International Journal of Trend in Scientific Research and Development (IJTSRD)
Special Issue on Advancements and Emerging Trends in Computer Applications - Innovations, Challenges, and Future Prospects
Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Wildfire Prediction and Prevention

Yash Rajesh Wankhede

PG Student, Department of Science and Technology, G. H. Raisoni University, Amravati, Maharashtra, India

ABSTRACT

Forest fires and extreme wildfire events pose a major threat to ecosystems worldwide. This paper implements various machine learning algorithms for the prediction of forest fires in Northern Thailand, a region which is severely impacted by fire events and the resulting pollution. Using publicly available satellite data of fires and weather information, two prediction models, namely the Random Forest and Support Vector Machines, were developed and tested for their accuracy in forecasting forest fire occurrences. Initial results indicate that both models have an accuracy of approximately 60% in predicting forest fires. The real-time prediction data based on current weather conditions is further displayed in a dashboard. The online dashboard has been integrated with Project FIREfly which is a collaboration with Chiang Mai University and the University of Glasgow to visualize real-time data of forest fires. Through the integration of the predictive models, the online dashboard is able to show the probability of forest fires which improves situational awareness for emergency response services and enables them to take proactive measures in managing forest fires.

KEYWORDS: *Machine Learning, Random Forest, Support Vector Machine, Decision Tree, Weka.*

I. INTRODUCTION

Forests are one of the most important resources of the world's ecological balance. Also, it provides oxygen for people and natural living areas for animals. The world has been losing its forests rapidly as a consequence of wildfires and tree cutting uncontrollably. Wildfire as a natural disaster has severe effects on all living creatures. Also, it has extremely large economic and social consequences. During the last few decades, gigantic wildfires have occurred in various places of the world. Creating a trustworthy model to predict the size of the burned area in a forest fire is necessary to allocate resources optimally for fire departments. In this paper, ML models have been used to predict how much fire will grow using the dataset that includes wind speed, humidity, location information, temperature, etc. The output of prediction is the burning area and its unit hectares. A wildfire susceptibility map for the two fire seasons in the Liguria region in Italy was created and validated by using the Random Forest (RF) method [1]. The susceptibility map was investigated considering the dataset of mapped fire environments covering a 21-year period (1997-2017) and different environmental susceptibility factors. Also, the authors aim to compare the performance of ML models. The proposed model is better than the other models to predict areas which were affected by a fire. Hung Van Le and friends [2] suggested a novel deep neural network model for the prediction of wildfires in a tropical region. They proposed 3 hidden layers to create a wildfire susceptibility map for the Gia Lai province in Vietnam which is called deep neural computing.

A literature review covering 300 publications by the end of 2019 was investigated in [3] to show that ML methods can be used in wildfires. It is shown that the common methods are RF, NN, SVM, Decision trees. Stella and friends use machine learning to address the next day forest fire prediction problem. An ML methodilizing Tree Ensemble and NN, where a large parameter search procedure is performed through cross- validation, has been applied to determine powerful models that are expected to generalize fine on the new data[4] . Meteorological parameters such as temperature, average rains to understand scale of a forest fire can be used. These parameters have been used as a input values for these forecast models, such as long short-term memory (LSTM) backpropagation neural network (BPNN) and recurrent neural network (RNN). The experimental results show, the scale of fire can be predict at the onset onccurence with these informations [5]. ML techniques such as RF, SVM and Logistic Regression (LR) have been exploited to build susceptibility map and compared for the study area of Northern Iran. It was revealed that RF has the highest accuracy and suited for wildfire sensitivty evaluation[6]. Novel gradient boosting models have been applied to predict wildfire activity trained with loss function Extreme-Value theory have been exploited for generate loss function. In the study, the benchmarked against boosting scheme was designed and shown to provide a better proxy for test set performance than pure cross validation. Estimates are compared against reinforcement approaches with different loss functions[7]. BPNN, RNN and LSTM techniques have been applied to data set which include Alberta region meteorological parameters taken from Canadian National Fire Database (CNFDB) [8]. In the study, length of fire time have been exploited along with meteorological parameters to predict burning area.

