MACHINE LEARNERS

Shivam Kumar 180717

Lokendar Kumar Meena 180383

Sachin Kumar Sahoo 19205274

Aman Deep Gupta 18103264

Dhiraj Narsingrao Chandawar 20103042

Dinesh Ahlawat 20103043

Accident Risk Analysis

There has been large scale development in infrastructure specially road connectivity in the past few decades. But the traffic on these roads has increased even at greater pace. Thus, the situation needs to be managed. Accident prone zones, traffic risks, congestion points need to be removed. This can be done by studying the traffic flow pattern, congestion points, traffic characteristics, road attributes etc. Now, with improved technique in data processing and understanding using modern day computational means like machine learning, artificial intelligence, etc., such challenges can be taken care of to a great extent with accuracy. This project carries a brief traffic analysis for predicting accident risk. The traffic data of US which would be used for analysis is available in public domain the link for which is "https://osu.app.box.com/v/us-accidents-june20" The data is classified in four domains:

1.Traffic events:

These can directly or indirectly lead to loss of life. Under this we would focus on:

a. Accident: This frequent kind of event where two or more vehicles collide.

b. Event: Different types of events and rallies (like religious, political, cultural, protests, etc.) are held on roads or highways which heavily affect movement of vehicles and the traffic flow.

c. Construction: In a developing country like India, we can observe road maintenance, highway building, etc. which blocks traffic.

d. Description: Short description of the accident.

e. Broken Vehicle: There are situations when disabled vehicles are running on the road thus causing inconvenience in movement of traffic.

f. Lane Blocked: There are situations in different weather conditions like snow, excess precipitation or underlaying sewer lines when the road is blocked by authorities.

g. Congestion: It refers to the situation (like office hours) when the traffic is slower than the speed limit.

2. Weather: This has a major effect on traffic, be it excess precipitation in the rainy season or fog in the winter season. Under this, the parameters that would be considered for analysis are: Temperature, Wind Chill, Humidity, Visibility, Wind Direction, Wind Speed, Precipitation.

3. Point of Interest: This term includes important places or represent a particular point of interest in an area. We would include the following features under this:

a. Amenity: It refers to specific places such as headquarters, restaurants, libraries, colleges, etc.

b. Bump: Refers to hump to decrease the speed.

c. Crossing: Refers to crossing across roads for cyclists, pedestrians etc.

d. Give-way: A road sign, which displays the priority of passing.

e. Junction: Refers to any highway ramp, exit, or entrance.

f. No-exit: Indicates there is no travel further.

g. Others are Railways, Roundabout, Station, Stops, Turning Loop, etc.

4. Location: There are certain roads, highways, and bypasses, and overbridges where chances of happening an accident have always been greater than others. So, location is an important aspect of it and we would have the following attributes to get the details of location:
Number, Street, Side (left/right), City, State, Zip-code, Country.

5. Time: This is also a useful category to analyse traffic characteristics. Here we include:

a. Sunrise/Sunset

b. Civil Twilight

c. Nautical Twilight

d. Astronomical Twilight

Overall, total 49 attributes would be analysed. Ten attributes from weather, Thirteen attributes from point of Interest (POI), eight from location, ten from traffic events and four from time.

Analysis Model: To better utilize heterogeneous sources of data and perform traffic accident prediction, we propose a deep neural network analysis. The brief overview is shown in the figure below.

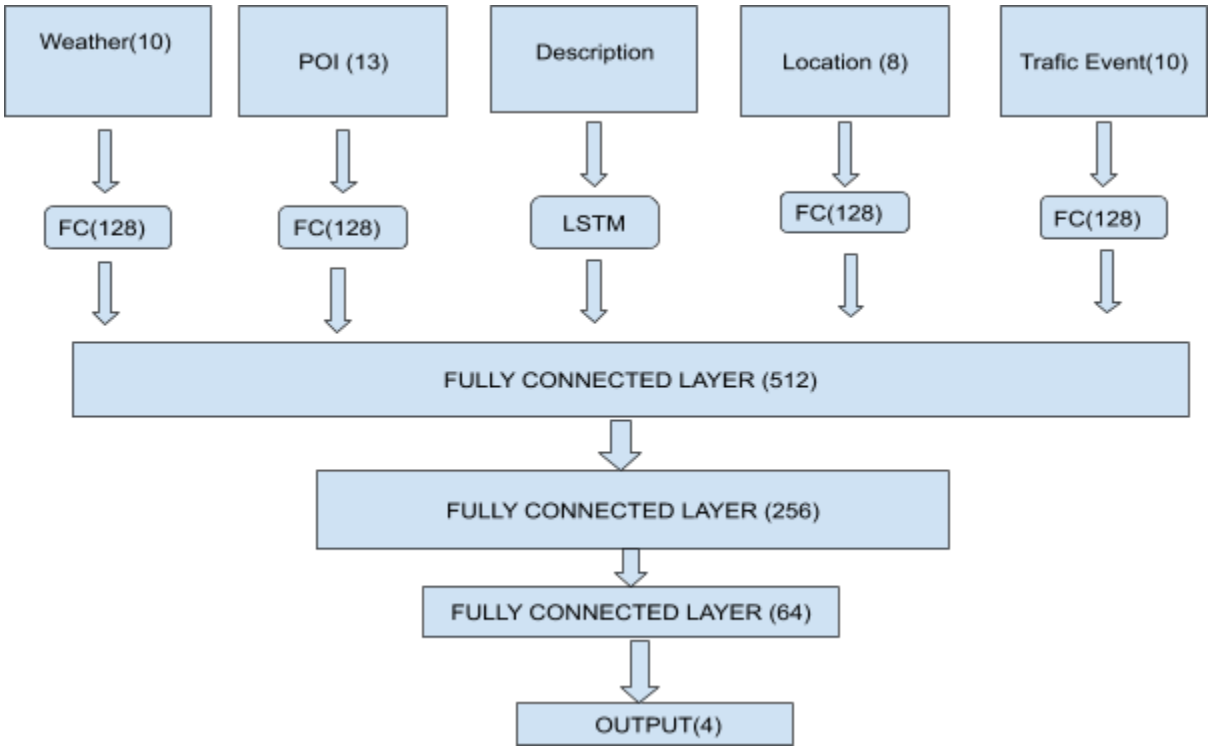

Figure: Flow Diagram showing model implementation on data

Fully-connected layer (FC): This is out-put of the above components to make the final prediction. We have four dense layers of size 512, 256, 64, and 4, respectively. Additionally, we batch normalization after the second and third layers. We will use ReLU as the activation function.