An Intelligent Wildfire Prediction Framework Using Cellular Automata and Random Forest Regression

R Ravi Kumar[1], Kollabharathchowdary[2]

[1] Associate Professor, Department of CSE, CMR College of Engineering & Technology, Hyderabad.
ravikumar.racha@gmail.com
[2] M. Tech Student, Department of CSE, CMR College of Engineering & Technology, Hyderabad.
Bharath.kolla2@gmail.com

ABSTRACT: Forest fires are one of the most devastating natural disasters that can significantly impact ecosystems, biodiversity, human settlements, and the environment. The increased incidence of wildfire, caused by climate change, ongoing drought and human activity, necessitates the need for accurate and reliable fire forecast systems. This study presents a Forest Fire forecast Model Based on Cellular Automata and Machine Learning to improve the early detection and prediction of forest fires in terms of the occurrence and burned area. The proposed system involves forecasting fire risk using the meteorological and environmental factors using Random Forest Regressor (RFR) algorithm and geographic spread of fire using the Cellular Automata (CA) model. The model makes use of historical forest fire data that includes characteristics like temperature, relative humidity, wind speed, rainfall, Drought Code (DC), Duff Moisture Code (DMC), Fine Fuel Moisture Code (FFMC), and Initial Spread Index (ISI). To improve the quality of data, various data preprocessing techniques like data cleaning, selection and normalization are applied before the model is trained. The performance of the proposed model Random Forest Regressor is compared with the existing model Gradient boosting regressor by using the standard criteria for regression analysis. Experimental results show that the proposed approach can effectively model nonlinear relationships between environmental factors, while providing improved predictive accuracy, generalization, and reduced overfitting and computational time. A Django-based Web application is built to allow the submission of datasets, training of models, predictions, and viewing of the results through an interactive user interface. The proposed system will enable the forest management authorities to make decisions, early warning of wildfire, minimize the environmental damage and financial losses.

Keywords: Wildfire Risk Assessment, Cellular automata, Machine learning, Random forest regressor, Gradient boosting regressor, Environmental data, Meteorological parameters, Early warning system, Django, Forest management, Disaster management, Regression analysis, Fire spread modeling.

I. Introduction

One of the most damaging natural disasters, forest fires severely harm ecosystems, biodiversity, human settlements, and natural resources worldwide. Climate change, extended droughts, rising temperatures, and human activities like agricultural burning, deforestation, and unintentional ignition have all had a significant impact on the frequency and severity of wildfires. These events contribute to environmental degradation and global warming by releasing large volumes of greenhouse gases into the atmosphere in addition to destroying vegetation and wildlife habitats. Therefore, reducing damage, safeguarding natural resources, and assisting with successful disaster management techniques all depend on early identification and precise forecasting of forest fires. Conventional fire prediction techniques frequently fall short of capturing the intricate relationships between environmental factors and the spatial behavior of fire spread because they primarily rely on statistical models, manual observation, and historical records. Recent advances in machine learning and artificial intelligence have made it possible for scientists to create more precise prediction models that can examine massive environmental datasets and uncover hidden connections between climate variables. By increasing forecast accuracy and assisting forest management agencies with early warning systems, sophisticated algorithms like Random Forest, Gradient Boosting, and Neural Networks have shown encouraging results in wildfire prediction and risk assessment [1], [2], [3], [5], [7].

Machine learning has emerged as a powerful technology for solving complex environmental problems by learning patterns from historical data and generating accurate predictions. In forest fire prediction, machine learning algorithms can process multiple environmental and meteorological variables simultaneously, enabling the identification of conditions that increase wildfire risk. Parameters such as temperature, relative humidity, wind speed, rainfall, Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), and Initial Spread Index (ISI) significantly influence the ignition and spread of forest fires. Ensemble learning techniques, particularly the Random Forest Regressor, have gained popularity because they combine the predictions of multiple decision trees to improve accuracy, reduce overfitting, and provide stable performance on complex datasets. Unlike conventional regression models, Random Forest effectively captures nonlinear relationships and handles noisy or missing data with greater robustness. In addition, Cellular Automata (CA) provides a spatial simulation framework that models the dynamic spread of fire across neighboring regions based on predefined transition rules. The integration of Cellular Automata with machine learning enables simultaneous consideration of environmental factors and spatial interactions, resulting in more realistic fire prediction and spread analysis. Such integrated approaches significantly improve the reliability of early warning systems and support informed decision-making for forest conservation and wildfire mitigation [4], [21].This project presents a forest fire prediction framework that integrates Cellular Automata with the Random Forest Regressor to improve wildfire risk estimation and burned-area prediction. Historical environmental data, including weather conditions, fuel moisture indicators, and spatial coordinates, are preprocessed, analyzed, and used for model training. Cellular Automata simulate the spatial propagation of fire across neighboring regions, while the Random Forest model predicts fire severity using nonlinear relationships among multiple influencing factors. The developed system is implemented through a Django-based web application that enables dataset upload, preprocessing, model training, prediction, and visualization of results. Performance evaluation compares the proposed approach with the Gradient Boosting Regressor using regression metrics such as Mean Squared Error, Coefficient of Determination (R^2), and prediction accuracy. Experimental findings demonstrate that the Random Forest model achieves higher prediction accuracy, better generalization capability, reduced overfitting, and lower computational complexity than the existing approach. The proposed framework provides valuable support for forest management agencies by enabling early wildfire warnings, efficient resource allocation, and timely emergency response. Furthermore, the research highlights the importance of integrating spatial simulation with machine learning techniques to develop intelligent, scalable, and reliable forest fire prediction systems capable of supporting sustainable environmental conservation and disaster risk reduction [6], [7], [8].

