2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) | 979-8-3503-9926-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICACITE57410.2023.10182818

Comparison of Forest Fire Prediction System using Machine Learning Alogrithms

ViswaBharathi
Dept. of CSE,
GITAM School of Technology,
GITAM (Deemed to be University),
Bengaluru, India
vmallika@gitam.edu

Chennuru PeddaReddy
GITAM School of Technology,
GITAM (Deemed to be University),
Bengaluru, India
pchennuru@gmail.com

***Abstract*—Fires in the forest are Uncontrolled flames, which gravely affect both normal and social resources, are one of the greatest harmful natural tragedies to the situation. The suggested approach makes advantage of a number of advancements, including machine learning techniques, which have a high probability of accurately representing the state of the forest ecosystem. Based on the strategy, we may be able to pinpoint the days that provide a substantial danger of forest fires and focus extra attention on forest fire prevention for forest guards. Forecasting future flames is part of the management of forest fires. It is crucial to the distribution of resources, the prevention of damage, and the pursuit of recovery. Using machine learning, this system is actively analyzing ways for predicting forest fires. We introduce a brand-new support vector machine-based forest fire risk prediction system. A forest's level of fire threat is predicted by the algorithm based on historical weather patterns and data. The algorithm's implementation makes use of the available data to correctly forecast the risk of a fire.**

***Keywords*— Forest Fires, Machine Learning Algorithm**

I. INTRODUCTION

One of the saddest tragedies in recent memory has been forest fires. Forest fires are caused by a number of variables, including deforestation and climate change, which have an ongoing effect on the environment. By acquiring satellite images of the area and notifying the proper authorities of a fire emergency, it is possible to put out forest fires. By the time the authorities hear about the issue, the damage caused by the fires to that specific sector will already be significant. Data connected to forests can be utilised to predict the regions where accidents might occur using data mining and machine learning techniques, providing an effective preventative strategy. There are numerous techniques utilised, both with and without Principal Component Analysis forecasts using techniques like Bagging, Boosting, Random Forest and K-Nearest Neighbors. When PCA was utilised, Gradient Bumping had the highest F-1score, while Logistic Regression had the highest score when PCA wasn't used Due to their superior temporal resolution across vast distances, geostationary cable distant detection schemes are a valuable device for locating and monitoring forest fires. These digital systems can gather, store, analyse, and display spatially referenced data, or information with a location.

II. LITERATURE SURVEY

Before applying the machine learning techniques, a case study must be conducted utilising the raw data set available and the existing methodology provided by the forest fire records. We talk about four fundamental traditional machine learning strategies and their algorithmic justifications based on the circumstances. Additionally provided is pre-processing of the required data. The fourth-place finisher is the k-mean clustering approach. Predefining the amount of collections and the early centroid theme is important because this learning process is unsupervised(s). Algorithm provides specifics on the mechanism.

Data mining can be used to forecast forest fires based on their past occurrences, which is one such effective method. For prediction, data mining needs a true and uncontaminated set of data. Before using the values for modelling, it is necessary to deal with any unresolved or numerous unknown values in the dataset. Prediction is based on the dataset about forest fires that is included in the UCI Machine Learning packages. proposed a comparable piece of work using the dataset to forecast the area affected by forest fires.

The current system is made up of sensors that can detect smoke and fire and data mining. In reality, climatic factors like temperature and wind are what start most forest fires, and numerous fire indices, including the forest Fire Weather Index, utilise this information. In this study, we investigate a Data Mining method for forecasting burned territory in forest fires. Real-world data gathered from the northeastern part of Portugal was used to test five different DM approaches, including Random Forests, four methods for feature collection (using three-dimensional, chronological, FWI mechanisms, and meteorological qualities). The model can forecast the burned area left behind by minor fires utilising the ideal four atmospheric inputs: disease, comparative stickiness, stream, and breeze.

III. PROPOSED SYSTEM

Here, after analyzing the outcomes of previous approaches, have to perform forest fires classification using Python and pandas operations The machine learning repository is where the information is gathered. The dataset, working environment, and predictive analytics methodology are all shown visually in an understandable manner. Data pre-processing is the primary step in the process of machine learning. The following step entails data cleansing, classification, modelling, and performance assessment. The random forest is used to improve the accuracy of the results.

