Actividad 2: Detección de Spam con Regresión Logística

Brayan Rodríguez Joseph Valencia Cruz 10 de septiembre de 2025

# Introducción

En este Actividad, desarrollamos un sistema para detectar correos electrónicos no deseados (spam) utilizando un modelo de aprendizaje automático basado en **Regresión Logística**. El objetivo fue clasificar automáticamente correos como *SPAM* (no deseados) o *HAM* (legítimos) a partir de un conjunto de datos de correos electrónicos. Este informe explica el proceso seguido, con énfasis en cómo funciona la Regresión Logística y los resultados obtenidos.

# Descripción del Proyecto

La actividad consistió en crear un programa en Python que procesa un conjunto de datos de correos electrónicos, los limpia, los convierte en un formato que el modelo pueda entender y entrena un modelo de Regresión Logística para clasificarlos. A continuación, se describen los pasos principales:

* 1. **Carga de datos**: Se cargó un archivo CSV con correos electrónicos y sus etiquetas (SPAM o HAM) y sus diferentes features. Los datos se mezclaron para evitar sesgos y se mostraron estadísticas básicas, como el número total de correos y la proporción de SPAM y HAM.
  2. **Preprocesamiento**: Se limpió el texto de los correos convirtiéndolo a minúsculas, eliminando signos de puntuación, palabras comunes (como “el” o “la”) y espacios innecesarios.
  3. **Vectorización**: Se transformó el texto limpio en números utilizando una técnica llamada TF-IDF, que asigna un peso a cada palabra según su importancia.
  4. **División de datos**: Los datos se dividieron en 80 % para entrenamiento y 20 % para prueba, manteniendo la proporción de SPAM y HAM.
  5. **Entrenamiento**: Se entrenó un modelo de Regresión Logística para aprender a clasificar los correos.
  6. **Evaluación**: Se midió el rendimiento del modelo con métricas como precisión, sensibilidad y F1-Score, además de una matriz de confusión.
  7. **Análisis**: Se identificaron las palabras más importantes para distinguir entre SPAM y HAM, y se mostraron ejemplos de predicciones.

# Regresión Logística

La **Regresión Logística** es un algoritmo de aprendizaje automático utilizado para clasificar datos en dos categorías, en este caso, SPAM o HAM. Funciona calculando la probabilidad de que un correo pertenezca a una categoría específica basándose en las palabras que contiene. Por ejemplo, si un correo tiene palabras como “gratis” o “oferta”, es más probable que sea SPAM.

 Es ideal para problemas de clasificación binaria (dos clases: SPAM o HAM).

 Proporciona probabilidades claras, lo que ayuda a entender cuánto confía el modelo en su predicción.

 Funciona bien con datos de texto transformados en números (como los obtenidos con TF-IDF).

 Es un estándar confiable en aprendizaje automático, fácil de implementar y con buenos resultados.

El modelo se entrenó con el 80 % de los datos y se probó con el 20 % restante. Durante el entrenamiento, el algoritmo ajustó sus parámetros para minimizar errores en la clasificación.

# Resultados

El modelo fue evaluado con varias métricas:

 **Precisión general (Accuracy)**: Mide el porcentaje de correos clasificados correctamente.

 **Precisión (Precisión)**: Indica qué proporción de correos clasificados como SPAM eran realmente SPAM.

 **Sensibilidad (Recall)**: Mide cuántos correos SPAM reales fueron detectados.

 **F1-Score**: Combina precisión y sensibilidad para dar una medida balanceada.

Además, se generó una **matriz de confusión** para mostrar cuántos correos fueron clasificados correcta o incorrectamente. También se analizaron las palabras más importantes, como “gratis” o “urgente” para SPAM, y palabras más comunes en correos legítimos, como “trabajo” o “reunión”. Finalmente, se mostraron ejemplos de predicciones para verificar cómo el modelo clasificaba correos específicos.

# Conclusión

Desarrollamos un sistema efectivo para detectar correos SPAM utilizando Regresión Logística. El proceso incluyó cargar y limpiar datos, transformarlos en números, entrenar un modelo y evaluar su rendimiento. La Regresión Logística demostró ser una herramienta poderosa para este problema, ya que es simple, interpretable y efectiva con datos de texto. Este proyecto nos permitió entender cómo el aprendizaje automático puede resolver problemas prácticos, como

filtrar correos no deseados, y cómo interpretar los resultados para mejorar el modelo en el futuro.