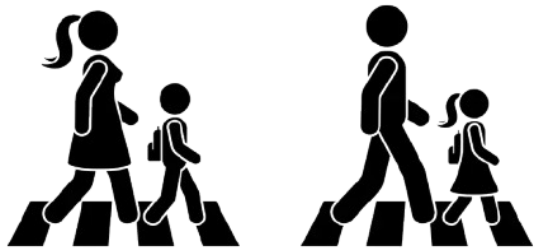
17th Urban Mobility India Conference & Expo 2024



PEDESTRIAN CRASH PREDICTION MODELLING ON A CORRIDOR BASED APPROACH USING MACHINE LEARNING TECHNIQUES

Paper ID: 0713

Ms. Soja R. Joseph
M. Tech Transportation Engineering
Jyothi Engineering College
Thrissur, Kerala

Co-Authors:

Mr. Sanjay Kumar V. S.
Principal Scientist & Head
Traffic Engineering & Safety Division
KSCSTE-NATPAC
Thiruvananthapuram, Kerala

Ms. Archana S.
Assistant Professor
Department of Civil Engineering
Jyothi Engineering College
Thrissur, Kerala

Ms. Jinumol K. R.
Project Fellow (M. Tech)
Traffic Engineering & Safety Division
KSCSTE-NATPAC
Thiruvananthapuram, Kerala

CONTENTS

- Introduction
- Literature Review
- Scope & Objectives
- Methodology
- Severity of Pedestrian Crashes: Thrissur District
- Demarcation of Pedestrian Vulnerable Corridors
- Pedestrian Vulnerable Corridors Based on Road Type
- Summary Statistics of Parameters
- Data Collection
- Pedestrian Crash Severity Prediction Models
- Feature Importance Analysis
- Performance Analysis
- Conclusion
- References

INTRODUCTION

- Globally, more than half of road crash fatalities are among **Vulnerable Road Users (VRU)**
  - VRUs : Those without the protection of an external shield
    - Pedestrians, cyclists, three-wheelers, and motorcyclists across all age groups
- Pedestrians are the most vulnerable group in road crashes
  - **Second-largest** group of reported fatalities among road user types, accounting for 23% (*Global Status Report on Road Safety 2023*)
- During the period 2018 to 2022, Kerala witnessed more than **5100 pedestrian fatalities** (*SCRB)*
- Studies on **pedestrian crash severity** modelling help better understand what factors contribute to injury severity
- The **corridor-level analysis** enhances understanding for developing effective countermeasures to improve pedestrian safety

LITERATURE REVIEW

Author	Findings
Tiwari, G. (2020)	Summary of the progress made in understanding pedestrian crash patterns over the last 120 years New research efforts are required to address pedestrian safety
Das, S. et al. (2020)	Identified factors like pedestrian volume, high approach speed, vehicle overtaking, footpath encroachment, on-street parking, wider road width, certain land-use types, inadequate sight distance, insufficient lighting, and absent police enforcement that contribute to pedestrian fatalities in hazardous road corridors
Al-Mahameed et al. (2019)	A corridor-level approach was adopted for study areas with a minimum of 100 residents per square mile, utilizing a grid method
Elalouf, A. et al. (2023)	The Extra Tree Classifier demonstrated the highest level of efficacy in predicting the seriousness of injuries sustained by pedestrians involved in traffic accidents among Support Vector Machine (SVM), Linear Support Vector Classifier (SVC), Decision Tree Classifier, Random Forest Classifier, Extra Tree Classifier, K-nearest Neighbors
Santos, K. et al. (2022)	Random Forest revealed to be a good approach for road traffic crash injury severity prediction followed by SVM, Decision Tree, and K-Nearest Neighbor
Meocci, M. et al. (2021)	Gradient boosting seems to be an appropriate model to fit classification models

SCOPE & OBJECTIVES

❖SCOPE

The scope of the study is limited to pedestrian crashes comprising fatal crashes, grievous injury crashes & minor injury crashes reported in Thrissur district in Kerala from 2018 to 2022.

