su esposo/compañero alguna vez la golpeó con el puño o con algo que pudo hacerle daño	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
24	D11V	Su esposo/compañero alguna vez la ha pasado o amarrado	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
25	D11W	Su esposo/compañero alguna vez la trató de estrangularla o quemarla	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
26	D11X	Su esposo/compañero alguna vez la amenazó con un cuchillo / pistola u otro tipo de arma	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
27	D11Y	Su esposo/compañero alguna vez la abofeteó con un cuchillo / pistola u otro tipo de arma	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
28	D11Z	Su esposo/compañero alguna vez ha utilizado la fuerza física para obligarla a tener relaciones sexuales aunque usted no quiera	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
29	D120	Su esposo/compañero alguna vez le obligó a realizar actos sexuales que Ud. no quería	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
30	D121	Su esposo/compañero alguna vez le retorcía el brazo o leó de su cabello	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
31	D122	CS: violencia física	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
32	D123	CS: violencia física	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
33	D124	CS: violencia física	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	
34	D125	CS: violencia física	N	1	0,3	0 1 2 3	No Frecuentemente Algunas veces Nunca	

N°	VARIABLE	DESCRIPCIÓN DE LAS VARIABLES	TIPO DE CARÁCTER	LONGITUD	RANGO DE VARIACIÓN	VALORES	DESCRIPCIÓN DE LAS ALTERNATIVAS	OBS
35	D108	Experimentó cualquier violencia menos severa (D105 A, B, C o D)	N	1	0,1	0	No	
36	D107	Experimentó cualquier violencia severa (D105 E, F o G)	N	1	0,1	1	Si (D105 A, B, C o D)	
37	D108	Experimentó cualquier violencia sexual (D105 H o I)	N	1	0,1	0	No	
38	D109	Cuánto tiempo después de haberse casado/unido con su (último) esposo (compañero) empacaron e suodier estas cosas	N	2	1-40, 95-98	95	Si (D105 H o I)	
39	D110A	A causa de su esposo/compañero alguna vez tuvo morosiones y/o diones	N	1	0,1	98	Antes de la separación/divorcio	
40	D110B	A causa de su esposo/compañero alguna vez ha tenido lesiones o desloaciones	N	1	0,1	99	Después de la separación/divorcio	
41	D110C	A causa de su esposo/compañero alguna vez tuvo que ir al médico o centro de salud por algo que él le hizo	N	1	0,1	99	No sabe	
42	D1110	A causa de su esposo/compañero alguna vez tuvo una herida, fueso rito, y dientes rotos, quemaduras	N	1	0,1	0	No	
43	D111E	A causa de su esposo/compañero alguna tuvo que ir al caratérero, fumero u otra persona por algo que él le hizo	N	1	0,1	0	No	
44	D110F	Acción CS: por el acto del manto	N	1	0,1	1	Si	
45	D110G	Acción CS: por el acto del manto	N	1	0,1	0	No	
46	D110H	Acción CS: por el acto del manto	N	1	0,1	0	No	
47	D111	Cualquier resultado por causa de su esposo/compañero	N	1	0,1	0	No	
48	D112	Alguna vez usad pego o maltrato físicamente a su esposo/compañero cuando no la estaba golpeando o maltratando físicamente	N	1	0,1	1	Si (D110A-C)	
49	D112A	Con qué frecuencia bebió físicamente a su esposo/compañero en los últimos 12 meses	N	1	0,2, 8	0	Nunca	
50	D113	Su esposo/compañero toma bebidas alcohólicas	N	1	0,1	1	Mucha frecuencia	
51	D114	Con qué frecuencia su esposo/compañero toma bebidas alcohólicas	N	1	0,2	2	Algunas veces	
52	D115B	Desde cuando tenía 15 años de edad fue maltratada físicamente por: madre	N	1	0,1, 8	0	No sabe	
53	D115C	Desde cuando tenía 15 años de edad fue maltratada físicamente por: padre	N	1	0,1, 8	1	Si	
54	D115D	Desde cuando tenía 15 años de edad fue maltratada físicamente por: hija	N	1	0,1, 8	0	No	
55	D115E	Desde cuando tenía 15 años de edad fue maltratada físicamente por: hijo	N	1	0,1, 8	1	Si	
56	D115F	Desde cuando tenía 15 años de edad fue maltratada físicamente por: hermana	N	1	0,1, 8	0	No	
57	D115G	Desde cuando tenía 15 años de edad fue maltratada físicamente por: hermano	N	1	0,1, 8	0	No	
58	D115H	Desde cuando tenía 15 años de edad fue maltratada físicamente por: otro pariente femenino	N	1	0,1, 8	0	No	
59	D115I	Desde cuando tenía 15 años de edad fue maltratada físicamente por: otro pariente masculino	N	1	0,1, 8	0	No	