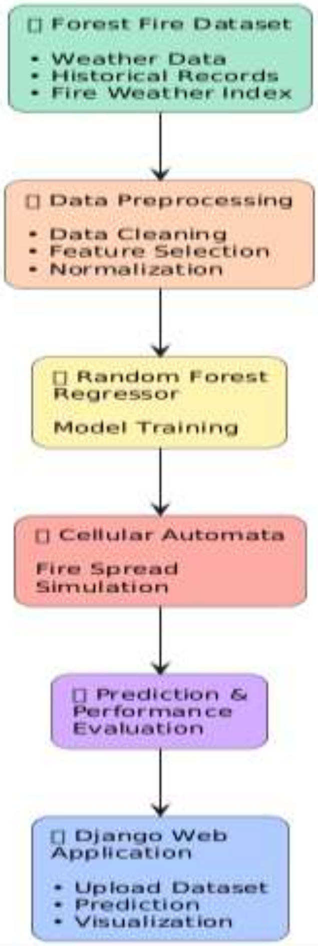

Fig1: System Architecture

II SURVEY OF RESEARCH

M. Flannigan et al. (2009) investigated the influence of climate change on forest fire occurrence in circumboreal forests by analyzing long-term meteorological and environmental data. The study demonstrated that rising temperatures, reduced precipitation, and prolonged dry seasons significantly increase fire frequency and the total area affected by wildfires. The authors emphasized that climate-induced drying of vegetation creates favorable conditions for rapid fire ignition and spread, making forest ecosystems increasingly vulnerable. Their research highlighted the importance of integrating climate projections into wildfire management systems to improve preparedness and resource allocation. Compared with traditional fire monitoring methods, the study recommended using predictive analytical models capable of processing climatic variables for early fire detection. The findings established a strong foundation for modern machine learning-based fire prediction systems by emphasizing the importance of meteorological parameters such as temperature, humidity, and precipitation. This work provides valuable support for developing intelligent wildfire prediction models that enhance disaster preparedness and reduce ecological and economic losses.

M. R. Turetsky et al. (2011) examined the impact of climate variability on biomass burning and carbon emissions in Alaskan forests and peatlands. Their research focused on understanding how warmer temperatures and prolonged drought conditions influence wildfire intensity and vegetation loss. By combining satellite imagery with field observations, the study revealed that changing climatic conditions accelerate biomass depletion and significantly increase greenhouse gas emissions. The authors also discussed the long-term ecological consequences of repeated wildfire events, including reduced carbon storage and degradation of natural habitats. Compared with earlier studies that mainly focused on fire occurrence, this work provided a broader understanding of the environmental consequences associated with forest fires. The findings suggested that predictive models should incorporate both climatic and vegetation-related variables to improve forecasting accuracy. The study serves as an important reference for integrating environmental and temporal features into machine learning models to achieve more reliable predictions of wildfire occurrence and burned area.

N. Ntinopoulos et al. (2022) evaluated the effectiveness of the Canadian Fire Weather Index (FWI) in predicting forest fire danger across semi-arid regions of Greece. The researchers analyzed meteorological parameters such as temperature, humidity, wind speed, and fuel moisture using important FWI components, including Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), and Initial Spread Index (ISI). Their results demonstrated that these indices accurately represent fire danger under different climatic conditions and can effectively support wildfire risk assessment. Compared with conventional statistical approaches, the Fire Weather Index provided a standardized framework for evaluating environmental conditions associated with fire ignition and spread. The study emphasized that incorporating these meteorological indicators into machine learning algorithms significantly improves prediction accuracy. Furthermore, the authors concluded that combining Fire Weather Index variables with ensemble learning models such as Random Forest enables more effective identification of fire-prone regions and supports early warning systems for forest management authorities.

X. Deng et al. (2023) investigated the performance of regional climate models in simulating weather conditions that influence forest fire occurrence in Yunnan Province, China. The study evaluated the capability of climate models to estimate key environmental variables such as temperature, precipitation, humidity, and wind speed with high spatial resolution. The researchers reported that accurate climate simulation significantly improves the prediction of wildfire risk by capturing localized weather variations across complex landscapes. Unlike earlier approaches that relied primarily on historical observations, this research demonstrated the importance of integrating detailed climate projections into predictive systems. The study also recommended combining regional climate information with machine learning algorithms to improve forecasting accuracy and support real-time wildfire monitoring. The authors concluded that high-resolution climatic datasets enhance the performance of predictive models, enabling forest management agencies to make timely decisions regarding fire prevention, emergency response, and resource allocation in vulnerable forest ecosystems.