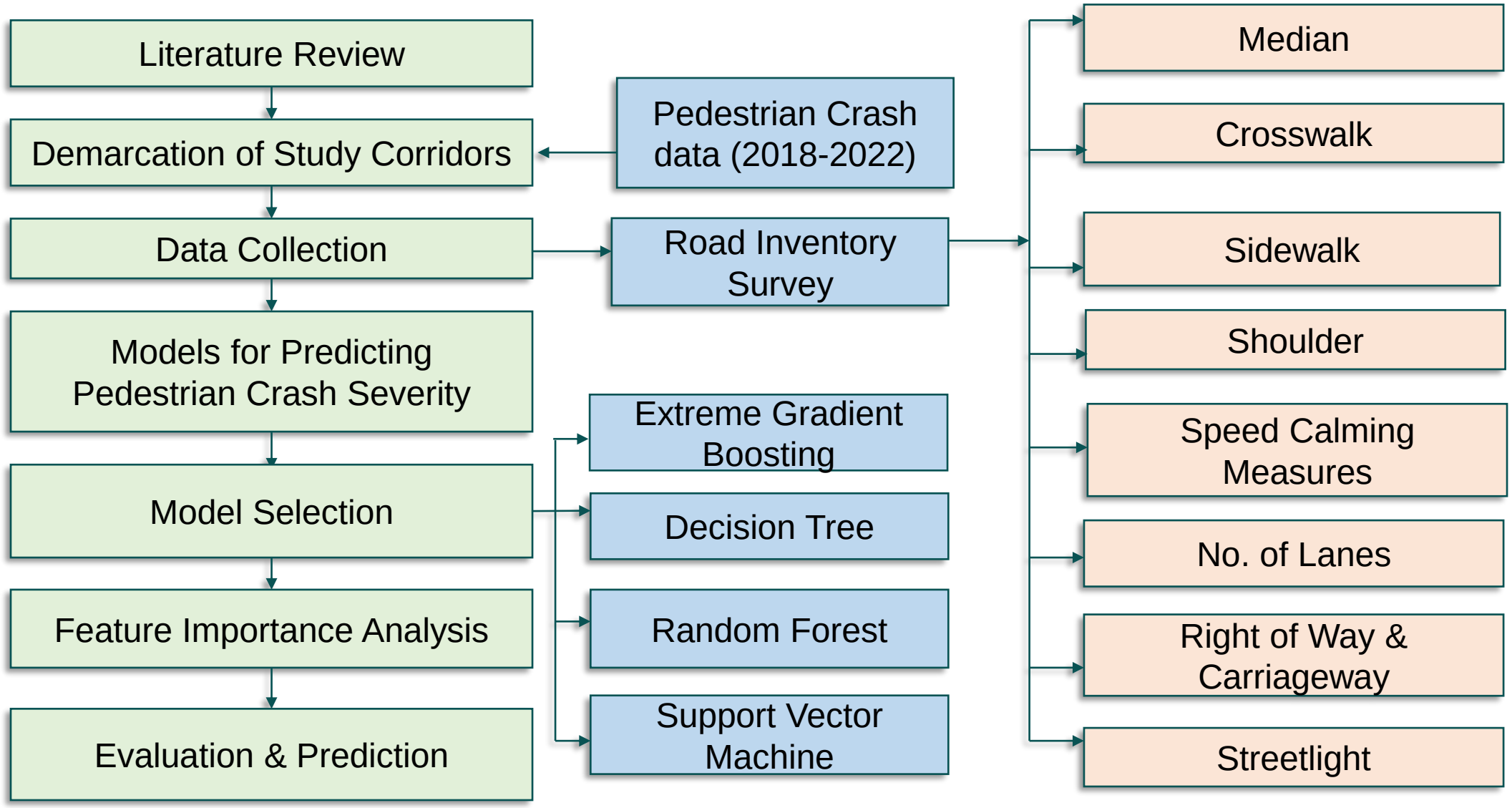
❖OBJECTIVES

- To demarcate the pedestrian vulnerable corridors in Thrissur district in Kerala
- To develop a pedestrian crash severity model using machine learning techniques

PEDESTRIAN VULNERABLE CORRIDORS

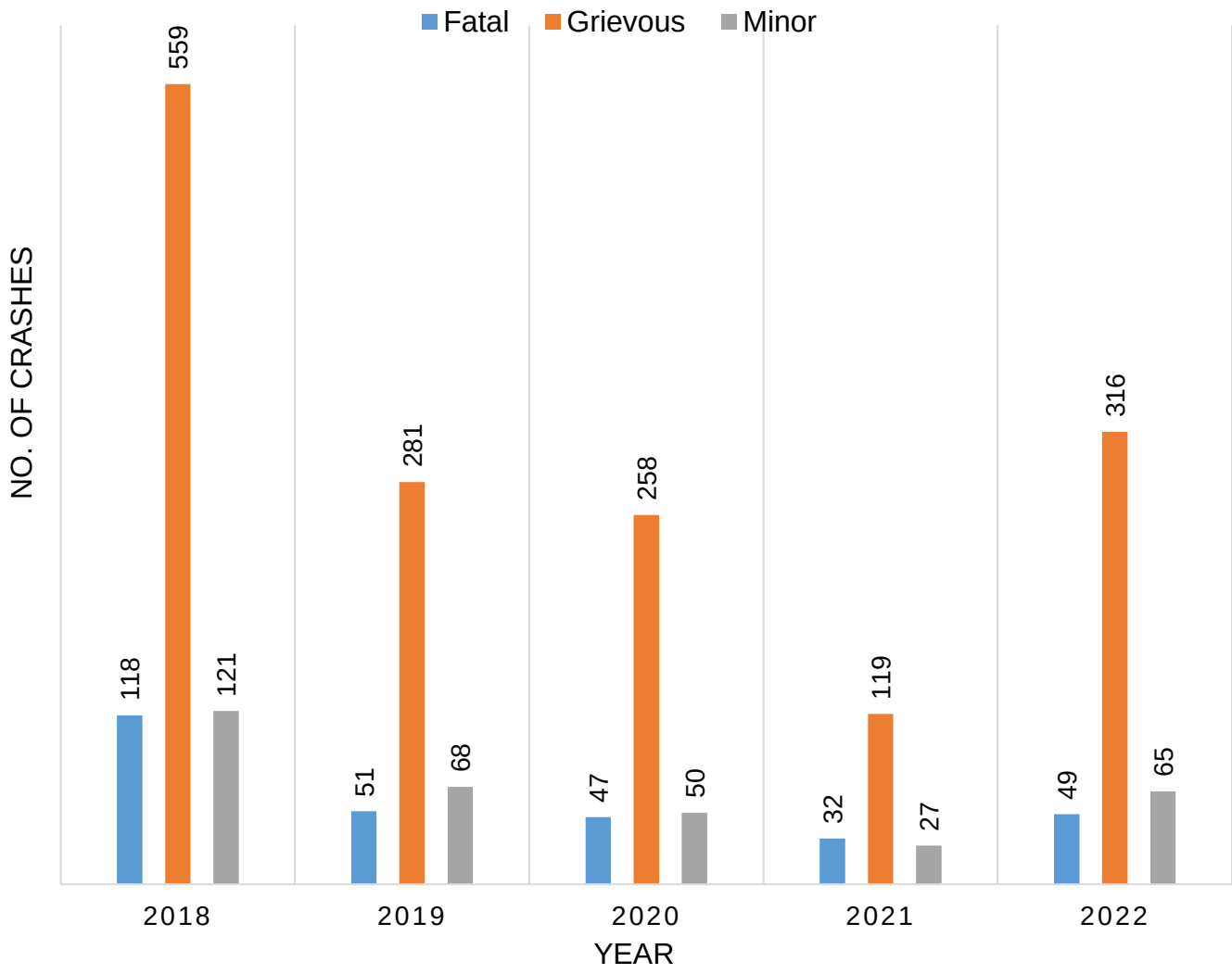
- Specific sections of roads that pose increased risks and safety concerns for pedestrians
- Characterized by conditions that make walking more hazardous and increase the likelihood of pedestrian related crashes