A. A. Existing System

Forest fires input details are used in this system. Using machine learning techniques, the user inputs are used to analyse forest fire prediction. K-Nearest Neighbors, Decision Trees, and Support Vector Machines are utilised to detect patterns in the data of heart disease patients collected from UCI Machine Learning. The accuracies and performances of these algorithms are compared. Early fires detection is

necessary when preparing the environment and climate. To analyse the data and offer various viewpoints on forest fires, ML algorithms are used.

B. Forest fires

In tropical, temperate, and boreal forests, fires can be either man-made or natural. Fuel, oxygen, and a heat source are the three requirements for a wildfire to burn. Trees, grasses, bushes, and even structures that are flammable and surround a fire are all regarded as fuel. A fire's strength varies according on the amount of fuel that is available in a given area. Air provides oxygen, which is necessary for a fire to ignite. Heat sources help the wildfire start by elevating the fuel's temperature. Lightning, smoking cigarettes or campfires, hot gusts, and even the Sun can start a wildfire.

IV. METHODOLOGY

A. Data Source

Forest fires are a major issue for the preservation of Tropical Forests. Knowing the frequency of forest fires over time can help avoid them. Brazil is home to the world's biggest jungle, the Amazon.

This dataset shows the number of forest fires in Brazil, broken out by state. The series spans around ten years (1998 to 2017). The data was collected from the Brazilian government's official website.

B. Using Alogrithms

The six primary algorithms employed in this system are described in this section.

1) Logistic Regression (LR)

Logistic Regression belongs to the most popular machine learning Algorithms.This method belongs to supervised Learning algorithm. By using this technique, the dependent variable can be predicted from a collection of independent factors. The outcome of a discrete dependent variable is predicted using this model. Therefore, requiring that the output be discrete. It can be either True (or) False.

An alternative structure involves replacing the inception net's first two convolutional layers with a modified version of the Szegedy, inception module. It is the objective of the chapter to expand workstations in which the demonstration is optimized through the stacking of numerous component figures.

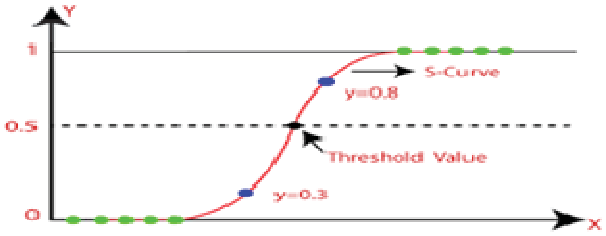


Fig. 1. LR

To avoid tall semantics, we recommend using 3x3 extended convolutions [3] instead of the first module's 5x5 convolution operation. The idea of using wider convolution operations

2) Support Vector Machine Algorithm (SVM)

For both regression and classification issues, Support Vector Machine (SVM) is used. This model is under the supervised Learning algorithm. Support vector machine selects the maximum points the that support in creation of a hyperplane. The hyperplane is the finest border. These maximum cases referred as a support vectors [5-9].

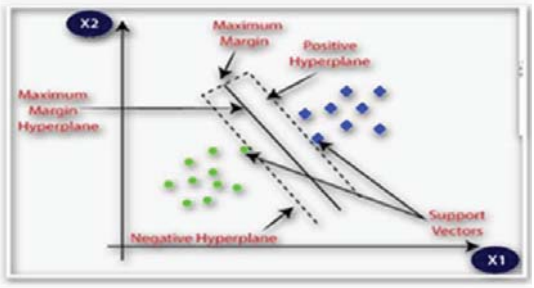


Fig. 2. SVM

3) K-Nearest Neighbor (KNN)

Although it may be used to solve both classification and regression issues, K-Nearest Neighbor is primarily employed for class-action problems. It is a member of the supervised machine learning algorithms family. The KNN is also called as a lazy learner algorithm. This model is easy compare to all other machine learning algorithms [10-13].