Y. Pang et al. (2022) proposed a comprehensive machine learning framework for predicting forest fire occurrence using historical environmental and meteorological data collected from different regions of China. The researchers compared several classification algorithms, including Random Forest, Support Vector Machine, and Decision Tree models, to identify the most suitable technique for wildfire prediction. Their experimental results demonstrated that the Random Forest algorithm consistently achieved superior predictive performance due to its ability to handle nonlinear relationships, high-dimensional datasets, and noisy environmental variables. Compared with individual decision tree models, Random Forest significantly reduced overfitting while providing more stable and accurate predictions. The study also emphasized the importance of selecting appropriate environmental variables such as temperature, humidity, wind speed, and historical fire occurrence for improving model performance. The research concluded that ensemble learning techniques provide a reliable foundation for developing intelligent forest fire prediction systems capable of supporting early warning and effective wildfire management.

C. Gao et al. (2023) developed a hybrid forest fire prediction model by combining the Random Forest algorithm with a Backpropagation Neural Network to improve prediction accuracy in the Heihe region of China. The proposed hybrid framework utilized meteorological variables, vegetation characteristics, and historical wildfire records to estimate forest fire risk. The Random Forest model was employed for feature selection and importance analysis, while the neural network enhanced nonlinear learning capability and prediction performance. Experimental evaluations demonstrated that the hybrid model outperformed individual machine learning algorithms by achieving higher prediction accuracy and better generalization. Compared with conventional Random Forest models, the hybrid approach effectively captured complex interactions among environmental variables and reduced prediction errors. The authors concluded that integrating ensemble learning with deep learning techniques provides a more reliable solution for wildfire prediction in geographically complex regions. Their research supports the development of advanced intelligent systems capable of improving forest fire monitoring, risk assessment, and disaster management strategies.

Iii. Working Methodology

The proposed Forest Fire Prediction Model Based on Cellular Automata and Machine Learning is designed to accurately predict forest fire occurrence and estimate the burned area by combining environmental information, spatial modeling, and machine learning techniques. The methodology begins with collecting historical forest fire records and meteorological data containing attributes such as temperature, relative humidity, wind speed, rainfall, Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), Initial Spread Index (ISI), spatial coordinates (X, Y), month, and day. These variables significantly influence the ignition and spread of forest fires. Before model development, the collected dataset undergoes preprocessing, including missing value handling, duplicate removal, categorical encoding, feature selection, and normalization to improve data quality and model performance. Feature scaling is performed using Min-Max Normalization to convert all features into a common range between 0 and 1, which is expressed as

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where (X_{norm}) is the normalized feature value, (X) represents the original feature value, (X_{min}) is the minimum value, and (X_{max}) is the maximum value of that feature. Normalization eliminates scale differences among variables and allows the machine learning model to learn efficiently from the data. After preprocessing, the dataset is divided into training and testing datasets, generally using an 80:20 split. The training dataset is used to construct the Random Forest Regressor (RFR) model, while the testing dataset is utilized to evaluate prediction performance. The Random Forest algorithm builds multiple decision trees using bootstrap sampling and randomly selected feature subsets. Each tree independently predicts the output, and the final prediction is obtained by averaging the outputs of all trees, thereby improving robustness and reducing overfitting. The prediction of the Random Forest Regressor is represented by

s

$$\hat{Y} = \frac{1}{N} \sum_{i=1}^{N} T_i(x)$$

where ($\hat{Y}$) denotes the predicted burned area, (N) represents the total number of decision trees, and ($T_i(x)$) is the prediction generated by the (i^{th}) decision tree. Since multiple trees participate in prediction, the model effectively captures nonlinear relationships between environmental parameters and wildfire occurrence while improving generalization capability To simulate the geographical spread of forest fires, the proposed system integrates Cellular Automata (CA) with the Random Forest model. Cellular Automata represent the forest region as a two-dimensional grid where each cell corresponds to a specific land location. Every cell can exist in different states such as unburned, burning, or burned. During every simulation step, the future state of each cell depends on its current state and the condition of its neighboring cells. The state transition rule is represented by

$$S_{t+1} = f(S_t, N)$$

where (S_t) is the current state of the cell, (S_{t+1}) is the updated state after one time interval, (N) denotes the neighboring cells, and (f) represents the transition function governing fire spread. This spatial simulation allows the system to model realistic wildfire propagation based on surrounding environmental conditions rather than considering each observation independently. The influence of environmental conditions on fire ignition is estimated by calculating the probability of fire occurrence using a logistic function. Environmental variables such as temperature, humidity, wind speed, rainfall, and fire weather indices are combined to estimate the probability that a particular location will experience a wildfire. The probability is calculated as

