**Pruebas del algoritmo Random Forest Classifier**

A continuación, se presentarán distintas pruebas del algoritmo con las cuales podremos revisar cual es el accuracy de nuestro modelo dependiendo de los distintos híper parámetros que utilicemos, así como las columnas disponibles, entre otras cosas.

Primera prueba:

Primero se nos pide el nombre del archivo y si queremos especificar columnas a usar o a quitar

Una captura de pantalla de un celular con texto e imagen

Descripción generada automáticamente con confianza media

Se nos pide también el número de columnas que utilizaremos (si elegimos la opción 1), así como que especifiquemos cuales serán estas

Texto

Descripción generada automáticamente con confianza media

También se nos pide especificar nuestra Y, y se nos despliega una lista con los posibles valores que puede tomar Y

Texto

Descripción generada automáticamente

También podemos especificar los híper parámetros de nuestro árbol

Texto

Descripción generada automáticamente con confianza media

Con esto se nos otorga un accuracy de nuestro modelo

Texto

Descripción generada automáticamente

También se nos despliega un pequeño gráfico que nos muestra la diferencia entre el valor real de train y la predicción que se hizo con los datos de train

Gráfico

Descripción generada automáticamente

Posteriormente se nos pide que asignemos valores a las columnas que especificamos serían nuestra X

Texto

Descripción generada automáticamente

Y por último nos arroja las probabilidades sobre las clases a las que puede pertenecer nuestra predicción y la clase a la que se predice que pertenece teniendo en cuenta las probabilidades

Interfaz de usuario gráfica, Texto, Aplicación, Sitio web

Descripción generada automáticamente

**Prueba con otro archivo (Iris) hacemos lo mismo que con el anterior:**

Primero se nos pide el nombre del archivo y si queremos especificar columnas a usar o a quitar

Texto

Descripción generada automáticamente

Se nos pide también el número de columnas que utilizaremos (si elegimos la opción 1), así como que especifiquemos cuales serán estas

Texto

Descripción generada automáticamente

También se nos pide especificar nuestra Y, y se nos despliega una lista con los posibles valores que puede tomar Y

Texto

Descripción generada automáticamente

También podemos especificar los híper parámetros de nuestro árbol

Texto

Descripción generada automáticamente con confianza media

Con esto se nos otorga un accuracy de nuestro modelo

Texto

Descripción generada automáticamente

También se nos despliega un pequeño gráfico que nos muestra la diferencia entre el valor real de train y la predicción que se hizo con los datos de train

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Posteriormente se nos pide que asignemos valores a las columnas que especificamos serían nuestra X

Texto

Descripción generada automáticamente

Y por último nos arroja las probabilidades sobre las clases a las que puede pertenecer nuestra predicción y la clase a la que se predice que pertenece teniendo en cuenta las probabilidades

Texto

Descripción generada automáticamente

Con el ejemplo anterior podemos variar los híper parámetros de nuestro árbol y ver cuanto eso repercute en nuestro accuracy

Texto

Descripción generada automáticamente

Y nuestra predicción tanmbien cambia a pesar de ser los mismos datos

Texto

Descripción generada automáticamente

**Ahora con híper parámetros diferentes un tanto más grandes**

**Imagen que contiene Texto

Descripción generada automáticamente**

Podemos apreciar que el accuracy se mantuvo, pero la gráfica cambio en cuanto a que predicciones erraron

Gráfico, Gráfico de líneas

Descripción generada automáticamente

En conclusión: los modelos dependen de muchas cosas, no solo del modelo en sí, sino también en los híper parámetros que nosotros coloquemos, como se encuentre nuestro dataset, si es que este tiene valores sesgados o no, es importante mencionar que para analizar datos de una forma más concisa se requiere el analizar los datos de forma profunda incluso antes de entra al modelo, realizar una limpieza adecuada y estandarizar valores, con la finalidad de ayudar a nuestro modelo a realizar mejores predicciones y dependiendo de dicha preparación podremos esperar unos resultados u otros.

**Prueba con otro data set**

**Interfaz de usuario gráfica, Texto

Descripción generada automáticamente**

**Texto

Descripción generada automáticamente**

Gráfico, Histograma

Descripción generada automáticamente

**Interfaz de usuario gráfica, Texto

Descripción generada automáticamente**