N°	VARIABLE	DESCRIPCIÓN DE LAS VARIABLES	TIPO DE CARÁCTER	LONGITUD	RANGO DE VARIACIÓN	VALORES	DESCRIPCIÓN DE LAS ALTERNATIVAS	OBS
60	D115J	Desde cuando tenía 15 años de edad fue maltratada físicamente por: ex pareja	N	1	0,1, 8	0	No	
61	D115K	Desde cuando tenía 15 años de edad fue maltratada físicamente por: novio actual	N	1	0,1, 8	0	No	
62	D115L	Desde cuando tenía 15 años de edad fue maltratada físicamente por: ex novio	N	1	0,1, 8	0	No	
63	D115M	Desde cuando tenía 15 años de edad fue maltratada físicamente por: madrastra	N	1	0,1, 8	0	No	
64	D115N	Desde cuando tenía 15 años de edad fue maltratada físicamente por: padrastro	N	1	0,1, 8	0	No	
65	D115O	Desde cuando tenía 15 años de edad fue maltratada físicamente por: suegra	N	1	0,1, 8	0	No	
66	D115P	Desde cuando tenía 15 años de edad fue maltratada físicamente por: suegro	N	1	0,1, 8	0	No	
67	D115Q	Desde cuando tenía 15 años de edad fue maltratada físicamente por: otra suegra	N	1	0,1, 8	0	No	
68	D115R	Desde cuando tenía 15 años de edad fue maltratada físicamente por: otro suegro	N	1	0,1, 8	0	No	
69	D115S	Desde cuando tenía 15 años de edad fue maltratada físicamente por: amiga	N	1	0,1, 8	0	No	
70	D115T	Desde cuando tenía 15 años de edad fue maltratada físicamente por: amigo	N	1	0,1, 8	0	No	
71	D115U	Desde cuando tenía 15 años de edad fue maltratada físicamente por: vecino	N	1	0,1, 8	0	No	
72	D115V	Desde cuando tenía 15 años de edad fue maltratada físicamente por: profesor	N	1	0,1, 8	0	No	
73	D115W	Desde cuando tenía 15 años de edad fue maltratada físicamente por: empleador	N	1	0,1, 8	0	No	
74	D115X	Desde cuando tenía 15 años de edad fue maltratada físicamente por: otro	N	1	0,1, 8	0	No	
75	D115Y	Desde cuando tenía 15 años de edad fue maltratada físicamente por: nadie más que conyuge	N	1	0,1, 8	0	No	
76	D115XA	Desde cuando tenía 15 años de edad fue maltratada físicamente por: extraño	N	1	0,1, 8	0	No	
77	D115XB	Desde cuando tenía 15 años de edad fue maltratada físicamente por: otro familiar femenino del esposo	N	1	0,1, 8	0	No	
78	D115XC	Desde cuando tenía 15 años de edad fue maltratada físicamente por: otro familiar masculino del esposo	N	1	0,1, 8	0	No	
79	D115XD	Desde cuando tenía 15 años de edad fue maltratada físicamente por: amigo	N	1	0,1, 8	0	No	
80	D115XE	Desde cuando tenía 15 años de edad fue maltratada físicamente por: la sociedad	N	1	0,1, 8	0	No	

*REC84DV_2024.SAV [ConjuntoDatos1] - IBM SPSS Statistics Editor de datos

Archivo Editar Ver Datos Transformar Analizar Gráficos Utilidades Ampliaciones Ventana Ayuda

Visible: 220 de 220 variables

	ID1	CASEID	D101A	D101B	D101C	D101D	D101E	D101F	D101G	D101H	D101I	D101J	D102	D103A	D103B	D103C	D103D
1	2024	325503101 2	No	No	No	No	No	No					0	No	No		No
2	2024	325503101 4															
3	2024	325503901 2															
4	2024	325504701 2	No	No	No	No	No	No					0	No	No		No
5	2024	325505001 1	No	No	No	No	No	No					0	No	No		No
6	2024	325505301 1	Si	No	No	No	No	No					1	No	No		No
7	2024	325508901 2	No	No	No	No	No	No					0	No	No		No
8	2024	325509301 1															
9	2024	325509301 2															
10	2024	325509701 2	No	No	No	No	No	No					0	No	No		No
11	2024	325600601 1	Si	No	Si	No	Si	Si					4	Nunca	Algunas v...		Algunas v...
12	2024	325600801 2	No	No	No	No	No	No					0	No	No		No
13	2024	325601301 2	Si	Si	Si	Si	Si	Si					6	Frecuent...	Frecuent...		Frecuent...
14	2024	325601301 3															
15	2024	325602801 1	Si	Si	Si	Si	No	No					4	Nunca	Nunca		Nunca
16	2024	325603201 3	No	No	No	No	No	No					0	No	No		No
17	2024	325605001 1															
18	2024	325605601 1															
19	2024	325701001 1															
20	2024	325702001 1															
21	2024	325707401 1															
22	2024	325710301 2															
23	2024	325712401 2															
24	2024	325712401 3															
25	2024	325800101 3	No	No	No	No	No	No					0	No	No		No
26	2024	325800601 3															
27	2024	325800601 4	No	No	No	No	No	No					0	No	No		No
28	2024	325804201 2	Si	Si	Si	Si	Si	Si					6	No	No		Algunas v...
29	2024	325806001 1	No	Si	No	No	No	No					1	No	No		No
30	2024	325806401 2															
31	2024	325806401 3	No	No	No	No	No	No					0	Algunas v...	No		Algunas v...
32	2024	325809301 1	No	No	No	No	No	Si	No				1	No	No		Algunas v...
33	2024	325810301 1	No	No	No	No	No	No					0	No	No		No

