Road traffic accident severity prediction using causal inference and machine learning

Nishtha Srivastava
Sardar Vallabhbhai National Institute of Technology
Surat, Gujarat, India
d20co005@coed.svnit.ac.in

Bhavesh N. Gohil
Sardar Vallabhbhai National Institute of Technology
Surat, Gujarat, India
bng@coed.svnit.ac.in

Suprio Ray
University of New Brunswick
Fredericton, Canada
sray@unb.ca

Abstract

The global rise in road traffic accidents presents substantial challenges across economic, societal, and public health domains, leading to millions of injuries and fatalities annually. Current studies on modeling and analyzing traffic accident frequency largely treat the issue as a classification task, primarily utilizing learning-based or ensemble methods. However, these approaches frequently neglect the intricate relationships among the multifaceted factors—such as road complexity, environmental conditions, driver behavior, and contextual elements—that contribute to traffic accidents and hazardous scenarios. We propose an approach that employs causal inference and causal Machine Learning (ML) techniques to predict accident severity and identify key causal factors. We evaluate our proposed approach with two datasets, from Ethiopia and UK. Given the inherent imbalance in these datasets, the Synthetic Minority Oversampling Technique (SMOTE) is utilized to achieve balanced data representation. Uplift modeling and causal inference methods are employed for severity prediction. Individual Treatment Effect (ITE) and Average Treatment Effect (ATE) are used to make interpretations of the predictions. Our research contributes to understanding and mitigating the impact of road traffic accidents through advanced causal analysis techniques, offering actionable insights for policymakers, urban planners, and public health officials globally.

CCS Concepts

• **Computing methodologies** → **Feature selection**; **Machine learning**; **Artificial intelligence**; • **Applied computing** → **Transportation**.

ACM Reference Format:
Nishtha Srivastava, Bhavesh N. Gohil, and Suprio Ray. 2024. Road traffic accident severity prediction using causal inference and machine learning. In *8th International Conference on Data Science and Management of Data (12th ACM IKDD CODS and 30th COMAD) (CODS-COMAD Dec '24), December 18–21, 2024, Jodhpur, India.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3703323.3703749

CODS-COMAD Dec '24, December 18–21, 2024, Jodhpur, India

ACM ISBN 979-8-4007-1124-4/24/12
https://doi.org/10.1145/3703323.3703749

1 Introduction

Road Traffic Accidents (RTAs) are a major global concern, resulting in millions of deaths and injuries annually. The 2018 Global Status Report on Road Safety reports a staggering 1.35 million RTA-related deaths each year [21]. To address this, various studies utilize predictive analytics to derive insights from crash data, enabling identification of patterns and key predictors for accidents [11, 25, 32, 37]. This includes classification algorithms, which uncover complex relationships among variables like infrastructure, driver behavior, and environmental factors, offering clues to reduce fatalities. Crash reports, generated after accidents, are rich in data, containing textual descriptions, figures, and numerical information to reconstruct events. These reports provide a complex interplay of factors—including infrastructure, behavior, environmental conditions, and vehicle attributes—that contribute to accidents [33]. However, the diversity and interconnection of these factors make it challenging to analyze causal relationships effectively. Traditionally, researchers have framed accident prediction as a classification task, applying machine learning (ML) algorithms to summarize and predict crash outcomes using predefined features [16, 29, 35]. While this approach has yielded significant insights, converting detailed textual data into numerical values can oversimplify and ignore crucial relationships among variables. Moreover, the quality of crash predictions depends heavily on data reliability and feature selection, as key predictors are essential to improve model accuracy. Few studies have incorporated causal inference in crash severity prediction [6, 12]. This research seeks to fill this gap by applying causal ML methods to accident data from Ethiopia and the UK, estimating causal effects and identifying key severity-influencing factors [8, 10, 13, 41]. Previous works, such as Aldhari et al. [2], relied on SHAP analysis to interpret feature importance, which may not capture variations in feature significance across different conditions, such as lighting and road type. Additionally, Chakraborty et al. [6] used Granger causality to select features, but this approach could miss specific factors critical to severe crashes [30]. Causal inference integrated with ML [41] offers a powerful alternative by focusing on cause-and-effect rather than mere correlations. Through estimates like Average Treatment Effect (ATE) and Individual Treatment Effect (ITE), causal inference can provide actionable insights, showing how specific interventions could lower accident rates. This approach can enhance traffic safety by learning from complex, unstructured accident data, providing precise guidance for accident prevention. We propose an approach based on causal inference and causal ML for RTE severity prediction. We leverage the CausalML [18] Python package for advanced ML algorithms for uplift modeling and causal inference. Our proposed approach ensures that predictions are not only statistically significant but also causally meaningful, leading to more effective and targeted accident prevention strategies. The potential applications of the proposed model extend beyond prediction, offering a versatile tool for stakeholders to make informed decisions, allocate resources efficiently, and implement targeted interventions, ultimately enhancing road safety. The primary contributions of this paper are:

- Analyzing unstructured accident data for both UK and Ethiopia
- Utilizing causal inference and ML to predict road accident injury severity in Ethiopia and UK.
- Investigating ITE and ATE to identify factors contributing to road accidents.
- Re-training the respective machine learning models with the most significant features and comparing its performance to the original model.

By performing this analysis and using causal ML, our study seeks to provide deeper insights into the factors affecting road accident severity, thereby contributing to the global effort to improve road safety. This paper is structured as follows: Section 2 reviews literature on causal inference and ML. Section 3 presents a comprehensive review of road accident prediction. Subsequent sections cover problem definition, proposed approach, dataset and analysis, evaluation, discussion, and concluding with Section 9.

2 Background

This section provides a background in causal inference, focusing on how to estimate treatment effects, specifically the Average Treatment Effect (ATE) and Individual Treatment Effect (ITE), using experimental or observational data. It introduces key concepts such as treatment variables, covariates, and outcome variables, and describes how ITE and ATE are mathematically defined to quantify the causal impact of interventions.

