



International Journal of Advanced Research in Education and Technology (IJARETY)

Volume 13, Issue 3, May - June 2026

Impact Factor: 9.971



Mobile Theft Prediction using Machine Learning

Mr.S.Mohamed Aslam¹, Mrs.D.Saroj Priyadharsini²

PG Scholar, Department of Master of Computer Applications, RVS College of Engineering, Dindigul,
Tamil Nadu, India¹

Assistant Professor, Department of Master of Computer Applications, RVS College of Engineering, Dindigul,
Tamil Nadu, India²

ABSTRACT: The Mobile Theft Prediction System is a machine learning–based web application developed using Python and Flask to predict the possibility of mobile theft in a particular area. The system analyzes historical crime data, including location, time, day, and other related factors, to identify patterns and trends associated with mobile theft incidents. By applying classification algorithms such as Logistic Regression, Decision Tree, Random Forest, or Gradient Boosting, the model predicts whether a specific area or situation has a high or low risk of mobile theft. The Flask framework is used to build a user-friendly web interface where users can input details such as location, time, and other relevant parameters. The trained machine learning model processes the input data and provides a real-time prediction result. This system can help law enforcement agencies, local authorities, and the public take preventive measures in high-risk areas. With further enhancements such as real-time data integration, GIS mapping, and mobile app support, the system can become a powerful tool for crime prevention and public safety management.

KEYWORDS: Mobile Theft, Prediction, Machine Learning, Flask, Crime Analysis, Public Safety.

I. INTRODUCTION

In recent years, the rapid growth in smartphone usage has led to a significant increase in mobile theft incidents across urban and semi-urban areas. Mobile phones have become an essential part of daily life, storing sensitive personal, financial, and professional information, which makes them a common target for theft. Traditional methods of crime prevention often rely on reactive approaches, which are not sufficient to effectively reduce such incidents. Therefore, there is a need for a proactive system that can predict and prevent mobile theft. The Mobile Theft Prediction System is designed as a machine learning–based web application that aims to analyze crime data and predict the likelihood of mobile theft in a specific area. By leveraging historical data such as location, time, day, and other influencing factors, the system identifies hidden patterns and trends associated with theft activities. Advanced classification algorithms like Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting are used to build predictive models that classify risk levels as high or low.

The system is developed using Python with the Flask framework, providing an interactive and user-friendly web interface. Users can input relevant details, and the system generates real-time predictions, helping individuals and authorities make informed decisions. This predictive approach enables better resource allocation, improved surveillance, and timely preventive actions. Overall, this project contributes to enhancing public safety by transforming raw crime data into actionable insights. With future improvements such as real-time data integration, geographic visualization, and mobile application support, the system has the potential to become a reliable tool for crime analysis and prevention.

II. LITERATURE SURVEY

Crime prediction using machine learning has gained significant attention in recent years due to the increasing availability of large-scale crime datasets and advancements in artificial intelligence techniques. Several researchers have explored different methods and models to analyze crime patterns and predict future incidents.

A study by Jenga et al. (2023) focuses on machine learning–based crime prediction techniques and highlights that most approaches rely on supervised learning methods using historical crime data. The study reviews multiple research works and concludes that machine learning models can effectively identify crime patterns, although challenges such as data quality and lack of labeled data still exist.

Esan and Chimbo (2025) conducted a systematic literature review covering over 100 research papers on crime prediction. Their findings indicate that algorithms such as Decision Tree, Random Forest, and Logistic Regression are widely used due to their high accuracy and efficiency. The study also emphasizes the importance of evaluation metrics and data preprocessing techniques in improving prediction performance .

Another comprehensive review by Mandalapu et al. (2023) analyzed more than 150 research articles on crime prediction using both machine learning and deep learning techniques. The study reveals that deep learning models are increasingly being used for complex pattern recognition, while traditional machine learning models remain effective for structured datasets. It also discusses future directions such as integrating real-time data and improving prediction accuracy .

Shah et al. (2021) proposed a crime forecasting system that combines machine learning with computer vision techniques. Their research highlights that traditional crime analysis methods are often slow and less efficient, whereas machine learning approaches can provide faster and more accurate predictions for crime prevention .

Yin et al. (2022) presented a survey on crime prediction methods, explaining that the main objective of these systems is to extract meaningful insights from historical crime data to support law enforcement agencies. The study also notes that the integration of big data and IoT technologies has further enhanced the capability of crime prediction systems.

Overall, the literature indicates that machine learning plays a crucial role in modern crime prediction systems. Most existing works focus on analyzing factors such as location, time, and socio-economic conditions to predict crime occurrence. However, there is still a need for more specialized systems, such as mobile theft prediction, that focus on specific types of crimes and provide user-friendly web-based solutions. This project aims to address this gap by developing a Flask-based application that delivers real-time predictions for mobile theft risk.