METHODOLOGY



SEVERITY OF PEDESTRIAN CRASHES: THRISSUR DISTRICT

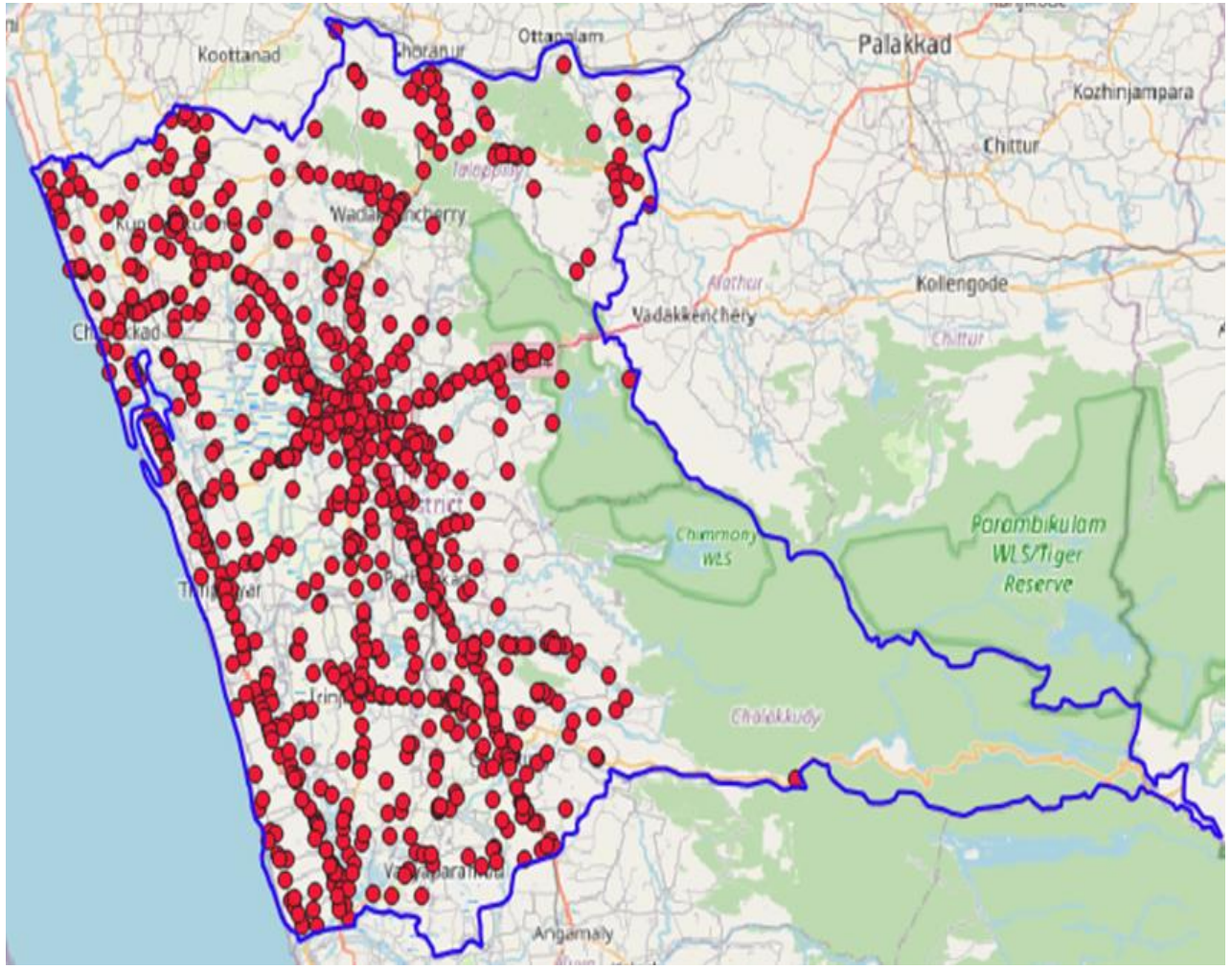
- 2nd ranked district in Kerala in terms of pedestrian fatalities from 2018 to 2022
- 2160 Pedestrian Crashes
  - ✓297 fatal crashes
  - ✓1,533 grievous injury crashes
  - ✓330 minor injury crashes



DEMARCATION OF PEDESTRIAN VULNERABLE CORRIDORS

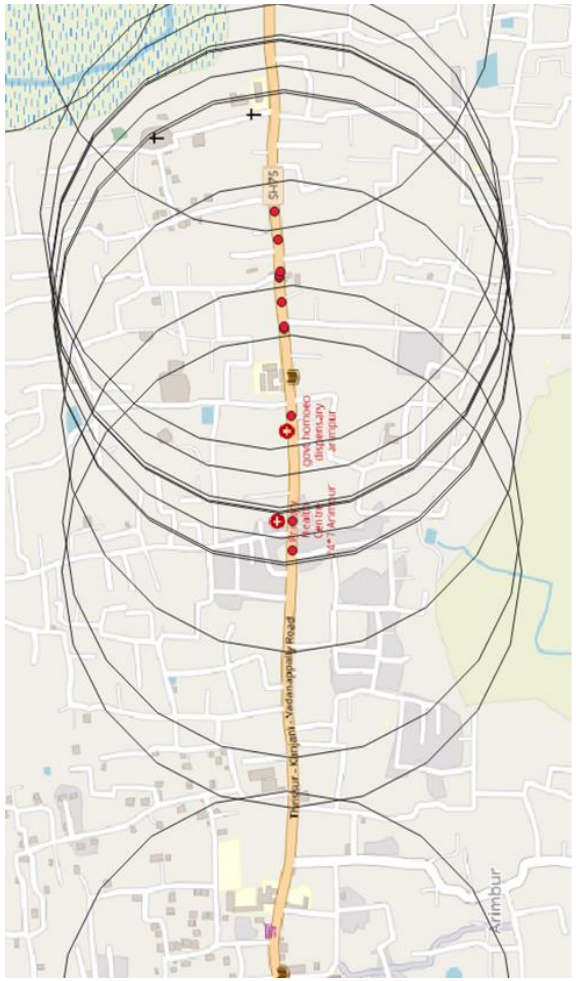
Steps involved :