Authors recommend that to have placed sensors that has massive resolution at the initial phase of fire to predict scale of wildfire. Different synthetic data generation techniques and different ML models have been applied the created synthetic dataset. Results have shown that SVM method has most accuracy to predict large forest fires Uncertainty is big problem to predict fire, in literature multi-fidelity technqies have became attractive from wildfire researchers, recently. Multi-fidelity techniques have been used to understand fire spread als Monte-Carlo and multi level Monte-Carlo simulation methods have been compared[10]. Not only weather parameters also smoke information has been used to predict wildfire events in early stage. LSTM has been applied with convolutional layers for smoke detection and reached high accuracy 97.8% Forecast future wildfires is vital point as well in forest fires management. Daily forest fire probability map forecast has been carried out using deep fully convolutinal neural network called AllConvNet. Authors estimated future burn probability map for next seven days using 2006-2017 wildfire period for Australia.

Predictive models such as RF, LR, Ridge Regression have been applied for estimate burning area size in The dataset contains parameters meausured in wildfires between 1911-2015 in the United States. RF algorithm has better performance than LR and Ridge Regression. Beşli and Tenekeci used the data obtained from the satellites for prediction. Forest fires were estimated using Normalized Difference Vegetation Index (NVDI), Land Surface Temperature (LST) and Thermal anomaly (TA) data calculated from satellite data. Decision trees were used to make predictions from the mentioned data. 70% of the data was used to be used as training and remaining as a test. The average performance of the applied method was determined by repeating the training and testing process 10 times with different data. In the experiments carried out, the fires were predicted correctly with an average sensitivity of 98.62%. The actual situation was determined with an average accuracy of 93.11% RF, linear regression, Stacked Regressor, NN, SVM and KNN algorithms have been used for forecast the burned areas with two different data set. Algorithms have been implemented on the Python environment. Also Data sets have taken from Kaggle and UCI, respectively. Performance of used ML algorithms compared each other. MAE and MAPE error metrics have been used to evaluate the performance of the models[15]. Logistic regression has been applied for predicting areas that can be burned using past meteorological parameters. This technique is easy to implement and also facilitates interpretation of the results obtained and possible duplication of the methodology in other regions or countries. [16]. Trucchia and friends proposed a study which is RF based. Their approach is about to obtain national susceptibility maps in Italy. Each pixel of the study are is classifed by the model. Experimental results show the ability of RF to notice the most senstive areas with defined factors.

II. RELATED WORK

Wildfires are among the most destructive natural disasters, causing environmental degradation, economic losses, and threats to human life. Several studies and technological advancements have been made in wildfire prediction and prevention. This section explores various research works and methodologies related to this field.

1. Machine Learning and AI-Based Approaches

Several studies have applied **machine learning (ML) and deep learning (DL)** techniques to wildfire prediction. Researchers have used historical wildfire data, meteorological conditions, and satellite imagery to train predictive models.

- **Liu et al. (2021)** developed a deep learning-based model using Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to analyze satellite images for wildfire detection.
- **Khorshidi et al. (2022)** implemented a Random Forest-based wildfire risk assessment model using climate and vegetation indices, achieving high accuracy in predicting fire-prone areas.

2. Satellite-Based Monitoring and Remote Sensing

Remote sensing technologies, including satellite imagery and Unmanned Aerial Vehicles (UAVs), have significantly contributed to wildfire monitoring and prediction.

- **NASA's Fire Information for Resource Management System (FIRMS)** provides near real-time satellite imagery for detecting active wildfires.
- **The European Space Agency (ESA)** utilizes the Sentinel-2 satellite to track vegetation dryness and detect early wildfire threats.

3. Weather and Climate-Based Models

Meteorological conditions play a crucial role in wildfire occurrence. Several studies have focused on climate- based prediction models.

