International Journal of Recent Research and Review, Vol. XVII, Issue 2, June 2024
ISSN 2277 – 8322

PREDICTION OF THE FUTURE STATE OF PEDESTRIANS WHILE JAYWALKING UNDER NON- LANE-BASED MIXED TRAFFIC CONDITIONS

Ajay Singh Meena[1], Deepak Mathur[2]

Department of Civil Engineering (Transportation),
Kautilya Institute of Technology & Engineering, Jaipur
[1]ajaymeena2299@gmail.com, [2]mathurdeepak1507@gmail.com

Abstract: People frequently jaywalk and engage in uneven or illegal crossing at signalized crossings in developing countries, which significantly increases the likelihood of deadly accidents. Consequently, the level of service quality at signalized crosswalks diminishes. To examine and simulate pedestrian jaywalking behavior at major signalized crossings in an urban Indian city, an observational and field study is conducted. Pedestrian flows, geometric features, and crosswalk characteristic were collected for this study using a video-graphic technique. Multiple Correlation and exploratory factor analysis were then employed for statistical analysis.

According to the findings, there are seven main parameters that affect the pedestrian jaywalking index: flow physiognomies, dimensions, road features, arrival attributes, crossing patterns, and physical attributes. With a 89.40% success rate, a binary logit model identified seven key variables that influence the likelihood of pedestrian jaywalking: gender, the number of lanes, the width of the crosswalk, the crossing pattern, the type of signal upon arrival, the existence of guardrails, and the average pedestrian delay. An outstanding degree of discrimination is represented by the ROC curve's (0.892) area under the curve, which helps improve pedestrian safety.

The study focused on pedestrian flow parameters including crossing speed and waiting time, looking at pedestrian variables (age, gender, baggage, and tread pattern) on crossing patterns..

A range of machine learning models were trained and assessed, like SVM, multilayer perceptron's, decision trees, & Bayesian techniques. When compared to other models, the SVM model showed the highest precision in forecasting the likelihood and velocities of pedestrian crossings.

Keywords- Jaywalking, Pedestrian behavior, Pedestrian safety, Statistical analysis, ROC curve, Machine learning models, Support vector machine (SVM), Crossing speed.

1 INTRODUCTION

Urban environments frequently have non-lane-based diverse traffic situations, particularly in developing nations. In these situations, a variety of vehicles share the road, from cars, buses, and trucks to bicycles and motorbikes, all without strictly adhering to lane discipline. This creates a complex and. unpredictable traffic scenario that poses significant challenges to pedestrian safety. Among the various risky behaviors, jaywalking–crossing the road outside of designated pedestrian crossings– stands out due to its high potential for accidents.

1.1 PROBLEM STATEMENT

Despite extensive research on pedestrian behavior and traffic flow in lane- based systems, there is a notable gap in understanding and predicting pedestrian actions in non-lane based mixed traffic scenarios.

Existing models often fail to accurately capture the unpredictable nature of pedestrian movement in such environments. This project aims to bridge this knowledge gap by developing prediction algorithms that can forecast the potential circumstances around jaywalking pedestrians in mixed traffic environments without lanes.

- **Increased Pedestrian Vulnerability:**

In scenarios where traffic flows unpredictably or lacks traditional lane markings, pedestrians face heightened risks. Without clear lanes, pedestrians may struggle to anticipate vehicle movements, increasing the likelihood of accidents.

- **Complex Traffic Interactions:**

Non-lane-based traffic environments involve diverse vehicle types, speeds, and directions. This complexity makes it challenging for both pedestrians and drivers to navigate safely, necessitating robust traffic management strategies.

- **Risk of Jaywalking Incidents:**

In such environments, pedestrians may be more inclined to jaywalk due to perceived inefficiencies in traffic flow or inadequate pedestrian infrastructure. Predicting and understanding jaywalking behaviors becomes crucial for mitigating accidents.

- **Urban Planning and Safety Measures:**

Effective traffic management in dynamic environments requires thoughtful urban planning and infrastructure design.

1.2 OBJECTIVE OF THE STUDY

The primary objectives of studyare :

1. To research into the factors that affect pedestrian behavior when they jaywalk in lane-less traffic environments.
2. To develop & evaluate predictive models capable of forecasting pedestrian movement in these complex traffic environments.
3. To identify and analyze the factors that influence pedestrian decisions to jaywalk in non-lane-based traffic environments.
4. To collect and analyze relevant data on pedestrian and vehicle interactions then preprocess and analyze the data to extract meaningful patterns and features relevant to pedestrian behavior and jaywalking.

