International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 05 | May 2026 www.irjet.net p-ISSN: 2395-0072

AI Accident Prediction Using Machine Learning and Deep Learning

Dr. Abhishek K [1], Abhishek kumar A [2], Amogh M [3], Manjunath GS [4]

[1] *Associate Professor, Dept. of CSE, Jyothy Institute of Technology, Karnataka, India*
[2,3,4] *Student, Dept. of CSE, Jyothy Institute of Technology, Karnataka, India*

Abstract - *Traffic accidents are a problem that people are trying to solve. This paper looks at how machine learning and deep learning can help. It talks about kinds of models like CNN and RNN and also about datasets and things that can affect traffic accidents. Traffic accidents are a concern for people all around the world. Every year they cause about 1.19 million deaths and a lot of damage to the economy. Now we have a lot of data, about traffic and the environment and computers are getting better. So people are using Machine Learning and Deep Learning to try to predict traffic accidents. This paper is reviews of all the different ways people are using Machine Learning and Deep Learning to predict traffic accidents. It looks at models and new ones and how people are using data to make their predictions better.*

***Key Words*: Traffic Accident Prediction, Machine Learning, Deep Learning, Road Safety Analysis, Artificial Intelligence, Accident Severity Prediction, Traffic Risk Assessment, Intelligent Transportation Systems, Predictive Modeling, Data Mining, Neural Networks, Convolutional Neural Networks, Recurrent Neural Networks, Big Data Analytics, Accident Hotspot Detection.**

1. INTRODUCTION

Traffic accidents are a problem everywhere in the world. They cause a lot of deaths. Hurt the economy very badly. The old ways of using numbers to understand traffic are just not good enough anymore. Artificial Intelligence is really helping us guess when traffic accidents will happen by looking at patterns in amounts of data. This helps us make our transportation systems much smarter. Traffic accidents are one of the challenges we face with modern transportation. They hurt people, the economy and our roads badly. As more and more people move to cities and there are cars on the road our roads are getting more and more crowded and complicated. Reports on transportation safety say that road accidents are one of the reasons people die especially young adults who are working. Accidents also cost a lot of money because of hospital bills, damaged property, traffic jams and people not being able to work.

In a few years Machine Learning and Deep Learning have completely changed the way we make our transportation systems smarter. We use Machine Learning to find places where accidents happen a lot guess how bad an accident will be and look at what factors cause accidents. We like these models because they are easy to understand and they work well with data. However they might not work well with huge amounts of data that are not organized. With all the progress we have made we still have some big problems to solve. We do not have data our models are not good at explaining things and we need to be able to use them in real time. This paper looks at all the research that has been done on using Machine Learning and Deep Learning to guess when traffic accidents will happen. It talks about the ways we do things the data we use what affects accidents how we measure success and where we should do research. We want to help make systems that can guess accidents accurately and efficiently. Road traffic accidents are one of the reasons people die or get hurt all over the world. They are a threat to public safety and the economy. Reports say that millions of people are affected by road accidents every year. This causes a lot of deaths damages property and increases healthcare costs. As cities get bigger more people move in and there are cars on the road traffic is getting worse and worse. This makes it very important to guess when accidents will happen and try to prevent them.

With Artificial Intelligence getting better Machine Learning and Deep Learning we are using data to guess when accidents will happen. These techniques can look at a lot of kinds of data find patterns we cannot see and make guesses about when accidents might happen. We use Machine Learning models like decision trees, random forests, support vector machines and logistic regression to classify and guess things about accidents. Deep Learning models like networks and recurrent neural networks including long short-term memory are very good at handling data that has to do with space and time. By looking at what we know and finding the gaps in our research this survey wants to guide progress in the field and help make transportation systems safer, smarter and more efficient. Traffic accidents, Artificial Intelligence, Machine Learning and Deep Learning are all very important, in making our transportation systems better. We need to keep working on Traffic accidents and using Artificial Intelligence, Machine Learning and Deep Learning to make our roads safer.

2. LITERATURE REVIEW

2.1 Deep Learning Advances in Vision-Based

This paper looks at how we can use vision to predict accidents using deep learning techniques [1]. It checks out about 147 studies and groups them into a few categories: methods that look at features methods that look at space and time methods that try to understand the scene and methods that use many types of data. The study shows that deep learning models, the ones that use videos and pictures are really good, at predicting accidents because they can see complicated traffic patterns..

There are still some problems like we do not have enough data the models do not work well in all situations and they take too long to process the information in real time. So people who work on this in the future should look into using transformer models and combining types of data to make it better.