Vista de datos Vista de variables

Abrir documento de datos IBM SPSS Statistics Processor está listo Unicode ON

*REC84DV_2024.SAV [ConjuntoDatos1] - IBM SPSS Statistics Editor de datos

Archivo Editar Ver Datos Transformar Analizar Gráficos Utilidades Ampliaciones Ventana Ayuda

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	ID1	Número	4	0	Año	Ninguno	Ninguno	10	Derecha	Nominal	Entrada
2	CASEID	Cadena	18	0	Identificación d...	Ninguno	Ninguno	20	Izquierda	Nominal	Entrada
3	D101A	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
4	D101B	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
5	D101C	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
6	D101D	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
7	D101E	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
8	D101F	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
9	D101G	Número	1	0	CS control	[0, No]...	9	7	Derecha	Nominal	Entrada
10	D101H	Número	1	0	CS control	[0, No]...	9	7	Derecha	Nominal	Entrada
11	D101I	Número	1	0	CS control	[0, No]...	9	7	Derecha	Nominal	Entrada
12	D101J	Número	1	0	CS control	[0, No]...	9	7	Derecha	Nominal	Entrada
13	D102	Número	2	0	Número de pro...	Ninguno	99	6	Derecha	Nominal	Entrada
14	D103A	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
15	D103B	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
16	D103C	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
17	D103D	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
18	D103E	Número	1	0	Abuso emocion...	[0, No]...	9	7	Derecha	Nominal	Entrada
19	D103F	Número	1	0	Abuso emocion...	[0, No]...	9	7	Derecha	Nominal	Entrada
20	D104	Número	1	0	Hubo violencia ...	[0, No]...	9	6	Derecha	Nominal	Entrada
21	D105A	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
22	D105B	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
23	D105C	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
24	D105D	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
25	D105E	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
26	D105F	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
27	D105G	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
28	D105H	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
29	D105I	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
30	D105J	Número	1	0	Su esposo/com...	[0, No]...	9	7	Derecha	Nominal	Entrada
31	D105K	Número	1	0	CS: violencia f...	[0, No]...	9	7	Derecha	Nominal	Entrada
32	D105L	Número	1	0	CS: violencia f...	[0, No]...	9	7	Derecha	Nominal	Entrada
33	D105M	Número	1	0	CS: violencia f...	[0, No]...	9	7	Derecha	Nominal	Entrada
34	D105N	Número	1	0	CS: violencia f...	[0, No]...	9	7	Derecha	Nominal	Entrada
35	D106	Número	1	0	Exoneración cu...	[0, No]...	9	6	Derecha	Nominal	Entrada

Vista de datos Vista de variables

IBM SPSS Statistics Processor está listo Unicode ON

1. Instalar y cargar el paquete

```
install.packages("haven") # ejecutar solo una vez
```

```
library(haven)
```

2. Importar archivo SPSS (.sav)

```
datos_endes <- read_sav("ruta/del/archivo/ENDES_2024.sav")
```

Ejemplo en Windows:

```
datos_endes <- read_sav("C:/Users/Usuario/Documents/ENDES_2024.sav")
```

Ejemplo en Mac/Linux:

```
datos_endes <- read_sav("/home/usuario/ENDES_2024.sav")
```

3. Verificar la estructura de los datos

```
str(datos_endes)
```

```
head(datos_endes)
```

◆ Convertir etiquetas SPSS a factores

SPSS suele guardar variables con labels. Para árboles de decisión es recomendable convertirlas:

```
datos_endes <- as.data.frame(datos_endes)
```

```
datos_endes[] <- lapply(datos_endes, function(x) {  
  if (is.labelled(x)) as_factor(x) else x  
})
```

