

Primera entrega de proyecto

POR:

Felipe Carlos Martínez Mármol

Juan Camilo Tabares Henao

Jorge Antonio Franco Vasquez

MATERIA:

Introducción a la inteligencia artificial

PROFESOR:

Raul Ramos Pollan



UNIVERSIDAD DE ANTIOQUIA

FACULTAD DE INGENIERÍA

MEDELLÍN 2022

1. Planteamiento del problema.

En cualquier parte del mundo, muchas personas son víctimas cada año en accidentes de tránsito, ya sea por imprudencia o confusiones. Lo cierto es que los datos estadísticos muestran que se han convertido en una gran problemática para los países del mundo, ya que no sólo deben lidiar con la circulación de tránsito, sino también con la salud de las personas afectadas.

La importancia del buen manejo de los datos puede ser vital en situaciones que impliquen riesgo en la seguridad de las personas como por ejemplo los accidentes automovilísticos, mediante este dataset se buscará predecir la accidentalidad teniendo en cuenta las causas y consecuencias, para poder trabajar desde la prevención y reducir la tasa de accidentalidad.

2. Dataset

El dataset que se va a utilizar proviene de Kaggle, el cual está compuesto por un archivo CSV con datos de accidentes de tránsito ocurridos en Reino Unido desde el año 2005 al 2014, recolectados por el gobierno del Reino Unido.

El dataset contiene un registro de más de 1.8 millones de accidentes, sin embargo para el proyecto solo se usarán los datos más recientes, es decir los correspondientes al año 2014, los cuales son 146322 accidentes en total, y contiene la siguiente información:

- **Accident_Index** Índice de accidentes
- **Location_Easting_OSGR** Ubicación Este
- **Location_Northing_OSGR** Ubicación de Norte
- **Longitude** Longitud del lugar de accidente
- **Latitude** Latitud del lugar del accidente
- **Police_Force** No. de Fuerza Policial
- **Accident_Severity** Severidad del accidente en una escala de 1 a 5

- **Number_of_Vehicles** - Número de vehículos involucrados en el accidente.
- **Number_of_Casualties** - Número de víctimas
- **Date** - Fecha
- **Day_of_Week** - Día de la semana
- **Time** - Hora
- **Local_Authority_(District)** Autoridad Local (Distrito)
- **Local_Authority_(Highway)** Autoridad Local (Carretera)
- **1st_Road_Class** - Tipo de la 1ra carretera
- **1st_Road_Number** - Número de la 1ra carretera
- **Road_Type** - Tipo de carretera
- **Speed_limit** - Límite de velocidad
- **Junction_Control** - Control en la intersección
- **2nd_Road_Class** - Tipo de la 2da carretera
- **2nd_Road_Number** - Número de la 2da carretera
- **Pedestrian_Crossing-Human_Control** - Control humano de peatones
- **Pedestrian_Crossing-Physical_Facilities** - Instalaciones físicas para el cruce de peatones
- **Light_Conditions** - Condición de iluminación el día del accidente

- **Weather_Conditions** - Condiciones meteorológicas el día del accidente
- **Road_Surface_Conditions** - Condiciones de la superficie de la carretera en un punto accidental
- **Special_Conditions_at_Site** - Condiciones especiales en el sitio
- **Carriageway_Hazards** - Peligros de la calzada
- **Urban_or_Rural_Area** - Área urbana o Rural
- **Did_Police_Officer_Attend_Scene_of_Accident** - ¿El oficial de policía asistió a la escena del accidente?
- **LSOA_of_Accident_Location** - “Lower Layer Super Output Area” es un sustituto para la locación geográfica de longitud y latitud
- **Year** - Año del evento accidental

3. Métricas

La principal métrica que se utilizará en el modelo de predicción de accidentes de tránsito es la Raíz del Error Cuadrático Medio o Root Mean Square Error (RMSE) y se calcula de la siguiente manera

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

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. En los resultados se busca que mientras menor sea el error, más adecuado es el modelo.

4. Desempeño

Dados los parámetros ingresados como la posición y otras condiciones como el clima o la hora del día, se pretende predecir el número de víctimas que pueda generar un accidente de tránsito y con ello poder implementar medidas de seguridad en las diferentes ciudades del Reino Unido.

La expectativa del modelo de predicción de accidentes de tránsito está muy ligado a la ubicación geográfica y a la redes viales, el objetivo es obtener un esquema y modelo de análisis que pronostique los lugares más críticos donde pueden ocurrir confusiones o problemas de tránsito que podría provocar pérdidas humanas o colisiones.

5. Bibliografía

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