- **Jolly et al. (2015)** analyzed the impact of climate change on global wildfire risk and developed the **Fire Weather Index (FWI)** to assess fire-prone conditions.
- **Abatzoglou & Williams (2016)** linked increasing wildfires in the western United States to rising temperatures and drought conditions.

4. IoT and Sensor-Based Wildfire Detection Systems

The Internet of Things (IoT) and sensor networks have been used to develop **real-time wildfire detection and prevention** systems.

- **Ahmed et al. (2021)** developed an IoT-based early fire detection system using temperature, humidity, and smoke sensors.
- **López et al. (2022)** proposed a wireless sensor network (WSN) combined with AI algorithms to detect wildfire outbreaks with high accuracy.

III. DATA AND SOURCES OF DATA

Accurate and reliable data is crucial for developing an effective wildfire prediction and prevention system. The data used in this project can be categorized into satellite imagery, meteorological data, historical wildfire records, geospatial data, and sensor-based real-time data. Below, we discuss the sources and significance of each type of data.

1. Satellite Imagery and Remote Sensing Data

- Satellite-based remote sensing provides large-scale wildfire monitoring and prediction. High-resolution images help detect fire-prone areas, vegetation dryness, and active wildfires.

Sources:

- **NASA's Fire Information for Resource Management System (FIRMS):** Provides near real-time fire detection data using MODIS (Moderate Resolution Imaging Spectroradiometer) and VIIRS (Visible Infrared Imaging Radiometer Suite). (Website: https://firms.modaps.eosdis.nasa.gov)
- **European Space Agency (ESA) Sentinel-2:** Captures high-resolution satellite images useful for analyzing vegetation indices and fire susceptibility. (Website: https://sentinels.copernicus.eu)

2. Meteorological and Climate Data

Weather conditions such as temperature, wind speed, humidity, and precipitation play a significant role in wildfire occurrence and spread.

Sources:

- **National Oceanic and Atmospheric Administration (NOAA):** Provides global climate and weather data, including temperature and wind patterns. (Website: https://www.noaa.gov)
- **Indian Meteorological Department (IMD):** Offers real-time weather data and historical climate records for wildfire-prone areas in India. (Website: https://mausam.imd.gov.in)

IV. RESEARCH METHODOLOGY

This paper employs a quantitative research design that targets the forested areas of Chiang Mai, Thailand. It utilises meteorological data, including temperature, pressure, wind, gust, and humidity, as inputs for the ML models, namely Random Forest and SVM. Figure 1 presents an overview of the implemented methodology and tools used for this work. As detailed below, an important considers the data collection and initial analysis to derive probabilities of forest fires. The model training uses these data to develop the ML models considered in this work. Finally, the models are deployed and integrated with live weather data in a dashboard developed for Project FIREfly.