2.1 Causal Inference

Causal ML provides tools to estimate the Average Treatment Effect (ATE)[23] , or Individual Treatment Effect (ITE) [1], from experimental or observational data. Specifically, given co-variates W, it quantifies the causal effect of intervention T on outcome Y for individuals characterized by observed features X , $X \subseteq W$, without imposing strict assumptions on the model's structure. Figure 1 shows an example of causal graph.

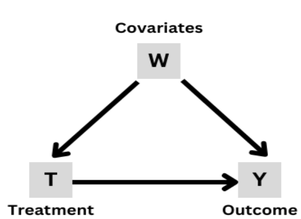

Figure 1: Example of a causal graph [38]

where,

- **Treatment effect (T):** Change in outcome if there is some change in the treatment variable.
- **Covariates (W):** Variables that are related to both treatment and outcome.
- **Outcome(Y) :** Output
- Independent variables, $X \subseteq W$

By treating all the features as a treatment variable, one by one, two types of treatment effects are calculated:

(1) Average Treatment Effect (ATE)
(2) Individual Treatment Effect (ITE).

2.2 Individual Treatment Effect (ITE)

ITE [36] for a particular feature shows the treatment effect for all the instances of the feature. The range for the treatment effect value is between 0 and 1. For no treatment, the value is 0. The Individual Treatment Effect (ITE) for a particular feature can be mathematically expressed as follows:

$$\text{ITE}(x_i) = P(Y|T = 1, X = x_i) - P(Y|T = 0, X = x_i)$$

where:

- $\text{ITE}(x_i)$ represents the Individual Treatment Effect for the i-th class of the feature x.
- $P(Y|T = 1, X = x_i)$ is the probability of the outcome Y given the treatment $T = 1$ and feature $X = x_i$.
- $P(Y|T = 0, X = x_i)$ is the probability of the outcome Y given no treatment $T = 0$ and feature $X = x_i$.

The treatment effect value ranges between 0 and 1, with a value of 0 indicating no treatment effect.

2.3 Average Treatment Effect (ATE)

ATE [36] for a particular feature is derived by calculating mean of ITE values for that feature. The feature with the highest ATE value, may have the highest causal effect for that given treatment variable. ATE for a particular feature is derived by calculating the mean of ITE values for that feature. It can be mathematically expressed as follows:

$$\text{ATE}(x) = \frac{1}{N} \sum_{i=1}^{N} \text{ITE}(x_i)$$

where:

- $\text{ATE}(x)$ represents the Average Treatment Effect for the feature x.
- N is the number of classes or instances of the feature x.
- $\text{ITE}(x_i)$ is the Individual Treatment Effect for the i-th class or instance of the feature x.

3 Literature Review

Recent advancements in crash severity prediction have focused on applying machine learning and big data analysis. Najada et al. [15] analyzed accident causes using Hong Kong's transportation data, finding Random Forest superior to Naïve Bayes and PART algorithms. Similarly, Richard and Ray [26] utilized big spatial data to study accidents in Fredericton and Laval, Canada, identifying key factors like weather, vehicle count, and accident type. Their

approach integrated big data and spatial analysis to model accident severity, though data inconsistencies between cities limited findings. Hamzah Al et al. [20] emphasized pre-processing to improve data reliability, noting common challenges like data imbalance in injury severity predictions, which are typically framed as binary classification problems [16, 29, 35] or sometimes as multiclass classifications [19, 22]. Additionally, text data in crash reports [17] can aid predictions but may result in information loss when converted numerically. Techniques such as SVM [39, 40], Logistic Regression [3, 5], and Bayesian networks [7] face bias from data imbalance. In Ethiopia, studies [9, 34] emphasize statistical methods and Decision Trees, highlighting a need for robust studies. Further, Hu et al. [12] proposed a Granger causality and Graph Convolutional Network method to improve crash risk prediction, enhancing interpretability through causal inference.
There is still enough room for improvement, as these studies do not explain causal relations between the features after the prediction. In a previous work [31] researchers applied causal inference for rail transit delay prediction. To our knowledge, this is the first work to utilize causal ML for accident severity prediction.

4 Problem definition

Causal ML (CML) approaches aim to help discover cause-and-effect relationships from observational data, which can provide deeper insights into the factors influencing accident severity. The problem of predicting accident severity can be mathematically framed as a supervised multi-class classification task. The function g is trained on a subset $\mathcal{D}_{\text{train}}$ and evaluated on subsets $\mathcal{D}_{\text{test}}$ and $\mathcal{D}_{\text{validate}}$. The objective is to find the optimal function g^* that minimizes a loss function $\mathcal{L}$ over the training data, expressed as $g^* = \arg\min_g \mathcal{L}(\mathcal{D}_{\text{train}}, g)$, with cross-entropy loss commonly used for such tasks. The prediction for a new instance $\mathbf{x}_{\text{new}}$ is given by $z_{\text{pred}} = g^*(\mathbf{x}_{\text{new}})$, and the model's performance is evaluated on $\mathcal{D}_{\text{test}}$ using metrics such as ITE and ATE.

5 Our proposed approach

Figure 2 illustrates the proposed model. In Step 1, raw traffic accident data is pre-processed for efficiency. Step 2 applies sampling techniques to address class imbalance in accident severity. In Step 3, the data is divided into components X (covariates), Y (outcome), and T (treatments). After pre-processing, the dataset is trained using an uplift tree classifier, and ITE and ATE scores are calculated to identify top features for model retraining. Figure 3 shows the causal graph for accident prediction, where *Y = Outcome (accident severity)*, *X = Covariates (weather, address, road, driver, vehicle)*, and *T = Treatment (spatial, temporal)*. Algorithm 1 outlines a procedure for processing and analyzing data using causal machine learning techniques. It starts by loading and pre-processing data, handling missing values, and converting categorical variables to binary. Next, it addresses class imbalance through resampling techniques like SMOTE. The data is then split into covariates, outcome, and treatment variables for model training using an uplift tree classifier. Feature analysis involves calculating individual and average treatment effects (ITE and ATE) to identify significant features. The model is then retrained using these top features, and its accuracy is evaluated. This structured approach ensures robust and interpretable causal insights. Detailed steps are explained in Sections 5.1, 5.2, 5.3 , 5.4, 5.5, and 5.6.