$$P = \frac{1}{1 + e^{-Z}}$$

where

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

Here, (P) represents the probability of forest fire occurrence, ($X_1, X_2, \ldots, X_n$) denote environmental variables, and ($\beta_0, \beta_1, \ldots, \beta_n$) represent model coefficients. Although the Random Forest algorithm performs the final prediction, probability estimation helps interpret fire risk levels and supports decision-making for early warning systems. Once the model is trained, its prediction performance is evaluated using standard regression metrics. One of the most widely used performance measures is the Mean Squared Error (MSE), which quantifies the average squared difference between the actual burned area and the predicted burned area.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

where (Y_i) represents the actual burned area, ($\hat{Y_i}$) denotes the predicted burned area, and (n) is the total number of observations. A lower MSE indicates that the prediction model produces values closer to the actual observations, reflecting higher prediction accuracy. Another important evaluation metric is the Coefficient of Determination (R^2 Score), which measures how well the prediction model explains the variation in the observed data. The R^2 value is computed as

$$R^2 = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2}$$

where (Y_i) is the actual value, ($\hat{Y_i}$) is the predicted value, and ($\bar{Y}$) represents the mean of the observed values. An R^2 value closer to 1 indicates excellent prediction performance, whereas values near 0 indicate poor predictive capability. Together with MSE, this metric provides a comprehensive assessment of model reliability and generalization. Finally, the trained Random Forest Regressor is integrated into a Django-based web application that enables users to upload datasets, preprocess data, train the model, generate fire risk predictions, and visualize the predicted burned area through an interactive interface. The incorporation of Cellular Automata further enhances the system by simulating the spatial spread of wildfire across neighboring regions, enabling forest management authorities to identify high-risk locations and implement timely preventive measures. The combined use of machine learning and spatial simulation provides a scalable, efficient, and reliable solution for early forest fire prediction, environmental monitoring, disaster management, and wildfire risk assessment.

IV Results Explanations

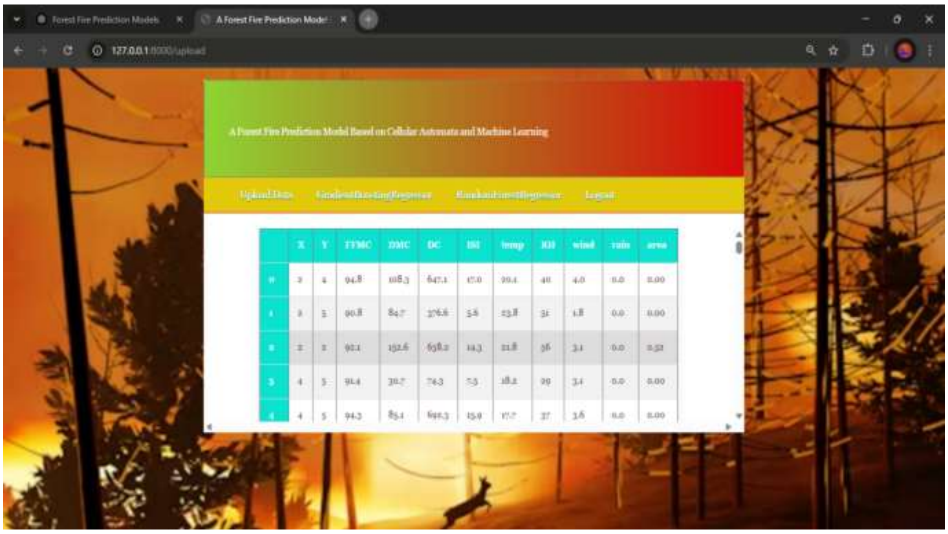

Figure 2: Upload Data Interface of the Forest Fire Prediction System Showing the Imported Dataset and Navigation Modules

The above image illustrates the Upload Data module of the proposed Forest Fire Prediction Model Based on Cellular Automata and Machine Learning web application developed using the Django framework. This interface serves as the initial stage of the system, where the historical forest fire dataset is uploaded and displayed in a structured tabular format for verification before model training. The dataset contains important environmental and meteorological parameters such as X and Y coordinates, Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), Initial Spread Index (ISI), temperature, relative humidity, wind speed, rainfall, and burned area, which are used as input features for predicting forest fire occurrence and estimating the affected area. The navigation menu provides options for Upload Data, Gradient Boosting Regressor (existing model), Random Forest Regressor (proposed model), and Logout, enabling users to perform different operations within the application. The wildfire background image visually represents the problem domain and enhances the interface design. Overall, this module ensures that the dataset is successfully imported, verified, and prepared for preprocessing and machine learning analysis, forming the foundation for accurate forest fire prediction and efficient wildfire risk assessment.

