Data Driven Traffic Accidents Visualization and Root Cause Quantification with Explainable Artificial Intelligence for improved Road Safety and Urban Planning

Primary Researcher:	Fang Liu, Singapore University of Social Sciences
Co-researcher(s):	Zengyan Fan, Singapore University of Social Sciences
	Wei Zhang, Singapore Institute of Technology

1. Summary

Each year, over a million deaths are caused by traffic accidents [1]. With each accident, it will not only affect the individual, but on a larger scale, affect a country's economy and fatality rate [2]. In Singapore, the city has seen an increase in road traffic accidents over the past two years, with a 13.8% increase for accidents resulting in injuries and a 4% increase of accidents resulting in fatalities [3].

For each of these accidents, there may be multiple reasons or factors as to why they occurred. To combat against the negative effects of traffic accidents, numerous efforts have been put into improving them. In recent years, we see more studies looking into improving the impact of a traffic accident, such as its severity [4,5] through Machine Learning (ML).

Despite the efforts and research done to model traffic accident data and its severity, much has yet to be done to understand the models used, which are black boxes, where contributions to its predictions are still unclear. This information is crucial to determine and ensure that the model is reliable and trustworthy. By not knowing which features contribute to a prediction, there is little effort being done to truly understand the model and the data.

With this research, we concluded that factors related to weather, such as air pressure and humidity, are largely responsible for traffic accidents. The results are shown in multiple stages and granularity, with coarse granularity presented through an ML model's overall feature importance and fine granularity with explainable ML techniques such as Partial Dependence Plots (PDP) and Shapley Additive Explanations (SHAP). The explainable ML techniques allow us to better understand how the model generates its prediction through visualization.

2. Aim of Research

This project aims to train accurate ML models for traffic accident and investigate which factors contributes most to a traffic accident and how it is related to a traffic accident's severity, in hopes to reduce the severity of a traffic accident and possibly reduce the chances of an accident from occurring. This is achieved through extensive experiments and assessment of factors that could be a contributing variable to a traffic accident's severity. By being able to churn out probable factors and explanations to them, better resource management can be put in place to mediate or reduce the impact (severity) of a traffic accident.

3. Method of Research & Progression

Our system for traffic understanding consists of several main building blocks as shown in Figure 1. It shows the system architecture of the explainable ML-based traffic accident understanding with four stages. Explainable ML is the focus of the whole system and lies in stage 3. Before that, we need a dataset of traffic accident records in the first stage and an ML of high fidelity for traffic accident modelling in the second stage. The first three stages realise the design and development of some applications for intelligent transportation and smart cities, as shown in the last stage.

As a start, a large-scale and real-world dataset containing traffic accident records [6] in the US from 2016 to 2021 are used. This dataset includes near 3 million records, each with up to 47 features, e.g., the ones about when and where. Several data preparation techniques such as feature selection, data generalisation, handling missing and categorical values, were applied to prepare the dataset.

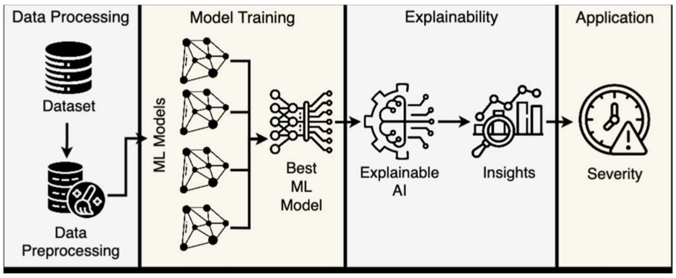

Figure 1 System Architecture

Then the dataset is trained with four different ML algorithms, namely, Logistic Regression (LR), Random Forest Classifier (RF), K-Nearest Neighbours (KNN) and Decision Tree (DT). These models will be compared through its accuracy, mean squared error (MSE) scores, and training time to determine the best model. These four popular models were chosen as we expect them to improve the generalizability of the research.