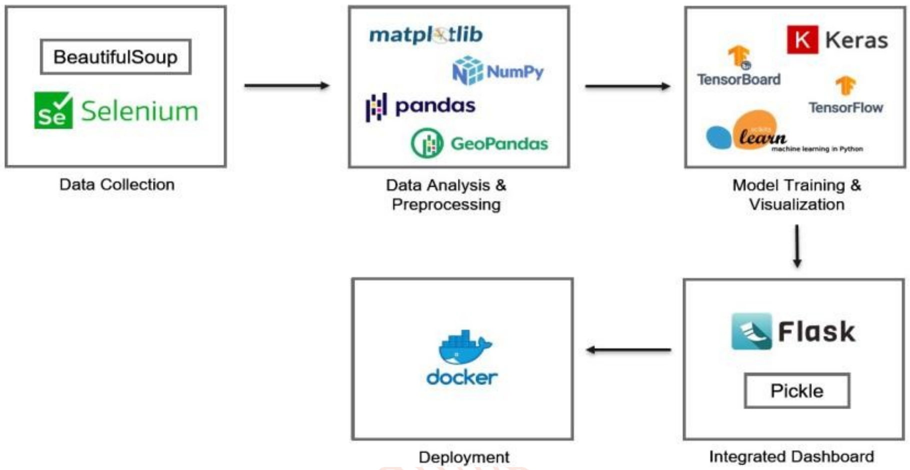

Fig.1 Overview of methodology

4.1. Data Collection

The primary dataset for this study is sourced from NASA FIRMS (Fire Information for Resource Management System), specifically utilizing the yearly summaries of active fire records dataset, measured by VIIRS S-NPP (Visible Infrared Imaging Radiometer Suite on the Suomi National Polar- Orbiting Partnership satellite) [13]. The VIIRS S-NPP dataset encompasses all historical fire occurrences categorized by country and location, thus encapsulating all fire incidents within Thailand from 2012 to 2021. VIIRS S-NPP was preferred over MODIS due to its enhanced spatial resolution, providing a detailed record of fires occurring around the world.

To extract data pertinent to the Chiang Mai region, the geographical boundaries of the area were manually traced out via Google Maps as shown in Figure 2, providing the necessary longitude and latitude ranges to be exported into a CSV file format for effective filtering of the FIRMS's provided dataset. This geospatial analysis pipeline, implemented in Python using GeoPandas, converts the Well- Known Text (WKT) strings to Shapely geometries and organizes them into a GeoDataFrame with a specified coordinate reference system (CRS). The Chiang Mai boundary is projected to Web Mercator (EPSG:3857) for compatibility with a basemap. These steps allow the original VIIRS S-NPP dataset to be crossreferenced against the delineated boundary, resulting in a refined dataset containing only the fire incidents within the Chiang Mai region.

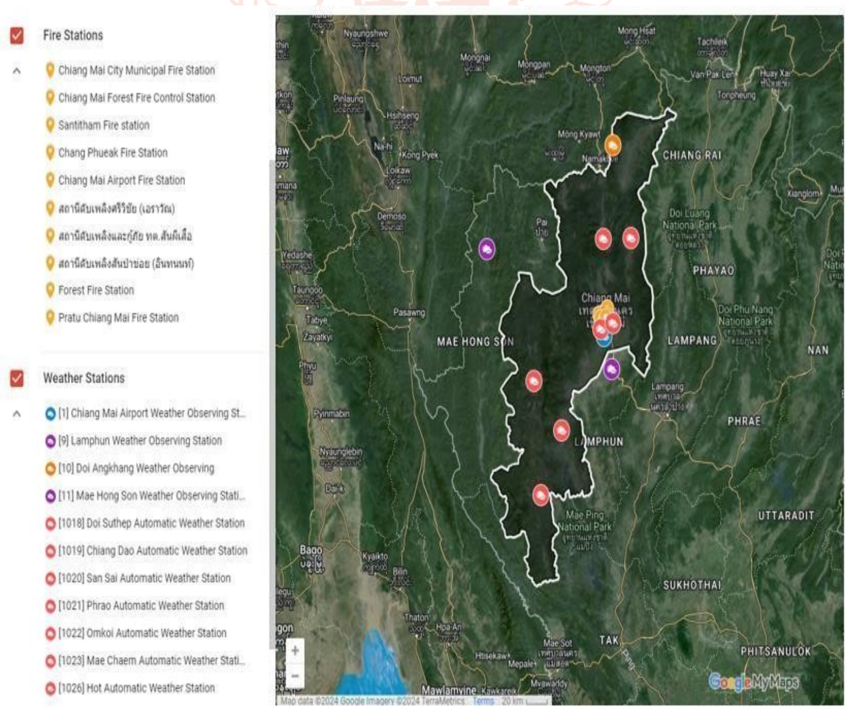

Fig.2 Chiang Mai polygon data

It is noted that the TMD data exhibits gaps for the years 2013, 2015, 2016 and 2021. To ensure completeness and accuracy in the meteorological dataset, the missing data are supplemented from Wunderground, a reputable commercial weather service. This data collection process involves utilizing a Python script designed to enhance efficiency and automate the extraction of data from the different sources. Initially, the script reads a list of URLs from a file. It then initiates a request to the respective websites and parses the HTML response. The script is tailored to filter weather data based on a specific date range and temperature unit, ensuring precision and relevance to the requirements. The script then extracts the target data from the HTML tables and systematically saves the data to a CSV file which can be used for training. Figure 3 shows a comparison of the data extracted from TMD and Wunderground for the dry season in 2012. It is apparent that both sources capture comparable data during that period which was also observed for other periods where overlapping data was available (not shown here). This provides confidence that the data stitching approach in this work provides reliable data for the ML training.

