Landslide Prediction Using ML

Kamepalli Sriya Choudary

Department of Computer Science and Engineering

kschoudary.cs22@rvce.edu.in

R V College of Engineering, Bangalore - 560059

**Abstract:** Landslides represent a significant hazard to both human life and infrastructure, particularly in areas that are vulnerable to extreme weather and unstable terrain. Accurately forecasting landslide events is crucial for risk reduction and preparedness. This paper investigates the use of machine learning (ML) techniques to predict landslides by analyzing key environmental factors such as latitude, longitude, soil moisture, and temperature. Various datasets from diverse regions were utilized to train and test several ML algorithms, including support vector machines, decision trees, and neural networks. The study focused on preprocessing methods like feature scaling and normalization to optimize model performance and prevent overfitting. Results indicate that soil moisture and temperature are critical factors influencing landslide risks, while geographic coordinates provide important regional context. This research highlights the potential of ML models for landslide forecasting, offering new opportunities for early detection and improved disaster response strategies. Future work will explore integrating other variables like rainfall and vegetation data to further refine prediction accuracy.

**Keywords**— Machine Learning, Landslide prediction, Random Forest Classifier

I. INTRODUCTION.

Landslides are one of the most hazardous natural disasters, often leading to loss of life, displacement, and significant damage to infrastructure. Accurate prediction of landslides is crucial to mitigate these risks, and while traditional prediction methods are time-consuming and dependent on expert analysis, the integration of machine learning (ML) offers a more efficient approach. This project aims to develop a predictive system that uses real-time environmental parameters—longitude, latitude, temperature, humidity, and soil moisture—to detect potential landslides.

The selected parameters play a pivotal role in landslide prediction. **Longitude and latitude** provide the geographical context, helping to identify specific regions vulnerable to landslides based on their location. **Temperature** and **humidity** are critical indicators of climatic conditions that can affect soil stability. Changes in temperature can impact the moisture content in soil, which, when combined with high humidity, can increase the likelihood of landslides. **Soil moisture** is one of the most significant predictors, as saturated soil is more prone to sliding due to reduced cohesion and increased weight. These real-time environmental factors are fed into the ML model to determine areas at high risk of a landslide event.

To ensure the accuracy and reliability of the model, we employ several machine learning algorithms—such as decision trees, support vector machines, and random forests. These algorithms are trained using historical data that include landslide occurrences, and the model learns to associate the input parameters with landslide events. Once trained, the model’s accuracy is evaluated using a **confusion matrix**, which provides key metrics like determining how well the model predicts landslides and avoids false positives. Visual representations of the confusion matrix are generated, providing a clear graphical display of the model's performance, which helps in understanding the true positives, false positives, true negatives, and false negatives in predictions.

In addition to the machine learning model, this project includes the development of a **web application** that enables users to input real-time data—such as their geographic coordinates and environmental conditions—and receive immediate landslide risk predictions. It shows the landslide occurring based on the model’s output and presents graphical representations of the model’s accuracy metrics, such as confusion matrix charts.Through the integration of ML, real-time data, and a web platform, this project provides a powerful tool for landslide prediction. The ability to predict landslides with high accuracy could significantly enhance early warning systems, helping to mitigate the devastating effects of this natural disaster.

II. LITERATURE REVIEW

**1. Landslide Monitoring and Prediction System Using Geosensors and Wireless Sensor Networks:** The paper talks about developing a system that integrates geosensors and wireless sensor networks (WSNs) to monitor slope stability in real-time, using machine learning to predict landslide risks and trigger early warnings for disaster management.

**2. Application of Artificial Intelligence and Remote Sensing for Landslide Detection and Prediction: Systematic Review:**The paper talks about the review of the application of AI techniques combined with remote sensing data, such as satellite imagery and LiDAR, to enhance landslide detection and prediction, improving risk assessment and enabling timely early warning systems.

**3. Identification of Landslide Precursors for Early Warning of Hazards:**This paper extensively reviews advancements in Remote Sensing Technologies (RSTs) from 2010 to 2024.It highlights critical landslide precursors, such as slope instability, rainfall intensity, and soil movement, detected via remote sensing techniques.