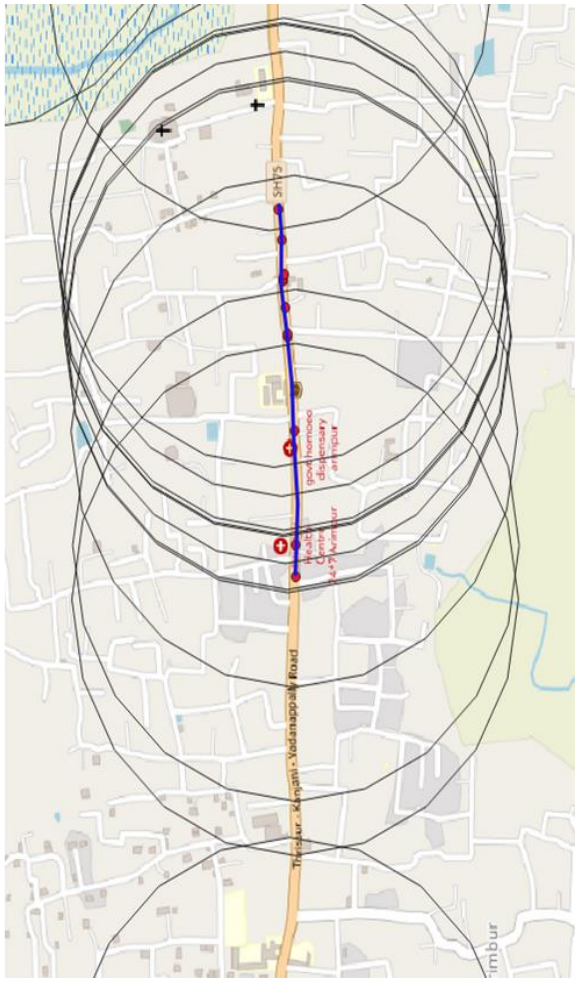
- Importing & Cleaning of crash data
- Buffer of 500m radius generated for each crash location
- Connect nearby pedestrian crash points until there are no crashes within a 500m radius
- Repeat the procedure for locations with a minimum of two crashes
- Single crashes within a buffer eliminated

