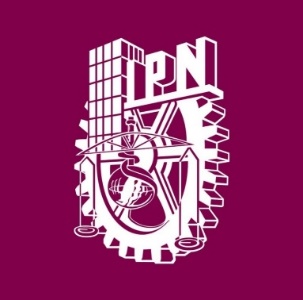
**INSTITUTO POLITÉCNICO NACIONAL**

**ESCUELA SUPERIOR DE CÓMPUTO**

**Ejercicio de Lab8: Naive Bayes**

**Alumno: Delgado Acosta Luis Bernardo**

**Franco Calderas Sergio Alberto**

**Grupo: 5BV1**

**Clasificador Naive Bayes**

El clasificador **Naive Bayes** es un algoritmo de aprendizaje supervisado basado en la teoría de probabilidad de Bayes y en la suposición de independencia entre características. Se llama “Naive” o ingenuo porque asume que las características son independientes entre sí, una simplificación que rar vez es cierta en la práctica, pero que hace que el modelo sea más rápido y eficiente.

En esta práctica se usa el clasificador Gaussian Naive Bayes de la librería de Sklearn. En este modelo, se asume que cada característica sigue una distribución normal dentro de cada clase. Esto significa que los valores de cada característica se distribuyen alrededor de una media específica con cierta varianza, formando la clásica “curva en forma de campana”.

**Implementación**

**Texto

Descripción generada automáticamente**

*Img.1 Implementación de GaussianNB*

Dentro de cada método de validación (Hold-Out, 10 Fold-Cross, Leave-One-Out) se implementó la clasificación por Naive Bayes.

Recibe el conjunto de datos completo como un DataFrame, en *test\_size* especifica la proporción de datos que se destinarán al conjunto de prueba. Por defecto, se utiliza el 30% de los datos de prueba y el 70% para el entrenamiento. Se separa las características y las etiquetas, posteriormente se calcula el número total de observaciones, el número total que se usarán para el conjunto de prueba, se crea una lista de índices para mezclarlos aleatoriamente. Se divide los índices de entrenamiento y prueba para hacer sus listas correspondientes. Finalmente, se entrena el modelo, y genera predicciones para el conjunto de prueba usando el modelo entrenado. Y así en los demás respectivos métodos de validación.

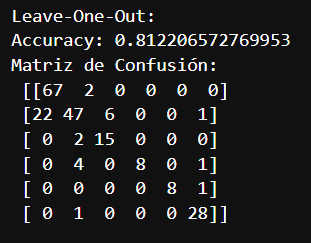
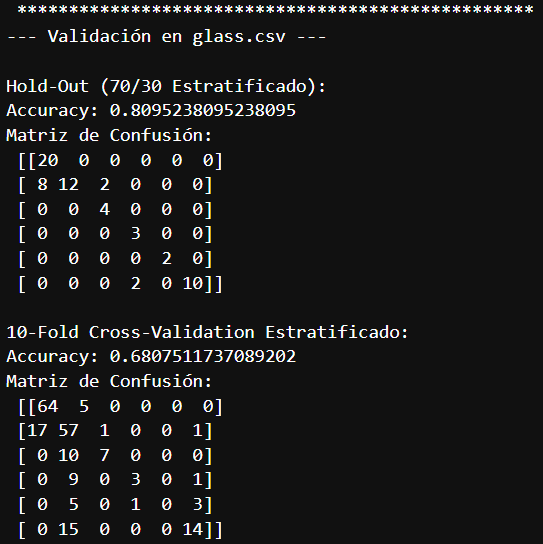
**Resultados**

Texto

Descripción generada automáticamente

*Img. 2 Resultados de raisin*

Las precisiones en los tres métodos son cercanas, todas alrededor del 82%, lo que indica que el clasificador Naive Bayes mantiene un rendimiento consistente y estable en los diferentes métodos de validación.



*Img. 3 Resultados de glass*

En el método de 10-Fold-Cross- Validation muestra dificultad para clasificar las clases, lo que sugiere que Naive Bayes tiene problemas al generalizar en estas clases debido a la escasez de ejemplos.

Texto

Descripción generada automáticamente

*Img. 4 Resultados de iris*

Como se observó en los resultados el clasificador Naive Bayes tiene la suposición de que las características siguen una distribución normal este puede no ser realista en ciertos casos, lo que afecta la precisión del modelo. Asume que las características son independientes entre sí. Esto reduce la precisión si las características están correlacionadas como el caso del data set de glass.

Enlaces a los repositorios

Bernardo: <https://github.com/BernardoD07/Machine-Learning-ESCOM.git>

Sergio: https://github.com/SergioCalderas/Trabajos-Machine-Learning.git