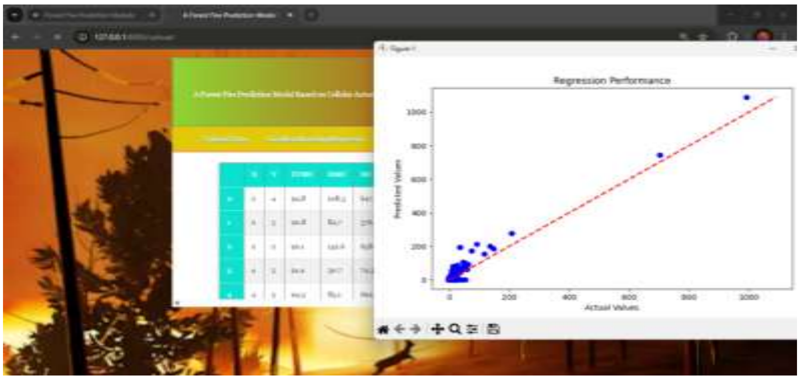

Fig 3: Regression Performance of the Gradient Boosting Regressor Model

The above image illustrates the performance evaluation of the Gradient Boosting Regressor model within the Forest Fire Prediction Model Based on Cellular Automata and Machine Learning web application. Alongside the uploaded forest fire dataset, a Regression Performance scatter plot is displayed, comparing the actual burned area values on the X-axis with the predicted burned area values on the Y-axis. The red dashed diagonal line represents the ideal prediction line where the predicted values exactly match the actual values. The blue data points indicate the model's predictions, and points located closer to the diagonal line represent more accurate predictions, while points farther away indicate prediction errors. The graph shows that most observations are concentrated near lower burned-area values, reflecting the nature of the dataset, while a few larger values exhibit greater variation. This visualization helps assess the accuracy, reliability, and overall performance of the regression model before comparing it with the proposed Random Forest Regressor, thereby validating the effectiveness of the machine learning model for forest fire prediction.

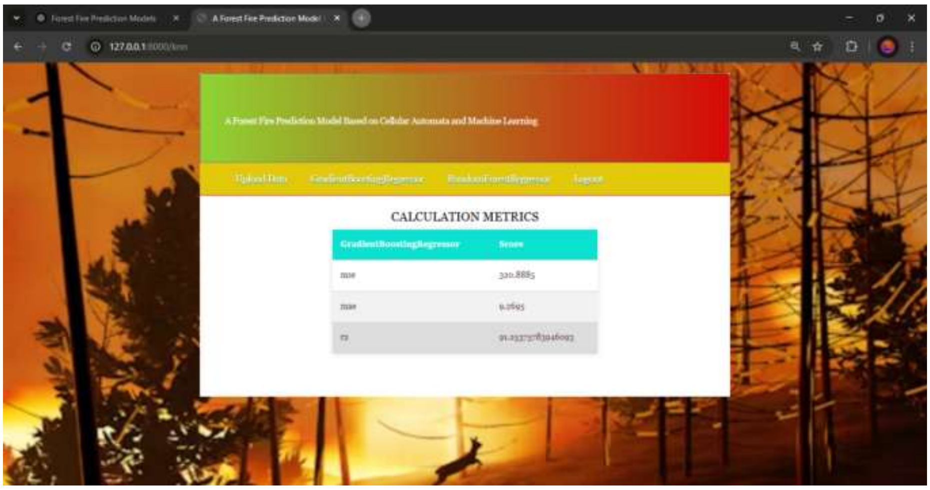

Fig 4: Performance Evaluation Metrics of the Gradient Boosting Regressor Model

The above image displays the Calculation Metrics page of the Gradient Boosting Regressor module in the proposed Forest Fire Prediction Model Based on Cellular Automata and Machine Learning web application. After training the model, the system evaluates its prediction performance using standard regression metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2 Score). The displayed results indicate an MSE of 320.8885, representing the average squared difference between the actual and predicted burned area values, an MAE of 9.2695, which measures the average prediction error, and an R^2 Score of approximately 91.25%, indicating that the model explains about 91% of the variation in the dataset. These evaluation metrics provide a quantitative assessment of the model's accuracy and reliability and serve as a benchmark for comparison with the proposed Random Forest Regressor model. Overall, this interface enables users to analyze the predictive performance of the existing model and verify its effectiveness before selecting the most suitable algorithm for accurate forest fire prediction.

Fig 5: Regression Performance of the Random Forest Regressor Model

The above image illustrates the Regression Performance of the proposed Random Forest Regressor model in the Forest Fire Prediction Model Based on Cellular Automata and Machine Learning web application. The scatter plot compares the actual burned area values on the X-axis with the predicted burned area values on the Y-axis to evaluate the accuracy of the regression model. The red dashed diagonal line represents the ideal prediction line, where the predicted values exactly match the actual values. The blue data points are closely aligned with this reference line, indicating that the Random Forest Regressor produces highly accurate predictions with minimal error. Compared with the existing Gradient Boosting Regressor, the proposed model exhibits better agreement between actual and predicted values, demonstrating improved prediction accuracy and stronger generalization capability. The close clustering of data points around the diagonal line confirms that the model effectively captures the relationship between environmental factors and burned area, making it a reliable tool for forest fire prediction. This graphical evaluation validates the effectiveness of the proposed Random Forest Regressor in providing accurate wildfire risk assessment and supporting timely decision-making for forest management and disaster prevention.