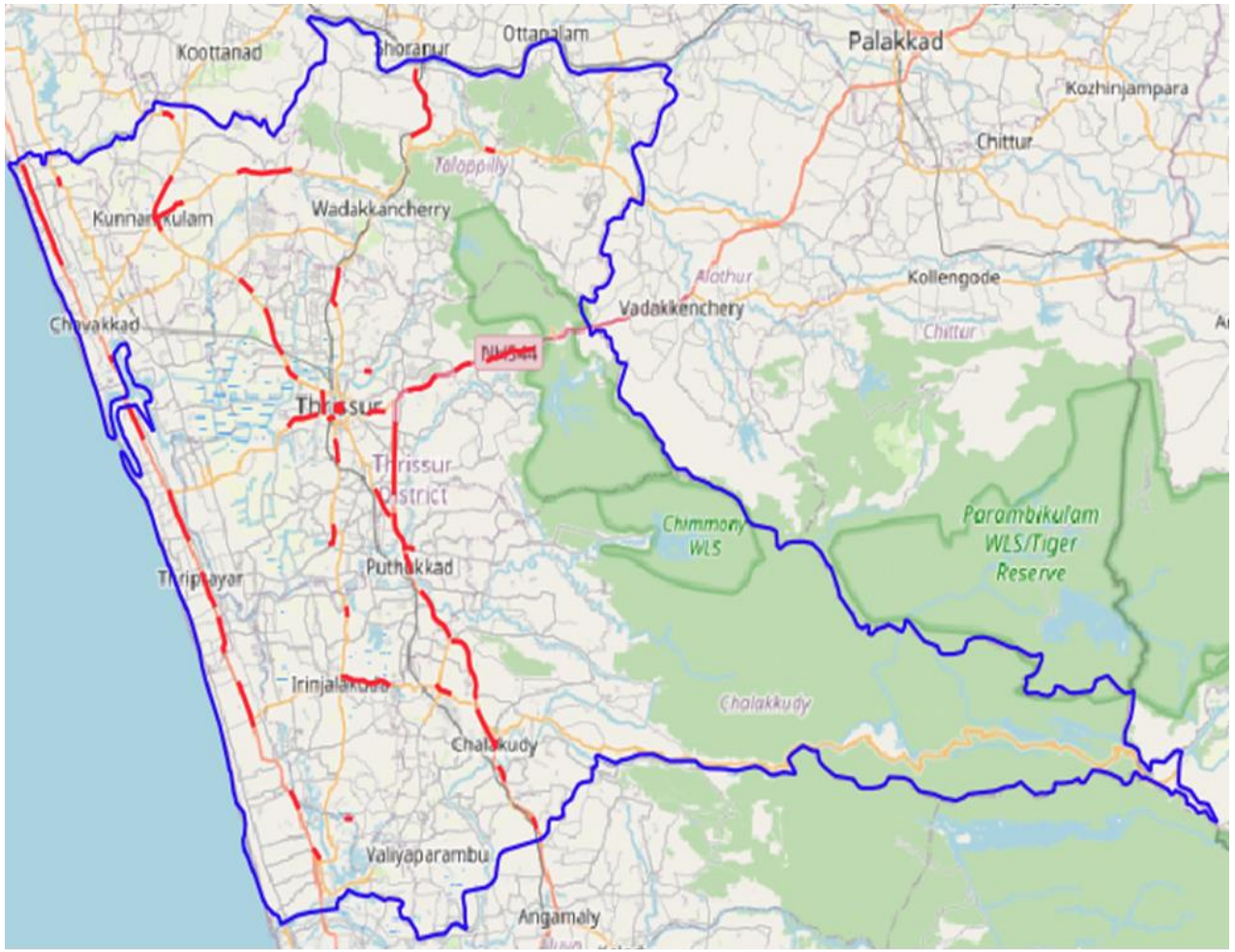


Imported crash coordinates in QGIS

DEMARCATION OF PEDESTRIAN VULNERABLE CORRIDORS







Buffer Generation

Pedestrian Vulnerable Corridors

PEDESTRIAN VULNERABLE CORRIDORS BASED ON ROAD TYPE

Total No. of Pedestrian Vulnerable Corridors: 57

Road Type	No. of Corridors	No. of Crashes	Corridor with Highest Crashes
NH	13	446	Thripayar – Valapad (79)
SH	30	708	Mapranam Center – Thellapilly (48)
MDR	14	326	East Fort Junction (37)

SUMMARY STATISTICS OF PARAMETERS

Parameters	Summary Statistics
Road Category	NH: 23%, SH: 58%, MDR: 19%
Number of Lanes	Single-lane: 16%, Two-lane: 58%, Four-lane: 16%, Six-lane: 11%
Right of Way (m)	5-45
Carriageway width (m)	3-20
Presence of Median	Available in 25% of corridors
Presence of Shoulder	Available in 85% of corridors
Presence of Sidewalk	Available in 25% of corridors
Presence of Crosswalk	Available in 42% of corridors
Presence of Speed Calming Measures	Available in 8% of corridors
Presence of Streetlight	Available in 68% of corridors
Severity	Fatal:14%, Grievous injury: 71%, Minor injury: 15%
Time of Day	Day: 55%, Night: 45%
Gender	Male: 71%, Female: 29%
Age Group	0-18yrs: 7%, 19-24yrs:3%, 25-50yrs: 29%, 51-75yrs: 54%, 75+yrs: 7%
Type of Vehicle	Two-wheelers: 58%, Cars: 23%, Three Wheeler: 7%, Bus: 5%, LCV: 5%, SCV: 1%, Others:1%
Crash Location	Bus stop: 26% ,Open area: 20%, Industrial area: 18% , Religious place: 10%, Residential area: 9%, Hospital zone: 7% , Institutional area: 5%, Others: 5%
Weather Condition	Sunny: 85%, Cloudy: 8%, Rainy:4%, Misty: 3%

DATA COLLECTION

Parameters collected from road inventory survey

PARAMETERS	DESCRIPTION
No. of Lanes *(Khanum H. et al., (2023), Mukherjee. D et al. (2020),etc)*	Single lane, Double lane, Four lane, Six lane
Median *(Meocci M. et al. (2021),R. Kraidi et al., (2020),etc)*	Present, Absent
Right of way *(Khanum H. et al., (2023), Mukherjee. D et al. (2020),etc)*	Total ROW of road section(m)
Carriageway *(Khanum H. et al., (2023), Mukherjee. D et al. (2020),etc)*	Carriageway width of road section (m)
Shoulder *(K. Haleem et al. (2017), Elalouf A. et al. (2023),etc)*	Present, Absent
Sidewalk *(K. Haleem et al. (2017), Elalouf A. et al. (2023),etc)*	Present, Absent
Crosswalk (Elalouf A. et al. (2023), Meocci M. et al. (2021), etc)	Present, Absent
Speed calming measures *(Ma. Y et al. (2020), Chakraborty. A et al. (2019),etc)*	Rumble strips, Speed humps, Both, Absent
Streetlight *(Elalouf et al. (2023), Riccardi et al. (2022))*	Present, Absent

Parameters extracted from crash data & FIR data

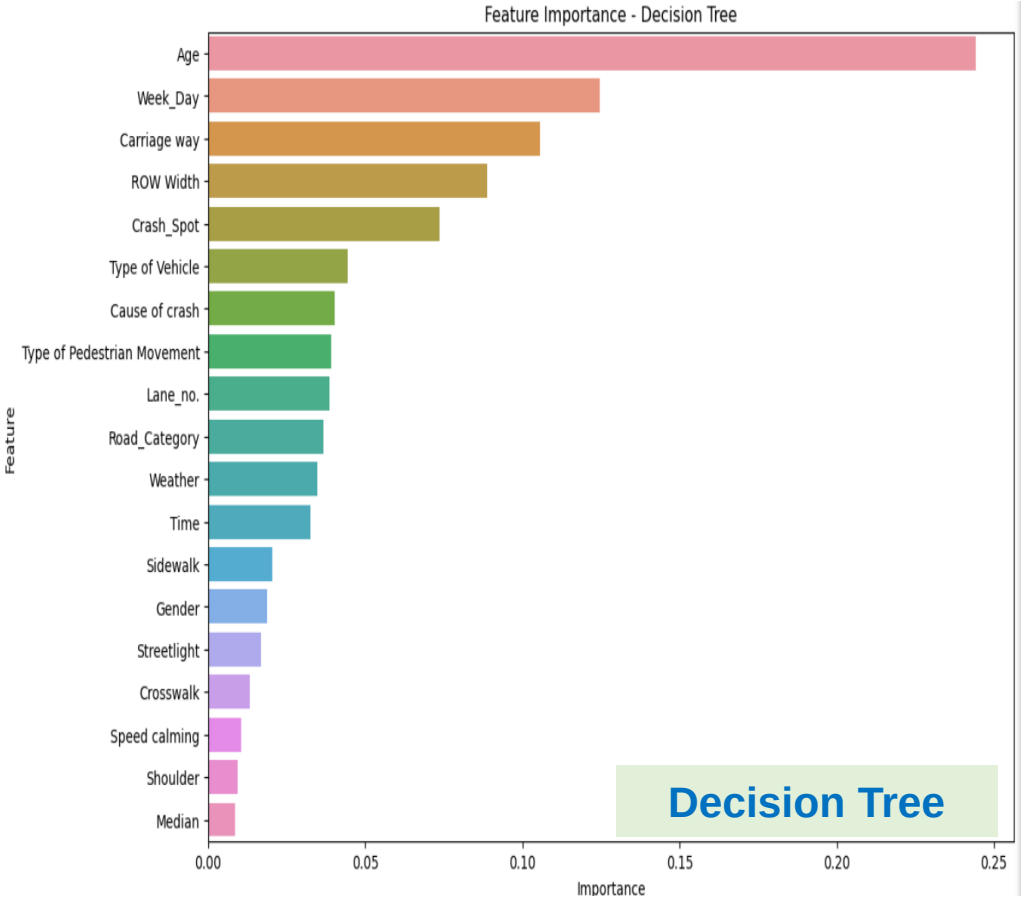
PARAMETERS	DESCRIPTION
Month *(Tao W. et al. (2022), Komol R. M. M.(2021), etc)*	January to December
Day of Week *(Tao W. et al. (2022), Khanum H. et al., (2023), etc)*	Monday to Sunday
Time of Day *(Meocci M. et al. (2021), Khanum H. et al., (2023),etc)*	Day, Night
Crash Location *(Khanum H. et al., (2023), Komol R. M. M.(2021, etc)*	Residential area, Bus stop, Hospital zone, Industrial area
Gender *(Das S. et al. (2021), Meocci M. et al. (2021), etc)*	Male, Female
Age (*Meocci M. et al. (2021),Tao W. et al.(2022), etc)*	<18, 19-24, 25-50, 51-75, >75
Area Type *(Meocci M. et al. (2021), Komol R. M. M.(2021), etc)*	Urban, Rural

