Project Proposal

**Project titl**e: US Accidents Trend Analysis and Classification

**Project Content:**

Today's public safety concerns demand extensive research and analysis of real-time traffic and accident data to forecast the likelihood of incidents. By alerting the public, accident risk prediction can dramatically increase public safety. In light of this, our team has decided to examine accident data from the United States to find any probable patterns in the occurrence of accidents. Individuals and organizations striving to reduce the risk of accidents may find this analysis to be a valuable source of information. It would be possible for these groups to make targeted improvements to those conditions if they were aware of the factors that raise the likelihood and severity of motor vehicle accidents. These initiatives not only have the potential to save lives but also money and enable the best possible use of resources. The US Accidents Trend Analysis and Classification project intends to find hidden and interesting trends within the US accidents dataset. The records include information about the date, time, location, severity, weather conditions, and other factors involved in the accident. The analysis starts with collecting data from the data source available on Kaggle and cleaning.

Further correlations, PCA, and feature extraction will be done on the dataset to retain the most desirable attributes of data. Next, a detailed Exploratory Data Analysis will be done to check for the trends in data at the County and State levels. In addition, by analyzing the data, our objective is to study possible patterns and trends in the occurrences of accidents, identify high-risk zones, and develop strategies to reduce the number of accidents and improve overall road safety. Based on the analysis models will be developed to predict and classify the accidents. We intend to implement a couple of Machine Learning algorithms that best fit the data to classify the accident records. Finally, the results will be evaluated over several metrics and presented with a detailed report.

**Description:**

This dataset on car accidents in the USA includes data from all 49 states. Several APIs that offer streaming traffic incident (or event) data were used to gather the accident data, which covered the period from February 2016 to December 2021. These APIs disseminate traffic information that has been gathered by several organizations, including the US and state departments of transportation, law enforcement organizations, traffic cameras, and traffic sensors embedded in the road networks. This dataset currently contains roughly 2.8 million accident records.