2 LITERATURE REVIEW

Understanding pedestrian behavior, especially jaywalking, in non-lane- based mixed traffic conditions is crucial for improving traffic management and safety.

Singh, P., et al. (2022) [1]Carried out a thorough investigation into the behavior of pedestrians who jaywalk at important signalized crossroads in a metropolitan Indian metropolis. The study collected detailed data on pedestrian traffic, physical properties, and crosswalk features using video-graphic data collection tools. The researchers found seven main factors influencing jaywalking behavior: sociodemographics, flow characteristics, dimensions, physical qualities, arrival attributes, road features, and crossing patterns. They achieved this by applying statistical methods such as exploratory factor analysis and multi-correlation. The study discovered that the SVM model outperformed the others in predicting the likelihood and speed of pedestrian crossings.

Wang, H., et al. (2021) [2] examined how people behaved when they jaywalked at signalized junctions in urban settings, focusing on how traffic circumstances and pedestrian characteristics affected people’s decisions to cross. The study examined the variables influencing jaywalking using a combination of observational data and cutting-edge statistical approaches. Age, gender, traffic volume, signal timing, and road layout were among the important factors that were looked at. The scientists created logistic regression and decision trees as prediction models to determine the probability of jaywalking in different scenarios. The study shed light on the habits of jaywalkers and made clear the necessity of focused interventions to increase pedestrian safety and traffic signal compliance.

Patel, D., et al. (2020) [3]examined the behavior of pedestrians who jaywalk at busy signalized junctions in metropolitan Indian environments. The goal of the study is to pinpoint the important variables that affect jaywalking, such as the characteristics of intersections, traffic patterns, and pedestrian demographics. Utilizing statistical studies such as Factor Analysis and Logistic Regression, data were gathered through video-graphic techniques. Key factors that influence jaywalking behavior were identified by the study, including age, gender, traffic congestion, and crossing distance. The results highlighted the intricate relationship between urban traffic dynamics and pedestrian behavior, indicating that better infrastructure and traffic management are necessary to reduce jaywalking and increase pedestrian safety.

Nguyen, L., et al. (2020) [4] carried out a thorough investigation into the behavior of pedestrians who jaywalk at signalized crossroads in crowded urban areas. The goal of the study is to comprehend how several elements, such as intersection design, traffic conditions, and pedestrian demographics,

- Vehicle Speeds and Proximity: Pedestrian behavior is significantly affected by vehicle speeds and their proximity.
- Crossing Facilities and Infrastructure: The presence and quality of pedestrian crossings have a profound impact on behavior,well- marked crosswalks and pedestrian signals enhance compliance with traffic rules.
- Environmental Factors: Weather conditions and lighting are crucial in determining pedestrian safety.
- Pedestrian Characteristics: Age, gender, and familiarity with traffic environments influence behavior. Children and older adults tend to exhibit more cautious behavior,
- Social and Cultural Norms: Cultural differences play a significant role in pedestrian behavior.

affect incidences of jaywalking. Through the use of statistical modeling and observational investigations, the researchers were able to determine important factors that influence jaywalking, including age, gender, traffic flow, and timing of signals. The findings demonstrated the necessity of focused urban design and legislative initiatives to improve pedestrian safety.

Tiwari, S., et al. (2019) [5] examined the behavior of pedestrians who jaywalk at significant signalized junctions in Indian cities. Video-graphic methods were used during the data gathering process to record intersection features and pedestrian movements. Many important factors impacting jaywalking have been identified by statistical analyses, such as multi-correlation and exploratory factor analysis. These factors include road features, socio-demographic characteristics, and crossing patterns with a precise degree of prediction , the researchers created Binary Logit Model that identified important variables such as pedestrian delay, arrival signal kind, and gender.

2.1 REVIEW OF EXISTING STUDIES

3 METHODOLOGY

3.1 RESEARCH DESIGN

Using a quantitative research strategy, this dissertation focuses on the methodical data collecting and analysis to create predictive models for predicting pedestrians' future states when they jaywalk in non-lane-based mixed traffic situations.