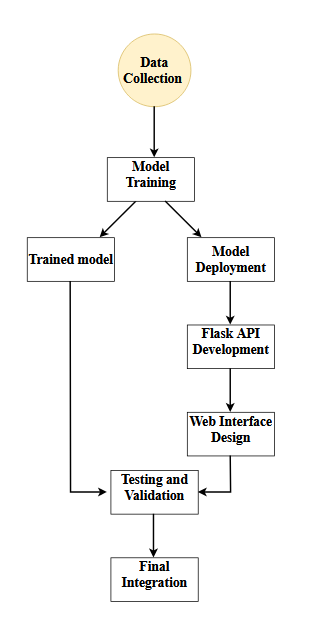
**4. Machine Learning Forecasting and Early Warning Models for Geological Hazards:**The paper explores various machine learning models to predict geological hazards, including landslides.It highlights the potential of supervised learning algorithms (eg: Random Forest) accurately forecasting landslide occurrences.Special attention is given to minimizing false positives and optimizing predictions for better reliability.

**5. Artificial Intelligence Techniques for Landslides Prediction Using Satellite Imagery:**

This article summarizes the performances of different classification techniques from recent literature followed by comparison and discussion with respect to accuracy. Based on the gap identified an effective prototype of the landslide classification approach is proposed. A slightly modified version of the deep learning model.

**6. Machine learning and artificial intelligence models development in rainfall-induced landslide prediction:**This paper aims to build landslide event prediction models using several AIML algorithms.The algorithms were trained with two different methods. The input of the algorithms was data obtained from the global satellite mapping of precipitation satellite observation.Each algorithm provided some model candidates with different parameter settings for each method. The best model was then chosen from each method.

III. METHODOLOGY



**Data Collection and Preprocessing**: The dataset includes sensor readings such as vibration, humidity, soil moisture, temperature, latitude, longitude.​

**Model Training**: The dataset is split into training and testing sets. A Random Forest Classifier is trained using the sensor features to predict the binary outcome.​

**Model Deployment**: The trained machine learning model is serialized and saved using pickle. Label encoders and scalers are also saved for consistent data preprocessing during inference​

**Flask API Development**: Flask API is created to handle requests from the web frontend.​

**Web Interface Design**: A user friendly web interface is developed using HTML, CSS, JavaScript​

**Testing and Validation**: The system is rigorously tested using the Postman tool for API requests and various frontend user inputs t ensure robustness and accuracy.​

**Final Integration**: The web interface and prediction system are integrated to provide a seamless and scalable landslide prediction solution.

IV. SOFTWARE REQUIREMENTS

**Programming Languages**:

TensorFlow or PyTorch for training models.

OpenCV for handling climate images.

Pandas and NumPy for data manipulation.

Matplotlib/Seaborn for visualizing climatic trends.

**Database**:

PostgreSQL or MongoDB for storing geospatial data and rainfall/climate statistics.

**APIs**:

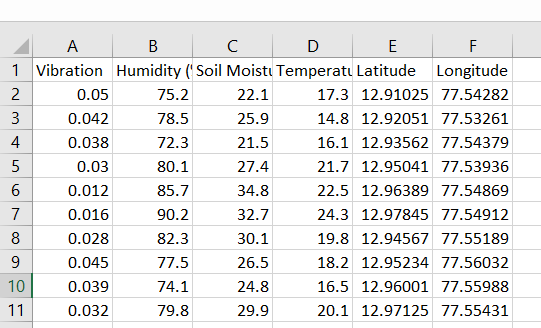
Google Maps or OpenStreetMap APIs for geospatial visualization.

Weather APIs (like OpenWeatherMap) for real-time climatic data.

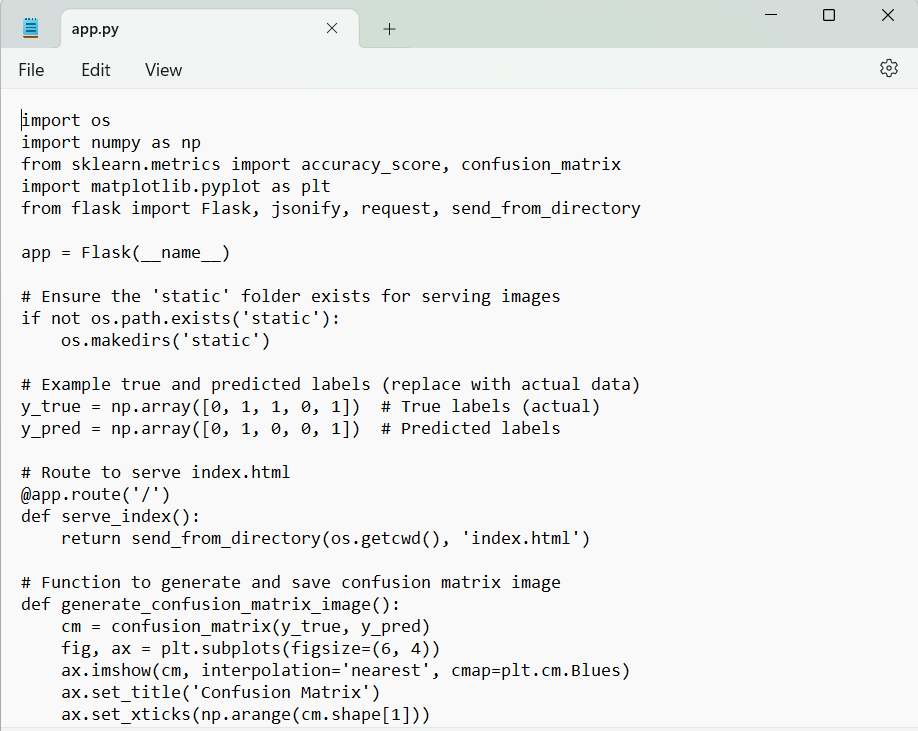
**Development Tools**:

