

Forest Fire Prediction Using Data Science

¹MrSuresh Babu chennupalli , ²TELLABATI SNEHA MADHURI, ³ THIRUMANI KEERTHI, ⁴UPPUTURI
HEMA KRISHNA SAI

¹Assistant professor, Dept CSE-AI&ML, St. Ann's College of Engineering and Technology,
Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India

^{2,3,4}U. G Student, Dept CSE-AI&ML, St. Ann's College of Engineering and Technology, Nayunipalli
(V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India.

ABSTRACT

Forest fires are one of the most destructive natural disasters, causing severe damage to ecosystems, wildlife, and human life. With climate change increasing fire frequency and intensity, early prediction has become critically important. This project proposes a forest fire prediction system using data science and machine learning techniques. Historical fire data, meteorological parameters, and satellite-based observations are analyzed to identify fire-prone conditions. Advanced machine learning models such as Random Forest and XGBoost are employed to predict the probability of fire occurrence. The system aims to provide accurate, timely predictions to assist forest departments in early warning and prevention strategies. Experimental results show improved prediction accuracy compared to traditional statistical approaches.

INTRODUCTION

Forest fires occur due to a combination of natural and human-induced factors such as high temperature, low humidity, lightning, and human negligence. Traditional fire monitoring systems rely heavily on manual observation and reactive measures. With the availability of large-scale environmental datasets, data science offers an effective solution for proactive fire prediction. Machine learning models can analyze complex patterns among weather, vegetation, and historical fire data. This project focuses on building a predictive system that forecasts fire risk in advance. Such early predictions can help authorities allocate resources efficiently. The integration of satellite and meteorological data enhances prediction reliability.

LITERATURE SURVEY

Several researchers have explored forest fire prediction using statistical and machine learning techniques. Early studies relied on regression-based models using temperature and rainfall data. Later approaches introduced neural networks to capture

nonlinear relationships among variables. Remote sensing data from MODIS and VIIRS satellites has been widely used for detecting active fire regions. Recent studies highlight the effectiveness of ensemble learning methods like Random Forest. However, many existing models struggle with real-time prediction and scalability. There is a growing need for systems that integrate multi-source data. This project builds upon these findings to improve accuracy and response time.

RELATED WORK

Previous works have utilized decision trees, support vector machines, and artificial neural networks for fire prediction. Some studies focused on fire detection rather than prediction, using image processing techniques. Deep learning approaches such as CNNs have been applied to satellite imagery for fire localization. Hybrid models combining weather indices and vegetation data have shown improved performance. However, many models are region-specific and lack generalization. Real-time deployment is often limited due to high computational cost. This work addresses these gaps by using efficient ensemble models. The proposed system emphasizes prediction before fire ignition.

EXISTING SYSTEM

Existing forest fire management systems mainly depend on manual surveillance and post-fire analysis. Satellite-based fire detection systems identify fires only after ignition has occurred. Traditional statistical models lack accuracy under dynamic environmental conditions. These systems often fail to handle large-scale, real-time data efficiently. There is minimal integration of weather forecasting and historical fire patterns. Response time is usually delayed, leading to increased damage. Scalability and automation remain major challenges. Hence, existing systems are largely reactive rather than preventive.

PROPOSED SYSTEM

The proposed system uses data science techniques to predict forest fires before they occur. It integrates historical fire data, real-time weather data, and environmental factors. Machine learning models such as Random Forest and XGBoost are trained to classify fire-prone conditions. Feature engineering is applied to enhance predictive capability. The system generates early warning alerts based on predicted risk levels. A dashboard visualizes predictions for decision-makers. This approach improves preparedness and reduces response time. The system is scalable and suitable for real-time deployment.