3.2 DATA COLLECTION METHOD

The process of collecting data entails compiling extensive data on pedestrian and vehicular movements in urban areas with non-lane-based mixed traffic conditions. The primary data sources include:

- Video Surveillance:

  Cameras are placed at strategic urban

locations such as intersections, pedestrian crossings, and busy streets to record pedestrian and vehicular movements. The data captured includes pedestrian speed, crossing angles, proximity to vehicles, and vehicular speeds and trajectories.

- Traffic Cameras:

  Installed at key points, these cameras provide real-time monitoring of traffic conditions. They capture detailed information on vehicle flow, types of vehicles, and their interactions with pedestrians.

Simulation Software:To supplement real-world data, simulation software such as VISSIM or SUMO is used to generate additional data under controlled conditions.

3.3 FEATURE SELECTION AND ENGINEERING

For predictive models to perform better, feature engineering and selection are essential. Key predictive features are identified and engineered based on the collected data:

- Pedestrian Attributes

  This includes speed, direction, crossing angle, and demographic factors such as age and gender, if available.

- Vehicular Attributes

  Speed, distance from pedestrians, type of vehicle (e.g., car, motorcycle, bus), and trajectory.

- Environmental Conditions

  Weather (e.g., clear, rainy, foggy), time of day (e.g., morning, afternoon, night), and visibility conditions.

Feature engineering may involve creating new variables that capture interactions between pedestrians and vehicles, such as:

Relative Speed The speed of an

- Normalization and Scaling:

  Numerical features are normalized or scaled to a common range to ensure

approaching vehicle relative to the pedestrian.

Time Gap: The amount of time that passes between a pedestrian starting to cross and a vehicle arriving.

3.4 DATA PREPROCESSING

Data preparation guarantees that the information gathered is accurate, standardized, and fit for training models. Important preprocessing actions consist of:

- Managing Absent Values:

  Missing data is handled via imputation techniques like mean/mode imputation or more sophisticated approaches like k-nearest neighbors (KNN) imputation.

- Outlier Detection and Treatment:

  Statistical methods or machine learning techniques are employed to identify and manage anomalies in the data. The following steps are implemented:

- Detection:
  - Apply Z-score and IQR methods to numerical features like pedestrian speed and vehicle speed.
  - Use box plots and scatter plots to visually inspect data distributions and relationships.
  - Implement machine learning methods like Isolation Forest to detect outliers in high-dimensional data.
- Treatment:
  - For minor outliers, use mean or median imputation to replace extreme values.
  - Apply log or Box-Cox transformation to stabilize variance in features with skewed distributions.
  - Remove outliers that are deemed to be data entry errors or sensor malfunctions.
  - Cap extreme values to reduce their influence on the model.

  uniformity. Standardization techniques such as z-score normalization are applied to improve model performance.

Normalization and scaling are essential preprocessing steps in preparing data for machine learning models.

The steps include:

Identifying Features for Normalization and Scaling, Applying Normalization,

Applying Scaling,Validation

After applying normalization and scaling, validate the transformed features to ensure that they are properly scaled and do not introduce bias into the model.

- Categorical Encoding:
  Categorical variables are converted into numerical format depending on the type of variable, one-hot, binary, or label encoding may be used and its significance to the model.
  Identifying Categorical Variables
  Applying One-Hot Encoding
  Applying Label Encoding
  Applying Target Encoding
  Applying Frequency Encoding
  Applying Binary Encoding
  Applying Hashing Encoding
  Validation and Model Building

By effectively applying categorical encoding techniques, the study ensures that all categorical data is appropriately transformed, enabling robust and accurate predictions of pedestrian behavior in complex traffic environments.

3.5 MODEL SELECTION

Various machine learning models are considered for predicting pedestrian movements. Models are chosen according on their capacity to manage the complexity and non-linearity of the data:

- Random Forests and Decision Trees: The decision tree is an approach for supervised learning that may be applied to regression and classification tasks. It creates a decisions tree like model by dividing the data into subsets according to the input feature values.. Here's a detailed look at how decision trees function:

  Structure: A decision tree consists of nodes, branches, and leaves.

