



UNIVERSITAT
Carlemany



Procesamiento de Lenguaje Natural (NLP) como apoyo a la Misión de Policía Nacional del Ecuador

Autor: Jorge Alberto Mendoza Bravo.

Titulación: Bàtxelor en Data Science.

Tutor: Yolanda Colom Torrens.





UNIVERSITAT
Carlemany

De:

 Planeta Formación y Universidades

OBJETIVOS DE DESARROLLO SOSTENIBLE



- **ODS 11: Ciudades y Comunidades Sostenibles** - El uso de tecnologías mejora la gestión de la información en el ámbito de seguridad.
- **ODS 16: Paz, Justicia e Instituciones Sólidas** - Apoya la reducción de la violencia y acceso a la justicia mediante el aumento de la capacidad de respuesta de la Policía Nacional.
- **ODS 9: Industria, Innovación e Infraestructura** - Fomenta la innovación y el desarrollo de infraestructura tecnológica.
- **ODS 17: Alianzas para Lograr los Objetivos** - Impulsa la colaboración entre la Policía Nacional, instituciones académicas y desarrolladores.

Contexto: El Ecuador vive una crisis de seguridad. El presente trabajo de investigación busca desarrollar un sistema que apoye a la Policía Nacional. Considerando el potencial de la información. Se pretende usar herramientas de Procesamiento del Lenguaje Natural (NLP) para clasificar grandes cantidades de texto y obtener datos útiles.



Fuente: Propia Operativo Móvil

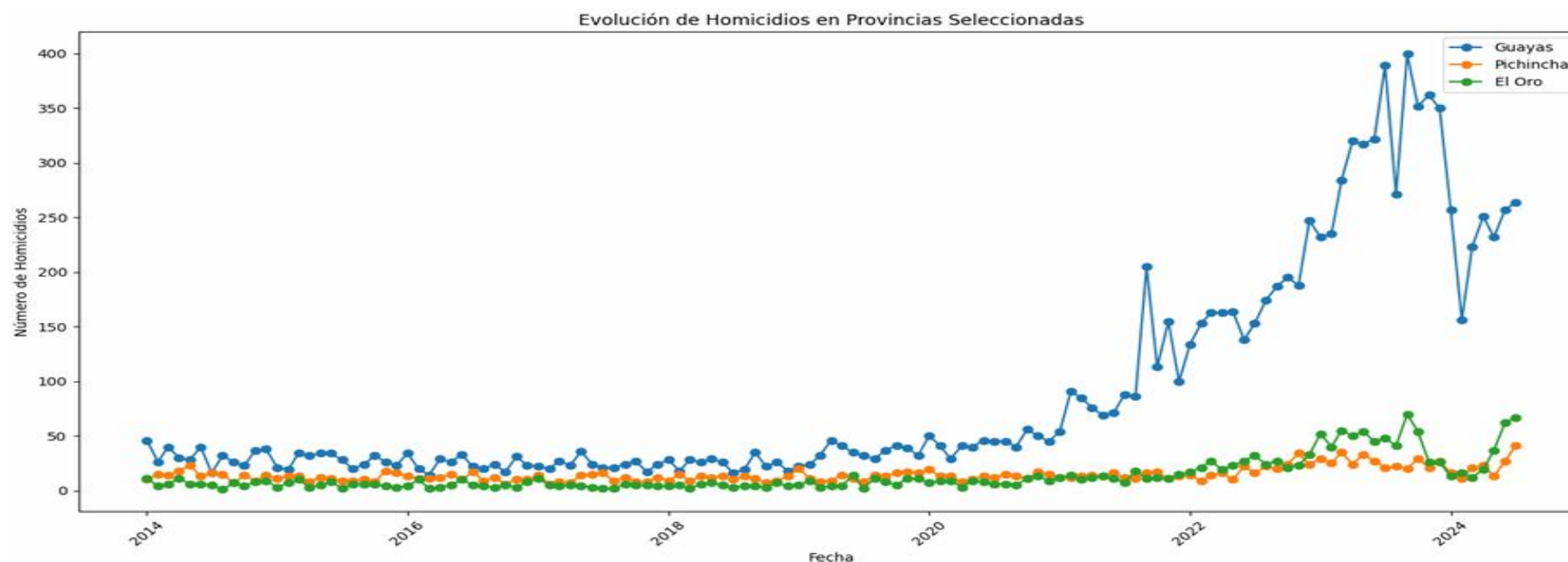


Gráfico Propio, Datos INEC, generado con Matplotlib



Objetivo General:

- Desarrollar una **herramienta de Procesamiento de Lenguaje Natural** (NLP) con el fin de apoyar a la misión de la Policía Nacional del Ecuador.