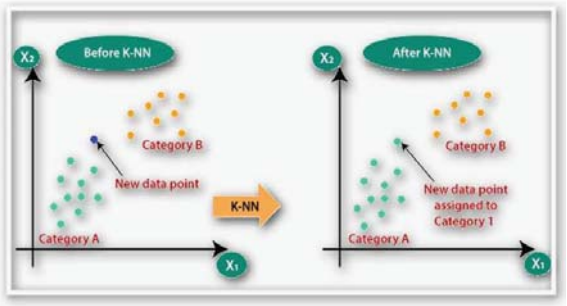


Fig. 3. KNN

4) Random Forest Algorithm

Both regression and classification issues can be solved with the Random Forest approach. It is mainly based on the idea of group (or) class learning, which is nothing but the combination of the multiple classifiers to solve the complicated problems and to improve the performance of the algorithm. To improve the dataset predicted accuracy, random forest algorithm uses a greater number of decision trees on many subsets of original dataset.

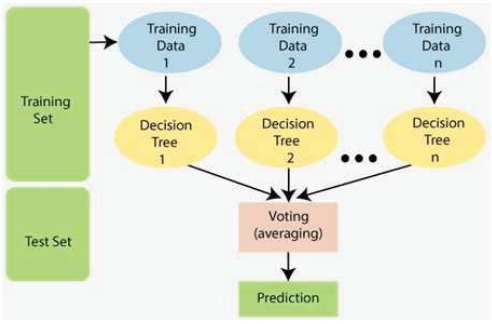


Fig. 4. Random Forest

V. GRADIENT BOOSTING CLASSIFIER

Gradient Boosting Machine is the powerful boosting algorithm compare to all different algorithms. Both continuous and categorical variables can be predicted using

the gradient boosting. GBM is the concept of ensemble learning, the boosting technique combines many easy models to make the final output.

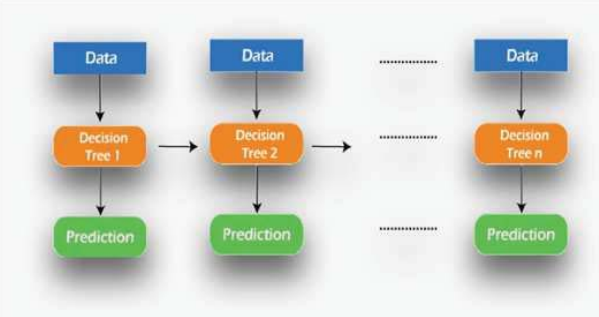


Fig. 5. GBM

A. Decision Tree

Decision trees are used to solve classification and regression issues; however, they are most frequently employed to solve classification issues. Each leaf node in the tree-structured classifier represents an outcome, with internal nodes representing the dataset properties, branching representing the decision rules. It has mainly two nodes the decision node and the leaf node. Multiple branches are found on decision nodes. The leaf nodes exist alone without any further branches.[15-16].

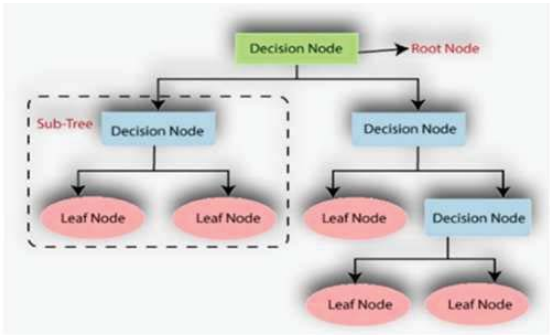
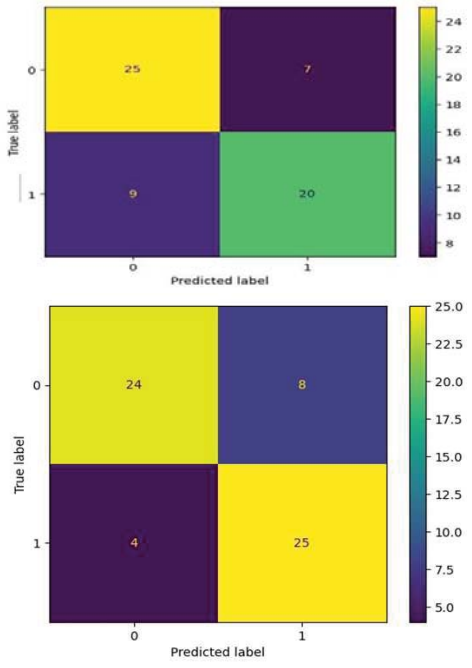