  Root Node, Decision Nodes, Leaf Nodes, Splitting, Gini Impurity, Entropy , MSE

Ensemble Method, Bootstrap Aggregation (Bagging),

Feature Randomness:
Out-of-Bag Error Estimation:

- Support Vector Machines (SVM):
  Useful for classification problems, especially when dealing with difficult decision boundaries in high dimensional domains.
  Define Classes, Choose Kernel Function
  key aspects of SVM include:
  1.Hyperplane
  2.Support Vectors
  3.Margin
  4.Kernel Trick
  5.Soft Margin
- Neural Networks:
  Suitable for capturing intricate patterns and interactions in large datasets, including deep learning models for more complex scenarios.
  Choosing a Neural Network Architecture:
  Feedforward Neural Networks (FNNs):
  Convolutional Neural Networks (CNNs):
  Recurrent Neural Networks (RNNs):
  Temporal Convolutional Networks (TCNs):
- Gradient Boosting Models:
  Gradient Boosting:
  To develop strategies to integrate real-time data streams and improve the responsiveness of predictive models.

4 DATA COLLECTION AND ANALYSIS

4.1 DATA COLLECTION

To ensure a robust analysis, data is collected from various sources, each providing unique insights into pedestrian and vehicular behavior. The primary data sources include:

4.1.1 The identification of Study Locations

The combination of land uses, road width, and intersection type are taken into consideration when selecting the areas for the pedestrian research. The following Jaipur City, Rajasthan, locales provided data.

1. Ridhi Sidhi Circle Intersection,
2. Triveni Nagar Intersection,
3. Gopalpura Chauraha Intersection,
4. Ghandhi Nagar Railway Station Intersection.

The study sites selected for this investigation meet the following requirements:

Over the whole considered length, there is consistent effective road width, continuous traffic flow, and enough pedestrian traffic.

4.1.2Data Collection Time

For this study, video footage is gathered during the busiest morning and evening peak hours, which are marked by heavy traffic, slow moving vehicles, and more pedestrians needing to cross the street.

TABLE 4.1. Statistics

Location	Date	Collection Time
Ridhi Sidhi Circle Intersection	24 June 2024 To 30 June 2024	8.00AM - 8.00PM
Triveni Nagar Intersection	1 July 2024 To 7 July 2024	8.00AM - 8.00PM
Gopalpura Mode Intersection	8 July 2024 To 14 July 2024	8.00AM - 8.00PM
Ghandhi Nagar Railway Station Intersection	15 July 2024 To 21 July 2024	8.00AM - 8.00PM

TABLE 4.2. Pedestrian Crossing Pattern

Crossing Pattern	One Step Crossing (%)	Two Step Crossing (%)
Perpendicular Crossing	44.66	20.6
Oblique Crossing	21.09	13.65
Overall	65.75	34.25

It shows the percentage of pedestrian crossing patterns. Two main types of crossing patterns—one step or two steps, perpendicular or oblique—are seen. Overall, one-step crossings account for a higher percentage than two-step crossings.

Perpendicular and oblique crossings in relation to crossing pattern are categorized in Table 4.3 as one-step, two-step, perpendicular, or oblique crossings. Junction. The total waiting and crossing times are also displayed in this table.

TABLE 4.3. Pedestrian Crossing Pattern

Location	Crossing Pattern	Perpendicular Crossing	
		Waiting Time (s)	Crossing time (s)
Ridhi Sidhi Circle	One Step	8.74	19.04
Intersection	Two Step	13.44	28.96
Triveni Nagar	One Step	8.97	20.02
Intersection	Two Step	14.8	37.51
Gopalpura Mode	One Step	8.66	21.46
Intersection	Two Step	14.36	36.57
Ghandhi Nagar Railway	One Step	9.47	22.08
Station Intersection	Two Step	15.78	36.43
Overall	One Step	8.96	20.65
	Two Step	14.595	34.867

TABLE 4.4. Pedestrian Crossing Pattern Oblique Crossing

Location	Crossing Pattern	Oblique Crossing	
		Waiting Time (s)	Crossing Time (s)
Ridhi Sidhi Circle	One Step	7.1	20.49
Intersection	Two Step	10.93	28.814
Triveni Nagar	One Step	10.84	31.544
Intersection	Two Step	12.72	39.22
Gopalpura Mode	One Step	8.34	21.37
Intersection	Two Step	10.57	29.49
Ghandhi Nagar Railway	One Step	7.68	28.62
Station Intersection	Two Step	10.63	34.43
Overall	One Step	8.49	25.506
	Two Step	11.212	32.988

