

entrega 1 - inteligencia artificial

1. Planteamiento del problema

El problema predictivo que vamos a resolver es Seoul Bike Trip Duration Prediction, en el cual trataremos de estimar la duración de un viaje en bicicleta teniendo en cuenta datos atmosféricos.

con estas estimaciones se podrían dar un idea general a los turistas y demás usuarios que alquilen cíclicas por cuánto tiempo podría alquilarla y con esta información calcular si dispone del suficiente dinero para alquilarla, puede observar si es más barato alquilar un bicicleta o tomar otro medio de transporte y también puede mirar cuanto tiempo se demora de ir de un lugar a otro si decide ir en bicicleta

2. Dataset

El dataset que vamos a utilizar es el siguiente: [enlace](#). el cual tiene 26 columnas y x (descargar el dataset) instancias, nuestro equipo agrega las 4 faltantes columnas.

ID- ID del viaje

Duration- la duración del viaje

Distance- la distancia del viaje

PLong - longitud recogida de la cicla

PLatd - latitud recogida de la cicla

DLong - longitud - entrega de la cicla

DLatd - latitud - entrega de la cicla

Haversine - distancia entre dos puntos

Pmonth - mes de recogida de la cicla

Pday - dia recogida de la cicla

Phour - hora recogida de la cicla

Pmin - minuto recogida de la cicla

PDweek - semana recogida de la cicla

Dmonth - mes entrega de la cicla

Dday - dia entrega de la cicla

Dhour - hora entrega de la cicla

Dmin - minuto entrega de la cicla

DDweek - semana entrega de la cicla

Temp - temperatura

Precip - precipitación

Wind - velocidad del viento

Humid - humedad

Solar - radiación solar

Snow - nevadas

GroundTemp - temperatura del suelo

Dust - concentración de polvo promedio

y las categorías que vamos a agregar son las siguientes

Speed - velocidad del viaje

Climate - clima del viaje(frio, calido, caliente)

TPrecip - tipo de precipitación(Ilovizna, lluvia, aguacero)

Ttravel - (largo, mediano, corto) se calcula por la distancia del viaje

3. Al ser una regresión utilizaremos como métrica de machine learning la MSE (Error cuadrático medio), MSE básicamente mide el error cuadrado promedio de nuestras predicciones. Para cada punto, calcula la diferencia cuadrada entre las predicciones y el objetivo y luego promedia esos valores.

Cuanto mayor sea este valor, peor es el modelo. Nunca es negativo, ya que estamos cuadrando los errores de predicción individuales antes de sumarlos, pero sería cero para un modelo perfecto. La MSE se define por la ecuación.

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Donde y_i es el resultado real esperado y \hat{y}_i es la predicción del modelo.

4. Desempeño

Con las predicciones de esta inteligencia artificial se espera que las personas empiecen a utilizar métodos de viajes alternativos como la cicla, ya que podrán comparar con precisión los tiempos de viajes que se demoraría en cicla con otros medios de transportes teniendo en cuenta que los otros medios de transporte como en carros o buses se podrían llegar a demorar mas por el tráfico.