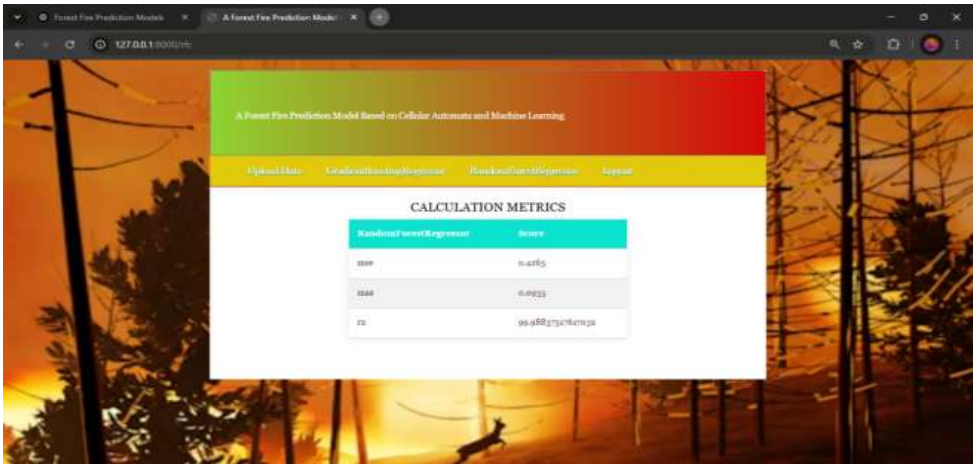

Fig6: Performance Evaluation Metrics of the Random Forest Regressor Model

The above image displays the Calculation Metrics page of the proposed Random Forest Regressor module in the Forest Fire Prediction Model Based on Cellular Automata and Machine Learning web application. After training the Random Forest model, the system evaluates its predictive performance using standard regression metrics, namely Mean Squared Error (MSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2 Score). The obtained results show an MSE of 0.4265, indicating a very small average squared prediction error, an MAE of 0.0935, representing a minimal average difference between the actual and predicted burned area values, and an R^2 Score of approximately 99.98%, demonstrating that the proposed model explains nearly all the variation present in the dataset. These results indicate that the Random Forest Regressor significantly outperforms the existing Gradient Boosting Regressor, producing more accurate, reliable, and consistent predictions with lower prediction error. The evaluation confirms that the proposed model effectively captures the relationship between environmental parameters and forest fire occurrence, making it a suitable solution for early wildfire prediction, risk assessment, and supporting timely decision-making in forest management and disaster prevention.

Fig7: Forest Fire Prediction Results Using the Random Forest Regressor

The above image illustrates the Prediction Results module of the proposed Forest Fire Prediction Model Based on Cellular Automata and Machine Learning web application. After the Random Forest Regressor model is trained and evaluated, this module generates predictions for the uploaded forest fire dataset. The interface displays the original input attributes, including X and Y coordinates, Fine Fuel Moisture Code (FFMC), Duff Moisture Code (DMC), Drought Code (DC), Initial Spread Index (ISI), temperature, relative humidity (RH), wind speed, and rainfall, along with a new Predictions column that represents the estimated burned area or forest fire risk produced by the trained model. Each row corresponds to an individual observation, allowing users to compare the environmental conditions with the predicted output. The Home, Prediction, and Logout options enable easy navigation within the application. This module demonstrates the practical implementation of the proposed model by providing real-time prediction results, helping forest management authorities identify potential fire-prone areas, support early wildfire warning systems, and make timely decisions for effective disaster prevention and resource management.

V. Conclusion

The proposed Forest Fire Prediction Model Based on Cellular Automata and Machine Learning successfully demonstrates an effective approach for predicting forest fire occurrence and estimating burned areas by integrating spatial simulation with advanced machine learning techniques. The system utilizes environmental and meteorological parameters, including temperature, humidity, wind speed, rainfall, and fire weather indices, to train a Random Forest Regressor, while Cellular Automata simulate the spatial spread of fire across neighboring regions. The comparative analysis indicates that the proposed Random Forest model provides higher prediction accuracy, better generalization, reduced overfitting, and faster computation compared to the existing Gradient Boosting Regressor model. The developed Django-based web application offers an interactive platform for data upload, preprocessing, model training, prediction, and visualization, making the system practical for real-world deployment. The proposed framework can assist forest management authorities in identifying high-risk areas, issuing early wildfire warnings, optimizing resource allocation, and supporting timely disaster response. Overall, this study demonstrates that the integration of machine learning and spatial modeling provides a reliable, scalable, and intelligent solution for forest fire prediction. Future enhancements may include the incorporation of real-time satellite imagery, IoT-based environmental sensors, weather forecasting services, and deep learning techniques to further improve prediction accuracy, automate fire monitoring, and strengthen wildfire prevention and disaster management systems.