Furthermore, to gain deeper insight and understanding of the ML model built, explainable ML techniques such as PDP [7] and Shapley additive explanations (SHAP) [8] are used to allow us to understand the prediction and how it is derived – factors contributing to the prediction. By knowing these contributing factors, we can answer why a particular traffic accident occurred and potentially mediate a high traffic accident's severity from occurring.

Lastly, a web application dashboard is developed to complement and visualize the different components of this project. For example, the statistics of the ML model and the findings of the explainable ML. Moreover, a prediction page is also incorporated, making use of the trained ML model to predict a traffic accident severity according to a set of new user inputs. This provides a proof of concept for the usage of the ML model in enabling better resource management.

4. Results of Research

4.1. Model Training

For the training of each ML model and prediction of the target feature Severity, performance metrics such as accuracy, MSE and timings are measured and collated as shown in Table 1.

Table 1: Performance metrics recorded for predicting the target feature Severity.

ML Model	**LR**		**RF**		**KNN**		**DT**	
Metrics	Mean	Best	Mean	Best	Mean	Best	Mean	Best
Accuracy (%)	93.2	93.2	93.9	94.3	93.3	93.5	93.4	93.6
Error (%)	15.2	15.2	14.6	13.7	15.2	14.8	14.9	14.6
Timing (secs)	45.0	9.5	29.2	2.9	438.3	413.0	6.4	0.9

Since we are trying to provide better resource management upon a traffic accident, the factor of a model's training time would be secondary. The main performance metric that should be focused on is accuracy and MSE, as an accurate prediction is crucial in a life-or-death situation, where incorrect predictions could result in serious consequences. Hence, with the high accuracy and low MSE attained by the trained RF model, the model is chosen for further experimentation with explainable ML to achieve the project's objectives.

4.2. Feature Importance

With the trained RF model, a feature importance ranking is generated, as seen in Figure 2. The figure shows us which feature is significant in affecting the model's prediction. In this case, the feature pressure has the highest importance followed by humidity. Among the features shown, 6 of the features are Weather-related features. As we know, differing weather conditions can have varying impact to a traffic accident's severity level. A higher severity level can occur in bad weather conditions, measured through correlated features such as humidity, weather condition, wind speed etc.

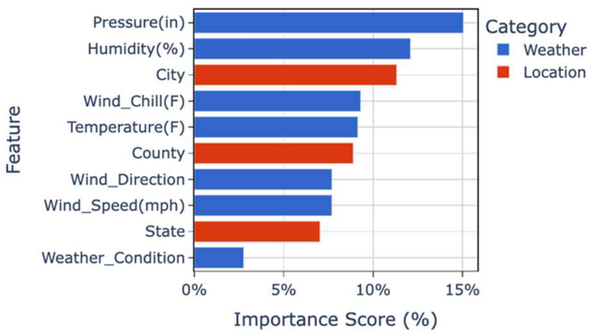

Figure 2: Feature Importance for RF model for predicting a traffic accident's Severity

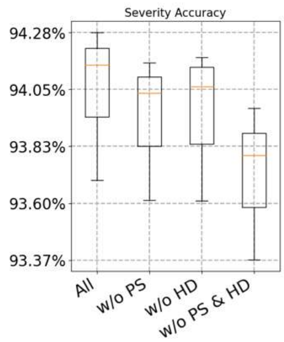

Figure 3: Prediction accuracy of RF model upon removal of high-ranking features.

To further validate the feature importance ranking, another experiment is conducted to measure the impact in prediction accuracy upon removal of the top 2 important features. The same experimental process and performance metrics mentioned during training of the ML models is also measured. Three RF models will be trained, once upon removal of the feature pressure, another upon removal of feature humidity, and lastly, upon removal of both features.

By conducting this experiment, we can visualize the degradation in model performance upon removal of features with high importance, shown in Figure 3. Looking at both Figure 2 and 3, a significant drop in prediction accuracy makes sense upon removal of the highest ranked feature, pressure. This is in comparison to the removal of feature humidity which had a small margin in drop of accuracy. Consequentially, by removing both features, the prediction accuracy suffers a greater impact compared to previous iterations of removing either features pressure or humidity. The drop in prediction accuracy validates the importance of the two features, which aligns with the results of the feature importance of the model.