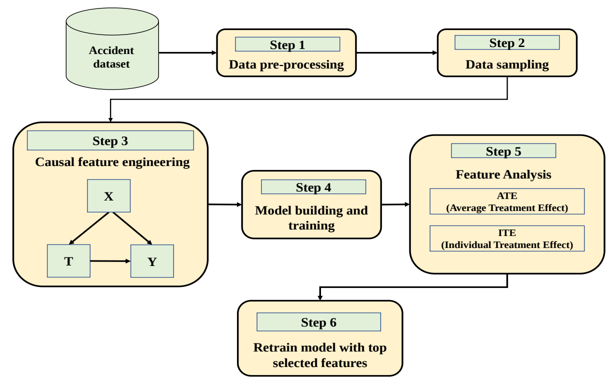

Figure 2: Proposed model

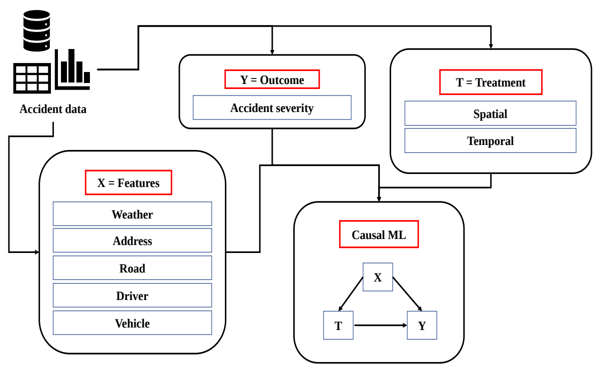

Figure 3: Causal graph for accident prediction

5.1 Data pre-processing

The first step is to clean and organize the raw accident dataset. This process may include handling missing values, removing duplicates, normalizing features, and encoding categorical data, to ensure the dataset is ready for analysis.

5.2 Data sampling

After pre-processing, the dataset is sampled to select representative data points. This step can involve splitting the data into training, validation, and testing sets, or balancing the dataset if it has class imbalances. SMOTE was used to balance the data by generating synthetic samples, creating interpolations of minority class instances to improve representation and support effective model training.

5.3 Causal feature engineering

This step focuses on identifying causally relevant features. It involves creating variables X, T, and Y where:

- X: Variables that are related to both treatment and outcome variables.

Algorithm 1 Prediction of traffic accident severity using causal ML

Data Pre-processing
Load data:
$D \leftarrow \text{load\_data}(\text{"}\textit{accident\_records.csv}\text{"})$
Handle missing values and convert categorical to binary:
$D' = \text{one\_hot\_encode}(D)$
Data Sampling
Address class imbalance:
$r = \frac{n_{\text{minor}}}{n_{\text{total}}}$
Apply SMOTE:
$D_{\text{balanced}} = \text{SMOTE}(D')$
Model Training
Train uplift model:
$\hat{f}(X, T) \leftarrow \text{UpliftTreeClassifier}(X, Y, T)$
Feature Analysis
Calculate ITE and ATE. Select top features:
$F_{\text{top}} = \{f \mid \text{ATE}(f) > \tau\}$
Model Retraining
Retrain model with top features:
$\hat{f}_{\text{retrained}} \leftarrow \text{UpliftTreeClassifier}(X_{F_{\text{top}}}, Y, T)$
Evaluate model accuracy.

- T: Treatment variable.
- Y: Outcome variable.

Causal relationships are established among these features to understand the impact of interventions or treatments on outcomes.

5.4 Model building and training

In the next step, an uplift tree classifier [28] is applied to train the dataset. Specifically designed for causal modeling, this classifier maximizes outcome differences between treated and untreated groups, enabling a clearer analysis of causal effects. By comparing projected outcomes for individuals who receive the intervention with those who do not, this approach enhances understanding of the treatment's impact. The `Uplift Tree Classifier`, available in the *CausalML* library, estimates Individual Treatment Effect (ITE) and Average Treatment Effect (ATE) at the subgroup level, allowing for a more precise assessment of feature contributions to treatment outcomes and improving causal inference accuracy. Key concepts of uplift tree classifier are mentioned next:

(1) **Treatment and Control Groups**: The dataset is divided into two groups—one receiving the treatment and the other as a control group. The classifier predicts the differential response or "uplift" that the treatment induces on the outcome.
(2) **Splitting Criteria**: Uplift trees use specific splitting criteria to maximize the difference in responses between the treatment and control groups rather than the standard decision tree criteria (like Gini or entropy). There are various ways to measure this difference:
   - **DeltaDeltaP ($\Delta\Delta P$)**: This criterion maximizes the difference between the probabilities of a positive outcome in the treatment and control groups across branches of a split [27].
   - **Divergence Measures**: Alternatives like Kullback-Leibler (KL) divergence [14] and Euclidean distance are also used to enhance splits where treatment and control outcomes diverge significantly.
(3) **Tree Construction**: Each split attempts to separate the population into branches where the treatment effect is maximized in one direction (e.g., positive uplift) and minimized in the other (e.g., no uplift or negative). Nodes in the tree are created based on criteria that account for the difference in response distributions between treatment and control branches. This enables the identification of population subgroups that exhibit varying degrees of sensitivity to the treatment.
(4) **Evaluation of Uplift**: Uplift is typically evaluated through uplift curves and metrics such as the *Area Under the Uplift Curve (AUUC)* [27]. These metrics assess how effectively the model segments the population by incremental response.