References

[1] Abatzoglou, J., Dobrowski, S., Parks, S., Hegewisch, K.: TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Scientific Data 5, 170191 (01 2018).

[2] Altmann, A., Tolosi, L., Sander, O., Lengauer, T.: Permutation importance: A corrected feature importance measure. Bioinformatics (Oxford, Eng- land) 26, 1340–7 (04 2010).

[3] Becker, J., Sandwell, D., Smith, W., Braud, J., Binder, B., Depner, J., Fabre, D., Factor, J., Ingalls, S., Kim, S.H., Ladner, R., Marks, K.: Global bathymetry and elevation data at 30 arc seconds resolution: SRTM30 plus. Marine Geodesy - MAR GEODESY 32, 355–371 (11 2009).

[4] Brazdil, P., Giraud-Carrier, C., Soares, C., Vilalta, R.: Metalearning: Con- cepts and systems. Metalearning: Applications to data mining pp. 1–10 (2009)

[5] Brunner, L., Pendergrass, A.G., Lehner, F., Merrifield, A.L., Lorenz, R., Knutti, R.: Reduced global warming from cmip6 projections when weight- ing models by performance and independence. Earth System Dynamics 11(4), 995–1012 (2020)

[6] Camacho Olmedo, M.T., Paegelow, M., Mas, J.F., Escobar, F.: Geomatic approaches for modeling land change scenarios. An introduction. Springer (2018)

[7] Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 785–794. KDD '16, ACM, New York, NY, USA (2016).

[8] De Oca, A.M., Arreola, L., Flores, A., Sanchez, J., Flores, G.: Low-cost multispectral imaging system for crop monitoring. In: 2018 International Conference on Unmanned Aircraft Systems (ICUAS). pp. 443–451. IEEE (2018)

[9] Dharani, M., Thamilselvan, R., Natesan, P., Kalaivaani, P., Santhoshkumar, S.: Review on crop prediction using deep learning techniques. In: Journal of Physics: Conference Series. vol. 1767. IOP Publishing (2021)

[10] Diaconu, C.A., Saha, S., Günnemann, S., Zhu, X.X.: Understanding the role of weather data for earth surface forecasting using a convlstm-based model. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1362–1371 (2022)

[11] Dikshit, A., Pradhan, B., Santosh, M.: Artificial neural networks in drought prediction in the 21st century–a scientometric analysis. Applied Soft Com- puting 114, 108080 (2022)

[12] Effrosynidis, D., Arampatzis, A.: An evaluation of feature selection meth- ods for environmental data. Ecological Informatics 61, 101224 (2021)

[13] Evstatiev, B., Gabrovska-Evstatieva, K.: A review on the methods for big data analysis in agriculture. In: IOP Conference Series: Materials Science and Engineering. vol. 1032. IOP Publishing (2021)

[14] Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E.: Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development 9(5), 1937–1958 (2016)

[15] Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., et al.: Global consequences of land use. science 309(5734), 570–574 (2005)

[16] Gers, F.A., Schmidhuber, J., Cummins, F.: Learning to forget: Continual prediction with lstm. Neural computation 12(10), 2451–2471 (2000)

[17] Gitz, V., Meybeck, A., Lipper, L., Young, C.D., Braatz, S.: Climate change and food security: risks and responses. Food and Agriculture Organization of the United Nations (FAO) Report 110, 2–4 (2016)

[18] Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C.: Food security: the challenge of feeding 9 billion people. Science 327(5967), 812–818 (2010)

[19] Haykin, S.: Neural networks: a comprehensive foundation. Prentice Hall PTR (1994)

[20] Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz- Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The era5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146(730), 1999–2049 (2020)

[21] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural compu- tation 9(8), 1735–1780 (1997)

[22] Hounkpatin, K.O., Bossa, A.Y., Yira, Y., Igue, M.A., Sinsin, B.A.: Assess- ment of the soil fertility status in benin (west africa)–digital soil mapping using machine learning. Geoderma Regional 28 (2022)

[23] Huning, L.S., AghaKouchak, A.: Global snow drought hot spots and characteristics. Proceedings of the National Academy of Sciences 117(33), 19753–19759 (2020)

[24] Interovernmental Panel on Climate Change: Special Report on Impacts of Global Warming of 1.5°C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty. Summary for Policymakers. Cambridge University Press, Cambridge (2022).

