

Inteligencia Artificial
Entrega 1

Estudiantes
Jean Carlos Julio Rodriguez
Anamaria Romero Carvajal

Docente
Raul Ramos Pollan

UNIVERSIDAD DE ANTIOQUIA
1 8 0 3

Universidad de Antioquia
Facultad de Ingeniería
2023-2

1.Planteamiento del problema

Los accidentes de tránsito son una problemática global que afecta a la seguridad y la salud de las personas. Los datos estadísticos muestran que este problema se ha convertido en una gran preocupación para los gobiernos, ya que implica un alto costo social y económico.

El buen manejo de los datos puede ser una herramienta valiosa para la prevención de accidentes de tránsito. Mediante el análisis de datos se pueden identificar las causas y consecuencias de los accidentes, lo cual puede ayudar a desarrollar políticas y medidas de prevención efectivas.

En este proyecto se utilizará un conjunto de datos de accidentes de tránsito para predecir la accidentalidad. El análisis de los datos se centrará en las causas y consecuencias de los accidentes, con el objetivo de identificar factores de riesgo que puedan ser abordados desde la prevención.

2.Elección de dataset

El conjunto de datos que se utilizará para este proyecto proviene de Kaggle, compuesto por un archivo CSV con datos de accidentes de tránsito ocurridos en Reino Unido entre 2005 y 2014, recopilados por el gobierno de la misma localidad. El conjunto de datos contiene un registro de más de 1,8 millones de accidentes, de los cuales se utilizarán datos más recientes, correspondientes al año 2014, que suman un total de 146.322 accidentes. El conjunto de datos contiene la siguiente información:

- 1) Accident_Inde.
- 2) Location_Easting_OSGR.
- 3) Location_Northing_OSGR.
- 4) Longitude.
- 5) Latitude.
- 6) Police_Force.
- 7) Accident_Severity.
- 8) Number_of_Vehicles.
- 9) Number_of_Casualties.
- 10) Date.
- 11) Day_of_Week.
- 12) Time.
- 13) Local_Authority_(District).
- 14) Local_Authority_(Highway).
- 15) 1st_Road_Class.
- 16) 1st_Road_Number.
- 17) Road_Type.
- 18) Speed_limit.
- 19) Junction_Control.
- 20) 2nd_Road_Class.
- 21) 2nd_Road_Number.
- 22) Pedestrian_Crossing-Human_Control.

- 23) Pedestrian_Crossing Physical_Facilities,Light_Conditions.
- 24) Weather_Conditions.
- 25) Road_Surface_Conditions.
- 26) Special_Conditions_at_Site.
- 27) Carriageway_Hazards.
- 28) Urban_or_Rural_Area.
- 29) Did_Police_Officer_Attend_Scene_of_Accident.
- 30) LSOA_of_Accident_Location.
- 31) Year.

3.Métrica

La métrica principal que se utilizará para evaluar el rendimiento del modelo de predicción de accidentes de tránsito es la raíz del error cuadrático medio (RMSE). calculado de la siguiente manera:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Donde RMSE es la raíz cuadrada del promedio de la suma de diferencias cuadradas entre los valores observados en la serie y los esperados según el modelo de tendencia. Dónde y_i corresponde a la serie observada e \hat{y}_i a la serie estimada, y n el número de datos totales. Los valores de RMSE más bajos indican que el modelo es más preciso.

4.Bibliografía

- “Road Accident (United Kingdom (UK)) dataset,” *Kaggle*, May 28, 2022. <https://www.kaggle.com/datasets/devansodariya/road-accident-united-kingdom-uk-dataset?resource=download>