◆ Identificar valores perdidos (missing)

```
colSums(is.na(datos_endes))
```

◆ Seleccionar variables relevantes (ejemplo ENDES violencia)

```
vars <- c("V527", "V529", "V602", "V633B", "V633D", "V744B", "V744D")
```

```
datos_modelo <- datos_endes[, vars]
```

1. Cargar librerías necesarias

Manejo de datos

```
library(tidyverse)
```

Árboles de decisión

```
library(rpart)
```

```
library(rpart.plot)
```

Evaluación del modelo

```
library(caret)
```

```
library(pROC)
```

2. Cargar y preparar la base de datos

Cargar base de datos (ejemplo)

```
datos <- read.csv("ENDES_2024.csv")
```

Convertir variable respuesta a factor (ejemplo: violencia = Sí / No)

```
datos$violencia <- factor(datos$violencia, levels = c(0,1),  
  labels = c("No", "Si"))
```

Seleccionar variables relevantes

```
datos_modelo <- datos %>%
```

```
  select(violencia, edad, nivel_educativo, estado_civil,  
    region, nivel_riqueza)
```

3. División de datos: entrenamiento y prueba

```
set.seed(123)
```

70% entrenamiento – 30% prueba

```
trainIndex <- createDataPartition(datos_modelo$violencia,  
  p = 0.7, list = FALSE)
```

```
train_data <- datos_modelo[trainIndex, ]
```

```
test_data <- datos_modelo[-trainIndex, ]
```

4. Entrenamiento del árbol de decisión

```
modelo_arbol <- rpart(  
  violencia ~ .,  
  data = train_data,  
  method = "class",  
  control = rpart.control(  
    cp = 0.01,      # complejidad del árbol  
    minsplit = 20,  # mínimo de observaciones para dividir  
    maxdepth = 5    # profundidad máxima  
  )  
)
```

5. Visualización del árbol

```
rpart.plot(  
  modelo_arbol,  
  type = 2,  
  extra = 106,  
  fallen.leaves = TRUE,  
  main = "Árbol de decisión - Violencia de género (ENDES 2024)"  
)
```

6. Predicción sobre el conjunto de prueba

```
pred_clase <- predict(modelo_arbol, test_data, type = "class")  
pred_prob <- predict(modelo_arbol, test_data, type = "prob")[,2]
```

7. Matriz de confusión y métricas

```
confusionMatrix(pred_clase, test_data$violencia, positive = "Si")
```

Esto te da automáticamente:

Accuracy

Sensibilidad

Especificidad

Valor predictivo positivo y negativo

8. Curva ROC y AUC

```
roc_obj <- roc(test_data$violencia, pred_prob)
```

```
plot(roc_obj, col = "blue", lwd = 2,  
     main = "Curva ROC - Árbol de decisión")
```

```
auc(roc_obj)
```

9. Importancia de variables

```
importancia <- modelo_arbol$variable.importance
```

importancia

10. Árbol podado ()

```
printcp(modelo_arbol)
```

```
cp_optimo <- modelo_arbol$cpstable[which.min(modelo_arbol$cpstable[, "xerror"]), "CP"]
```

```
arbol_podado <- prune(modelo_arbol, cp = cp_optimo)
```

```
rpart.plot(arbol_podado,  
  type = 2,  
  extra = 106,  
  main = "Árbol de decisión podado")
```

```

otro Código R
# Instalar si no las tienes
install.packages(c("rpart", "rpart.plot", "caret", "readxl" ))
library(readxl)
# Leer la primera hoja del Excel
datos <- read_excel("D:/G.xlsx")
# Convertir a data.frame
datos <- as.data.frame(datos)
# Ver estructura
str(datos)
# Convertir todas las variables a factor
datos[] <- lapply(datos, as.factor)
# Asegurar que la variable objetivo sea factor
datos$violencia <- factor(datos$violencia, levels = c("no", "si"))
library(caret)
set.seed(123)
indice <- createDataPartition(datos$violencia, p = 0.7, list = FALSE)
train <- datos[indice, ]
test <- datos[-indice, ]

library(rpart)
# usa un ÁRBOL DE CLASIFICACIÓN basado en el algoritmo CART
modelo_arbol <- rpart(
  violencia ~ .,
  data = train,
  method = "class",
  control = rpart.control(
    cp = 0.01,
    minsplit = 5,
    maxdepth = 5 ))
library(rpart.plot)
rpart.plot(
  modelo_arbol,
  type = 3,
  extra = 104,
  fallen.leaves = TRUE,
  box.palette = "Blues",
  main = "Árbol de clasificación de violencia")
pred <- predict(modelo_arbol, test, type = "class")
confusionMatrix(pred, test$violencia)
modelo_arbol$variable.importance

```