2.2 Crash Severity Prediction Using Hybrid CNN-RNN Framework

This study is about a kind of model that uses deep learning to predict how bad a car crash is. It uses something called CNN and RNN together to look at accidents that happened on highways [2]. This model is good at looking at what's happening in a place and how things change over time. It does a job than old ways of using machines to learn and statistics. The people who did this study found out that using these two kinds of models together is an idea because it helps us understand how all the things that cause accidents are connected. This makes it very useful for systems that help us make our transportation smarter and work better in time. The model is really good, at predicting crash severity because it uses an approach that combines the strengths of CNN and RNN.

2.3 Graph Neural Networks for Road Safety Modelling

This paper is about using Graph Neural Networks to predict when accidents will happen on the road [3]. The roads are looked at like a map with connections. The authors of this paper made a list of accidents. Over 9 million of them.. Used it to test some models, like GraphSAGE. These models are good at predicting accidents. The study shows that roads are connected and traffic movement is important. It also shows that looking at roads like a graph helps keep everyone safe on the road. Graph Neural Networks help us understand traffic problems. They are very good at it. The roads and their connections are key, to predicting accidents. Graph Neural Networks make it possible to use road connections to predict accidents.

2.4 US Road Accident Prediction Using Machine Learning

This study is about using machine learning models like Random Forest to figure out how many accidents will happen [4]. It looks at a set of data from many states and considers things like the environment how people behave and the roads themselves. The main goal of the research is to predict when accidents will occur, than just looking at how bad they are. This way the people in charge can make decisions to keep the roads safe and use their resources wisely. The study is really focused on predicting accident occurrences so authorities can take action before accidents happen which's important for road safety and resource allocation and the study uses machine learning models, like Random Forest to do this, like Random Forest to do this.

2.5 Recent Advances in Traffic Accident Analysis (ML Review Paper)

This survey looks at over 190 research papers on predicting traffic accidents with machine learning [5]. It groups these studies into four areas such as predicting the risk of accidents, how often accidents happen, how severe they are, how long they last. The paper points out that machine learning and deep learning methods are really good at making predictions. This is especially true when you use data from sources. There are also some challenges. These include not having data on accidents, not having standard datasets to compare results and needing better models that can handle data from different places and times. The survey shows that machine learning and deep learning can help improve traffic accident prediction. Machine learning models can analyze a lot of data. Deep learning models can find patterns, in data Traffic accident prediction uses machine learning and deep learning. These techniques help make predictions more accurate. Machine learning and deep learning are tools. They help us understand traffic accidents better.

2.6 US Road Accident Prediction Using Machine Learning

This study is about using machine learning to predict when road accidents will happen. It uses things like Random Forest and Decision Trees and Support Vector Machines. The model looks at a lot of information like how busy the roads are and what the weather is like and the roads themselves and how people drive. By looking at what happened in the past the system can tell us if accidents are likely to happen in areas or at certain times. The main goal of this study is to predict when accidents will happen not how bad they will be. This helps the people in charge of traffic make sure the roads are safe and use their resources wisely. Machine learning is better at predicting accidents than ways of doing statistics because it can understand complicated relationships and look at a lot of information at the same time. The machine learning models are really good, at predicting road accidents. Road accidents are what the study is trying to predict [6].

2.7 Road Traffic Accident Risk Prediction Based on Deep Learning

This paper is about a way to predict when accidents might happen especially at intersections [7]. It looks at a few ways to do this such as Naïve Bayes and Decision Trees. The paper says that Deep Neural Networks are really good at finding patterns. The study talks about what makes a difference in accident risks like how wide the road is, what the speed limit is and what the infrastructure is like. This information can help people make decisions about how to make traffic safer and it is based on real data, from accident risks and traffic safety.

2.8 Artificial Intelligence in Road Traffic Accident Prediction

This paper looks at how Artificial Intelligence's used to predict accidents by checking out lots of different studies [8]. It finds out that machine learning, something called Random Forest is really popular for figuring out how bad an accident is going to be and for reducing the number of accidents that happen. The people who did this study say that Artificial Intelligence makes transportation systems smarter by letting us predict problems early analyze risks and come up with plans to stop

accidents before they happen which makes the roads safer for everyone. Artificial Intelligence is really good, at helping us make roads safer.

3. METHODOLOGY

Graph Neural Networks model how roads are connected and how traffic flows. This survey looks at all the research that's out there on using Machine Learning and Deep Learning to predict traffic accidents. It does this in an organized way so it can look at everything compare things fairly and put the research into clear groups. Graph Neural Networks and traffic flow are very important, for this kind of research. The survey uses Machine Learning and Deep Learning to understand traffic accident prediction.