Objetivos específicos

- Identificar emociones y la relevancia del texto.
- Evaluar el sistema en un entorno controlado.
- Corroborar su eficacia al detectar información útil.
- Documentar instantáneamente los reportes.

Relevancia

del

proyecto:

Este sistema como herramienta de procesamiento de texto busca **reducir los tiempos** de respuesta de la Policía Nacional del Ecuador y puede ser de utilidad en múltiples aplicaciones de tipo preventivas, investigativas y de inteligencia.

- **Herramientas para la clasificación del texto para NLP.** [Speech and Language processing](#) 3ra edición de Jurafsky y Martin. Contiene algoritmos fundamentales, explicaciones del funcionamiento de diferentes métodos para la clasificación de texto, aplicaciones del NLP y anotaciones de la estructura lingüística.
- **Modelos BERT de redes neuronales.** Elegí utilizar este tipo de modelos pre-entrenados debido a su destacada capacidad de captar el significado de una palabra en relación con las palabras que la rodean. Devlin et al. 2018 / Google.
- **Tokenización.** Consiste en dividir el texto en partes pequeñas para procesar los datos de forma mecánica. Es importante y para el proyecto el texto se divide en tokens por palabra para ayudar al análisis de entidades e identificación de palabras clave, así como por oración para el análisis de sentimiento según el significado de cada palabra dentro de la oración.

Análisis de Entidades

En **Guayaquil, Ecuador**, los lobos atacaron violentamente el **15/03/2025** con pistolas y c4. La policía reportó que un vehículo con placa **XYZ-789** fue utilizado en el asalto. Fue un día horrible y caótico para la ciudad. El festival en **Quito, Ecuador**, el **20/04/2025** fue un éxito total. **Los organizadores** usaron un vehículo con placa **ABC-123** para transportar equipos, y los fuegos artificiales incluyeron dinamita controlada. ¡Todos disfrutaron de un día increíble! **La lluvia** arruinó todo el día ayer. Fue un completo desastre, nadie salió de casa y el ambiente estaba lleno de tristeza. En **Esmeraldas, Ecuador**, el grupo **farc** planea un ataque con **ak-47** y tnt para el **25/05/2025**. **La situación es alarmante** y las autoridades están desesperadas por detenerlos. Hoy fue un día maravilloso lleno de alegría. **Todos** estaban contentos y compartiendo momentos felices sin preocupaciones. **Los choneros están** causando problemas otra vez con pistolas y dinamita. Es una situación terrible que no parece tener fin. En **Manta, Ecuador**, el **30/01/2025**, un equipo de seguridad usó rifles para proteger un evento importante. **Todo salió perfectamente** y **la gente** aplaudió el gran trabajo.

Análisis de entidades NER con es_core_news_sm captura de pantalla del aplicativo

Análisis de Sentimiento

El narcotráfico alimenta una violencia sin precedentes en Ecuador, que cerró 2024 con una tasa de 38 homicidios por cada 100.000 habitantes. (1 estrellas)

Esta mañana desayune un delicioso bolón con huevo frito. (5 estrellas)

Ayer asesinaron a Raulito, uno de los integrantes del GDO Tiguerones. (1 estrellas)

Mañana hará buen clima. (5 estrellas)

Hace unos instantes detuvieron a alias Charles, quien habría asaltado una gasolinera en horas de la mañana, uno de sus cómplices habría sido visto portando una AK-47 en los alrededores, se logró captar imágenes del hecho por medio de cámaras del ECU-911. (1 estrellas)

Análisis de sentimiento (token por oración) captura de pantalla del aplicativo

- **Etiquetas NER:** de la Librería spaCy se utiliza el modelo “**es_core_news_sm**”