5.5 Feature analysis

The purpose of feature analysis is to quantify how each feature, when varied, impacts accident severity. This is particularly useful in causal models, as it allows us to go beyond correlation to understand causation. For example, factors like road surface conditions, driver's age, and weather could have distinct causal impacts on accident severity, which is captured through their ATE or ITE values. To calculate ITE, for each feature in the dataset, we treat that feature as the treatment variable (T) and designate the remaining features as covariates (W). This ensures that we examine the treatment effect of each feature individually while accounting for the effects of all other variables.

5.6 Retraining model

Model retraining is typically done to improve model performance and ensure it remains relevant with changing data patterns, especially when new data influences predictive accuracy. Retraining occurs after identifying top features that impact accident outcomes. Here, significant features are selected based on metrics like Average Treatment Effect (ATE) and Individual Treatment Effect (ITE) scores. Once identified, these features guide the model retraining process, optimizing it with the most relevant inputs and enhancing accuracy. Retraining frequency can vary, but it usually depends on the rate at which new data is generated or existing patterns change. Models could be retrained periodically, such as quarterly or semi-annually, to maintain optimal performance.

6 Dataset and analysis

In this section, we describe the datasets utilized in our study and outline the feature analysis methods used to analyze the data.

6.1 Dataset

For this study, we use 2 accident datasets:

- Ethiopia
- UK

Details of the dataset are provided in the next section.

6.1.1 Ethiopia. Addis Ababa, the capital of Ethiopia, hosts many continental and international organizations, including the African Union, making it a key diplomatic city. However, like many African cities, it struggles with low motorization rates, as noted in a 2023 study by Ambo et al. [4]. Ethiopia has approximately 1,138,365 registered vehicles, with about 70% in Addis Ababa. This research analyzes a dataset from the Sub-city Police Departments (2017–2020), featuring 32 attributes and 12,316 instances. We used 70% (8,621 instances) for training and 30% (3,695 instances) for testing. Details of the dataset features are listed in Table 1. In addition to the features in Table 1, the remaining features pertain to the address.

6.1.2 UK. The UK government data service gathers information on traffic accidents through detailed police records, which are recorded in a format called STATS 19 [24]. This comprehensive dataset, covering the years 2010 to 2012, includes 32 different features. Accident severity is categorized into three levels: fatal (value 1), serious (value 2), and slight (value 3). In 2010, there were 154,414 reported casualties from traffic accidents in the UK, followed by 151,474 in 2011, and 1,45571 in 2012. Table 1 lists 22 of these features. Besides those listed, the remaining features are related to the address information. Specifically, 70% of the dataset was used for training, allowing the model to learn from a majority of the data, while 30% was reserved for testing to assess the model's predictive.

Accident records for the Ethiopia and UK datasets were first encoded in Excel, then converted to CSV, and preprocessed. Initial data processing involved removing outliers and handling missing values. In the Ethiopia dataset, missing values were labeled as "Unknown," and categorical gaps were filled. The data was further transformed into a binary format (0,1) using Pandas dummy variables, optimizing it for efficient model training. Both datasets displayed a marked class imbalance. In Ethiopia, 84.56% of cases were minor, 14.15% serious, and 1.28% fatal, while 80.31% of UK records indicated major severity. To balance the data, SMOTE was applied to generate synthetic samples, interpolating minority samples to enhance representation and support effective model training. For causal machine learning, data was organized into X (covariates like weather, road, and driver details), Y (accident severity), and T (treatment variables across spatial and temporal factors). This structured framework supported using an uplift tree classifier, revealing key factors influencing accident severity. After the data pre-processing stage, the accident dataset is trained using uplift tree classifier [28].

6.2 Feature analysis

Feature analysis is divided into 2 sections:

(1) Feature analysis for Ethiopia
(2) Feature analysis for UK

6.2.1 Feature analysis for Ethiopia. In this study, feature analysis is conducted by employing ITE and ATE, leveraging the causal inference capabilities of the Python library *UpliftTreeClassifier* [27, 42]. As shown in Figure 4, when the road surface condition is "*snow*" or "*flood over 3 cm deep*", then the ITE value is maximum (with value 1) as discussed in Table.4. Similarly, when the light condition is "*darkness -lights unlit*" then the ITE value is maximum as shown in Figure 5. Same trend can be observed in Figure 6, when the features value of driving experience is "*unknown*" or there is "*no license*", then the ITE value is maximum i.e. 1. As shown in Figure 7, when day of the week is Monday, then the ITE value is maximum.

6.2.2 Feature analysis for UK. Figure 8 shows ITE bar plots for the number of vehicles. The number of vehicles, ranging from 1 to 6, exhibits a narrow and more consistent ITE range, suggesting

Table 1: Features of Ethiopia and UK accident datasets

Feature Category	Feature Name	Dataset
Time Features	Time of the accident	Ethiopia
	Day of the week	Ethiopia
	Date of accident	UK
	Year of occurrence of the accident	UK
Spatial Features	Longitude	UK
	Latitude	UK
	Urban or rural area	UK
	Junction type like traffic signal	Ethiopia
	Junction detail	UK
Road Features	Road type	UK
	Category of the lane	Ethiopia
	Road surface type	Ethiopia
	Road surface condition	UK
	Speed limit	UK
	Light condition like daylight, etc.	Ethiopia, UK
	Weather condition like snowing, etc.	Ethiopia, UK
Driver Features	Age of the driver	Ethiopia
	Gender of the driver	Ethiopia
	Educational level of the driver	Ethiopia
	Vehicle driver relation	Ethiopia
	Driving experience of the driver	Ethiopia
Accident Features	Accident index	UK
	Number of Vehicles	UK
	Number of Casualties	UK
	Type of collision	Ethiopia
	Vehicle movement like going straight, etc.	Ethiopia
	Cause of accident like drunk driving, etc.	Ethiopia
	Accident severity like fatal injury, etc.	Ethiopia, UK

