



Excelencia que trasciende

DELVALLE
GRUPO EDUCATIVO

Inteligencia artificial

Proyecto 3

Mauricio Julio Rodrigo Lemus Guzmán – 22461

a. Análisis Exploratorio de Datos (EDA)

El dataset utilizado contiene mensajes de texto etiquetados como "SPAM" o "HAM". Se cargó el archivo spam_ham.csv, que incluía dos columnas principales: Label y SMS_TEXT.

Se identificaron las siguientes características:

- Los mensajes HAM son notablemente más largos que los SPAM en promedio.
- Ciertas palabras como "free", "win", "call", "now" y URLs son frecuentes en mensajes SPAM.
- Palabras comunes en mensajes HAM tienden a ser más conversacionales y personales.

b. Limpieza de Datos

La limpieza se realizó mediante:

- **Tokenización:** División del texto en palabras individuales.
- **Pasar a minúsculas:** Para uniformidad.
- **Eliminación de stop words:** Se removieron palabras comunes del inglés que no aportan información predictiva.
- **Eliminación de puntuación y dígitos.**
- **Stemming:** Reducción de palabras a su raíz gramatical usando PorterStemmer.
- **Detección de URLs:** Se agregó un token especial url_detected cuando el mensaje contenía enlaces.

c. Modelo Probabilístico (Naïve Bayes)

El modelo se basa en la siguiente fórmula bayesiana:

Donde:

- $P(S)$: Proporción de mensajes SPAM en los datos de entrenamiento.
- $P(W|S)$: Frecuencia relativa de una palabra en mensajes SPAM.
- $P(W|H)$: Frecuencia relativa de una palabra en mensajes HAM.

d. Pruebas de Rendimiento

El modelo fue evaluado usando el 20% del dataset como conjunto de prueba.

Se calcularon las siguientes métricas a partir de la matriz de confusión:

- **Precisión:** Proporción de mensajes clasificados como SPAM que efectivamente lo eran.
- **Recall (Sensibilidad):** Proporción de mensajes SPAM que fueron correctamente identificados.
- **F1-Score:** Promedio armónico entre precisión y recall.

También se exploraron diferentes valores de **threshold** para determinar la clasificación final, observando que un valor de 0.5 balancea bien entre precisión y recall.

e. Discusión de Resultados

Los resultados obtenidos muestran un rendimiento competitivo del clasificador bayesiano, a pesar de su simplicidad. Algunas observaciones:

- La eliminación de stop words y uso de stemming ayudaron a generalizar mejor sobre palabras clave.
- El token especial `url_detected` aportó poder predictivo en SPAM.
- El modelo es sensible a palabras con alta frecuencia en un solo grupo, lo cual es deseado.

Limitaciones observadas:

- Palabras nuevas o con poca frecuencia no aportan suficiente información.
- La asunción de independencia entre palabras (Naïve Bayes) no siempre se cumple, pero el rendimiento sigue siendo aceptable.

Conclusión

El proyecto demostró la efectividad de modelos probabilísticos simples como Naïve Bayes en tareas de clasificación de texto. La limpieza adecuada de los datos, el diseño del modelo y la interfaz de usuario contribuyeron a un resultado funcional, interpretable y didáctico.