Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias de la Computación

Máquinas de aprendizaje

Reporte: Un ejemplo de XGBoost

Docente: Abraham Sánchez López

Alumno	Matrícula
Taisen Romero Bañuelos	202055209

Un ejemplo de XGBoost

La práctica es una continuación de lo que hemos visto de XGBoost. Al inicio se hace mención sobre PCA para reducir columnas (características), se menciona que no es tan necesario pero me parece curioso que al final, con la tabla de importancia de características, veamos que tal vez si sea buena idea hacer una reducción de características. Me parece interesante porque normalmente en el EDA uno ve de cierta forma el futuro de la práctica porque los modelos suelen ser una corroboración de lo que se encuentra con ligeros ajustes, sin embargo, en este caso al final vemos un leve replanteamiento de los pasos a seguir. En este sentido, lo más sencillo sería eliminar las clases que tengan una importancia menor a cierto umbral (por ejemplo, 0.01), pero creo que valdría la pena hacer un análisis de las clases.

Basándose en la matriz de confusión vemos que al modelo le cuesta reconocer la diferencia entre las clases 2 y 3, posiblemente porque se parezcan entre sí. De hecho, podemos corroborar la hipótesis de que se parecen porque en el gráfico t-SNE los puntos correspondientes a la clase 2 están muy cerca a los de la clase 3, lo que sugiere que están relacionados entre sí.

```
Confusion Matrix and Statistics

          Reference
Prediction    1    2    3    4    5
         1 5567   10    2    1    0
         2   15 3761   21    0    0
         3    2   18 3388   14    0
         4    0    0   22 3190    4
         5    0    1    0   11 3595
```

Quizá en lugar de eliminar las variables de importancia menor a 0.01 podríamos hacer un PCA a las clases 2, 3 y 4 para que se puedan diferenciar más. Aunque también podríamos eliminar unas cuantas características de importancia irrelevante pero esto sería mejor hacerlo después del PCA porque si diferenciamos las características puede darse el caso de que adquieran una importancia diferente.

Otra alternativa sería hacer una rejilla para combinar parámetros, aunque dado su costo computacional y de tiempo dejaría esta opción hasta el final.