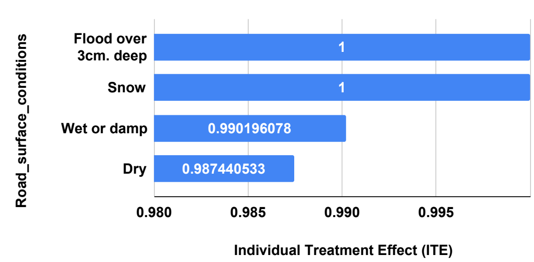

Figure 4: ITE plot for road surface conditions (Ethiopia)

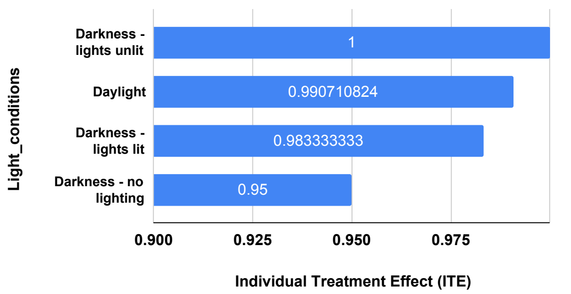

Figure 5: ITE plot for light conditions (Ethiopia)

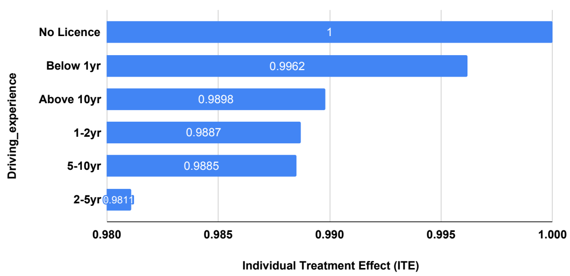

Figure 6: ITE plot for driving experience (Ethiopia)

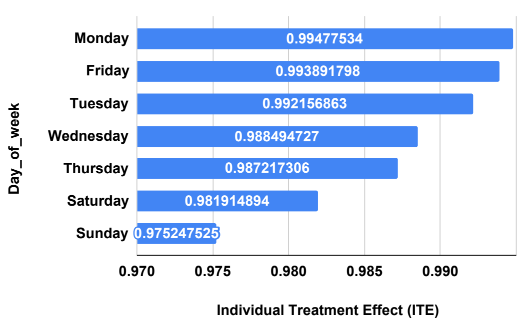

Figure 7: ITE plot for day of week (Ethiopia)

a predictable influence. These differences underscore the need for targeted interventions and tailored models to address the unique influence of feature effectively.

The results of this analysis, including the detailed ITE and ATE scores for the features under investigation, are comprehensively presented in Section 7.2.

6.3 Feature rankings

The feature rankings for Ethiopia and UK dataset, as shown in Table 2, obtained using ATE and ITE scores, are applied for model retraining. Post-retraining results are discussed in Section 7.2 for both datasets.

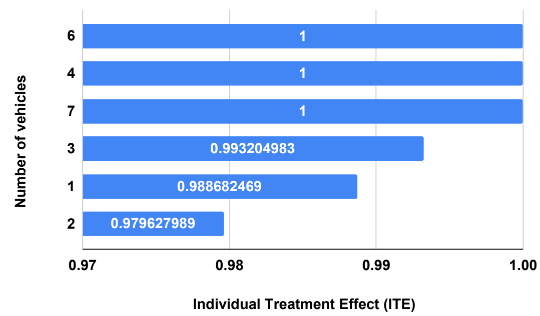

Figure 8: ITE plot number of vehicles (UK)

Table 2: Top selected features for Ethiopia and UK datasets

Sr. No.	Ethiopia	UK
1	Age of the driver	Number of vehicles
2	Educational level	Spatial attributes like Longitude, Latitude
3	Driving experience	Temporal attributes like day, month, and year
4	Road features	Weather features
5	Weather features	Light condition
6	Number of Vehicles	Road features like Road type, Road surface condition, and Junction details

7 Evaluation

In this section, we present the evaluation of our proposed approach, assessing its performance through various metrics and comparing it to existing methods to determine its effectiveness and suitability for the given problem.

7.1 Experimental setup

We used accident data from Ethiopia and the UK to evaluate the model's performance, conducting experiments in Python 3.12.3 on a server with a 3.31 GHz Intel Xeon CPU and 16 GB RAM. For causal ML implementation, we utilized the CausalML library, which provides uplift modeling and causal inference techniques.

7.2 Results based on ATE

The ATE values for features in the UK and Ethiopia accident dataset, shown in Table 3, reveal varied impacts on accident severity. The comparison of ATE values between the UK and Ethiopia datasets reveals significant differences in influential factors. The Table 3 presents the ATE values for various features related to accident datasets in the UK and Ethiopia. In the UK, the most significant feature is the "Number of Vehicles," which has an ATE of 1, indicating a strong influence on accident outcomes. Following closely are "Latitude" (ATE=0.99491) and "Longitude" (ATE=0.98811), suggesting that geographical factors are crucial in understanding accident patterns. Other notable features include "Date of accident" (ATE=

Table 3: Average Treatment Effect for Ethiopia and UK datasets

Feature Name (UK)	ATE	Feature Name (Ethiopia)	ATE
Number of Vehicles	1	Area accident occurred	0.914981
Latitude	0.994912	Weather conditions	0.897513
Longitude	0.988112	Road alignment	0.894179
Date of accident	0.974912	Types of Junction	0.884675
Time of accident	0.972911	Driving experience	0.868067
Day of week	0.964122	Education level	0.867068
Year of occurrence of the accident	0.961121	Lanes or Medians	0.866712
Light condition	0.897513	Day of week	0.864212
Weather condition	0.897513	Number of vehicles	0.851645
Road surface condition	0.884675	Road surface type	0.824984
Junction detail	0.872681	Age band of driver	0.823707
Speed limit	0.867068	Road surface condition	0.795527
Police Force	0.795527	Light conditions	0.784809
Urban or rural area	0.784809	Sex of driver	0.745574

0.97491) and "Time of accident" (ATE= 0.97291), highlighting the importance of temporal variables. In Ethiopia, the leading feature is "Area accident occurred" with an ATE of 0.914981, indicating its substantial impact on accidents. Other critical features include "Weather conditions" (ATE = 0.897513) and "Road alignment" (ATE= 0.894179), which reflect the importance of environmental and road conditions. The dataset also emphasizes "Types of Junction" (ATE= 0.884675) and "Driving experience" (ATE= 0.868067), pointing to the relevance of driver and road characteristics.