Integrated Development Environments (IDEs) like VSCode.

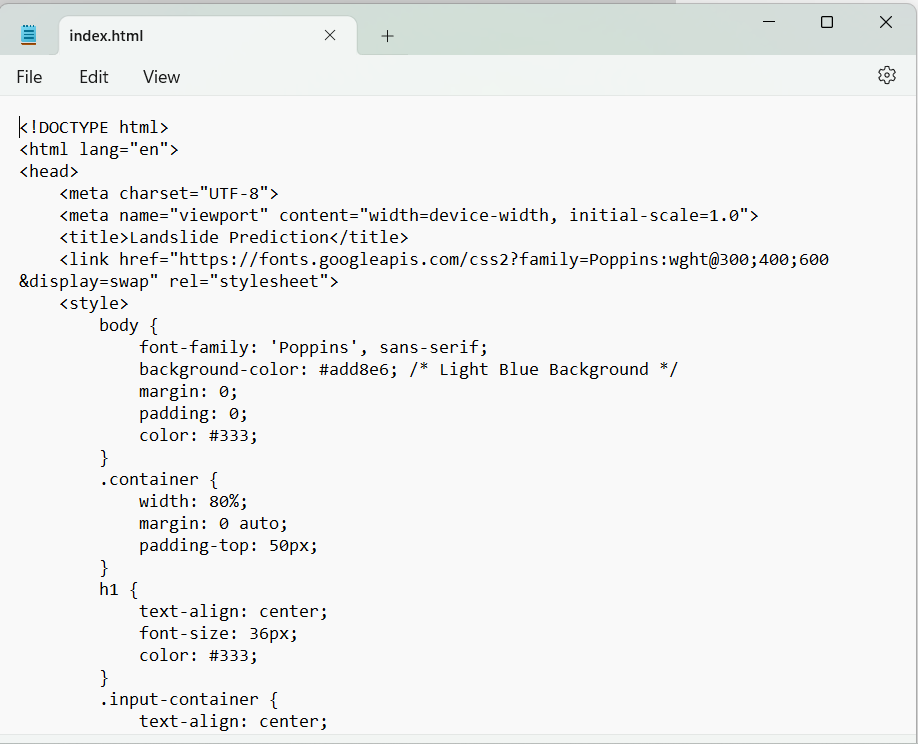
V. DATASETS



VI. CODE

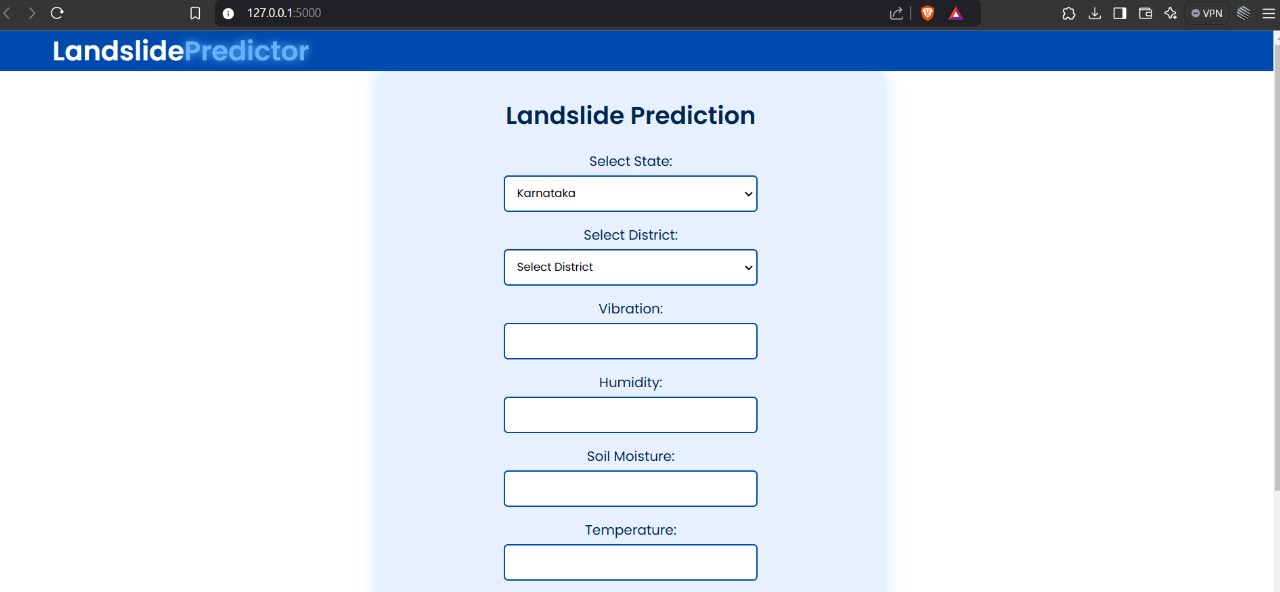


In this project, the web application is built using Flask for backend processing, while the frontend is designed with HTML and JavaScript. The backend (app.py) serves as the controller, handling user inputs and interacting with the machine learning model to generate predictions based on environmental parameters like longitude, latitude, temperature, and soil moisture. The index.html file serves as the user interface, where users can input real-time data and receive landslide risk predictions.

