

ABSTRACT

Anupriya Goyal(anupriya,50287108)
Sravanthi Adibhatla(sadibhat,50288587)

Topic: Predicting Explosions and remote violence

Problem statement:

We are trying to perform data and predictive analysis on the recently adverse social unrest in the society - Explosions and Remote Violence. Unfortunately every day we hear news related to violence, so we want to perform predictive analysis and exploratory data analysis through Apache SparkSQL. This analysis will be useful in predicting the upcoming events that might happen in future in the countries which are more prone to remote violence. This will help the government to take necessary steps to avoid such events.

Reason for using Spark is: in comparison to MapReduce and other Apache Hadoop components, the Apache Spark API is very friendly to developers, hiding much of the complexity of a distributed processing engine behind simple method calls.

Dataset:

We are collection data related to our topic from ACLED(www.acleddata.com), where we will be collecting data from different countries and regions for the Explosions/ remote violence.