Overall, the ATE values reveal that both geographical and temporal factors are vital in the UK, while environmental conditions and road features play a more significant role in Ethiopia's accident dynamics.

7.3 Results based on ITE

Table 4 summarizes the treatment effects of various features related to road accidents in Ethiopia, highlighting key factors influencing accident occurrences and severity. The Table 4 summarizes key features influencing road accidents in Ethiopia, focusing on their treatment effects as indicated by ITE scores. The top five features include the age band of drivers, with those aged 31-50 having the highest ITE value of 0.9888, indicating a strong correlation with accidents. Accident locations such as hospitals and schools score 1.0, highlighting high-risk zones. Mondays and Fridays exhibit the highest accident occurrences with ITE scores of 0.9948 and 0.9939, respectively. Additionally, unlicensed drivers and illiterate individuals both have an ITE score of 1.0, underscoring the critical link between education and road safety. ITE is vital for quantifying the impact of these factors and guiding targeted interventions to reduce accidents.

7.4 Comparison with baselines

Table 5 provides a comparison of accuracy with and without feature selection using causal ML post-application of uplift tree classifier for UK and Ethiopia dataset. The analysis uses XGBoost, a powerful gradient boosting framework, as the baseline model for its high performance and efficiency. In the UK dataset, the accuracy without selecting features was 92.40%, which improved to 94.11% with selected features, resulting in an increase of 1.71%. The percentage improvement in accuracy is approximately 1.85%. This indicates that while feature selection enhanced the model's performance, the increase was modest. Conversely, the Ethiopia dataset showed a more significant improvement. Its accuracy without feature selection stood at 91.20%, rising to 95.17% with selected features, marking an increase of 3.97%. The percentage improvement here is about 4.31%, demonstrating that the selected features played a crucial role in enhancing predictive capability. Overall, both datasets benefited from feature selection, but the Ethiopian dataset exhibited a more pronounced improvement

8 Discussion

This section highlights key findings on factors influencing road traffic accident severity, vehicle involvement, and temporal/spatial influences. We discuss potential biases in UK and Ethiopia datasets and explore the scalability and real-time applications of our models to inform future road safety interventions.

8.1 Effect of number of vehicles

The "Number of Vehicles" feature is critical in the UK dataset, with an Average Treatment Effect (ATE) of 1. This indicates that multiple vehicles significantly influence accident outcomes. As the number of vehicles involved increases, both the severity and complexity of incidents rise, resulting in greater casualties and damage. Multiple vehicle collisions often create a cascade effect, where the initial impact triggers a series of collisions, complicating rescue efforts and overwhelming emergency services. Such incidents incur higher medical costs and longer recovery times, leading to broader societal impacts. Moreover, these accidents frequently contribute to traffic congestion and delays, affecting areas beyond the crash site. The involvement of several vehicles often correlates with higher speeds, adverse weather, and risky driving behaviors, further complicating the situation. Understanding these dynamics informs policy decisions, emphasizing the need for effective traffic management and safety measures to reduce vehicle density and enhance road safety across the UK.

Table 4: Summary of Ethiopia accident and treatment effects

Sr. No	Name of Feature	ITE Key	ITE Value
1	Age_band_of_driver	31-50	0.988797932
		Under 18	0.987679671
		18-30	0.986486486
		Over 51	0.983758701
2	Area_accident_occured	Hospital areas	1
		School areas	1
		Residential areas	0.992805755
		Office areas	0.992211838
		Recreational areas	0.991967871
		Industrial areas	0.991701245
		Market areas	0.945945946
3	Day_of_week	Monday	0.99477534
		Friday	0.993891798
		Tuesday	0.992156863
		Wednesday	0.988494727
		Thursday	0.987217306
		Saturday	0.981914894
		Sunday	0.975247525
4	Driving_experience	No Licence	1
		Below 1yr	0.996296296
		Above 10yr	0.989840348
		1-2yr	0.988711195
		5-10yr	0.988568588
		2-5yr	0.981120201
5	Educational_level	Illiterate	1
		Writing & reading	0.990384615
		Junior high school	0.988840263
		Elementary school	0.987538941
		Above high school	0.985981308
		High school	0.983799705
6	Light_conditions	Darkness - lights unlit	1
		Daylight	0.990710824
		Darkness - lights lit	0.983333333
		Darkness - no lighting	0.95
7	Road_allignment	Escarpments	1
		Steep grade up	1
		Tangent road	1
8	Road_surface_conditions	Flood over 3cm. deep	1
		Snow	1
		Wet or damp	0.990196078
		Dry	0.987440533
9	Road_surface_type	Gravel roads	1
		Asphalt roads	0.987938766
		Earth roads	0.984693878
10	Number_of_vehicles	6	1
		4	1
		7	1
		3	0.993204983
		1	0.988682469
		2	0.979627989

Table 5: Comparison of accuracy with and without feature selection using causal ML post-application of uplift tree classifier for UK and Ethiopia dataset

Dataset	Accuracy without selecting features (%)	Accuracy with the selected features (%)
UK	92.40	94.11
Ethiopia	91.20	95.17

8.2 Effect of temporal and spatial treatments

In our study, we analyze spatial and temporal treatments to assess their causal effects on road traffic accident severity. Temporal variables include the time of the accident (ATE = 0.9729), day of the week (ATE = 0.9641 for the UK, 0.8642 for Ethiopia), and accident date (ATE = 0.9749). These factors help us understand how timing influences severity, especially during peak hours. For spatial treatments, we examine longitude and latitude (ATE = 0.9881), urban versus rural areas (ATE = 0.7848), and road alignment (ATE = 0.8942). Our causal model isolates these effects, confirming the significant influence of spatial and temporal factors on accident outcomes, guiding future road safety interventions.