Fig. 6. Decision Tree

7. Confusion matrices for different algorithms



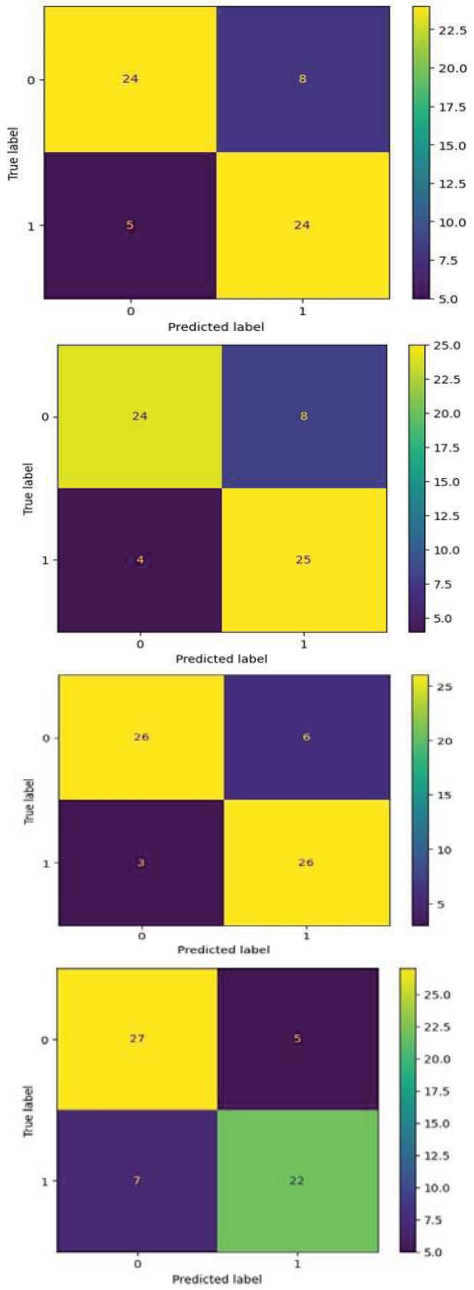


Fig. 7. Confusion Matrics

VI. Result And Analysis

The outcomes demonstrate how a data set and six machine learning algorithms were utilised to forest fire prediction

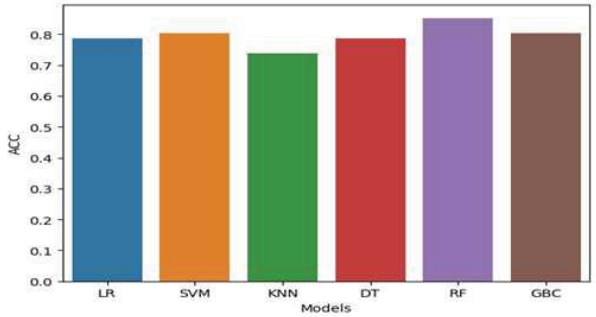


Fig. 7. Prediction of six algorithms

VII. Conclusion And Future Scope

We may increase the accuracy of random forest and the boosting approach by fine-tuning the parameters and adding other variables such as forest vegetation, forest cover, tree type, and buildup index.

This study primarily seeks to create a predictive model using climatic factors. We can distinguish between minor and major flames by identifying the charred area. The project's future work can be accomplished by developing probabilistic models that, by applying specific conditions, can pinpoint the fire's source.

On the basis of forest fire predictor can be predicted using the six machine learning algorithms KNN, SVM, GBC, DT, LR, and RF. The Random Forest algorithm success rate of 88% top out of six algorithms.