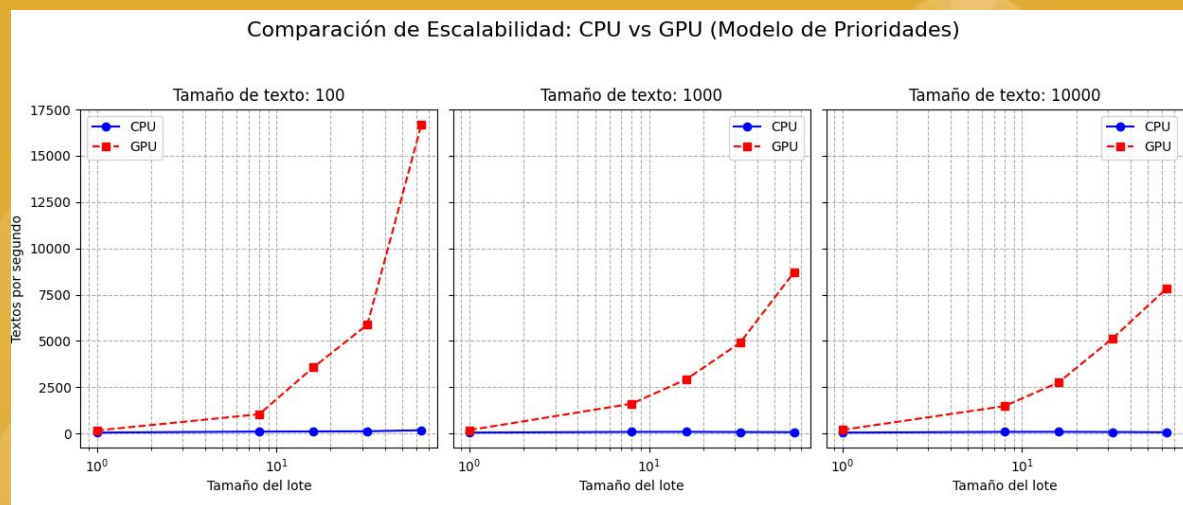
PERSON: Personas, incluidas las ficticias.
NORP: Nacionalidades o grupos religiosos o políticos.
FAC: Edificios, aeropuertos, carreteras, puentes, etc.
ORG: Empresas, agencias, instituciones, etc.
GPE: Países, ciudades, estados.
LOC: Ubicaciones no GPE, cadenas montañosas, cuerpos de agua.
PRODUCT: Objetos, vehículos, alimentos, etc. (No servicios).
EVENT: Huracanes, batallas, guerras, eventos deportivos, etc., con nombre propio.
WORK_OF_ART: Títulos de libros, canciones, etc.
LAW: Documentos nombrados que han sido convertidos en leyes.
LANGUAGE: Cualquier idioma con nombre.
DATE: Fechas absolutas o relativas, períodos de tiempo.
TIME: Momentos del día menores a 24 horas.
PERCENT: Porcentajes, incluido “%”.
MONEY: Valores monetarios, incluyendo unidad.
QUANTITY: Medidas, como peso o distancia.
ORDINAL: “primero”, “segundo”, etc.
CARDINAL: Números que no entran en otra categoría.

Extracción de Información y Contextualización:

Técnicas para reconocimiento de entidades (Nadeau & Sekine, 2007) y priorización de alertas (Adesokan & Madria, 2023) permiten filtrar y dar relevancia a los datos extraídos, adaptándose a distintos escenarios.

Aplicación en Redes Sociales y Gestión de Crisis:

- Imran et al. (2013, 2015) y Gómez Torres & Salazar Morales (2024) muestran cómo extraer y analizar información en tiempo real **durante eventos de emergencia**.
- La arquitectura Big Data (Martínez-Castaño et al., 2018) facilita el procesamiento de **grandes volúmenes de datos** en plataformas sociales.



Pruebas de escalabilidad con CPU y GPU captura de pantalla

Las pruebas de escalabilidad han dado resultados positivos en cuanto a la viabilidad de uso del aplicativo para su aplicación en el análisis en tiempo real de la información en condiciones con grandes volúmenes de datos como en análisis de redes sociales.

Capacidad actual 5,150 textos por segundo en GPU (batch_size=32, text_size=10000) sin afectar el servicio sin errores de memoria, manteniendo un uso eficiente de recursos.

Aspectos Legales y Éticos:

La implementación deberá **alinearse con normativas vigentes** (Constitución de la República del Ecuador) y consideraciones **éticas** en el uso de Big Data (Zwitter, 2014).

La validez y **transparencia** de los modelos se complementa con estrategias explicativas, promoviendo la confianza en su uso en contextos de **seguridad**.