8.3 Potential bias

Using datasets from both Ethiopia and the UK strengthens our study, though we acknowledge certain limitations and biases. These datasets reflect distinct traffic conditions and socio-economic contexts, with Ethiopia showing low motorization and less developed roads, while the UK represents a highly motorized, structured traffic environment. Temporal differences add further complexity: the UK data spans 2010-2012, while the Ethiopian data covers 2017-2020, meaning advancements in vehicle technology and traffic laws over time may affect the comparability of results.

8.4 Scalability

This study emphasizes using causal inference and machine learning to predict road traffic accident severity, which is especially useful for dynamic traffic management and reducing accident impact. Applying these models in real time could help authorities identify and prioritize high-risk zones, enabling more effective decision-making. Combined with live traffic data, these models could prompt timely actions, like rerouting traffic, adjusting speed limits, or dispatching emergency services more efficiently.

9 Conclusion

In this paper, we present a comprehensive causal machine learning approach for traffic accident analysis that integrates data preprocessing, sampling, causal inference, and feature analysis. The process starts with data preparation, followed by SMOTE to balance classes, enhancing model performance. Causal analysis divides data into covariates, outcomes, and treatments, allowing a focused study of factors influencing accident severity. An uplift tree classifier then assesses how various features affect outcomes, with Individual and Average Treatment Effects (ITE and ATE) highlighting factors like driver demographics and road conditions. Our future work will focus on model scalability and adaptation to regional variations, as well as improving data quality for broader applicability.

References

[1] Jason Abrevaya, Yu-Chin Hsu, and Robert P Lieli. 2015. Estimating conditional average treatment effects. *Journal of Business & Economic Statistics* 33, 4 (2015), 485–505.

[2] Ibrahim Aldhari, Meshal Almoshaogeh, Arshad Jamal, Fawaz Alharbi, Majed Alinizzi, and Husnain Haider. 2022. Severity prediction of highway crashes in Saudi Arabia using machine learning techniques. *Applied Sciences* 13, 1 (2022), 233.

[3] Rabia Emhamed AlMamlook, Keneth Morgan Kwayu, Maha Reda Alkasisbeh, and Abdulbaset Ali Frefer. 2019. Comparison of machine learning algorithms

