

Proyecto #2
CC3085 Inteligencia Artificial

De manera individual, realice un proyecto donde implemente los conceptos vistos de clase de probabilidad y redes bayesianas para construir un clasificador de SPAM/HAM.

Entrega

- Trabajo escrito: **3 de mayo 11:59 PM**
- Formato: PDF
- Presentación: **3 de mayo, presencial, periodo de clase**

Trabajo escrito

Elabore un reporte detallado sobre los pasos necesarios para la correcta construcción de un filtro SPAM/HAM usando bayes.

La estructura mínima de su reporte debe comprender:

a. Análisis de datos exploratorio (EDA).

Características sobre el data set que considere relevantes o interesantes para explicar la diferencia sobre los grupos de SPAM/HAM.

b. Limpieza de datos

Explique los pasos y beneficios de todos los procesos de limpieza a los que someta su información:

Ex. Eliminar stop words, lemmatization, stemming.

c. Modelo

Explique el paso a paso de los cálculos necesarios para calcular la probabilidad que el texto permanezca a SPAM o HAM.

Utilice la formulas en el anexo de este documento.

d. Pruebas de rendimiento.

Considerando que genero un modelo utilizando un subconjunto de datos de entrenamiento, mida el rendimiento de su modelo sobre el dataset de prueba.

Basándose en los valores de la matriz de confusión, obtenga valores de **recall, precisión y f1-score** de su modelo.

e. *Discusión de resultados*

Discuta el significado de los valores de las métricas obtenidas en las pruebas de rendimiento. Explique el impacto de las decisiones que tomó desde el proceso de limpieza y los cálculos del modelo y plantee el impacto positivo o negativo que cada uno de estos pudo tener en el rendimiento de su modelo.

Siéntanse en la libertad de agregar fases adicionales en caso lo consideren necesario.

Presentación en vivo

En conjunto con la programación del clasificador, debe construir un módulo que cumpla con lo siguiente:

- Acepta el ingreso de un prompt tipo texto.
- Retorna la posibilidad de que el texto anterior sea SPAM y también las palabras que reconoció del texto.

Rubrica

Trabajo escrito	EDA	10 pts
	Limpieza de datos	10 pts
	Modelo	10 pts
	Pruebas de rendimiento	20 pts
	Discusión	20 pts
Presentación presencial	Resultados	30 pts

Anexo

Probabilidad de que un texto sea SPAM dado que contiene la palabra W:

$$P(S|W) = \frac{P(W|S)P(S)}{P(W|S)P + P(W|H)P(H)}$$

Donde,

$P(S|W)$ es la probabilidad de que un texto sea SPAM dado que contiene la palabra W.

$P(W|S)$ es la probabilidad de la palabra W aparezca en un texto que es SPAM.

$P(W|H)$ es la probabilidad de la palabra W aparezca en un texto que es HAM.

$P(S)$ es la probabilidad de que cualquier texto sea SPAM.

$P(H)$ es la probabilidad de que cualquier texto sea HAM.

Probabilidad de que un texto sea SPAM dado que contiene las palabras W_1 a W_n :

$$P(S|\mathbf{W}) = \frac{P_1 P_2 \dots P_n}{P_1 P_2 \dots P_n + (1 - P_1)(1 - P_2) \dots (1 - P_n)}$$

Probabilidad de cada palabra

$$P(W) = \frac{\frac{N_{w,s}}{N_s}}{\frac{N_{w,h}}{N_h} + \frac{N_{w,s}}{N_s}}$$

Donde,

$N_{w,s}$ es la cantidad de correos SPAM que contienen la palabra W .

$N_{w,h}$ es la cantidad de correos HAM que contiene la palabra W .

N_s es la cantidad de correo de SPAM

N_h es la cantidad de correos de HAM.