Uso Policial y Seguridad Pública:

Arenas Tamayo et al. (2023) y Berg & Bledsoe (2024), que contextualizan la crisis de seguridad y la necesidad de **respuestas inmediatas**.

Herramientas para **detectar discursos de odio** y alertas tempranas (Burnap & Williams, 2015; Ribeiro et al., 2018) ayudan a **prevenir la escalada de conflictos** en redes sociales, ofreciendo sistemas que se integran a protocolos policiales.

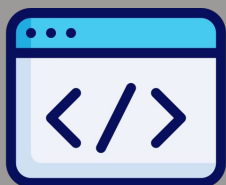
Fases de Desarrollo

Preparación del sistema:



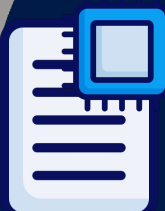
Se emplea Python y todas las librerías consideradas para el proyecto

Construcción de interfaz gráfica con Streamlit:



Descripción, entrada de texto, resultados, reporte.

Entrenamiento y evaluación

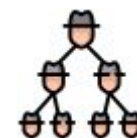


Cargar modelos y datos limpios, entrenar, evaluar, ajustar, integrar

(precisión, recall, F1-score)

Mejoras

creación de diccionarios para el reconocimiento de entidades en el ámbito



Pruebas de Precisión:

Desempeño de los modelos (precisión, recall, F1-score).

Casos reales: Implementación del sistema en escenarios de monitoreo de seguridad.



Análisis de Sentimientos: Identifica emociones (positivas, neutrales o negativas) para calificar cada oración.

Análisis de prioridad: Asigna categorías a mensajes, priorizando información relevante.



Análisis de entidades: Además de NER, identifica, a base de diccionarios, la mención de armas, explosivos, drogas, vehículos, grupos terroristas, nombres y alias de criminales

Conclusiones:

- Se ha logrado crear la herramienta a base de las técnicas de NLP de manera exitosa con potencial de apoyo a la policía
- Es posible identificar emociones y prioridades con éxito con un porcentaje aceptable
- Se ha evaluado el sistema en un entorno controlado
- Mediante las pruebas prácticas se ha logrado clasificar información útil
- Se han generado los reportes en formato PDF de forma instantánea.

NLP Policía Nacional - Ecuador

Ingresa texto relacionado con Ecuador

Nombre del archivo PDF (sin extensión)

reporte_novedades

Analizar

Generar Reporte PDF

Captura de pantalla interfáz de usuario.

Reporte de Novedades - 18/04/2025

Análisis de Novedades

Un informe de inteligencia policial establece que el grupo narcocriminal Los Lobos tiene injerencia en 18 de las 35 cárceles del país. En ciertos centros, ha protagonizado amotinamientos y crímenes en medio de una guerra contra Los Choneros para tener el control del sistema carcelario. Lea también: Menores envueltos en crímenes, el trágico resultado del abandono estatal Las cárceles están militarizadas desde enero de este año, allí, los lobos, habían instalado sus centros de operación criminal. Ecuavisa tuvo acceso a un informe militar que calcula que la banda criminal Los Lobos tiene entre siete mil y ocho mil integrantes; mientras que trabajos de inteligencia policial consideran que tienen presencia en 18 provincias

Tabla de Novedades Registradas

ID	Tipo Alerta	Ubicación	Identidades	Fechas	Vehículo	Armas/Explosivo
1	lobos - Novedad	Los Lobos	N/A	N/A	N/A	N/A
2	choneros - Novedad	N/A	N/A	N/A	N/A	N/A
3	lobos - Novedad	N/A	Menores	N/A	N/A	N/A
4	lobos - Novedad	Ecuavisa, Los Lobos	N/A	N/A	N/A	N/A

Captura de pantalla - reporte en PDF.



Aplicativo en tiempo real:
en red

Proyecciones Futuras:

- **Procesar partes policiales en tiempo real** (análisis de entidades, clasificación y estudio del delito y estructuras criminales)
- **Análisis de redes sociales** (prevención de eventos de conmoción masiva para asignación de talento humano acorde a la necesidad)
- **Acelerar peritajes** (reducir el tiempo de procesamiento de grandes volúmenes de medios digitales recuperados para investigación)
- **Investigación criminal y localización** de personas requeridas por la ley.





UNIVERSITAT
Carlemany

De:

 Planeta Formación y Universidades

Gracias por su atención