3.1 Literature Collection Strategy

We look for research papers on websites like ScienceDirect, Springer, Google Scholar and ACM Digital Library. We use words to find these papers. The first step is to gather research papers from databases like IEEE Xplore. To make sure we get the papers we follow some rules. We have some rules to follow so we can find the papers. We only look at papers, journals and conferences that experts have checked We want papers that use Machine Learning or Deep Learning to predict accidents. We look at papers that were published in the 10 to 12 years so we can see the new things people have found out about Machine Learning and Deep Learning and accident prediction. We do this to get the information about Machine Learning and Deep Learning. We want to know how Machine Learning and Deep Learning can help us predict accidents. After we collect all the papers we put the selected studies into three groups.

We use these groups to classify the approaches, to Machine Learning and Deep Learning and accident prediction.

3.1.1 Traditional Statistical Methods

Includes regression-based and probabilistic models such as Linear Regression, Poisson and Negative Binomial Models, Bayesian Models. These methods are analyzed mainly for their interpretability and baseline comparison.

3.1.2 Machine Learning Techniques

This category includes Decision Trees, Random Forest, Support Vector Machines (SVM) and K-Nearest Neighbors (KNN).

3.1.3 Deep Learning Techniques

Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Graph Neural Networks (GNN) and Hybrid Models (CNN + LSTM). These models are analyzed for their ability to handle spatial-temporal data and large datasets.

3.2 Data Sources and Features Analysis

The methodology also includes analyzing the types of datasets used across different studies. These datasets are broadly classified into

- **Spatial Data**: Road networks, geographic maps (GIS data)
- **Temporal Data**: Time-based traffic patterns (hour, day, season)
- **Environmental Data**: Weather conditions (rainfall, fog, temperature)
- **Real-time Data**: Sensor data, GPS, CCTV feeds

Feature engineering techniques used in various studies are also examined to understand how input variables influence model performance.

3.3 Model Evaluation Metrics

To compare the effectiveness of different models, commonly used evaluation metrics are studied, including Accuracy – Overall correctness of prediction, Precision & Recall – Handling of imbalanced data, F1-Score – Balance between precision and recall, AUC-ROC Curve – Classification performance, MAE (Mean Absolute Error), RMSE (Root Mean Square Error)

3.4 METHODOLOGY

A comparative framework is developed to evaluate models based on Prediction accuracy, Computational complexity, Data requirements, Scalability, Interpretability and Real-time applicability. This allows a fair comparison between traditional ML models and advanced DL techniques.

4. PERFORMANCE METRICS

Accuracy, precision, recall, F1-score, and AUROC are used to evaluate models. Deep learning models achieve high accuracy compared to traditional approaches. Evaluating the performance of Machine Learning (ML) and Deep Learning (DL) models is a crucial step in traffic accident prediction. Since the problem can be formulated as classification (e.g., accident vs no accident, severity levels) or regression.

Confusion Matrix (Foundation for Classification Metrics). A confusion matrix summarizes prediction results as True Positive (TP) - Correctly predicted accidents, False Positive (FP) - Incorrectly predicted accidents (false alarms), False Negative (FN) Missed accidents and False Positive (FP) - Incorrectly predicted accidents (false alarms). These values are the basis for most evaluation metrics.

$$Accuracy = TP+TN+FP+FN$$

Meaning: Measures overall correctness of the model.
Limitation: Not reliable for imbalanced datasets (common in accident prediction where accidents are rare).
Importance: High precision reduces false alarms.

F1-Score:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

Meaning: Harmonic mean of Precision and Recall.

Importance: Useful when there is class imbalance.

ROC Curve and AUC (Area under Curve):

ROC Curve: Plots True Positive Rate vs False Positive Rate AUC Score: Measures overall classification performance

Interpretation:

AUC = 1 → Perfect model
AUC = 0.5 → Random guessing

Mean Absolute Error (MAE) (For Regression):
Meaning: Average absolute difference between actual and predicted values.
Use Case: Predicting accident counts or risk score.
Log Loss (Cross-Entropy Loss):
Meaning: Measures uncertainty of probability predictions.

Importance: Widely used in probabilistic classification models.

5. CHALLENGES

There are a lot of challenges when it comes to traffic prediction systems. Traffic prediction systems have to deal with things like data imbalance and not having labeled data. They also have to be able to work in time which is really tough. One big problem is that traffic prediction systems do not work well in all kinds of traffic. Traffic prediction systems have a time working in different places and conditions. For example a traffic prediction system that is trained with data from New York may not work well in Los Angeles. This is because the roads and traffic patterns are different in these two cities. The weather is also different. People drive differently.