PARAMETERS	DESCRIPTION
Speed Limit (kmph)*(Tao W. et al. (2022), Yang L. Et al. (2022), etc)*	<40, 40-60, 60-80, >80, Absent
Weather Condition (*Khanum H. et al., (2023), Komol R. M. M.(2021), etc)*	Cloudy, Rainy, Misty, Sunny
Road Category *(Meocci M. et al. (2021), Yang L. Et al. (2022), etc)*	NH,SH,MDR
Vehicle Type *(Khanum H. et al., (2023), Yang L. Et al. (2022), etc)*	Two Wheeler, Car, Bus, LCV ,SCV, Three Wheeler, Others
Type of pedestrian movement *(Meocci M. et al. (2021), Komol R. M. M.(2021), etc)*	Along the road, Across the road
Cause of crash *(Komol R. M. M.(2021))*	Over speeding, Alcohol/Drugs, Rash & Negligence
Crash Severity *(Komol R. M. M.(2021), Elalouf et al. (2023), etc)*	Fatal, Grievous, Minor

PEDESTRIAN CRASH SEVERITY PREDICTION MODELS

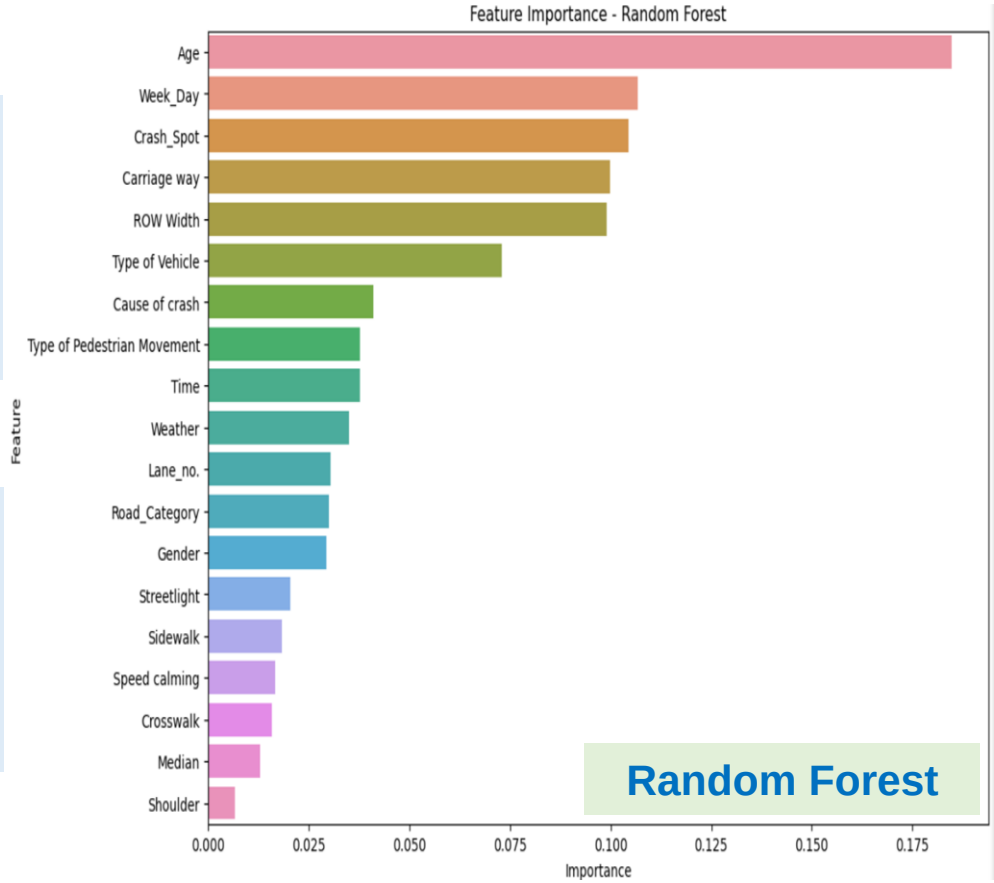
- Sample Size : 1480
- Development Environment: Jupyter Notebook
- Language used: Python
- Model Selection: Decision Tree, Random Forest, SVM, XGBoost
- Feature Importance Analysis
- Validation and Prediction