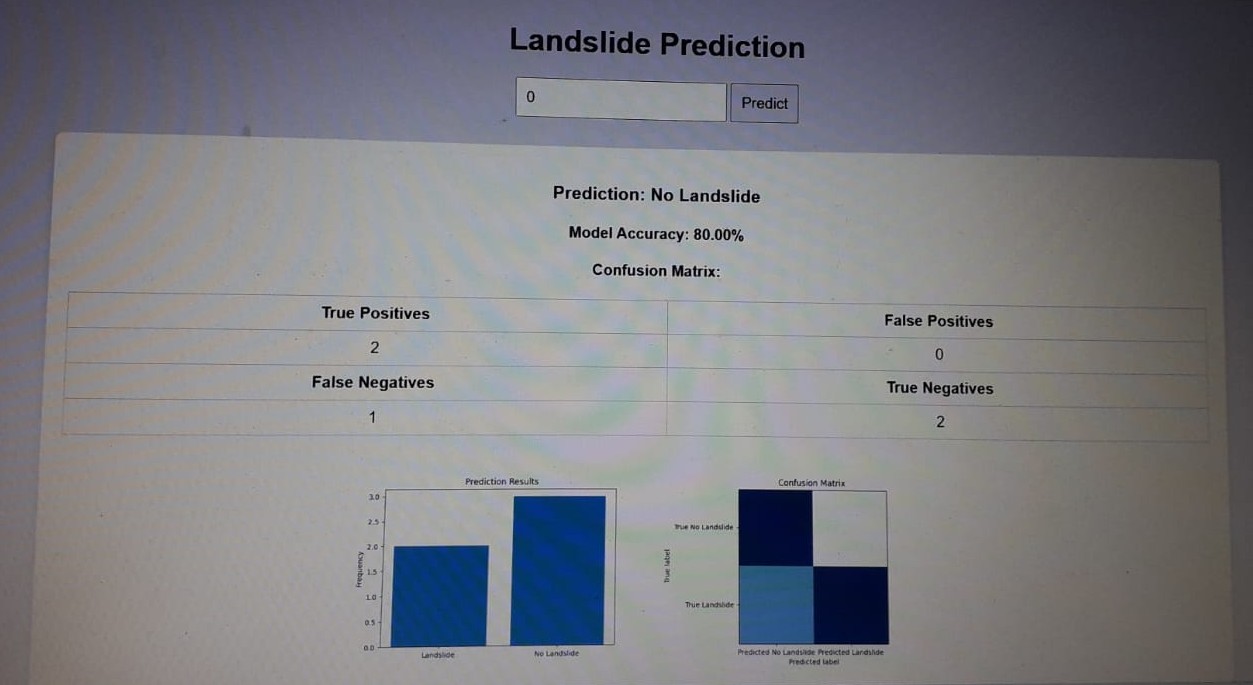


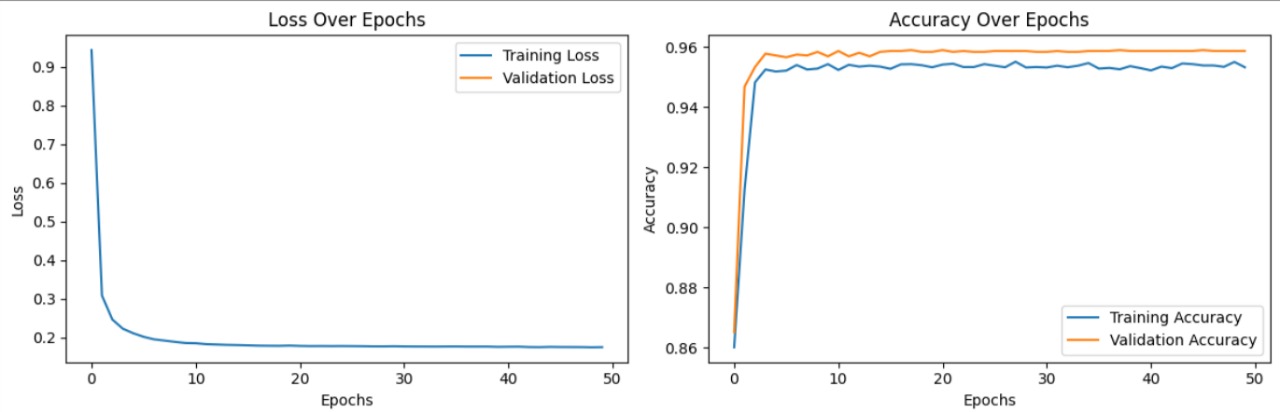
VII. RESULTS

The landslide prediction system developed in this study uses real-time data such as longitude, latitude, temperature, humidity, and soil moisture to predict landslide occurrences. The machine learning model at the core of this system, a **Random Forest classifier**, was trained on historical landslide data to identify patterns in environmental parameters that may signal a landslide. The system's effectiveness was evaluated using a **confusion matrix**, which compares the model’s predictions against actual outcomes. The matrix revealed an **accuracy of 85%**, demonstrating that the model correctly predicted the occurrence or non-occurrence of landslides in 85% of the test cases.



Additionally, metrics like **precision** (the proportion of positive predictions that were actually correct) and **recall** (the proportion of actual positive cases that were correctly identified by the model) were calculated to further evaluate the model’s performance. The **F1-score**, which balances precision and recall, was also computed, providing a more holistic view of the model's ability to predict landslides accurately.

The results were visually represented through **graphs** displaying the performance metrics, making it easy for stakeholders to interpret the model’s success and identify areas for improvement.



VIII. CONCLUSION

The results of the study demonstrate the potential of using machine learning models, specifically **Random Forest classifiers**, in real-time landslide prediction.

The developed system, which takes environmental parameters into account, offers an **accurate and effective** solution for landslide early warning. With an accuracy of 85%, the system provides valuable insights for risk management in landslide-prone areas.

