

PRIMERA ENTREGA PROYECTO DE INTELIGENCIA ARTIFICIAL

Integrante: Juan José Ramírez Cuervo

1) Problema predictivo a resolver:

Dados los datos obtenidos mediante resonancia magnética nuclear, se quiere predecir las interacciones entre dos átomos en una molécula, es decir, la constante de acoplamiento escalar. Este experimento es útil para comprender mejor la estructura y la dinámica de las moléculas en áreas como la ciencia ambiental, la ciencia farmacéutica y la ciencia de materiales.

2) Dataset a utilizar:

<https://www.kaggle.com/competitions/champs-scalar-coupling/data>

Este dataset tiene más de 130000 muestras de las interacciones entre los átomos, 47 columnas de las cuales algunas son cualitativas al indicar el índice del átomo y se eliminarán el 5% o más de los datos al evaluarlos.

3) Métricas de desempeño requeridas (de machine learning y de negocio)

Los datos se evalúan con el método Mean Absolute Error, se calculan para cada tipo de acoplamiento escalar y luego se promedian entre tipos, de modo que una disminución del 1 % en MAE para un tipo proporciona la misma mejora en la puntuación que una disminución del 1 % para otro tipo. .

4) Primer criterio sobre cuál sería el desempeño deseable en producción.

Para esta métrica, el MAE para cualquier grupo tiene un piso de $1e-9$, por lo que la puntuación mínima (mejor) posible para predicciones perfectas es de aproximadamente -20,7232.