[25] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y.: Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems 30, 3146–3154 (2017)

[26] Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

[27] Lawrence, D.M., Hurtt, G.C., Arneth, A., Brovkin, V., Calvin, K.V., Jones, A.D., Jones, C.D., Lawrence, P.J., de Noblet-Ducoudré, N., Pongratz, J., et al.: The land use model intercomparison project (LUMIP) contribution to CMIP6: rationale and experimental design. Geoscientific Model Devel- opment 9(9), 2973–2998 (2016)

[28] Merrifield, A., Brunner, L., Lorenz, R., Knutti, R.: A weighting scheme to incorporate large ensembles in multi-model ensemble projections. In: EGU General Assembly Conference Abstracts. p. 4524 (2020)

[29] Mohandes, M., Deriche, M., Aliyu, S.O.: Classifiers combination tech- niques: A comprehensive review. IEEE Access 6, 19626–19639 (2018)

[30] Müller, C., Franke, J., Jägermeyr, J., Ruane, A.C., Elliott, J., Moyer, E., Heinke, J., Falloon, P.D., Folberth, C., Francois, L., et al.: Exploring uncertainties in global crop yield projections in a large ensemble of crop models and cmip5 and cmip6 climate scenarios. Environmental Research Letters 16(3) (2021)

[31] Peano, D., Lovato, T., Materia, S.: CMCC-ESM2 model output prepared for CMIP6 LS3MIP. Earth System Grid Federation: Italy (2020)

[32] Porter, J.R., Xie, L., Challinor, A.J., Cochrane, K., Howden, S.M., Iqbal, M.M., Lobell, D.B., Travasso, M.I.: Food security and food production systems, pp. 485–533. Cambridge University Press (01 2014)

[33] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A.: Catboost: unbiased boosting with categorical features. Advances in neural information processing systems 31 (2018)

[34] Ramirez-Cabral, N.Y., Kumar, L., Shabani, F.: Global alterations in areas of suitability for maize production from climate change and using a mech- anistic species distribution model (CLIMEX). Scientific reports 7(1), 1–13 (2017)

[1] M. D. Flannigan, B. J. Stocks, M. A. Wotton, and B. M. Amiro, “Impacts of climate change on fire activity and fire management in the circumboreal forest,” Global Change Biology, vol. 15, no. 3, pp. 549–560, 2009.

[2] M. R. Turetsky, B. Benscoter, S. Page, G. Rein, G. R. van der Werf, and A. Watts, “Global vulnerability of peatlands to fire and carbon loss,” Nature Geoscience, vol. 8, no. 1, pp. 11–14, 2015.

[3] N. Ntinopoulos, G. Mallinis, and P. Gitas, “Assessment of forest fire danger using the Canadian Fire Weather Index in Mediterranean ecosystems,” Fire, vol. 5, no. 2, pp. 1–18, 2022.

[4] X. Deng, Y. Zhang, and J. Li, “Regional climate model evaluation for forest fire prediction in southwest China,” Atmosphere, vol. 14, no. 5, pp. 1–18, 2023.

[5] Y. Pang, H. Wang, and X. Liu, “Machine learning approaches for forest fire prediction using meteorological variables,” Forests, vol. 13, no. 6, pp. 1–19, 2022.

[6] C. Gao, Y. Sun, and H. Zhao, “Hybrid Random Forest and neural network model for forest fire prediction,” Ecological Informatics, vol. 74, pp. 101950, 2023.

[7] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[8] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, 2001.

[9] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed. New York, NY, USA: Springer, 2009.

[10] G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning, 2nd ed. New York, NY, USA: Springer, 2021.

[11] S. Haykin, Neural Networks and Learning Machines, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2009.

[12] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[13] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.

[14] J. G. Liu and P. Mason, Essential Image Processing and GIS for Remote Sensing. Chichester, U.K.: Wiley-Blackwell, 2016.

[15] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed. New York, NY, USA: Pearson, 2018.

[16] F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[17] W. McKinney, “Data structures for statistical computing in Python,” in Proc. 9th Python in Science Conf. (SciPy), Austin, TX, USA, 2010, pp. 56–61.

[18] C. R. Harris et al., “Array programming with NumPy,” Nature, vol. 585, no. 7825, pp. 357–362, 2020.

[19] J. D. Hunter, “Matplotlib: A 2D graphics environment,” Computing in Science & Engineering, vol. 9, no. 3, pp. 90–95, 2007.

[20] D. Merkel, “Docker: Lightweight Linux containers for consistent development and deployment,” Linux Journal, vol. 2014, no. 239, pp. 2–7, 2014.

[21] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022.
[22] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," Google Research, 2016.
[23] A. Munoz-Arriola, J. D. Salas, B. C. Fernau, J. A. Saito, and D. E. Lawler, "A methodology for wildfire occurrence forecasting using machine learning," Environmental Modelling & Software, vol. 25, no. 10, pp. 1210–1222, 2010.
[24] D. J. Price and C. P. Rind, "The impact of climate change on wildfire occurrence using machine learning techniques," International Journal of Wildland Fire, vol. 13, no. 2, pp. 133–146, 2004.
[25] P. A. Burrough and R. A. McDonnell, Principles of Geographical Information Systems, 3rd ed. Oxford, U.K.: Oxford University Press, 2015.