RoadIQ: A Data based approach to understand and model Road Accident Risk Prediction.

Anuj Singh, Student No: 5305926, Email: A.R.Singh@student.tudelft.nl

Style/type: Research Draft (Will try to actually implement and complete experiments as much as possible to resemble a Research Paper)

Abstract: About 1.35 million people die each year as a result of road traffic crashes and an additional 30-40 million are severely injured or disabled. Not only do road accidents cause human loss, but are also a major cause for economic damage in every developed as well as developing country. Road traffic safety officials can facilitate traffic safety through various means such as law enforcement, education and awareness about the topic and by controlling traffic at specific high risk and crowded areas. But forementioned statistics are testimony to the fact that the current methods of enforcing transportation safety aren't the most efficient and that there is still a huge scope for improvement. This paper aims at designing a system to predict the risk or probability of traffic accidents based on various spatio-temporal factors such as a detailed timestamp indicating the time and day of the week, coordinates of various regions in the geographic domain of research, clubbed with detailed road traffic information such as Annual Average Daily Traffic (AADT) for various vehicle types, population densities of regions, weather conditions, road types and various other informative dimensions. Machine learning algorithms that learn from historical data can alert users about the risk of an accident by fetching useful user data in real time and thus avoid the dangerous areas for that time. It is also useful in the context of better urban planning to create efficient systems of road maintenance and traffic management particularly with the upcoming advent of automated cars. This research also suggests how the satellite images of regions and roads scraped using Google Maps API can be used to enhance the predictive power of the structured data with the help of Deep Learning methods such as Convolutional Neural Networks (CNN). The data used for all the analysis and modelling has been obtained from the Department of Transport of Great Britain.

Objectives and Contributions: 1. Understand how the features in the data influence the accident of frequency. 2. Use Population density data of Great Britain from 2011 census to use with original tabular data for better insight. 3. Use Google Static Maps API for satellite images of location coordinates in Great Britain. 4. Build Learning based (Deep Learning/ Machine Learning) based model for mixed data types for optimal prediction.

Research questions: 1. How do the spatio-temporal, population density and traffic flow-based factors and variables influence road traffic accident frequency? 2. How to mathematically model and predict the probability of a road accident using historical data based on the discussed factors?

Structure of sections (ToC):

5. Results
6. Future Work

References:

[1] Abdel-Aty, M., Chen, C., Radwan, E., Brady, 1999a. Analysis of accident-involvement trends by drivers age in Florida. ITE Journal on the Web (Feb. 1999), pp. 69–74.

[2] Abdel-Aty, M., Chen, C., Radwan, E. 1999b. Using conditional probabilities to explore the driver age effect in accidents. ASCE Journal of Transportation Engineering 125(6).

[3] Agent, K., Deen, R., 1975. Relationship between roadway geometrics and accidents. Transportation Research Record 541, 1–11.

[4] Yannis, G., Dragomanovits, A., Laiou, A., La Torre, F., Domenichini, D., & Richter, T. et al. (2017). Road traffic accident prediction modelling: a literature review. Proceedings Of The Institution Of Civil Engineers - Transport, 170(TR5). Retrieved from https://trid.trb.org/view/1489250

[5] Chen, C., 1997. Statistical Analysis of the Effect of Demographic and Roadway Factors on Traffic Crash Involvement. M.S. thesis, Department of Civil Engineering, University of Central Florida

[6] Fridstrom, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., Thomsen, L., 1995. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. Accident Analysis and Prevention 27 (1), 1–20.