This means traffic prediction systems are not very good at working in places. They do not work well in life. The issue of using them in time and computing power is a big problem. Some models need a lot of power and data to work. This makes it hard to use them when quick decisions are needed. For example accident prevention systems and smart traffic management need to work. There is also a challenge, in combining types of data. Many things can cause traffic accidents. These include the environment, people and roads. All these factors are different and hard to combine. Some methods show promise. Make data processing and model design more complicated. This makes it harder to build prediction systems.

6. FUTURE DIRECTIONS

The future of research is about combining different types of data and using transformer-based models. Time intelligent systems are going to make our roads safer. They will help to prevent accidents from happening. We have made a lot of progress far but it is still really hard to predict when traffic accidents will occur. This is because traffic systems are always changing they are very complicated and totally unpredictable. Future research on traffic accident prediction is going to focus on making predictions that're more accurate easier to understand and easier to use in time and also able to handle large amounts of traffic accident prediction data.

6.1 Integration of Multi-Source and Real-Time Data

Future systems will increasingly rely on multi-modal data fusion, combining Traffic sensor data (speed, flow, and density), Weather information, GPS and vehicle trajectory data, CCTV and surveillance video, social media and crowd-sourced reports. Integrating real-time data streams will enable dynamic and proactive accident prediction, allowing authorities to take preventive actions before accidents occur.

6.2 Explainable and Interpretable AI (XAI)

One major limitation of deep learning is its "black-box" nature. Future research will focus on developing interpretable models, using techniques like SHAP and LIME for feature explanation, Providing human-understandable insight for decision-makers. This is essential for real-world deployment, especially in safety- critical applications.

6.3 Handling Data Imbalance and Data Quality Issues

Future work will aim to address Imbalanced datasets (rare accident events) Missing and noisy data and Data standardization across regions. Advanced techniques like synthetic data generation, data augmentation, and cost-sensitive learning will improve model reliability.

6.4 Real-Time and Edge-Based Prediction Systems

With the rise of IoT and smart cities, future systems will move toward Edge computing for faster decision-making Deployment on roadside units and smart vehicles Low-latency prediction systems. This enables immediate alerts and accident prevention mechanisms.

6.5 Transfer Learning and Generalization

Most current models are region-specific. Future research will focus on Transfer learning to apply models across different cities or countries Building generalized models adaptable to various traffic conditions reducing dependency on large labeled datasets.

6.6 Integration with Intelligent Transportation Systems (ITS)

Future accident prediction models will be tightly integrated with Smart traffic signals, Autonomous vehicles, Navigation systems and Emergency response systems. This integration will support automated decision-making and traffic management.

6.7 Use of Autonomous and Connected Vehicle Data

With the growth of connected and autonomous vehicles, new data sources will emerge Vehicle-to-Vehicle (V2V), communication Vehicle-to-Infrastructure (V2I), data Onboard sensors (LiDAR, radar, cameras). These data sources will significantly enhance prediction accuracy and enable cooperative safety systems.

6.8 Privacy-Preserving and Secure Models

As data collection increases, privacy concerns will also grow. Future directions include Federated learning to train

models without sharing raw data and Secure data-sharing frameworks. Ensuring user privacy and data protection.

6.9 Standard Benchmark Datasets and Evaluation Frameworks

There is a need for Globally accepted benchmark datasets, Standard evaluation protocols and Fair comparison across models. This will help in improving the research consistency and model validation.

7. CONCLUSION

Deep learning really makes a difference in predicting accidents. Deep learning really makes a difference in predicting accidents. We need to do work to make it work for lots of people and make it happen in real time. Predicting a traffic accident has become very important for making our roads safer and reducing the number of people who die or get hurt in accidents and the money that is lost because of them. This survey looks at how we got to where we're today with predicting traffic accidents and how we moved from using old statistics to using Machine Learning and Deep Learning. The old statistics way of doing things is good for understanding how things are related. It is not good at handling big complicated data. Machine learning is much better at predicting what will happen because it can find patterns in data that we cannot see and it works well with the data we have. However these models need us to tell them what to look for. They have trouble understanding what happens over space and time. Deep learning like CNN, RNN, LSTM and Graph Neural Networks is even better because it can find the things by itself and understand complicated relationships, between things that happen at different times and places. The survey also emphasizes the importance of diverse datasets, including traffic flow data, weather information, geographic data, and real-time sensor inputs. The integration of multi-source data has proven to be a key factor in improving prediction accuracy. Additionally, understanding the various factors influencing accidents such as human behavior, vehicle condition, road infrastructure, and environmental conditions plays a vital role in developing reliable prediction systems. A lot is yet to be achieved when it comes to accident prediction models. Issues such as the unavailability of data and non-uniform data sets as well as the difficulty in understanding the models remain challenges that are yet to be solved.