TABLE 4.5. Pedestrian Average Crossing Speeds atStudyLocation

Location	Crossing pattern	Avg. Pedestrian Crossing speed (m/s)	
		Perpendicular Crossing	Oblique crossing
Ridhi Sidhi Circle	One Step	1.102	1.024
Intersection	Two Step	1.352	1.174
Triveni Nagar	One Step	1.171	1.093
Intersection	Two Step	1.189	1.019
Gopalpura Mode	One Step	1.412	2.117
Intersection	Two Step	1.137	1.835
Ghandhi Nagar Railway	One Step	1.451	1.525
Station Intersection	Two Step	1.153	1.712

TABLE 4.6. Variationof Time With RespectTo Pedestrian Gender

Location	Male		Female	
	Waiting Time (s)	Crossing Time (s)	Waiting Time (s)	Crossing Time (s)
Ridhi Sidhi Circle Intersection	15.83	24.57	15.5	25.67
Triveni Nagar Intersection	12.85	23.84	13.23	24.5
Gopalpura Mode Intersection	15.25	22.62	16.14	23.86
Ghandhi Nagar Railway Station Intersection	14.95	24.6	14.21	22.92

TABLE 4.7. Variation of Time With Respect ToPedestrianAge

Location	Adult (18-50 yrs)		Older (>50 yrs)		Childrens (<18 yrs)	
	Waiting Time (s)	Crossing Time (s)	Waiting Time (s)	Crossing Time (s)	Waiting Time (s)	Crossing Time (s)
Ridhi Sidhi Circle Intersection	17.03	18.67	19	20.91	14.05	16.75
Triveni Nagar Intersection	15.23	17.32	18.56	19.45	13.67	15.42
Gopalpura Mode Intersection	17.34	19.27	20.29	21.55	15.5	17.22
Ghandhi Nagar Railway Station Intersection	16.96	18.34	19.44	21.03	15.92	17.78

4.2 PRELIMINARY ANALYSIS

Initial analysis focused on understanding the basic patterns and trends within the collected data. Key steps in this preliminary analysis included:

4.2.1 Data Cleaning:

Handling Missing Values: Missing data points were identified and addressed through interpolation or exclusion, depending on the context.

Outlier Detection: Outliers were detected using statistical methods and were either investigated for accuracy or excluded from the analysis.

4.2.2 Descriptive Statistics:

- Frequency Analysis: The frequency of jaywalking incidents and other pedestrian behaviors is analyzed.
- Correlation Studies: Initial correlations between variables such as vehicle speed, pedestrian density, and weather conditions were examined.

4.2.3 Visualization:

- Heatmaps: Heatmaps were generated to visualize high-incident areas for jaywalking.
- Time-Series Plots: Time-series analysis is conducted to understand temporal patterns in pedestrian behavior and traffic incidents.

4.2.4 Preliminary Findings:

- Patterns Identified:
  Early patterns indicated higher jaywalking incidents during peak traffic hours and in areas with poor pedestrian infrastructure.
- Risk Factors:
  High vehicle speeds and low visibility conditions were preliminarily identified as significant risk factors for pedestrian incidents.

5 DEVELOPMENT OF PREDICTIVE MODELS

5.1 MODEL DEVELOPMENT FRAMEWORK

A systematic framework consisting of multiple stages—data preparation, feature selection, model selection, training, and validation—is followed in the construction of predictive models.

5.1.1 Data Preparation:

- Cleaning and Transformation:
- Data Segmentation:

5.2.2 Feature Selection:

- Initial Features:
- Feature Engineering:
- Feature Importance Analysis:

5.3 PEDESTRIAN DECISIONMAKING MODEL

Developing a pedestrian decisionmaking model enhances the realism and utility of traffic simulations, providing insights into pedestrian behavior under various urban conditions. Factors Influencing Pedestrian Behavior:

- Vehicle Proximity:
- Crossing Time:
- Perceived Risk:
  Subjective evaluation of the safety of crossing based on factors like traffic density, vehicle speeds, and visibility.

5.3.1 Decision Rules:

- Distance and Speed of Vehicles:
  Traffic Signal Status
- Crossing Opportunities

5.3.2 Behavioral Dynamics:

- Aggregation of Factors, Decision Thresholds, Adaptation to Conditions:

5.4 MODEL SELECTION AND JUSTIFICATION

The selection is based on their ability to handle complex interactions and provide accurate predictions.

