**Algorithms for Predicting Traffic Accidents Severity in France**

**Mogondo M. Edwin**

**B.E Chemical Engineer**

**IBM Data Science Program, Coursera E-Learning**

**Houston, TX**

**Email:** [emogondo@gmail.com](mailto:emogondo@gmail.com)

**ABSTRACT: Road accidents are among the most critical challenges that is facing humanity as they lead to many deaths, injury, economic losses and fatalities each year. There is an overwhelming need for accurate models in transportation industry and government. This investigation was conducted to help all emergency responders in France to predict the severity of an accident based on the information given and to know what kind of resources (money, labor and equipment) needed to be deployed. There are many machine learning techniques that can be applied. This project used Logistic Regression (LR) and Random Forest (RF) on 2005 through 2016 France traffic accident data. The findings of this investigation show that RF can be very promising in predicting accident severity. RF has shown predicted performance with 71.6% than that of LR with 65.1%.**

1. **INTRODUCTION**

**Road accidents leads to deaths, injury, and property damage resulting in huge loss at both social and economic levels. According to research conducted by World Health Organizations (WHO) in 2018, approximately 1.35 million die each year as a result of road accidents. Moreover, road traffic accidents injuries are the leading cause of death for children and young adults between age 5 and 29 years. Furthermore, road accidents cost most countries about 3% of their gross domestic product (GDP). The 2030 Agenda set by WHO is geared toward cutting the number of accidents by half.**

**As the demand of vehicle increases the world today the number of injury and death from traffic accidents is projected to go up. According to HAL archive of 2014, road accidents costed France about 35.7 and 50 billion Euros, and that number is expected to rise. Clearly there is a problem in transport systems that needs to be solved, thanks to Machine Learning Algorithms and Big Data tools that can readily be applied to collect this information that be useful in helping solve these challenges.**

**Supervised learning techniques and to be specific classification methods are among commonly used methodology to mine traffic data. These methods are trained on known datasets where factor that contribute to accidents are known. These machines then are used to make smart decisions to predict accidents severity in real time and which in turn can eradicate avoidable accidents.**

**The goal of this project is to achieve accuracy in predicting traffic accident severity by studying known factors. These include** road category, traffic regime, number of traffic lanes, surface condition, and infrastructure to predict “grav” which stands for severity in the known dataset. Additionally, this study aims at helping department of emergency to allocate resources efficiently and thus save time, money and proper allocation of equipment.