for predicting traffic accident severity. In *2019 IEEE Jordan international joint conference on electrical engineering and information technology (JEEIT)*. IEEE, 272–276.
[4] Tefera Bahiru Ambo, Jian Ma, Chuanyun Fu, and Eskindir Ayele Atumo. 2023. Investigating road conditions of crash blackspots in Addis Ababa, Ethiopia: a random parameters negative binomial model. *International Journal of Crashworthiness* (2023), 1–12.
[5] Stefan Candefjord, Azam Sheikh Muhammad, Pramod Bangalore, and Ruben Buendia. 2021. On Scene Injury Severity Prediction (OSISP) machine learning algorithms for motor vehicle crash occupants in US. *Journal of Transport & Health* 22 (2021), 101124.
[6] Meghna Chakraborty, Timothy J Gates, and Subhrajit Sinha. 2023. Causal analysis and classification of traffic crash injury severity using machine learning algorithms. *Data science for transportation* 5, 2 (2023), 12.
[7] Cong Chen, Guohui Zhang, Rafiqul Tarefder, Jianming Ma, Heng Wei, and Hongzhi Guan. 2015. A multinomial logit model-Bayesian network hybrid approach for driver injury severity analyses in rear-end crashes. *Accident Analysis & Prevention* 80 (2015), 76–88.
[8] Victor Chernozhukov, Christian Hansen, Nathan Kallus, Martin Spindler, and Vasilis Syrgkanis. 2024. Applied causal inference powered by ML and AI. *arXiv preprint arXiv:2403.02467* (2024).
[9] Zewditu Abdissa Denu, Mensur Yassin Osman, Telake Azale Bisetegn, Gashaw Andargie Biks, and Kassahun Alemu Gelaye. 2021. Prevalence and risk factors for road traffic injuries and mortalities in Ethiopia: systematic review and meta-analysis. *Injury prevention* 27, 4 (2021), 384–394.
[10] Stefan Feuerriegel, Dennis Frauen, Valentyn Melnychuk, Jonas Schweisthal, Konstantin Hess, Alicia Curth, Stefan Bauer, Niki Kilbertus, Isaac S Kohane, and Mihaela van der Schaar. 2024. Causal machine learning for predicting treatment outcomes. *Nature Medicine* 30, 4 (2024), 958–968.
[11] Nicholas Fiorentini and Massimo Losa. 2020. Handling imbalanced data in road crash severity prediction by machine learning algorithms. *Infrastructures* 5, 7 (2020), 61.
[12] Junjie Hu, Jun Bai, Jiayu Yang, and Jaeyoung Jay Lee. 2024. Crash risk prediction using sparse collision data: Granger causal inference and graph convolutional networks approaches. *Expert Systems with Applications* (2024), 125315.
[13] Jean Kaddour, Aengus Lynch, Qi Liu, Matt J Kusner, and Ricardo Silva. 2022. Causal machine learning: A survey and open problems. *arXiv preprint arXiv:2206.15475* (2022).
[14] Kingo Kobayashi et al. 2002. *Mathematics of information and coding*. Vol. 203. American Mathematical Soc.
[15] P Sree Lakshmi, Monika Saxena, Sakshi Koli, Kapil Joshi, Khairul Hafezad Abdullah, and Durgaprasad Gangodkar. 2022. Traffic response system based on data mining and internet of things (IoT) for preventing accidents. In *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*. IEEE, 1092–1096.
[16] Zhengjing Ma, Gang Mei, and Salvatore Cuomo. 2021. An analytic framework using deep learning for prediction of traffic accident injury severity based on contributing factors. *Accident Analysis & Prevention* 160 (2021), 106322.
[17] Amir Mehdizadeh, Miao Cai, Qiong Hu, Mohammad Ali Alamdar Yazdi, Nasrin Mohabbati-Kalejahi, Alexander Vinel, Steven E Rigdon, Karen C Davis, and Fadel M Megahed. 2020. A review of data analytic applications in road traffic safety. Part 1: Descriptive and predictive modeling. *Sensors* 20, 4 (2020), 1107.
[18] Aleksander Molak. 2023. *Causal Inference and Discovery in Python: Unlock the secrets of modern causal machine learning with DoWhy, EconML, PyTorch and more* (1. ed.). Packt Publishing, Birmingham. https://amzn.to/3RebWzn.
[19] Bhaven Naik, Li-Wei Tung, Shanshan Zhao, and Aemal J Khattak. 2016. Weather impacts on single-vehicle truck crash injury severity. *Journal of safety research* 58 (2016), 57–65.
[20] Hamzah Al Najada and Imad Mahgoub. 2016. Anticipation and alert system of congestion and accidents in VANET using Big Data analysis for Intelligent Transportation Systems. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*. 1–8. https://doi.org/10.1109/SSCI.2016.7850097
[21] World Health Organization et al. 2018. Global status report on road safety.
[22] Mohamed Osman, Rajesh Paleti, Sabyasachee Mishra, and Mihalis M Golias. 2016. Analysis of injury severity of large truck crashes in work zones. *Accident Analysis & Prevention* 97 (2016), 261–273.
[23] Judea Pearl. 2010. An introduction to causal inference. *The international journal of biostatistics* 6, 2 (2010).
[24] Robert Raeside and David White. 2004. Predicting casualty numbers in Great Britain. *Transportation research record* 1897, 1 (2004), 142–147.
[25] Md Adilur Rahim and Hany M Hassan. 2021. A deep learning based traffic crash severity prediction framework. *Accident Analysis & Prevention* 154 (2021), 106090.
[26] Rene Richard and Suprio Ray. 2017. A tale of two cities: Analyzing road accidents with big spatial data. In *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, 3461–3470.
[27] Piotr Rzepakowski and Szymon Jaroszewicz. 2010. Decision trees for uplift modeling. In *2010 IEEE International Conference on Data Mining*. IEEE, 441–450.
[28] Piotr Rzepakowski and Szymon Jaroszewicz. 2012. Decision trees for uplift modeling with single and multiple treatments. *Knowledge and Information Systems* 32 (2012), 303–327.
[29] Anshuman Sharma, Zuduo Zheng, Jiwon Kim, Ashish Bhaskar, and Md Mazharul Haque. 2020. Is an informed driver a better decision maker? A grouped random parameters with heterogeneity-in-means approach to investigate the impact of the connected environment on driving behaviour in safety-critical situations. *Analytic methods in accident research* 27 (2020), 100127.
[30] Ali Shojaie and Emily B Fox. 2022. Granger causality: A review and recent advances. *Annual Review of Statistics and Its Application* 9, 1 (2022), 289–319.
[31] Nishtha Srivastava, Bhavesh N. Gohil, and Suprio Ray. 2024. Rail transit delay forecasting with Causal Machine Learning. In *Proceedings of 1st ACM SIGSPATIAL International Workshop on Spatiotemporal Causal Analysis* (Atlanta, GA, USA) *(STCausal '24)*. New York, NY, USA, 1–10.
[32] Lukuman Wahab and Haobin Jiang. 2020. Severity prediction of motorcycle crashes with machine learning methods. *International journal of crashworthiness* 25, 5 (2020), 485–492.
[33] Jianyu Wang, Shuo Ma, Pengpeng Jiao, Lanxin Ji, Xu Sun, and Huapu Lu. 2023. Analyzing the risk factors of traffic accident severity using a combination of random forest and association rules. *Applied Sciences* 13, 14 (2023), 8559.
[34] Awtachew Berhe Woldu, Abraham Aregay Desta, and Tewolde Wubayehu Woldearegay. 2020. Magnitude and determinants of road traffic accidents in Northern Ethiopia: a cross-sectional study. *BMJ open* 10, 2 (2020), e034133.
[35] Chengcheng Xu, Zijian Ding, Chen Wang, and Zhibin Li. 2019. Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes. *Journal of safety research* 71 (2019), 41–47.
[36] Guandong Xu, Tri Dung Duong, Qian Li, Shaowu Liu, and Xianzhi Wang. 2020. Causality learning: A new perspective for interpretable machine learning. *arXiv preprint arXiv:2006.16789* (2020).
[37] Zekun Yang, Wenping Zhang, and Juan Feng. 2022. Predicting multiple types of traffic accident severity with explanations: A multi-task deep learning framework. *Safety science* 146 (2022), 105522.
[38] Liuyi Yao, Sheng Li, Yaliang Li, Hongfei Xue, Jing Gao, and Aidong Zhang. 2019. On the estimation of treatment effect with text covariates. In *International Joint Conference on Artificial Intelligence*.
[39] Rongjie Yu and Mohamed Abdel-Aty. 2014. Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Safety science* 63 (2014), 50–56.
[40] Zihe Zhang, Qifan Nie, Jun Liu, Alex Hainen, Naima Islam, and Chenxuan Yang. 2024. Machine learning based real-time prediction of freeway crash risk using crowdsourced probe vehicle data. *Journal of Intelligent Transportation Systems* 28, 1 (2024), 84–102.
[41] Yang Zhao and Qing Liu. 2023. Causal ML: Python package for causal inference machine learning. *SoftwareX* 21 (2023), 101294.
[42] Zhenyu Zhao, Yumin Zhang, Totte Harinen, and Mike Yung. 2022. Feature selection methods for uplift modeling and heterogeneous treatment effect. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 217–230.