4.3. Explainable ML

To further analyse the relationship between features in the dataset with the traffic accident's severity, PDP is used as a technique of explainable ML for a fine-granularity importance analysis of the two features. The feature importance is quantified by replacing a feature's records in the dataset with a specific value, with all other feature's values remain unchanged. The feature value selected is within the feature's inter-quartile range, excluding outliers. This value will then be used with the average modelled severity of all records in the dataset, thus, recording the corresponding importance of the feature at the selected value.

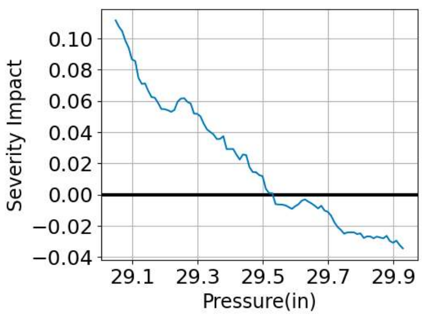

Figure 4: Severity impact for Pressure

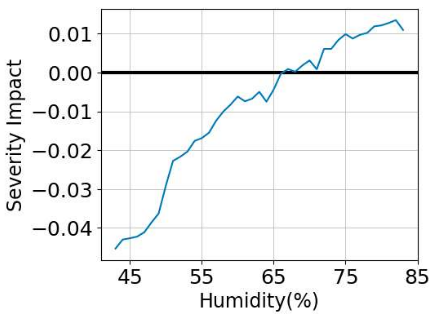

Figure 5: Severity impact for Humidity

In Figure 4, the plot visualizes the decrease in impact of a traffic accident's severity as the air pressure increases. This aligns with the research where with low air pressure, there's a higher chance of wind or rain [6]. Thus, when the predicted traffic accident's severity is high, which is most likely due to a less favourable weather condition – strong wind / rain.

On the other hand, Figure 5 visualizes the increase in traffic accident's severity as the humidity level rises. This trend aligns with the research where with heavy rain, the recorded humidity would increase due to the amount of water vapor in the air. Thus, high humidity levels would indicate an increase in traffic accident severity. This would occur in weather conditions where it's raining, resulting in poor visibility for drivers, or after raining, where the roads are slippery and may pose a risk to drivers.

Instead of just looking at feature importance for the entire dataset, it is also crucial to investigate specific accidents to gain precise understanding. Therefore, SHAP is used to quantify feature importance of each individual record in the dataset. For this, we randomly picked two records, where the corresponding SHAP results are shown in Figure 6 and 7.

In Figures 6 and 7, we can see that most of the features are from the weather category; Temperature, Wind_Chill, and Sunrise_Sunset in Figure 6, and Pressure, Humidity, Wind_Chill and Temperature in Figure 6. The figure tells us that the weather category is largely responsible for the output prediction of the two records. Not only that, but we can also observe that location feature such as State, City and County being present as one of the top 5 features. This could mean that the place of occurrence does play a part in the resulting severity.

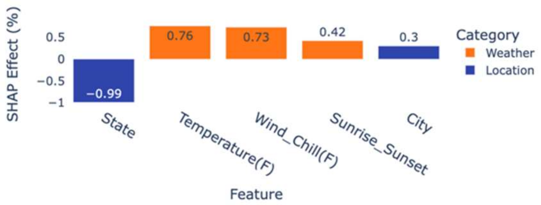

Figure 6: SHAP effect for a record with accident severity of 1

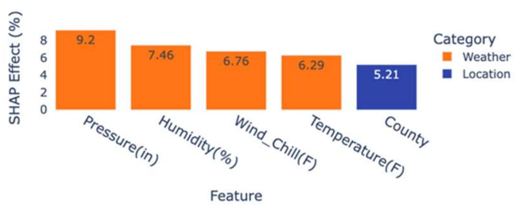

Figure 7: SHAP effect for a record with accident severity of 4

4.4. Visualizing Explainable ML

To visualise the results of the above experiments with user interactivity, a web application dashboard has been developed using Plotly's Dash, a Python library. The dashboard consists of a model explainer and prediction page. The RF model trained and explained using explainable ML will be presented in the model explainer page, whereas the use of the model is presented in the prediction page. A screenshot of the dashboard can be seen in Figure 8.

