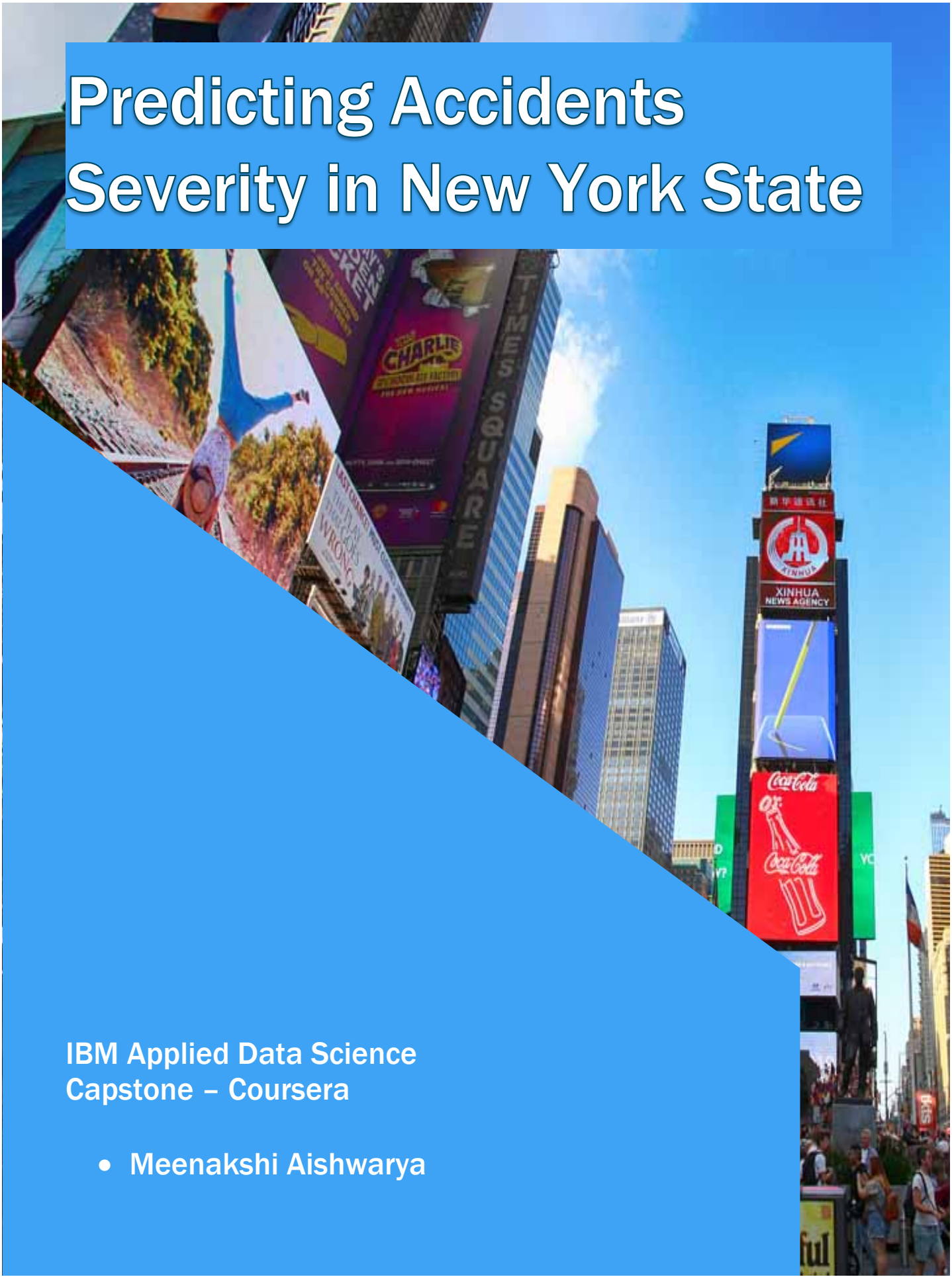
Predicting Accidents Severity in New York State



IBM Applied Data Science Capstone – Coursera

- Meenakshi Aishwarya

Introduction

In today's world the number of vehicles on road are incresing. Be it cars or two-wheelers, taxis, trucks, buses, etc., the number vehicles on the road is increasing by the day. While this shows a good scale of development, it also means that there are higher chances of road accidents. Road accidents can be fatal and very dangerous; thus we must do our best to avoid them.