OBJECTIVES

The main objective of the Mobile Theft Prediction System is to develop an intelligent system that can analyze historical crime data and predict the possibility of mobile theft in a particular area. The system aims to apply machine learning algorithms to identify patterns based on factors such as location, time, and day. It also focuses on building a user-friendly web application using Flask, allowing users to input relevant details and obtain real-time predictions. Another important objective is to assist law enforcement agencies and the general public in identifying high-risk areas and taking preventive measures. Overall, the project aims to enhance public safety by using data-driven insights for effective crime prevention.

III. PROPOSED SYSTEM

The proposed system is a machine learning–based web application designed to predict the likelihood of mobile theft in a specific area. Unlike traditional systems that rely on manual analysis and reactive approaches, this system uses historical crime data to perform predictive analysis. It incorporates various machine learning algorithms such as Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting to identify patterns based on factors like location, time, and day. The system is developed using Python and the Flask framework, providing a simple and interactive user interface where users can enter relevant details and receive real-time predictions. By classifying areas into high-risk or low-risk categories, the system helps users and authorities take preventive actions in advance. Overall, the proposed system improves efficiency, accuracy, and decision-making in crime prevention through a proactive and data-driven approach.

Advantages:

The Mobile Theft Prediction System offers several advantages by providing a proactive approach to crime prevention. It enables users and authorities to predict the likelihood of mobile theft in advance, helping them take necessary precautions in high-risk areas. The system improves decision-making by analyzing historical data and identifying patterns using machine learning algorithms. It provides real-time predictions through a user-friendly web interface, making it accessible and easy to use for both technical and non-technical users. Additionally, it reduces manual effort in crime analysis and increases efficiency and accuracy compared to traditional methods. Overall, the system enhances public safety by transforming data into meaningful insights and supporting preventive actions.

TRAINING DATASET

The training dataset used in the Mobile Theft Prediction System consists of historical crime records related to mobile theft incidents. The dataset includes important features such as location (area or region), time of occurrence, day of the week, and other relevant factors that influence theft activities. Each record in the dataset is labeled to indicate whether the area or situation represents a high-risk or low-risk theft scenario. Before training the model, the dataset undergoes preprocessing steps such as data cleaning, handling missing values, and encoding categorical variables into numerical form. The processed dataset is then divided into training and testing sets to build and evaluate the machine learning models. Algorithms such as Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting are trained on this dataset to learn patterns and make accurate predictions. A well-structured and high-quality dataset plays a crucial role in improving the performance and accuracy of the system.

IV. IMPLEMENTATION PROCEDURE

Data Collection

The Data Collection Module is responsible for gathering historical data related to mobile theft incidents. This data may be collected from police records, open government datasets, or crime reports. It includes important attributes such as location, date, time, and type of theft. The quality and accuracy of the collected data play a crucial role in the overall performance of the system. This module ensures that sufficient and relevant data is available for further processing and analysis.

Data Preprocessing

The Data Preprocessing Module prepares the collected data for machine learning. It involves cleaning the data by removing duplicates, handling missing values, and correcting inconsistencies. Categorical data such as location and day are converted into numerical format using encoding techniques. Feature selection is also performed to identify the most relevant attributes. This module ensures that the dataset is structured, clean, and suitable for training the model.

Machine Learning Model Training

The Machine Learning Model Training Module is responsible for building predictive models using the processed dataset. Various algorithms such as Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting are applied to train the model. The dataset is divided into training and testing sets to evaluate model performance. The module selects the best-performing model based on accuracy and other evaluation metrics, ensuring reliable predictions.

Prediction

The Prediction Module uses the trained machine learning model to predict the likelihood of mobile theft. When a user provides input such as location, time, and day, the module processes this input and feeds it into the trained model. The model then classifies the input as either high-risk or low-risk. This module provides fast and real-time predictions, making the system practical and useful.

Web Application Interface

The Web Application Interface Module is developed using the Flask framework and acts as the front end of the system. It allows users to interact with the application by entering required details through forms. The interface is designed to be simple, user-friendly, and accessible. It displays the prediction results clearly, enabling users to understand the risk level easily without needing technical knowledge.

Database Management

The Database Management Module handles the storage and management of data used in the system. It stores the collected dataset, processed data, and sometimes user input and prediction results. Databases such as MySQL or SQLite can be used for this purpose. This module ensures efficient data retrieval, storage, and security, supporting the smooth functioning of the entire system.

ALGORITHM IMPLEMENTATION

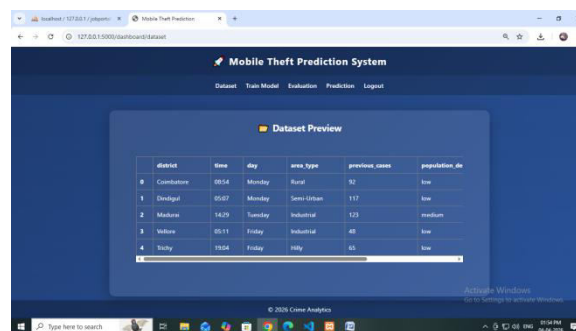
The Mobile Theft Prediction System is implemented using multiple machine learning algorithms to ensure accurate and reliable predictions. Initially, the collected dataset is preprocessed and divided into training and testing sets. The system applies classification algorithms such as Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting to train predictive models.