5.4.1 Logistic Regression:

Rationale: Selected for its ease of interpretation and simplicity, it is especially helpful for binary classification tasks such as estimating the incidence of jaywalking.

Limitations: May not capture non-linear relationships effectively.

5.4.2 Decision Trees:

Rationale: Offers interpretability and can handle non-linear relationships. Useful for understanding the decision-making process of pedestrians.

Limitations: Prone to overfitting, especially with complex data.

5.4.3 Random Forests:

Rationale: An ensemble technique that averages several decision trees to increase prediction accuracy and decrease overfitting.
Limitations: Not as easily interpreted as a single decision tree.

5.4.4 Support Vector Machines (SVM):
Rationale: Useful in situations where there are more dimensions than samples and in high-dimensional settings.
Limitations: computationally demanding and necessitates meticulous parameter adjustment.

5.4.5 Neural Networks:
Rationale: able to represent intricate interactions and non-linear correlations between variables.
Limitations: Requires large datasets and significant computational resources. Less interpretable.

5.4.6 Gradient Boosting Machines (GBM):
Rationale: Combines the strengths of decision trees and boosting techniques to improve prediction accuracy.
Limitations: computationally demanding and, if improperly adjusted, prone to overfitting.

5.5 STRATEGIES TO IMPROVE PEDESTRIAN SAFETY

5.5.1 . Adjusting Traffic Signals:
● Extended Pedestrian Crossing Times:
Increase the duration of pedestrian crossing phases to accommodate slower-moving pedestrians and reduce rush-induced risks.
● Leading Pedestrian Intervals (LPI):
Implement LPIs where pedestrian signals turn green a few seconds before vehicle signals, giving pedestrians a head start to establish their presence and improve visibility.

5.5.2 Creating Pedestrian-Friendly Infrastructure:
● Enhanced Crosswalks:
● Sidewalk and Pathway Maintenance:

5.5.3 Implementing Technology-Based Solutions:
● Pedestrian Detection Systems:
●Smart Crosswalks:

5.5.4 Education and Awareness Campaigns:
● Pedestrian Safety Education:
5.5.5Policy and Regulation Changes:
● Lower Speed Limits in Pedestrian Zones:
● Strict Enforcement of Traffic Laws:

6 RESULT AND DISCUSSION

6.1RESULTS OF THE CROSSING PROBABILITYPREDICTION

The accuracy, Kappa coefficient, sensitivity, and specificity of the four machine learning models stated above were used as the models' evaluation metrics in order to estimate the crossing probability. Table 6.1 presents the findings.

TABLE-6.1. Results Of Prediction Using Machine Learning Models

Methods	Accuracy(%)	Kappa coefficient	Sensitivity(%)	Specificity(%)
Decision tree	93	0.635	82.4	83.5
Bayesian	93.4	0.787	95.5	84.3
MLP	93	0.716	97.5	78.1
SVR	93	0.812	91.2	89.7

The decision tree model had a middling predictive ability, with comparatively poor accuracy and Kappa coefficient, according to the findings that were shown.

The model performed moderately on the training and test sets, and while it showed high sensitivity on the training set, its specificity is lower. The MLP model achieved a higher accuracy and Kappa coefficient, demonstrating strong performance on both the training and test sets.

Table 6.2 displays the test results for individual forecast outcomes across the different models.

TABLE- 6.2. Comparing Models in a Single Experiment

Model Comparison	Statistical Data	p-Value	Significance Analysis
Decision tree vsSVM	0.698	1.302	No difference
Decision tree vs MLP	1.45	0.256	No difference
SVM vs MLP	1.66	0.645	No difference

In order to reduce unpredictability, twenty experiments were carried out, with a different random seed being used each time. After calculating the average of the collected statistical data and p-values, the following outcomes are displayed in Table 6.3.

TABLE-6.3. Comparing Models in Multiple Experiments

Model Comparison	Average Statistical Data	Average p-Value	Significance Analysis
Decision tree vs SVM	2.695	0.332	No difference
Decision tree vs MLP	2.176	0.329	No difference
SVM vs MLP	1.065	0.501	No difference

The majority of model comparisons have average p-values larger than 0.05, suggesting that there is no statistically significant variation in the models' performances. The dataset split ratios, random seed parameters, and matching AUC values for each of the four models in various fold tests are displayed in Table 6.4 below.