4.2. Data Analysis and Preprocessing

Figure 3 presents the meteorological data which has been collected for the years 2012-2021 as discussed in Section 2.1. The data clearly indicates the dry seasons during December to March as reflected by low precipitation which appears to be consistent for every year in the series The relative humidity also reduces during that period with the lowest humidity during the months of March and April Temperature and wind speed tend to peak towards the end of the dry season in April presents the time history of fire occurrences within the Chiang Mai region extracted from the VIIRS S-NPP satellite data. The data exploration over the different years demonstrates that most fires in the region occur consistently during the months February to April. This corresponds to the dry season as observed in Figure 3. While the fire occurrence peaks in February for most years in the series, 2016 and 2019 show some interesting trends. In both years the rain season started later causing unusually dry April months (compared to other years). We can observe that for both these years, the fire occurrences lasted until April indicating the obviously strong correlation between precipitations, humidity versus likelihood of fire occurrences.

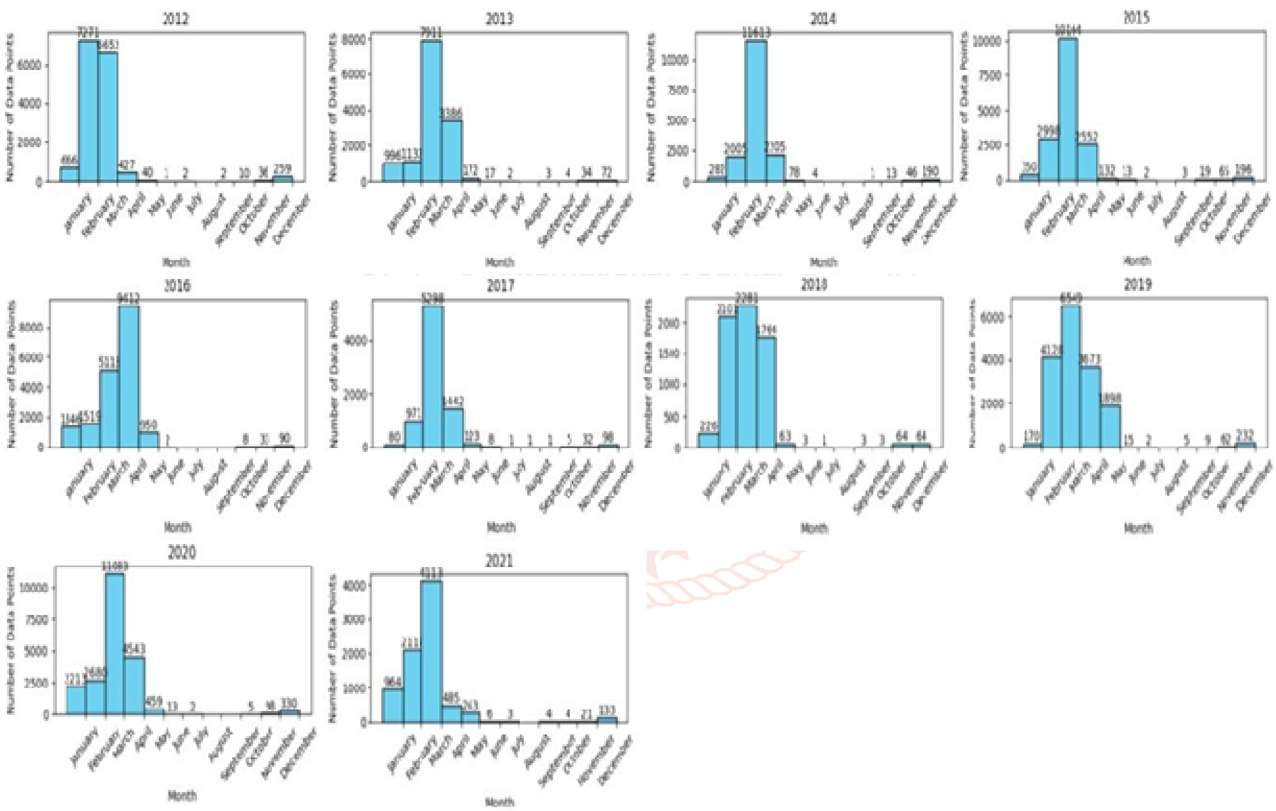

Figure.3. Distribution of fire points within Chiang Mai region across the months during 2012-2021

V. RESULTS AND DISCUSSION

Both the Random Forest and Support Vector Machine (SVM) algorithms were rigorously evaluated in the context of forest fire detection and prevention. The performance of these two approaches was systematically compared using a range of metrics suitable for assessing efficacy in imbalanced datasets [15]. Key evaluation metrics employed in this comparison include precision, recall, F1-score, and the Area under the ROC Curve (AUC-ROC). These metrics provide a comprehensive evaluation of the effectiveness of both Random Forest and SVM algorithms in addressing the critical task of forest fire detection. In Table 1, explanations of these key evaluation metrics are provided to understand their roles in assessing the performance of the Random Forest and SVM algorithms.

Table1: Key Evaluation Metrices

Metric	Description
Precision	Quantifies the proportion of true positive predictions among all positive predictions.
Recall	Measures the proportion of true positives among all actual positive instances.
F1-score	Harmonic mean of precision and recall, providing a balanced assessment of model Performance.
AUC-ROC	Evaluates the model's ability to distinguish between positive and negative instances.

