



**Facultad de Ciencias de la Administración**

**Carrera de Ingeniería en Sistemas y  
Telemática**

**PREPROCESAMIENTO DEL TEXTO DE LOS  
AUDIOS GENERADOS DE LAS  
CONVERSACIONES DEL ECU-911**

**Trabajo de Titulación previo a la obtención del  
grado de Ingeniero en Sistemas y Telemática**

**Autor:**

Pablo Andrés Molina Pinos

**Director:**

Ing. Marcos Patricio Orellana Cordero

**Cuenca – Ecuador**

**2024**

## **DEDICATORIA**

A mis padres, mi hermano y mis abuelos, por su apoyo constante en cada paso de mi vida y por darme la fortaleza para seguir adelante.

A mis amigos que he conocido durante este proceso de formación, que me han apoyado y transitado este camino junto a mí.

## **AGRADECIMIENTO**

A todos los docentes que he tenido durante mi carrera universitaria, cada uno me ha enseñado cosas valiosas a lo largo de mi formación académica y personal.

A mi tutor de este trabajo de titulación, Ing Marcos Orellana, por el apoyo durante el proceso de ejecución de este trabajo.

Al personal del LIDI, por haber solventado dudas y dado soporte durante el desarrollo de este trabajo de titulación.

## Índice de Contenidos

DEDICATORIA .....	i
AGRADECIMIENTO .....	ii
Índice de Contenidos .....	iii
Índice de Figuras.....	iv
Índice de Tablas .....	v
RESUMEN .....	vi
ABSTRACT .....	vi
1. Introducción.....	1
1.1 Objetivos.....	1
1.2 Marco teórico.....	1
1.2.1 Técnicas enfocadas en mejorar el rendimiento de modelos de lenguaje .....	1
1.2.2 Técnicas para formalización de texto .....	2
2. Revisión de literatura .....	4
2.1 Procesadores Gramaticales y Morfológicos .....	4
2.2 Etiquetado y formalización mediante diccionarios especializados .....	5
2.3 Modelos estadísticos para corregir errores gramaticales .....	5
3. Métodos .....	7
3.1 Comprender los datos.....	7
3.2 Preprocesar .....	7
3.3 Formalizar .....	8
3.3.1 Mapa de bits .....	10
3.4 Validar la formalización .....	12
4. Resultados.....	12
4.1 Clasificación de las llamadas .....	13
5. Discusión .....	16
6. Conclusión.....	18
7. Referencias .....	19

## Índice de Figuras

Figura 1 .....	7
Figura 2 .....	13
Figura 3 .....	14
Figura 4 .....	14
Figura 5 .....	15
Figura 6 .....	15
Figura 7 .....	16
Figura 8 .....	16

## Índice de Tablas

Tabla 1 .....	8
Tabla 2 .....	9
Tabla 3 .....	10
Tabla 4 .....	11

## RESUMEN

El ECU 911 se encarga de recibir las llamadas de emergencia y coordinar la respuesta a través de diversos servicios de emergencia. Todas estas interacciones permanecen almacenadas como respaldo de cada llamada. El presente trabajo propone el preprocesamiento de texto del contenido de dichas llamadas mediante técnicas de formalización de texto seleccionadas, como diccionarios de modismos, búsqueda en la RAE (Real Academia Española), correctores ortográficos y etiquetado de NER (Named Entity Recognition). Durante el proceso de formalización, se propone el uso de una herramienta auxiliar como es el mapa de bits, para facilitar la toma de decisiones y una correcta formalización. Finalmente, se evalúa el puntaje obtenido de un modelo de comparación semántica, donde se compara entre las llamadas con el texto transcrito informal, y cuando se compara el texto informal con el texto formalizado, con el fin de conocer si la formalización de texto mejora el rendimiento de modelos de Inteligencia Artificial.

**Palabras clave:** preprocesamiento, minería de datos, comparación semántica, texto formalizado, texto desestructurado. inteligencia artificial.

## ABSTRACT

ECU 911 is responsible for receiving emergency calls and coordinating responses through various emergency services. All these interactions are stored as backups for each call. This work proposes the text preprocessing of the content from these calls using selected text formalization techniques, such as idiomatic dictionaries, searches in the RAE (Royal Spanish Academy), spelling correctors, and Named Entity Recognition (NER) tagging. During the formalization process, the use of an auxiliary tool, such as a bitmap, is suggested to facilitate decision-making and proper formalization. Finally, the obtained score from a semantic comparison model is evaluated. This comparison is made between calls with informal transcribed text and when comparing informal text with formalized text, aiming to determine if text formalization improves the performance of Artificial Intelligence models.

**Keywords:** preprocessing, data mining, semantic comparison, formalized text, unstructured text, artificial intelligence.



Marcos Orellana  
0102668209



Pablo Andrés Molina Pinos  
0106725393