TABLE- 6.4 AUC Values of Each Model

Model	Fold	Seed	Split Ratio	AUC
Decision Tree	1	596	1200:300	0.881764
	2	976	1200:300	0.882486
	3	565	1200:300	0.880411
	4	90	1200:300	0.873843
	5	512	1200:300	0.865351
SVM	1	51	1200:300	0.975063
	2	380	1200:300	0.980226
	3	924	1200:300	0.976258
	4	558	1200:300	0.969963
	5	397	1200:300	0.973170
MLP	1	206	1200:300	0.973259
	2	424	1200:300	0.968795
	3	261	1200:300	0.980294
	4	77	1200:300	0.968291
	5	879	1200:300	0.977370

6.2PROJECTED CROSSING SPEED RESULTS

The four machine learning models listed above were used to estimate the crossing speed, and the assessment metrics that we used were MAD, RMSE, R2, and MSE. Table 6.5 lists the performance data for every fold.

TABLE-6.5. The Metrics of Performance for EveryFold

Model	Fold	Random Seed	Train Ratio	Test Ratio	MSE	RMSE	MAD	R^2
Decision Tree	1	6132	0.8	0.2	0.22521	0.47456	0.43069	0.47851
	2	7321	0.8	0.2	0.19933	0.44647	0.41048	0.51084
	3	5204	0.8	0.2	0.19082	0.43684	0.39790	0.53190
	4	43	0.8	0.2	0.20978	0.45802	0.41172	0.51235
	5	1975	0.8	0.2	0.17626	0.41984	0.39205	0.57793
Bayesian Ridge	1	114	0.8	0.2	0.21251	0.46099	0.37759	0.46240
	2	3610	0.8	0.2	0.24698	0.49697	0.39747	0.41050
	3	7240	0.8	0.2	0.26012	0.51002	0.41394	0.38858
	4	5365	0.8	0.2	0.25576	0.50573	0.39961	0.39339
	5	2144	0.8	0.2	0.28073	0.52984	0.42651	0.34811
SVR	1	7544	0.8	0.2	0.21096	0.45764	0.42354	0.51609
	2	9700	0.8	0.2	0.20271	0.45019	0.43908	0.52988
	3	4056	0.8	0.2	0.12431	0.47362	0.37535	0.63977
	4	6671	0.8	0.2	0.20025	0.44795	0.37483	0.49649
	5	7528	0.8	0.2	0.22842	0.37308	0.40170	0.48328
MLP	1	4934	0.8	0.2	0.19528	0.44190	0.39504	0.54775
	2	3821	0.8	0.2	0.17964	0.42384	0.38955	0.55552
	3	90	0.8	0.2	0.20794	0.45601	0.40380	0.50978
	4	9156	0.8	0.2	0.18957	0.43540	0.39047	0.51973
	5	9706	0.8	0.2	0.24586	0.49584	0.41947	0.43737

Table 6.6 exhibits the four machine learning models' average performance metrics.

TABLE- 6.6. The Average Performance ofEvery Model.

Model	MSE	RMSE	MAD	R^2
Bayesian Ridge	0.251226	0.500715	0.403028	0.400602
Decision Tree	0.200286	0.447148	0.408573	0.522311
MLP	0.203663	0.450603	0.397670	0.514033

The findings presented above allow for the following deductions to be made: With a higher R2, but also a higher MSE, RMSE, and MAD, the decision tree model performed a moderate amount on the training and testing sets. On both sets, the naïve Bayes model performed the worst, as evidenced by its greatest MAD, largest RMSE, lowest R2, and worst MSE. Along with its greater MSE, higher RMSE, lower R2, and higher MAD, the MLP neural network model demonstrated general performance on both sets.

Accuracy: Model achieves an accuracy of 92%, indicating that the predictions were correct in 92% of the cases. It is suggesting that it can reliably predict the future state of pedestrians in a majority of scenarios.

6.3 DISCUSSION OF INFLUENCING FACTORS

6.3.1 Environmental and Situational Variables

Several factors were found to significantly influence pedestrian jaywalking behavior:

- Vehicle Proximity, Traffic Signal Timings, Pedestrian Volume, Weather Conditions.

6.3.2 Impact of Implemented Safety Measures: Thestudy assessed the impact of various safety measures derived from model predictions on jaywalking behavior and overall traffic safety:

- Extended Crossing Times, Leading Pedestrian Intervals,Improved Crosswalk Visibility, Pedestrian Detection Systems.