5.1. Random Forest

The evaluation of the Random Forest algorithm for forest fire detection has yielded promising results as shown in Table 2. As per the classification report, the model achieved an accuracy of 59%, which reflects a moderate predictive performance typical for complex classification tasks like fire risk prediction. It is noteworthy that the precision for class 0 (representing no fire occurrences) is high at 97%, indicating a low rate of false positives.

However, the precision for class 1 (indicating fire occurrence) is comparatively lower at 7%, suggesting room for improvement in accurately identifying positive instances. In terms of recall scores, reflecting the model's ability to capture true positive instances, they are 59% for class 0 and 66% for class 1. The F1-score, a metric balancing precision and recall, is 73% for class 0 and 13% for class 1. The area under the receiver operating characteristic curve (AUC-ROC) is calculated at 0.6661, suggesting a moderate level of discrimination ability. In summary, while the random forest model performs well in predicting instances belonging to class 0 compared to class 1, the AUC-ROC value of 0.67 indicates that the model is only moderately effective in distinguishing between the two classes. Therefore, further enhancements are necessary to bolster its performance in accurately identifying fire occurrences.

5.2. Support Vector Machine (SVM)

In our examination of forest fire detection utilizing the Support Vector Machine (SVM) algorithm, we also analyzed performance metrics presented in Table 2. The SVM model generated an AUC-ROC score of 0.70, signifying a modest yet enhanced ability to differentiate between positive and negative instances compared to prior models like the Random Forest algorithm. Notably, there were substantial discrepancies in precision and recall scores: precision was marked at 8% and recall at 71% for class 1, whereas for class 0, precision stood at 98% with a recall of 61%. This highlights the model's struggle in accurately categorizing positive instances while minimizing false positives, a challenge akin to that observed with the Random Forest algorithm. The F1-score, gauging the equilibrium between precision and recall, was computed at 14% for class 1 and 75% for class 0, indicating suboptimal overall performance, albeit marginally improved compared to the Random Forest model. Additionally, the SVM model attained a 61% accuracy on the test dataset, aligning with the Random Forest model's performance, thus indicating limited efficacy in classifying both positive and negative instances.

