IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13240

PREDICTING WILDFIRES USING MACHINE LEARNING TECHNIQUES

Kavitha. R[1], Iraniyan Pandian[2], Kumaran. M[3]

Student M.E (CSE), Jaya Engineering College, Chennai, India[1]

Professor, Department of CSE, Jaya Engineering College, Chennai, India[2]

Professor, Department of CSE, Jaya Engineering College, Chennai, India[3]

Abstract: Forest fires have become one of the most serious issues. Forest fires have a significant influence on ecosystems and have a significant impact on greenhouse gas and aerosol levels in the atmosphere. Wildfires have devastated a large quantity of forest and wildlife as a result of these fires. Forest fires are caused by two major factors: global warming caused by an increase in the average temperature of the earth, and human irresponsibility. Predictions must be made to discover sections of land that have the potential to burn and lead to a large forest fire based on meteorological conditions in order to prevent forest fires. Our suggested system will focus on parameters such as temperature, humidity, and other variables that contribute to wildfires. There are a variety of fire detection algorithms available, each with its own approach to the problem.

Keywords: Training, satellites, wind speed, forestry, predictive models, decision trees, wind forecasting.

I. INTRODUCTION

Machine learning (ML) is a branch of research concerned with understanding and developing methods that "develop" - that is, approaches that use data to improve performance on a set of tasks. It is regarded as a component of artificial intelligence. Machine learning algorithms construct a model from sample data, referred to as training data, in order to make predictions or judgements without being explicitly programmed to do so. Machine Learning algorithms are utilised in a wide range of applications, including medical, email filtering, speech recognition, agriculture, and computer vision, where developing traditional algorithms to execute the required tasks would be difficult or impossible. However, not all machine learning is statistical learning. Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, and to uncover key insights in data mining projects.

Mathematical optimisation research provides methodology, theory, and application fields to the subject of machine learning. Data mining is a closely connected topic of research that focuses on exploratory data analysis via unsupervised learning. Some machine learning implementations use data and neural networks to replicate the operation of a biological brain. Machine learning is sometimes known as predictive analytics when used to commercial concerns.

The premise behind learning algorithms is that approaches, methods, and inferences that have proven successful in the past are likely to do so again in the future. These conclusions can be straightforward, such "since the sun has been rising every morning for the last 10,000 days, it will probably rise again." likewise for tomorrow morning. They can be subtle, like "Y% chance that undiscovered black swans exist because X% of families have geographically separate species with color variants."

Programs that use machine learning can do tasks without having them explicitly coded. Computers use available data to learn in order to do specific jobs. For straightforward jobs given to computers, it is easy to build algorithms that instruct the device how to carry out all the steps necessary to address the issue at hand; no learning is required on the part of the computer.

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. Additionally, it provides an excellent way for employees or business owners to present data to non-technical audiences without confusion. In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13240

II. DATA ANALYTICS

Data analytics (DA) is the process of examining data sets to find trends and draw conclusions about the information they contain. Increasingly, data analytics is done with the aid of specialized systems and software. Data analytics technologies and techniques are widely used in commercial industries to enable organizations to make more-informed business decisions. Scientists and researchers also use analytics tools to verify or disprove scientific models, theories and hypotheses.

As a term, data analytics predominantly refers to an assortment of applications, from basic business intelligence (BI), reporting and online analytical processing (OLAP) to various forms of advanced analytics. In that sense, it's similar in nature to business analytics, another umbrella term for approaches to analyzing data. The difference is that the latter is oriented to business uses, while data analytics has a broader focus. The expansive view of the term isn't universal, though: In some cases, people use data analytics specifically to mean advanced analytics, treating BI as a separate category.

Data analytics initiatives can help businesses increase revenue, improve operational efficiency, optimize marketing campaigns and bolster customer service efforts. Analytics also enable organizations to respond quickly to emerging market trends and gain a competitive edge over business rivals. Depending on the application, the data that's analyzed can consist of either historical records or new information that has been processed for real-time analytics. In addition, it can come from a mix of internal systems and external data sources.

At a high level, data analytics methodologies include exploratory data analysis (EDA) and confirmatory data analysis (CDA). EDA aims to find patterns and relationships in data, while CDA applies statistical techniques to determine whether hypotheses about a data set are true or false. EDA is often compared to detective work, while CDA is akin to the work of a judge or jury during a court trial -- a distinction first drawn by statistician John W. Tukey in his 1977 book Exploratory Data Analysis. Data analytics can also be separated into quantitative data analysis and qualitative data analysis. The former involves the analysis of numerical data with quantifiable variables. These variables can be compared or measured statistically. The qualitative approach is more interpretive -- it focuses on understanding the content of non- numerical data like text, images, audio and video, as well as common phrases, themes and points of view.