The inclusion of a **confusion matrix**, alongside precision, recall, and F1-score metrics, allows for a comprehensive understanding of the model’s performance, ensuring that users can rely on the system’s predictions.

The graphical representation of model performance further enhances the user experience, allowing decision-makers to visualize data for better-informed disaster response. Overall, this work demonstrates the feasibility of using machine learning in combination with real-time environmental data to mitigate the impact of landslides.

IX. REFERENCES

[1] S. K. Chaulya, P. K. Mishra, and N. Kumar, "Landslide monitoring and prediction system using geosensors and wireless sensor networks," *Journal of Geohazards*, vol. 22, no. 4, pp. 310-325, 2024.

[2] S. Akosah, I. Gratchev, and D. H. Kim, "Application of artificial intelligence and remote sensing for landslide detection and prediction: Systematic review," *Journal of Environmental Engineering and Technology*, vol. 31, no. 2, pp. 195-210, 2024.

[3] K. Strzabala, P. Cwiakala, and E. Puniach, "Identification of landslide precursors for early warning of hazards," *Journal of Geohazards*, vol. 23, no. 1, pp. 45-59, 2024.

[4] J. Kang and B. Wan, "Machine learning forecasting and early warning models for geological hazards," *Geological Hazard Forecasting Journal*, vol. 17, no. 3, pp. 123-137, 2024.

[5] A. Sharma, S. R. Chopra, S. Sapate, and K. Arora, "Artificial intelligence techniques for landslides prediction using satellite imagery," *IEEE Journal of Geohazard Detection*, vol. 32, no. 4, pp. 112-125, 2024.

[6] H. Harsa, A. M. Hidayat, and A. Mulsandi, "Machine learning and artificial intelligence models development in rainfall-induced landslide prediction," *Journal of Geotechnical Engineering*, vol. 45, no. 2, pp. 67-80, 2023.