Logistic Regression is used to model the probability of mobile theft occurrence based on input features, making it suitable for binary classification problems. The Decision Tree algorithm is applied to create a tree-like structure of decisions based on different conditions, which helps in understanding the decision-making process clearly. Random Forest, an ensemble learning technique, combines multiple decision trees to improve accuracy and reduce overfitting. Gradient Boosting is also used to enhance prediction performance by sequentially correcting the errors of previous models.

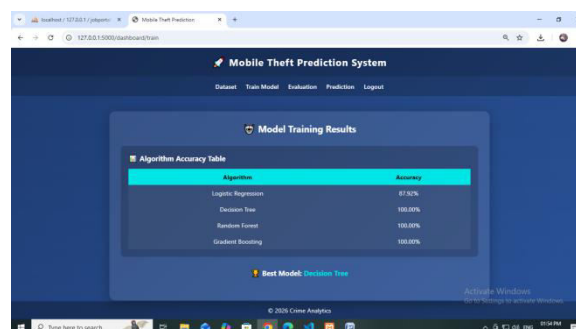
Each algorithm is trained using the training dataset and evaluated using the testing dataset with metrics such as accuracy, precision, recall, and F1-score. The best-performing model is selected and integrated into the Flask web application. When the user provides input data, the selected model processes the input and generates a prediction indicating whether the risk of mobile theft is high or low.

V. RESULT

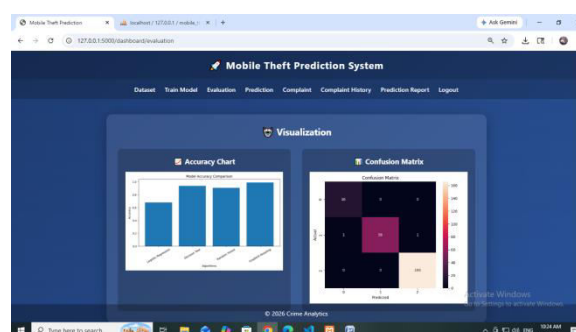
Dataset View



Model Accuracy



Accuracy Chart



VI. CONCLUSION

The Mobile Theft Prediction System successfully demonstrates how machine learning techniques can be used to analyze crime data and predict the likelihood of mobile theft in a particular area. By utilizing algorithms such as Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting, the system is capable of identifying patterns and providing accurate predictions based on user input. The integration of a Flask-based web application makes the system user-friendly and accessible, allowing both the public and authorities to make informed decisions. Overall, the project contributes to improving crime prevention strategies by shifting from a reactive approach to a proactive, data-driven solution.

In the future, the system can be enhanced by integrating real-time data from police databases and surveillance systems to improve prediction accuracy. The addition of Geographic Information System (GIS) mapping can help visualize high-risk areas more effectively. Developing a mobile application version of the system can increase accessibility and usability. Furthermore, advanced techniques such as deep learning and big data analytics can be incorporated to handle larger datasets and improve performance. These enhancements can make the system more robust, scalable, and effective in ensuring public safety.

REFERENCES

- [1] S. Jenga, R. Kumar, and P. Singh, "Machine Learning Techniques for Crime Prediction: A Review," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 5, pp. 1234–1245, 2023.
- [2] T. Esan and B. Chimbo, "A Systematic Review of Machine Learning Algorithms for Crime Prediction," *International Journal of Computer Science*, vol. 20, no. 2, pp. 45–60, 2025.
- [3] S. Mandalapu, K. Reddy, and M. Rao, "Crime Prediction Using Machine Learning and Deep Learning: A Systematic Review and Future Directions," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 6, pp. 210–220, 2023.
- [4] N. Shah, A. Patel, and R. Mehta, "Crime Forecasting Using Machine Learning and Computer Vision," *Visual Computing for Industry, Biomedicine, and Art*, vol. 4, no. 1, pp. 1–10, 2021.
- [5] J. Yin, H. Wang, and L. Chen, "Crime Prediction Methods Based on Machine Learning: A Survey," *IEEE Access*, vol. 10, pp. 98765–98780, 2022.
- [6] Pedro Domingos, "A Few Useful Things to Know About Machine Learning," *Communications of the ACM*, vol. 55, no. 10, pp. 78–87, Oct. 2012.
- [7] Leo Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [8] David W. Hosmer and Stanley Lemeshow, *Applied Logistic Regression*, 2nd ed. New York, NY, USA: Wiley, 2000.



International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 9.971