

Car Accident Severity Report



Student:

Rishabh Shinde

Course:

Applied Data Science Capstone

Introduction to the Problem

Weather acts through visibility impairments, precipitation, high winds, and temperature extremes to affect driver capabilities, vehicle performance (i.e., traction, stability, and maneuverability), pavement friction, roadway infrastructure, crash risk, traffic flow, and agency productivity. The table below, summarizes the impacts of various weather events on roadways, traffic flow, and operational decisions.

Prediction of future weather conditions has significant impact on social and economic areas of human life. By gathering the weather data, meteorology opens the possibility of analyzing significant patterns in large amounts of data. Over 150 thousand lives annually are claimed due to the climate changes in temperature and precipitation trends

Climate changes also affect traffic flow. By changing external conditions in which transport takes place and which affect the health or concentration of driver's unfavorable meteorological conditions can lead to traffic accidents, injuries, and death. World Health Organization indicates that the number of road traffic deaths is troublesome and has plateaued at 1.25 million per year.

This study tries to determine the correlation between weather conditions and traffic accident occurrences by analyzing collected data. Data analysis is one of the activities of data science focused on obtaining important information from collected data

Business Understanding

To reduce the frequency of car collisions in a community, an algorithm must be developed to predict the severity of an accident given the current weather, road and visibility conditions. When conditions are bad, this model will alert drivers to remind them to be more careful or possible change travel date, mode or time, so that it would be safer.

Target Audience

The target audience for this project would be:

- Daily commuters, who have travel on a regular basis for work. This project would warn them of possible dangers while driving based on the weather condition, road conditions and the time of the day
- It would also help the first responders like ambulance and firefighter services. The system could help them stay prepared in advance by indicating them of high chances of accidents that could take place.

Thus, ensuring anyone in need could receive their help without much delay

Data Understanding

The data has been retrieved and processed through various sources, and database.

The main source being the data-collisions csv file

Our predictor or target variable will be 'SEVERITYCODE' because it is used measure the severity of an accident from 0 to 5 within the dataset. Attributes used to weigh the severity of an accident are 'WEATHER', 'ROADCOND' and 'LIGHTCOND'.

SEVERITYCODE

Severity codes are as follows:

- 0- Little to no probability (Clear conditions)
- 1- Very Low probability (chance or damage property)
- 2- Low Probability - Chance of Injury
- 3- Mild Probability - Chance of Serious Injury
- 4- High Probability - Chance of Fatality

WEATHER

- Overcast: Overcast sky conditions occur when clouds cover all or most of the sky and cause low visibility conditions
- Rains: Heavy or moderate rainfall, which causes roads to be slippery
- Clear: Clear weather conditions

ROADCOND

This attribute refers to the road condition for a day

- Wet: Usually refers to wet and snowy conditions on a day
- Dry: Normal road conditions

LIGHTCOND

This attribute gives information of light conditions when the accident took place and will be useful in predicting in what conditions an accident is probable.

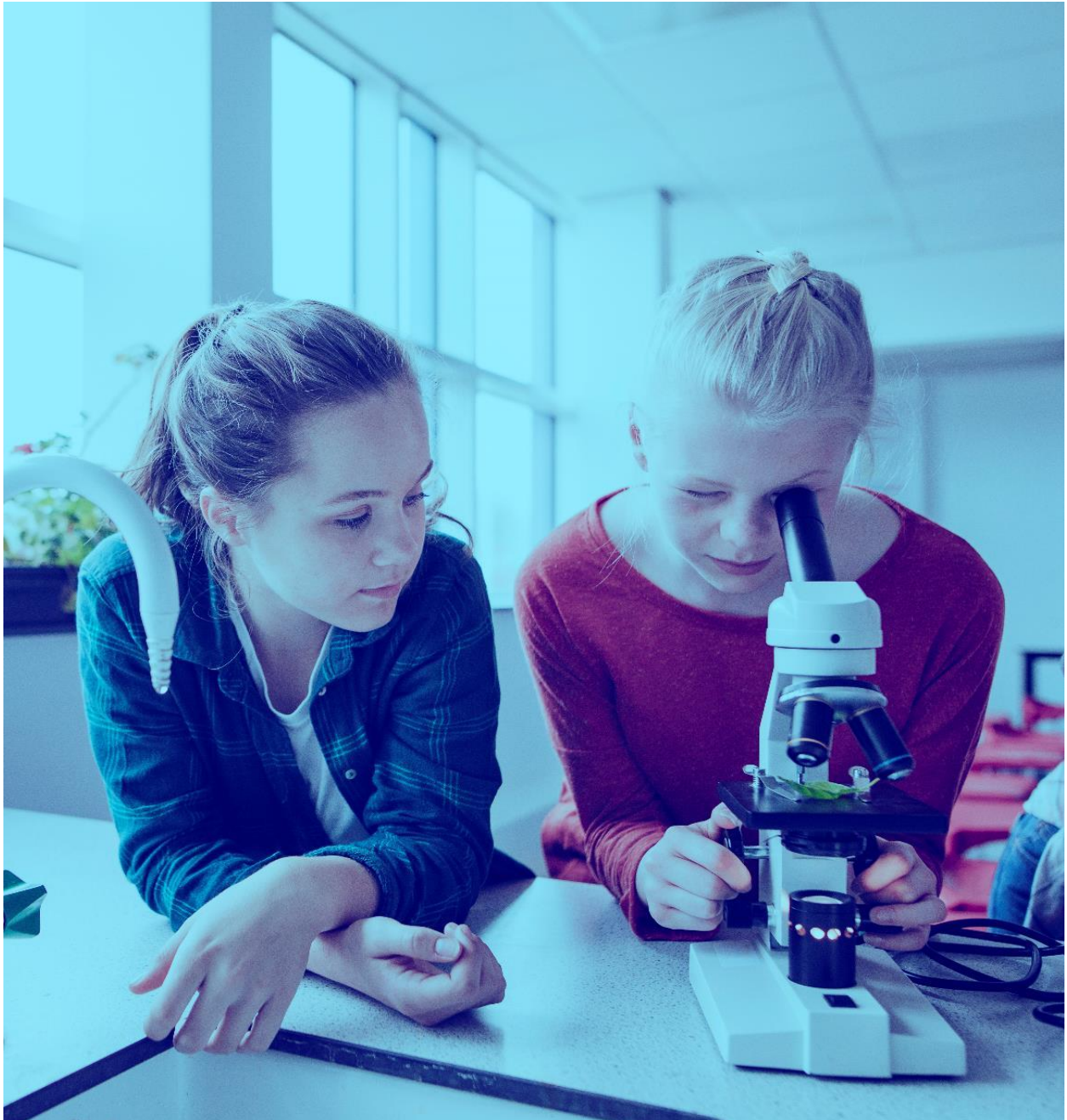
- Daylight: Indicates daylight conditions
- Dark with Street Lights on: Dark conditions but streetlight source was present
- Dark without Street Lights: Pitch dark conditions, only head lights to guide the way
- Dawn: Early Morning, day starts getting more light
- Dusk: late evening, light from the day reduces

Data Balancing and Cleaning

Data is unbalanced and cannot be directly used for analysis

In its original form, this data is not fit for analysis. For one, there are many columns that we will not use for this model. Also, most of the features are of type object when they should be numerical type.

We must use label encoding to covert the features to our desired data type.



Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo.

Heading 05 Lorem Ipsum

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit.

Heading 06 Sit Amet

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt.