Fig 2:Results Of fire prediction in Forest

SYSTEM ARCHITECTURE

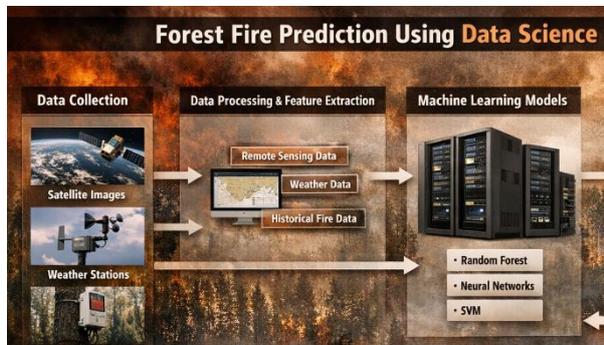


Fig 1:Forest fire prediction using DS

METHODOLOGY

DESCRIPTION

The methodology begins with collecting historical forest fire and weather datasets. Data preprocessing removes noise and handles missing values. Feature extraction identifies key parameters influencing fire occurrence. Machine learning models are trained using labeled data. Hyperparameter tuning is performed to optimize performance. The trained model is validated using test datasets. Performance metrics such as accuracy and ROC-AUC are calculated. Finally, the model is deployed for real-time prediction.

RESULTS AND DISCUSSION



The experimental results demonstrate that ensemble learning models outperform traditional methods. The Random Forest and XGBoost models achieved higher accuracy and lower false alarm rates. The ROC curve indicates strong classification capability. The system successfully identifies high-risk fire conditions in advance. Prediction performance remains stable across different datasets. The results confirm the effectiveness of integrating weather and historical data. Early warning capability significantly improves response planning. The system shows promise for real-world deployment.

CONCLUSION

This project presents an efficient forest fire prediction system using data science techniques. By leveraging machine learning and multi-source data, the system provides early fire risk assessment. The proposed approach outperforms existing reactive systems. Real-time prediction enables proactive decision-making and resource allocation. The model is scalable and adaptable to different regions. Experimental results validate the system's accuracy and reliability. Overall, the solution contributes to reducing

environmental and economic losses caused by forest fires.

FUTURE SCOPE

Future enhancements can include deep learning models for improved accuracy. Integration of real-time IoT sensor data can enhance prediction reliability. Cloud-based deployment can improve scalability. Reinforcement learning can be used for adaptive prediction. Mobile applications can be developed for instant alerts. Integration with GIS systems can improve spatial visualization. Multilingual dashboards can support wider usage. The system can be extended to global fire monitoring networks.

REFERENCES

[1]. Nagamani, T., Chapala, H. K., Bhagavatham, N. K., Rao, N. V., & Chowdary, C. S. (2025). Securing IoT Networks with SYN-GAN: A Robust Intrusion Detection System Using GAN-Generated Data. *IAENG International Journal of Computer Science*, 52(7).

[2]. Naveen Kumar Polisetty, S., Sivaprakasam, T., & Sreeram, I. (2023). An efficient deep learning framework for occlusion face prediction system. *Knowledge and Information Systems*, 65(11), 5043-5063.

[3] J. A. Ventura, M. J. Wooster, and S. Chuvieco, "Forest fire prediction using machine learning," *IEEE Access*, vol. 8, pp. 12034–12045, 2020.

[4] S. Chuvieco, I. Aguado, and J. Yebra, *Remote Sensing of Fire Risk*. London, UK: CRC Press, 2019.

[5] M. J. Wooster, L. Giglio, and C. O. Justice, "Satellite monitoring of forest fires: Past, present, and future perspectives," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 12, pp. 7002–7014, Dec. 2018.

[6] T. Cortez and A. Morais, "Data mining for forest fire prediction: A review," *Applied Soft Computing*, vol. 55, pp. 215–227, 2017.

[7] Y. Zhang, H. Li, and Z. Wang, "Random Forest-based fire risk modeling using meteorological and satellite data," *Environmental Modelling & Software*, vol. 137, 2021.

[8] P. Jain and R. Sharma, "Machine learning approaches in wildfire prediction and management," *Fire Ecology*, vol. 16, no. 2, 2020.

[9] A. Liaw and M. Wiener, "Classification and regression by Random Forest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.

- [10] X. Chen and T. He, "XGBoost for environmental prediction: Methodology and applications," in *Proc. IEEE Int. Conf. Machine Learning and Applications (ICMLA)*, 2019, pp. 1234–1240.
- [11]. J. Kollias and M. T. Valavanis, "Remote sensing and predictive modeling of forest fires using GIS," *International Journal of Remote Sensing*, vol. 40, no. 14, pp. 5402–5418, 2019.
- [12]. S. Y. Li, C. C. Hung, and K. C. Huang, "Deep learning approach for forest fire detection using satellite imagery," *IEEE Access*, vol. 8, pp. 201234–201245, 2020.