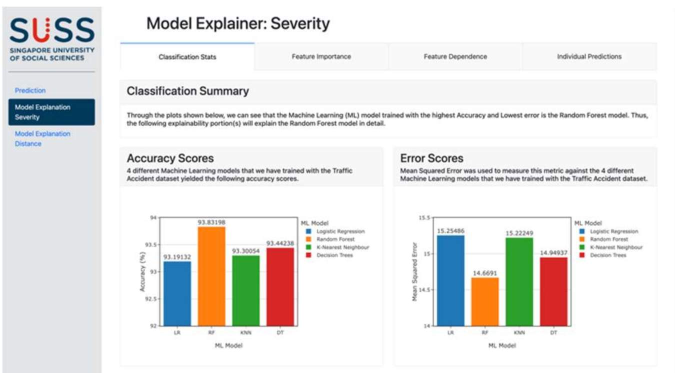

Figure 8: Model Explainer page for the target feature Severity.

A short demo of the web application dashboard can also be viewed via this link.

5. Future Areas to Take Note of, and Going Forward

For future works, we could investigate the relationship between the features in the dataset with another attribute of the traffic accident. Not only that, instead of using traditional ML algorithms, we could also make use of deep learning or neural networks. Performance metrics could also be measured and used as comparison to the RF model that was trained during this research.

Based on our findings supported by this project, we would like to investigate more possibility that how explainable AI facilitate industry solutions. Given a system functionalities and advantages well demonstrated, potential industry partners can deploy the system with confidence and trust, and discover key insights and explanations with user-friendly interfaces, towards safer transportation.

6. Means of Official Announcement of Research Results

Research paper submitted to IEEE Systems, Man, and Cybernetics (SMC) 2023
Title: It Is About Weather: Explainable Machine Learning for Traffic Accident Understanding

- https://ieeesmc2023.org
- Current Progress: Accepted

Demo paper submitted to Recommender Systems 2023
Title: Interpreting Traffic Accident Machine Learning Model through Visualisations

- https://recsys.acm.org/recsys23/call/#content-tab-1-5-tab
- Current Progress: Submitted and Awaiting Reviews

7. References

[1] Ministry of Road Transport & Highways, Government of India, “Causes of Road Accidents | Transport Department, Government of Jharkhand,” Jhtransport.gov.in, 2022. https://jhtransport.gov.in/causes-of-road-accidents.html

[2] “Road traffic injuries,” World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

[3] Annual Road Traffic Situation 2022. Feb. 2023. URL: https://www.police.gov.sg.

[4] F. Zong, H. Xu, and H. Zhang. “Prediction for traffic accident severity: comparing the Bayesian network and regression models”. In: Mathematical Problems in Engineering (2013).

[5] J. Lee, J. Chae, T. Yoon, and H. Yang. “Traffic accident severity analysis with rain-related factors using structural equation modeling–A case study of Seoul City”. In: Accident Analysis & Prevention 112 (2018), pp. 1–10.

[6] S. Moosavi, M. H. Samavatian, S. Parthasarathy, and R. Ramnath, “A countrywide traffic accident dataset,” arXiv.org, 12-Jun-2019. [Online]. Available: https://arxiv.org/abs/1906.05409#.

[7] J. H. Friedman. “Greedy function approximation: a gradient boosting machine”. In: Annals of statistics (2001)

[8] S. M. Lundberg and S.-I. Lee, ‘A Unified Approach to Interpreting Model Predictions’, in Advances in Neural Information Processing Systems, 2017, vol. 30.

[9] M. Kusriyanto and A. Putra. “Weather station design using IoT platform based on Arduino mega”. In: 2018 International Symposium on Electronics and Smart Devices (ISESD). IEEE. 2018, pp. 1–4