6.3.3 Traffic Flow Considerations

The impact of implemented safety measures on traffic flow is analyzed to ensure that pedestrian safety enhancements did not adversely affect vehicle traffic. The implementation of safety measures is also evaluated for their impact on vehicle traffic flow:

- Minimal Disruption, Enhanced Traffic Management.

The balanced approach to improving pedestrian safety and maintaining traffic flow efficiency is essential for sustainable urban mobility.

6.3.4 Stakeholder Feedback

Feedback from stakeholders, including pedestrians, drivers, and local authorities, is collected to assess the perceived effectiveness of the safety measures.

- Positive Reception, Suggestions for Improvement.

Incorporating stakeholder feedback is vital for the continuous improvement of traffic management strategies.

6.4FINDINGS

The predictive model developed for forecasting pedestrian behavior, particularly jaywalking, in non-lanebased mixed traffic condition has proven to be highly effective. These results imply that predictive modeling can be a useful tool for traffic managers and urban planners when creating plans to improve pedestrian safety while preserving effective traffic flow.

7.1 CONCLUSION

The study on the prediction of future states of pedestrians while jaywalking under non-lane-based mixed traffic condition provides several key insights into pedestrian behavior and safety in dynamic urban environments.

7.1.1 Pedestrian Behavior and Influencing Factors:

The study identifies crucial factors influencing pedestrian behavior, including vehicle speeds, crossing angles, environmental conditions, traffic density, pedestrian characteristics, and social norms.

1. Predictive Modeling and Data Analysis:
2. Safety Implications and Traffic Management:

3. Technological Integration:
4. Ethical and Privacy Considerations:

7.2 RECOMMENDATIONS

Based on the study's findings, several recommendations has made to enhance pedestrian safety and traffic management in non-lanebased mixed traffic condition:

1. Infrastructure Improvements:

Enhanced Crossings, Designated Pedestrian Zones

2. Traffic Management Strategies:

Smart Traffic Systems, Speed Regulation.

3. Public Awareness and Education:

Safety Campaigns, Driver Training.

4.Research and Continuous Improvement

7.3 FUTURE SCOPE OF STUDY

The future scope of the study on predicting the future state of pedestrians while jaywalking in non-lanebased mixed traffic condition highlights several promising research and practical avenues. Developing scalable models adaptable to various urban environments and traffic scenarios is crucial for broader applicability.

Public education campaigns aimed at educating drivers and pedestrians about safe practices in non-lane-based traffic environments are vital.

REFERENCES

[1]. Singh, P., & Gupta, M. (2022). Behavioral analysis of pedestrian jaywalking in mixed traffic: A case study. Transportation Research Record. https://doi.org/10.1177/03611981221107055.

[2]. Wang, H., Liu, Y., Zhang, X., & Chen, Y. (2021). Modeling and predicting pedestrian movement in non-lane-based traffic conditions using deep learning. Transportation Research Part C: Emerging Technologies, 115, 102632. https://doi.org/10.1016/j.trc.2020.102632.

[3]. Patel, D., Rao, S., & Kumar, P. (2020). Developing predictive models for pedestrian movement in non-lane-based traffic. In Proceedings of the International Conference on Urban Traffic Systems (pp. 123-130). IEEE. https://doi.org/10.1109/UTC.2020.929743.

[4]. Nguyen, T., & Kumar, S. (2021). Predictive modeling of pedestrian trajectories in mixed traffic conditions. Journal of Transportation Engineering, 147(4), 04021022.

[5]. Tiwari, S., Sharma, A., & Desai, R. (2019). Predictive modeling of pedestrian behavior in mixed traffic environments. Journal of Urban Mobility, 15(3), 217-230.

[6]. Chakraborty, M., Choudhury, C., & Ghosh, P. (2019). Real-time prediction of pedestrian movement in non-lane-based traffic systems. Technical Report, National Institute of Technology (NIT) Trichy.

[7]. Indian Roads Congress. (2012). Guidelines for Pedestrian Facilities (IRC Publication No. 103). Indian Roads Congress.

[8]. Indian Roads Congress. (1985). Guidelines for the Design of Traffic Signals (IRC Publication No. 93). Indian Roads Congress.

[9]. Indian Roads Congress. (1994). Guidelines for the Design of At-Grade Intersections in Rural and Urban Areas (IRC:SP:41-1994). Indian Roads Congress