Informe de Modelo “Naive Bayes” para Detección de Noticias

1. Objetivo del Proyecto

Se entrenó un modelo clasificador Naive Bayes, un modelo de aprendizaje supervisado, para diferenciar automáticamente entre noticias reales y falsas basándose en el texto de las mismas.

2. Metodología

• Datos: Se utilizaron un conjunto de 10 noticias previamente etiquetadas.

• Preparación: El texto fue convertido a vectores numéricos para que el modelo pueda procesarlo.

• Entrenamiento: Se usó el 80% de los datos para entrenar al modelo, que aprendió a asociar ciertas palabras con cada categoría (real o falso).

• Evaluación: El 20% restante de los datos se usó para probar el rendimiento del modelo con información que no había visto antes.

3. Resultados del Modelo

• Precisión Total (Accuracy): [Aquí pones el valor numérico, ej: 100%]

◦ Esta métrica indica el porcentaje de predicciones correctas que hizo el modelo en el conjunto de prueba.

• Matriz de Confusión: [Aquí insertas la matriz de 2x2 que generó tu código]

◦ La matriz muestra que el modelo clasificó correctamente [ej: 1 noticia 'real' y 1 noticia 'falsa'], sin cometer errores.

4. Clasificación de Nuevas Noticias

El modelo entrenado se utilizó para clasificar nuevas noticias, con los siguientes resultados:

• "Nuevo estudio demuestra que el café mejora la memoria" -> Clasificada como: [real/falso]

• "Expertos afirman que los gatos pueden hablar con humanos" -> Clasificada como: [real/falso]

5. Conclusión

El modelo clasificador Naive Bayes demostró ser muy eficaz para esta tarea con el conjunto de datos proporcionado. A pesar de su simplicidad, logra una alta precisión.

Es recomendable entrenarlo con un gran volumen de datos para poder aumentar su confiabilidad y precisión en distintos casos.