FEATURE IMPORTANCE ANALYSIS



Decision Tree:
Age, day, carriageway width, ROW, location, vehicle type, cause of crash

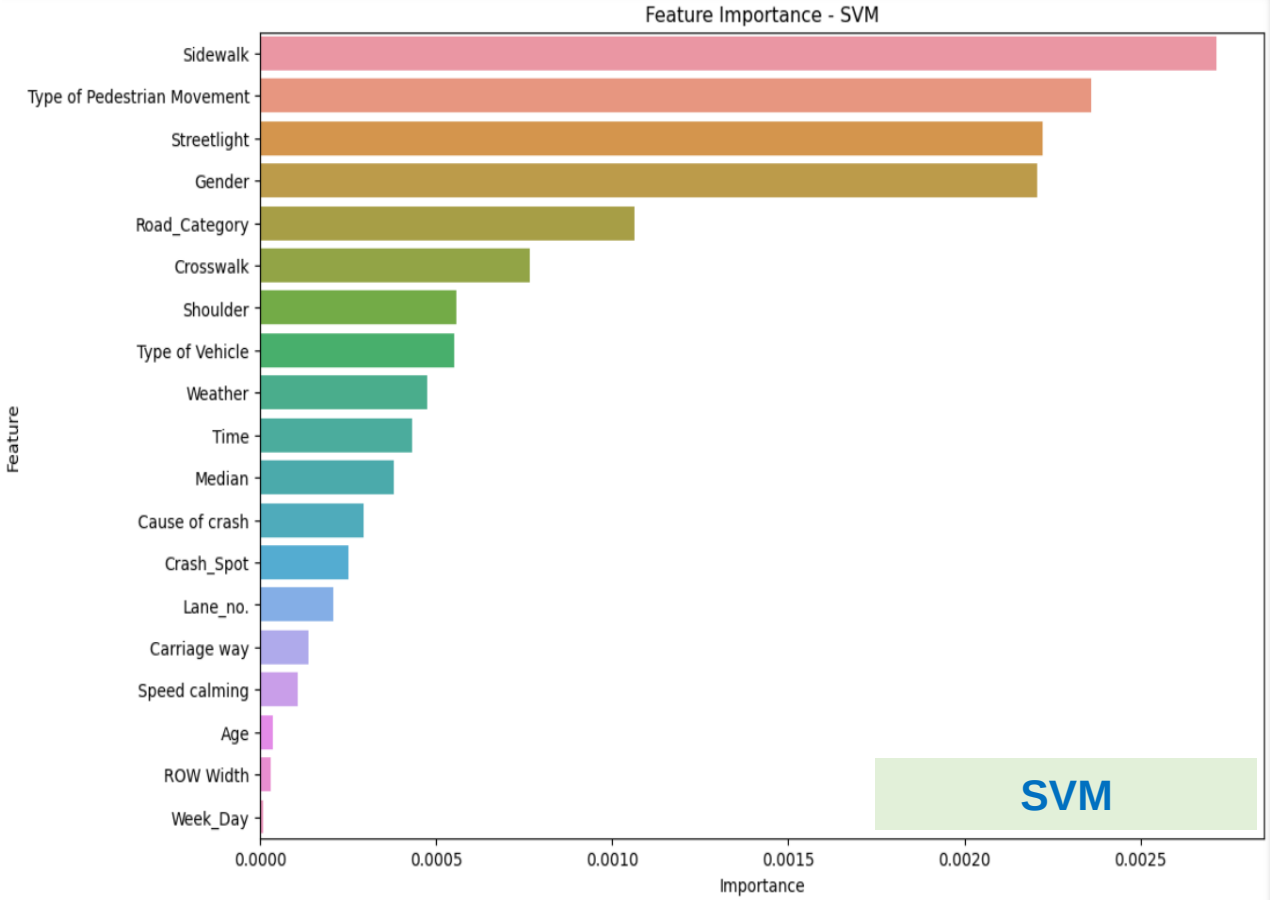
Random Forest:
Age, day, location, carriageway width, ROW, vehicle type, cause of crash