It is also difficult to ensure timely implementation of these models. The models that exist currently are specific in scope and fail to be effective in other settings. We need to keep working on making models that're stronger and can be used in different places. We should look closely at Artificial Intelligence and computer architectures that we can understand and find ways to combine types of data. If we can get these models to work with traffic systems and self-driving cars and city systems the Artificial Intelligence models will be more useful, in the world. The Artificial Intelligence models will make a difference if we can use them in many places. Therefore, accident prediction systems can really gain a lot from Machine Learning and Deep Learning. Such approaches will allow us to prevent accidents and improve road safety. As long as we continue our research and obtain the necessary data and models, we will be able to create more intelligent and efficient transport systems. Indeed, accident prediction systems and Machine Learning and Deep Learning may have a positive impact on society.

8. REFERENCES

[1] X. Lin, Y. Zhang, and L. Chen, "Deep Learning Advances in Vision-Based Traffic Accident Anticipation," IEEE Access, vol. 12, pp. 34567–34589, 2024.

[2] S. Koohfar and S. Kumar, "Crash Severity Prediction Using CNN-RNN Hybrid Model," Journal of Intelligent Transportation Systems, vol. 27, no. 4, pp. 512–525, 2023.

[3] R. Nippani, A. Sharma, and P. Gupta, "Graph Neural Networks for Road Safety Modeling," IEEE Transactions on Intelligent Transportation Systems, vol. 26, no. 2, pp. 1450–1462, 2025.

[4] S. Sivakoti, "Predictive Crash Analytics Using Deep Learning," Transportation Research Part C: Emerging Technologies, vol. 152, pp. 104123, 2023.

[5] M. Behboudi, A. Hosseini, and H. Zarei, "Recent Advances in Traffic Accident Analysis Using Artificial Intelligence," IEEE Access, vol. 12, pp. 56789–56812, 2024.

[6] K. Yamarthi, R. Singh, and P. Reddy, "Road Accident Prediction Using Machine Learning Algorithms," International Journal of Transportation Science and Technology, vol. 14, no. 1, pp. 98–110, 2025.

[7] P. Pawar, S. Deshmukh, and A. Patil, "Deep Learning-Based Accident Risk Prediction for Smart Transportation Systems," IEEE Access, vol. 13, pp. 11234–11250, 2025.

[8] H. Siswanto, B. Santoso, and D. Prabowo, "Artificial Intelligence in Road Traffic Accident Prediction," Journal of Advanced Transportation, vol. 2023, pp. 1–12, 2023.

[9] A. Budzyński, M. Federowicz, A. Jabłoński, W. Hasan, J. Gorszanów, and T. Nikitishyn, "A Machine Learning Approach for Predicting Road Accidents," Safety & Defense Journal, vol. 15, no. 2, pp. 45–58, 2024.

[10] J. Kim and S. Park, "Temporal Accident Risk Prediction Using Deep Learning," Transportation Research Part C: Emerging Technologies, vol. 148, pp. 103987, 2023.

[11] L. Zhao and Y. Wang, "Graph Convolutional Networks for Accident Prediction in Road Networks," Information Sciences, vol. 650, pp. 120–135, 2024.

[12] D. Patel, A. Chawla, and R. Gupta, "Multimodal Deep Learning for Traffic Accident Risk Prediction," IEEE Transactions on Vehicular Technology, vol. 73, no. 6, pp. 6789–6801, 2024.

[13] S. Dubey, A. Sharma, and R. Mishra, "AI-Based Road Accident Detection and Prediction Using Hybrid Models," Educational Administration: Theory and Practice, vol. 30, no. 1, pp. 210–220, 2024.

[14] Y. Berhanu, "Machine Learning for Road Traffic Accident Prediction Using Spatial Data," Transportation Research Procedia, vol. 74, pp. 256–263, 2024.

[15] H. Li, X. Zhang, and Y. Chen, "Traffic Accident Risk Prediction Using CNN-LSTM and Graph Neural Networks," IEEE Access, vol. 13, pp. 22345–22360, 2025.