Data warehouses typically store structured business data. However, most data in organisations is unstructured data, which occupies about 80 per cent of an enterprise's data [13], [14]. Therefore, it is vital to transform the traditional data warehouse into an efficient unstructured data warehouse. Big data refers to huge data sets characterised by larger volumes and greater variety and complexity, generated at a higher velocity than the normal operational data that an organisation has handled before. As more and more enterprises recognise the values and advantages associated with big data insights, the adoption of big data tools like Hadoop ecosystem is growing. Hence, utilising big data tools as an enhancement to the data warehouse to handle unstructured data besides structured one is a feasible and practical approach to resolve the limitation of the traditional data warehouse and potentially expand its adoption in organisations.

III. LITERATURE REVIEW

Preeti T, Dr. Suvarna Kanakaraddi, Aishwarya Beelagi, Sumalata Malagi, Aishwarya Sudi "Forest Fire Prediction Using Machine Learning Technique",2021 The main aim of Prediction of forest fire id expected to reduce the collision of forest fire in the future. There are so many fire detection algorithms are available with many different approaches towards the detection of fire in the forest. In the existing work processes the fire affected part is predicted and analyzed based on the satellite images. The meteorological parameters such as rain, temperature, wind and humidity were used To predict the occurrences of a forest fire in the proposed system.

B. Pratima Chaubey, Nidhi J. Yadav, Abhishek Chaurasiya "Forest Fire Prediction System using Machine Learning ",2020 The forest fires that are occurring in the forest region or wild land and are uncontrolled fires and cause significant and more damage to human and natural resources, which are one of the most dangerous disaster to the ecological environment .The existing systems use various technology like Machine learning techniques and Artificial Intelligence and Wireless network utilized for collecting 24- hour weather data continuously, which provides a higher chance to reflect perfectness of the status of forest environment. Depending on those system, we can decide which days have the highest possibility of catching a forest fires and danger and paid special attention to prevent forest fire for forest guards.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024
DOI: 10.17148/IJARCCE.2024.13240

C. Amira. A. Elsonbaty, Ahmed M. Elshewey "Forest fire Detection Using Machine Learning Technique",2020 Stated that Nowadays, Forest fires became one of the major important problems that cause huge damages to several areas around the globe. This paper displays predicting forest fire-prone areas using machine learning regression techniques. The data set used in this paper is presented within the UCI machine learning repository that consists of climate and physical factors of the Montesinos park which is present in Portugal. The research also proposes many machine learning approaches, linear regression, ridge regression and lasso regression algorithm with data set size of 13 features and 517 entries for each row. The accuracy of the linear regression algorithm gives higher accuracy than ridge regression and lasso regression algorithms.

D. Pranjali Bora, Sandeep Sharma, Sandeep Banerjee, K. Sudha, M. Sravanisai, "prevention part I: Prediction and web-based analysis",2022 Stated that this particular paper elaborates an approach to develop a system which predicts the catching of forest fires with accuracy. The paper introduces a new approach for predicting forest fires that includes a user-friendly web application for easy access. A FWI system is developed with the help of IBM Python Flask service to help in the predicting of forest fires. IBM Auto-AI service has also been used in predicting the onset of forest fires and present very useful information. The prediction results are displayed in a web application deployed on the IBM cloud. The system is expected in helping with the growing concerns over forest fires and reduce the impact of the disasters as much as possible around the globe.

E. Ayu Shabrina, Intan N. Wahyuni, Rifika Sadikin, Arninda L. Latifah "Evaluation of Random Forest model for forest fire prediction based on climatology over Borneo" 2019 States that Indonesia has entered an alarming condition related to catching of the forest fires. It has become a seasonal hazardous phenomenon in tropics in the Indonesia. As it is the largest tropical forest in Indonesia, Borneo is the most susceptible area to catch fire especially in dry condition. Forest fires are threatened by human activities, ecosystem and climate processes, but in Borneo only variable of climate can be quantified the research objective is to assess the effectiveness of the random forest model in predicting forest fires using satellite data of burned areas and climate variables as input. Prediction of forest fires is expected to reduce the influence of forest fires in the future Through an analysis of annual and spatial variability, it was found that the random forest model, incorporating all selected climate variables, effectively represents forest fire events across the Borneo region of Indonesia.

F. Salma A Sahel, Samar O Alosaimi, Muhammad Arif, Khloud K Alghamdi Mashael E Alsahafi, Maram A Alharthi and Maryam Arif "Role of Machine Learning in Forest Fire Management" 2021. Across the globe Forest fire disasters are recently getting lots of attention due to the climate change. Globally, changes in the climates are also affects rapidly changing of the fire patterns on Globe. Accurate information on the occurrence of fires, their impact on the environment, and their spread is essential for effective fire management. Predicting of fire activities in the forest is beneficial to the authorities to make efficient, optimal and sound decisions in fire management system. The objective of the paper is to provide a summary of the latest advancements in predicting forest fires, estimating their spread rate, detecting them, and mapping the areas that have been burned.

G. Muhammad Arif, Khloud K Alghamdi , Salma A Sahel1 , Samar O Alosaimi1 , Mashael E Alsahafi1 , Maram A Alharthi1 and Maryam Arif In this research work they principally introduced effective fire management. Accurate information is necessary for managing effectively. The focus of this paper is to provide an overview of the latest developments in predicting, detecting, estimating the spread rate, and mapping the burned areas of forest fires.