In this project, the prediction of the severity of the road accidents is done using a Machine Language Model.

Business Problem

The objective of this capstone project is to

1. Analyse accidents severity in New York State
2. To predict the severity of an accident in New York State using machine learning techniques.

The problem question is: What could be the severity of the accident if I go for a road trip today?

Data

To solve the problem, we need the following data:

- List of accidents in New York State
- Date, time, duration and location of each accident
- Cause and severity of each accident

In order to meet these criteria, the source of dataset selected is given below:

https://www.kaggle.com/sobhanmoosavi/us-accidents

the above dataset is a countrywide traffic accident dataset, which covers 49 states of the United States with 49 attributes. It has data from February 2016 to June 2020. The data set is in the form of csv file. So the dataset is should be prepared by removing the unnecessary data. The records from New York State should be selected and the features or attributes of dataset should be identified. The following features or attributes are selected from the dataset:

#	Attribute	Description	Nullable
1	ID	This is a unique identifier of the accident record.	No
2	Source	Indicates source of the accident report	No
3	TMC	A traffic accident may have a Traffic Message Channel (TMC) code which provides more detailed description of the event.	Yes
4	Severity	Shows the severity of the accident, a number between 1 and 4, where 1 indicates the least impact on traffic (i.e., short delay as a result of the accident) and 4 indicates a significant impact on traffic (i.e., long delay).	No
5	Start_Lng	Shows longitude in GPS coordinate of the start point.	No
6	Start_Lat	Shows latitude in GPS coordinate of the start point.	No

7	Distance	The length of the road extent affected by the accident.	No
8	Side	Shows the relative side of the street (Right/Left) in address field.	Yes
9	City	Shows the city in address field.	Yes
10	County	Shows the county in address field.	Yes
11	State	Shows the state in address field.	Yes
12	Timezone	Shows timezone based on the location of the accident (eastern, central, etc.).	Yes
13	Temperature(F)	Shows the temperature (in Fahrenheit).	Yes
14	Humidity(%)	Shows the humidity (in percentage).	Yes
15	Pressure(in)	Shows the air pressure (in inches).	Yes
16	Visibility(mi)	Shows visibility (in miles).	Yes
17	Wind_Direction	Shows wind direction.	Yes
18	Weather_Condition	Shows the weather condition (rain, snow, thunderstorm, fog, etc.)	Yes
19	Amenity	A POI annotation which indicates presence of amenity in a nearby location.	No
20	Bump	A POI annotation which indicates presence of speed bump or hump in a nearby location.	No
21	Crossing	A POI annotation which indicates presence of crossing in a nearby location.	No
22	Give_Way	A POI annotation which indicates presence of give_way in a nearby location.	No
23	Junction	A POI annotation which indicates presence of junction in a nearby location.	No
24	No_Exit	A POI annotation which indicates presence of no_exit in a nearby location.	No
25	Railway	A POI annotation which indicates presence of railway in a nearby location.	No
26	Roundabout	A POI annotation which indicates presence of roundabout in a nearby location.	No
27	Station	A POI annotation which indicates presence of Station in a nearby location.	No
28	Stop	A POI annotation which indicates presence of stop in a nearby location.	No
29	Traffic_Calming	A POI annotation which indicates presence of traffic calming in a nearby location.	No
30	Traffic_Signal	A POI annotation which indicates presence of traffic signal in a nearby location.	No
31	Turning_Loop	A POI annotation which indicates presence of turning loop in a nearby location.	No
32	Sunrise_Sunset	Shows the period of day (i.e. day or night) based on sunrise/sunset.	Yes
33	Hour	Shows the hour at accident	No

34	Weekday	Shows weekday of accident	No
35	Time_Duration(min)	Shows duration of accident	No