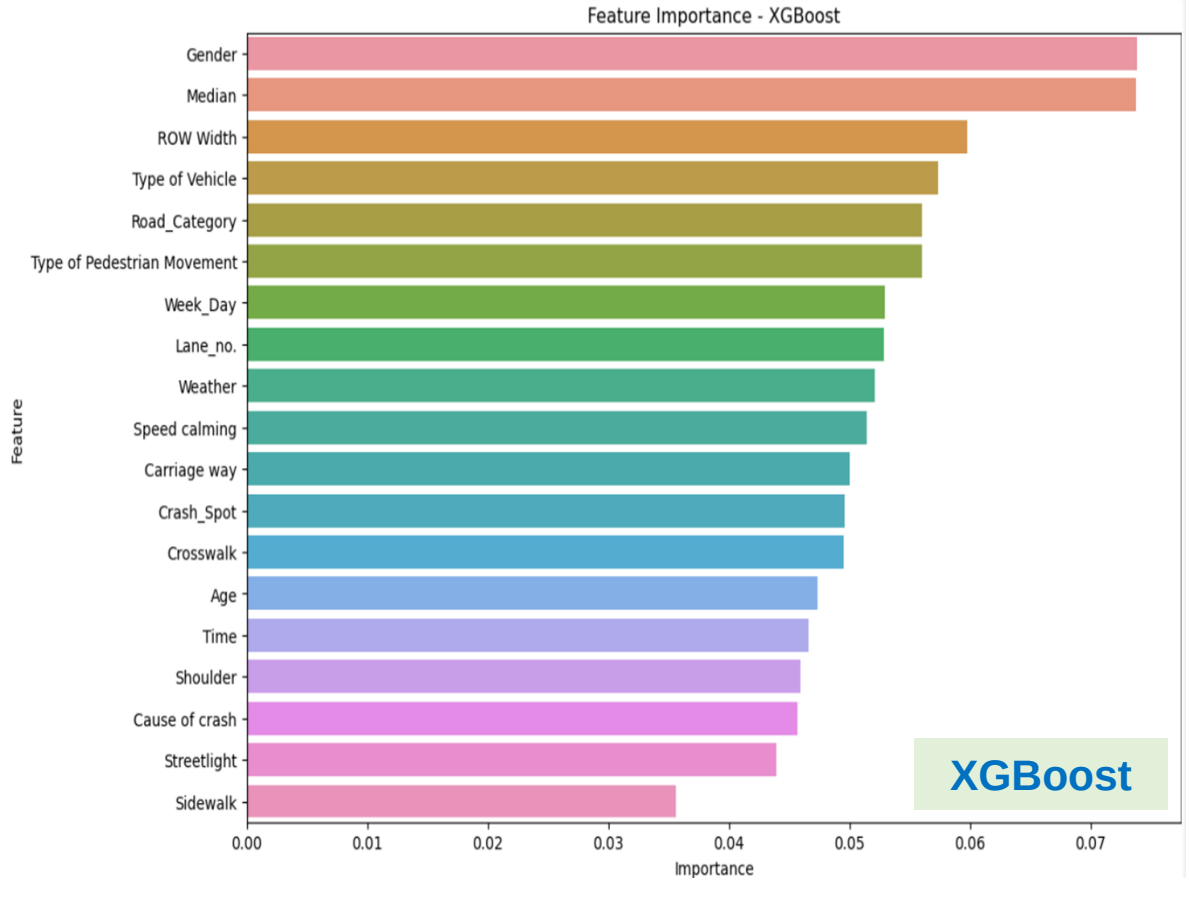


FEATURE IMPORTANCE ANALYSIS

SVM:
Sidewalks, type of pedestrian movement, streetlight, gender, road type, crosswalk, shoulder, vehicle type, weather condition, crash occurrence time

XGBoost:
Gender, median, ROW, vehicle type, road type, type of pedestrian movement





PERFORMANCE ANALYSIS

- SVM demonstrated the **highest accuracy** (0.68) and F1 score (0.54),
- Decision Tree had a balanced performance, with the highest precision (0.55) and a strong F1 score (0.55)
- Random Forest, offered higher accuracy (0.64) than the Decision Tree
- XGBoost had moderate accuracy (0.61) and precision (0.50), delivered comparable results in recall and F1 score (0.54)
- Decision Tree performed well across all metrics but was slightly outperformed by the SVM in accuracy and recall

Model	Accuracy	Precision	Recall	F1 score
Decision Tree	0.56	0.55	0.56	0.55
Random Forest	0.64	0.47	0.64	0.53
SVM	0.68	0.46	0.68	0.54
XGBoost	0.61	0.50	0.61	0.54

Support Vector Machine was the most effective model, balancing accuracy, recall, and F1 score despite its lower precision

CONCLUSION

- 57 pedestrian vulnerable corridors were demarcated in Thrissur district, Kerala, using GIS analysis
  - 1,480 pedestrian crashes, with 205 fatal, 1052 grievous injury crashes, and 223 minor injury crashes
- Grievous injury accounted for 71% of crashes, mostly occurring during the day (55%) and in rural areas (71%)
- Two-wheelers were involved in 58% of crashes, with pedestrians aged 51-75 years being the most affected
- Pedestrian vulnerable corridors included National Highways (13), State Highways (30), and Major District Roads (14) and reported 446, 708, and 326 pedestrian crashes respectively
- 19 independent variables related to road infrastructure, environmental conditions, and pedestrian movement patterns were considered for modelling
- Age, gender, crosswalks, sidewalks, vehicle type, and pedestrian movement type are common significant parameters across models
- SVM highlighted sidewalks, type of pedestrian movement, streetlight, road type, crosswalk, shoulder, vehicle type, weather condition, crash occurrence time as significant factors through feature importance analysis
- SVM model showed the highest accuracy (0.68) and F1 score (0.54), in predicting crash severity

REFERENCES

- Chakraborty, D. Mukherjee, and S. Mitra, “Development of pedestrian crash prediction model for a developing country using artificial neural network,” *International Journal of Injury Control and Safety Promotion*, vol. 26, no. 3, pp. 283–293, Jul. 2019.
- D. Mukherjee and S. Mitra, “Modelling risk factors for fatal pedestrian crashes in Kolkata, India,” *International Journal of Injury Control and Safety Promotion*, vol. 27, no. 2, pp. 197–214, Feb. 2020.
- Elalouf, S. Birfir, and T. Rosenbloom, “Developing machine-learning-based models to diminish the severity of injuries sustained by pedestrians in road traffic incidents,” *Heliyon*, vol. 9, no. 11, p. e21371, Nov. 2023.
- G. Tiwari, “Progress in pedestrian safety research,” *International Journal of Injury Control and Safety Promotion*, vol. 27, no. 1, pp. 35–43, Jan. 2020.
- K. Haleem, P. Alluri, and A. Gan, “Analyzing pedestrian crash injury severity at signalized and non-signalized locations,” *Accident Analysis & Prevention*, vol. 81, pp. 14–23, Aug. 2015.
- K. Santos, J. P. Dias, and C. Amado, “A literature review of machine learning algorithms for crash injury severity prediction,” *Journal of Safety Research*, Dec. 2021.
- S. Rankavat and G. Tiwari, “Pedestrians perceptions for utilization of pedestrian facilities – Delhi, India,” *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 42, pp. 495–499, Oct. 2016.
- W. Tao *et al.*, “An Advanced Machine Learning Approach to Predicting Pedestrian Fatality Caused by Road Crashes: A Step toward Sustainable Pedestrian Safety,” *Sustainability*, vol. 14, no. 4, pp. 24–36, Jan. 2022.
- WHO. 2023. *Global status report on road safety 2023*.

THANK YOU!....