H. MENG ZHANG, HAO LIANG, AND HAILAN WANG "A Neural Network Model for Wildfire Scale Prediction Using Meteorological Factor In the early stages of a wildfire, the predictive model allows fire rescuers to take necessary actions to reduce the extent of damage based on the predicted severity of the fire. Prediction models were developed using a back propagation neural network (BPNN), a long short-term memory (LSTM) network, and a recurrent neural network (RNN), taking meteorological factors as input values. Of these classification methods, LSTM exhibited the greater accuracy of 90.9%.

IV. PROPOSED METHODOLOGY

To design and develop a user-friendly web-based forest fire prediction system using Machine Learning.

The proposed system is depicted below. The user interface (UI) of the system comprises the following input fields for manual entry by the user:

Latitude of the Location
Longitude of the Location

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13240

Brightness Intensity
Day/Night Indicator
Fire Radiative Power (FRP)
Types of Forest Fire
Scan Level
Year of First Fire/Previous Fire Occurrence (YYYY format)
Month of First Fire/Previous Fire Occurrence (month value format)
Date of First Fire/Previous Fire Occurrence (DD format)

a. The user is required to manually input values into the aforementioned fields.

b. A "Predict" button is situated at the bottom of the webpage, enabling the user to initiate forest fire prediction.

c. Upon clicking the "Predict" button, an algorithm is triggered.

d. The algorithm is engineered to optimize outcomes with minimal time consumption. Please refer to [5] in the introduction for details regarding the algorithm's time complexity.

e. The algorithm utilizes a threshold value of 1 to forecast fire occurrence in a specific geographical location.

f. If the algorithm's outcome [2] surpasses the threshold value of 1, the system predicts a high probability of forest fire occurrence.

g. If the algorithm's outcome indicates a probability greater than the threshold, the system predicts a high likelihood of forest fire occurrence.

h. Conversely, if the algorithm's outcome suggests a probability lower than the threshold, the system predicts a low likelihood of forest fire occurrence.

i. The predicted outcome of the algorithm, along with precautionary measures, is displayed on the webpage

Advantages of Proposed system

- Provide security.
- Detailed processing of provided values.
  Provides proper precautions to prevent the recatching of forest fire

V. CONCLUSION AND FUTURE ENHANCEMENTS

Results and analysis will provide valuable insights into the strengths and limitations of the developed system, facilitating further improvements and refinements. To execute the classification algorithms, the tool used is flask webapp data examination. For classification procedure no more than a separation of data is particular from the loaded data. To choose a subset from innovative data, “Select attribute” are utilized by the operative. The preferred subset is then subjected to “X-Validation” operator. It develop the classification representation which is validated by the test data.

The objective of machine learning is to develop a model that can correctly forecast results from input data. The model would be trained using a dataset containing geographical factors (such as Latitude, levels, etc.) This is the case for predicting the forest fire using random forest regression. Once trained, the model can be used to make predictions on previously unexplored data. Metrics like MSE, RMSE and accuracy can be used to assess the model's accuracy.

The efficiency of the Random-forest regression model for predicting forest fire can be inferred from the prediction accuracy, Training time and feature importance. The model can be regarded as a successful tool if it has great precision and accuracy. The model can be regarded as a successful tool if it has great precision and accuracy.

The future work of the project would be the improvement of architecture for light and other weather scenarios. Also, can develop a model for small changes in climate in future. An algorithm for testing daily basis dataset instead of accumulated dataset could be of paramount Importance for further research.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13240

REFERENCES

[1]. Preeti T, Dr. Suvarna Kanakaraddi, Aishwarya Beelagi, Sumalata Malagi, Aishwarya Sudi "Forest Fire Prediction Using Machine Learning Technique",2021

[2]. Pratima Chaubey, Nidhi J. Yadav, Abhishek Chaurasiya "Forest Fire Prediction System using Machine Learning ",2020

[3]. Amira. A. Elsonbaty, Ahmed M. Elshewey "Forest fire Detection Using Machine Learning Technique",2020

[4]. Pranjali Bora, Sandeep Sharma, Sandeep Banerjee, K. Sudha, M. Sravanisai, "prevention part I: Prediction and web-based analysis",2022.

[5]. Ayu Shabrina, Intan N. Wahyuni, Rifika Sadikin, Arninda L. Latifah "Evaluation of Random Forest model for forest fire prediction based on climatology over Borneo" 2019.

[6]. Salma A Sahel, Samar O Alosaimi, Muhammad Arif, Khloud K Alghamdi Mashael E Alsahafi, Maram A Alharthi and Maryam Arif "Role of Machine Learning in Forest Fire Management" 2021.

[7]. Muhammad Arif, Khloud K Alghamdi, Salma A Sahel, Samar O Alosaimi, Mashael E Alsahafi, Maram A Alharthi1 and Maryam Arif .

[8]. MENG ZHANG, HAO LIANG, AND HAILAN WANG "A Neural Network Model for Wildfire Scale Prediction Using Meteorological Factor