

T2	Minería de Datos #12 Hernández Hernández Karla Guadalupe #18 Pérez Arellano Narda Teresa #25 Saldaña de la Fuente Ana Karen #29 Torres Rivera Paul	8°B	E#1
-----------	---	------------	------------

En esta práctica se mostrará el funcionamiento del método de clasificación Naive Bayes que consiste en el proceso de encontrar la hipótesis más probable, dada un conjunto de ejemplos de entrenamiento y un conocimiento a priori sobre la probabilidad de cada hipótesis.

Requerimientos

Sistema Operativo: Windows 8.1

Python 3.5.1

Anaconda 2.4.1

* Dataset: 10 instancias

* Prueba (nueva a instancia a clasificar)

* Dataset

El dataset que se utilizó en esta práctica es un archivo de texto plano (.txt) que contiene 10 instancias, cada una con 3 atributos:

- Fiebre (True/False)
- Dolor de cabeza (True/False)
- Fluido Nasal (Constipación/Normal/Escurrimiento)

Y 2 posibles clases:

- Gripe (Yes)
- No Gripe (No)

Estos datos representan 3 síntomas que se presentan en personas con gripe o no gripe.

* Prueba

Es la nueva instancia que se desea clasificar, donde se define el valor de cada uno de los tres atributos.

Fiebre, Dolor de Cabeza, Fluido Nasal

→ True, False, Constipación

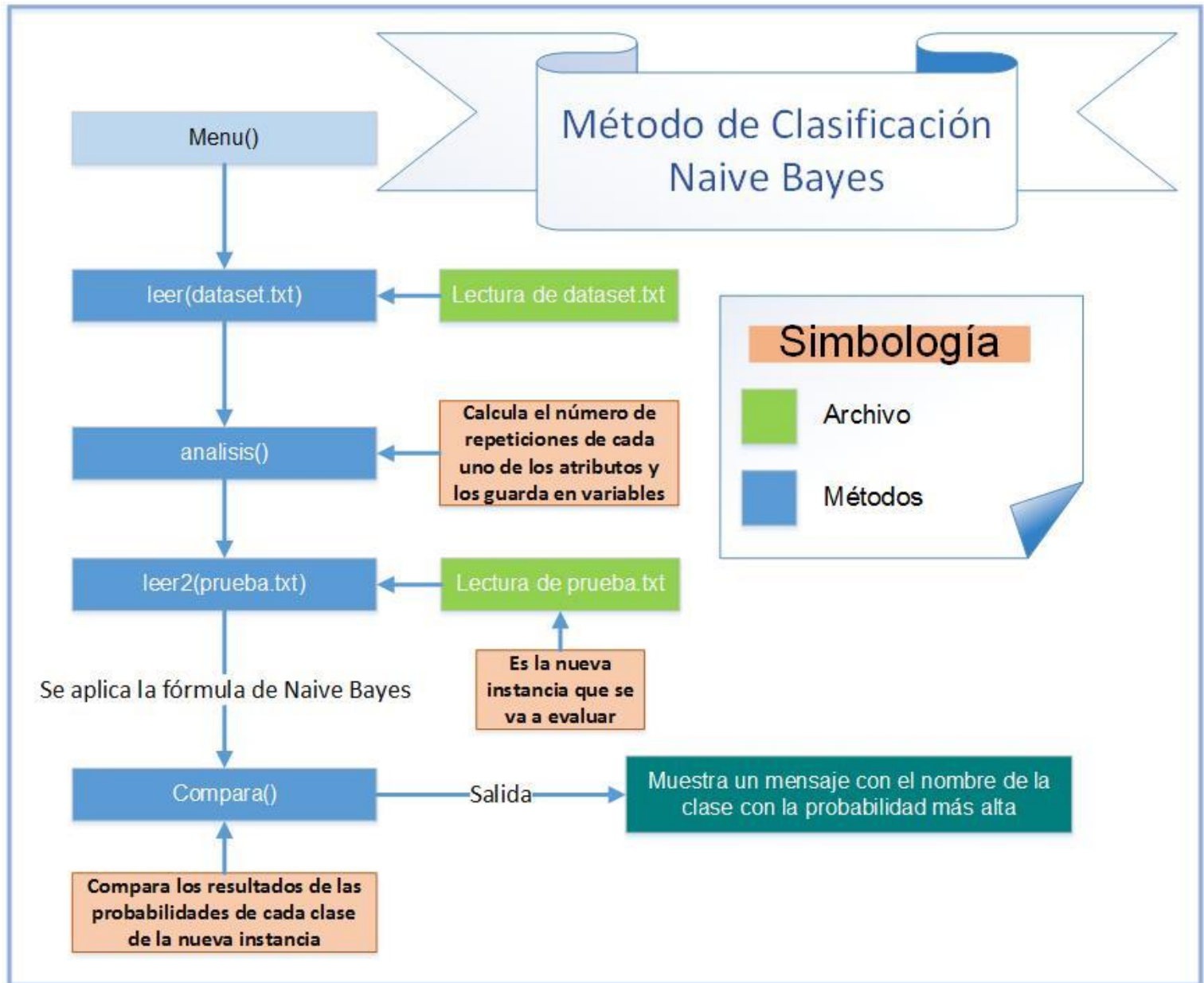
Procedimiento

1. A partir del dataset: calcular la probabilidad de cada una de las clases
2. Para cada clase realizar un recuento de los valores de atributos que toma cada instancia. Determinando así la probabilidad de que sucedan los valores de los atributos dado cada una de las clases
3. Clasificar una nueva instancia:
 - a. Para cada clase disponible, se determinan los valores de probabilidad de cada valor de los atributos de la nueva instancia
 - b. Aplicar la fórmula de Naive Bayes
 - c. Seleccionar la clase con la probabilidad más alta

Funcionamiento del programa

1. Ejecutar el programa
2. Insertar la ruta del archivo (.txt) que será el dataset
3. El programa mostrará en pantalla el número de elementos que contiene el dataset.
4. En la parte de prueba; insertar la ruta del archivo (.txt) de prueba
5. El programa evaluará y mostrará en pantalla las probabilidades de que suceda cada una de las clases. También mostrará un mensaje de con el nombre de la clase con la probabilidad más alta.

Diagrama



Código y archivos disponibles en: Github <https://github.com/poltores7/bayes>