References

[1] Diwakar Pant, Sandeep Verma, Piyush Dhulia, "A study on disaster detection and management using WSN in Himalayan region of Uttarakhand", Advances in Computing Communication & Automation (ICACCA)(Fall)2018 3rd International Conference on, pp. 1-6 2018.

[2] Thangavel, S., & Selvaraj, S. (2023). Machine Learning Model and Cuckoo Search in a modular system to identify Alzheimer's disease from MRI scan images. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 1-9. https://doi.org/10.1080/21681163.2023.2187239

[3] Evizal Abdul Kadir, Sri Listia Rosa, An Yulianti, "Application of WSNs for Detection Land and Forest Fire in Riau Province Indonesia", Electrical Engineering and Computer Science (ICECOS)2018 International Conference on pp. 25-28, 2019.

[4] Saravanan, T., & Saravanakumar, S. (2022). Enhancing investigations in data migration and security using sequence cover cat and cover particle swarm optimization in the fog paradigm. International Journal of Intelligent Networks, 3, 204-212.

[5] George E. Sakr, Imad H. Elhajj, George Mitri and Uche Chukwu C. Wejinya "Artificial Intelligence for Forest Fire Prediction", Advanced Intelligent Mechatronics ,Montréal, Canada, July 6-9, 2019

[6] Saravanan, T., & Nithya, N. S. (2019). Design of dynamic source routing with the aid of fuzzy logic for cross layered mobile ad hoc networks. International Journal of Engineering and Advanced Technology, 8(6), 4241–4247. https://doi.org/10.35940/ijeat.F9024.088619

[7] Zhu H, Gao D, Zhang S. A perceptron algorithm for forest firepredictionbased on wireless sensor networks. Journal on Internet of Things. 2019; 1(1):25-31.

[8] Saravanan, T., & Saravanakumar, S. (2022, February). Modeling an Energy Efficient Clustering Protocol with Spider Cat Swarm Optimization for WSN. In 2022 IEEE VLSI Device Circuit and System (VLSI DCS) (pp. 188-193). IEEE.

[9] Tehrany MS, Jones S, Shabani F, Martínez-Álvarez F, Tien Bui D. novel ensemble modeling approach for the spatial prediction of tropical forest fire susceptibility using LogitBoost machine learning classifier and multi-source geospatial data. Theoretical and Applied Climatology. 2019; 137(1):637-53.

[10] Saravanan, T., Ramalingam, D., Keerthika, K., & Sathish, T. (2022, December). Convolutional Neural Network-based Covid-19 Analysis with Internet of Things. In 2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE) (pp. 1-6). IEEE.

[11] Saravanan, T., Sathish, T., & Keerthika, K. (2022, December). Forecasting Economy using Machine Learning Algorithm. In 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT) (pp. 1-5). IEEE.

[12] Saravanan, T., & Nithya, N. S. (2018, December). Energy aware routing protocol using hybrid ANT-BEE colony optimization algorithm for cluster-based routing. In 2018 4th International Conference on Computing Communication and Automation (ICCCA) (pp. 1-6). IEEE.

[13] Saravanan, T., & Nithya, N. S. (2018, December). Energy Aware Routing Protocol Using Hybrid ANT-BEE Colony Optimization Algorithm for Cluster Based Routing. In 2018 4th International Conference on Computing Communication and Automation (ICCCA) (pp. 1-6). IEEE.

[14] Saravanan, T., & Nithya, N. S. (2020, December). Mitigation of attack patterns based on routing reliance approach in MANETs. In 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN) (pp. 387-392). IEEE.

[15] Saravanan, T., & Saravanakumar, S. (2021, December). Privacy Preserving using Enhanced Shadow Honeypot technique for Data Retrieval in Cloud Computing. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 1151-1154). IEEE.

[16] Saravanan, T., Jhaideep, T., & Bindu, N. H. (2022, April). Detecting depression using Hybrid models created using Google's BERT and Facebook's Fast Text Algorithms. In 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 415-421). IEEE.