

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

**FACULTAD DE INGENIERÍA ELÉCTRICA, ELECTRÓNICA,  
INFORMÁTICA Y MECÁNICA**

**ESCUELA PROFESIONAL DE INGENIERÍA INFORMÁTICA Y DE  
SISTEMAS**



**TESIS**

---

**IMPLEMENTACIÓN DE TÉCNICAS DE MINERÍA DE TEXTO PARA LA  
CLASIFICACIÓN DE TICKETS DE SOPORTE EN LA OFICINA DE  
TECNOLOGÍAS DE LA INFORMACIÓN DE EGEMSA**

---

Para optar al título profesional de:

**INGENIERO INFORMÁTICO Y DE SISTEMAS**

Presentado por:

**BR. ALEJO HIRPAHUANCA, FERNANDO**

Asesor:

**DR. PALOMINO OLIVERA, EMILIO**

**CUSCO – PERÚ**

**2022**

## RESUMEN

En EGEMSA cuando un colaborador de la empresa solicita un ticket de soporte mediante correo electrónico, llamada telefónica o de manera presencial; el personal de la oficina de Tecnologías de la Información (TI) revisa dicha solicitud y lo deriva a un especialista de soporte para su solución inmediata. A medida que se incrementaba la cantidad de solicitudes de servicio de TI, el personal responsable de atender los tickets, dedicaba mucho tiempo evaluando cada descripción del ticket y tratando de asignarle su categoría correspondiente. Como resultado de esta tarea manual, la mayoría de las veces el personal de la oficina de TI completaba con la etiqueta “Otros” o “Ninguno” para estos tickets; esto implicaba una dificultad crucial al momento de generar informes ya que no contemplaba el detalle de cada categoría.

Es por ello que se implementó las técnicas de minería de texto para la clasificación de los tickets de soporte, en efecto se construyó un modelo clasificador de tickets para categorizar el nuevo ticket en función de su descripción. Asimismo, el conjunto de datos que se utilizó fueron datos internos de la empresa importados desde su sistema de mesa de ayuda. Se recolectó alrededor de 3000 tickets de soporte, considerando el título, la descripción del ticket, la fecha y su categoría. Para el entrenamiento del modelo se dividió el 75% del total de tickets y para las pruebas el 25%. Al final del entrenamiento y pruebas, el modelo SVC lineal alcanzó una precisión del 75% al hacer predicciones para los nuevos tickets. La contribución de las técnicas de minería de texto para la clasificación de los tickets en la oficina de tecnologías de la información fue automatizar el proceso de clasificación, mejor obtención de informes de servicios de TI, y la satisfacción de los encargados de la oficina de TI.

Palabras clave: categorías, clasificación, datos no estructurados, minería de texto, modelos, preprocesamiento, técnicas, tickets de soporte.