Metric	Random Forest		Support Vector Machines	
	Class 0 – No Fire	Class 1 – Have Fire	Class 0 – No Fire	Class 1 – Have Fire
Precision	0.97	0.07	0.98	0.08
Recall	0.59	0.66	0.61	0.71
F1-score	0.73	0.13	0.75	0.14

5.3. Web Application Dashboard

To present the results of the machine learning algorithms to the concerned parties, a dashboard is employed to visualise the predicted fire spots, offering real-time insights through an intuitive design and interactive visuals. Upon entering the webpage, users will be directed to a landing page which contains three main sections, specifically the forecasted weather bar, a map dashboard and details section. Forecasted weather details are retrieved from an online weather service API, ensuring the inclusion of up- to-date meteorological information in the predictive analysis as shown in Figure 7. Given that the algorithm solely offers predictions, it is imperative to involve fire personnel or employ alternative methods, such as UAV surveillance [4], to verify the accuracy of the predictions and discern between true alarms and false positives. While the red markers represent forecasted fire occurrence upon successful check of the actual location, the fire personnel can update the database to indicate that the alarm has been verified. This is represented by the blue icons in fig.3

Figure 4. Landing page with map dashboard and integration of weather forecast for Chiang Mai.

Date	Longitude	Latitude	State	Probability	Alert	Nearest Station
25/10/23	18.183458	98.435448	Forecasted	51.2%	False	Chiang Mai City Municipal Fire Station
25/10/23	19.689024	98.977913	Cleared | False Alarm	72.0%	True	Chiang Mai City Municipal Fire Station
25/10/23	18.355873	98.370499	Forecasted	51.0%	False	Chang Phueak Fire Station

Figure 5. Landing page displays details for stakeholders

Accessible to the public, the landing page displays critical information including date, longitude, latitude and the fire prediction details This serves as an alert system, notifying community members of potential fire spots and facilitating early evacuation measured for the affected areas. The state input table is only accessible to fire personnel and administrators where the status of the alarm can be updated from 'Forecasted' to either 'False Alarm' or 'Cleared'. These updates are reflected simultaneously on the main landing page.

VI. CONCLUSION

In this paper, forest fire detection and prevention were explored using machine learning (ML) algorithms such as Random Forest and Support Vector Machine (SVM), emphasizing the potential role of ML-based strategies in safeguarding against such calamities. The significance of employing these algorithms for effective monitoring and management of forest fires, while concurrently emphasizing the involvement of the community in conservation efforts was also discussed. Through this approach, conservatories can assess the efficacy of existing measures and make necessary adjustments, while also fostering community engagement in proactive fire prevention strategies. Upon comparative evaluation, both Random Forest and SVM algorithms showed inherent strengths and weaknesses. Their accuracy hovered around 60%, indicating potential challenges in identifying fires. Random Forest in particular, demonstrated lower precision for class 1 and handling imbalanced data. Additionally, reliance on weather data solely from the Chiang Mai International Airport could impact the precision of both models, considering potential disparities in representing conditions across different geographic locations.

The integration of predictive capabilities into a dashboard to visualize fire hotspots and facilitate informed decision-making for forest management authorities was also discussed in the paper. Forest fire prediction capability can be enhanced by incorporating additional environmental variables, leveraging ensemble methods, or exploring deep learning approaches. By combining ML algorithms with community involvement, Project Firefly aims to establish a comprehensive framework that empowers individuals to play a vital role in preventing forest fires. This inclusive approach not only enhances the effectiveness of conservation measures but also fosters a sense of ownership and responsibility among community members towards protecting our natural resources.

VII. References

[1] 'Huge Forest areas destroyed by raging fires in Chiang Mai'. Accessed: Feb. 29, 2024. [Online]. Available: https://www.nationthailand.com/thailand/general/40035868.

[2] 'Forest fires blanket Chiang Mai in haze'. Accessed: Feb. 29, 2024. [Online]. Available: https://www.bangkokpost.com/thailand/general/2749424

[3] 'Forest fire monitoring and prevention using an UAV-based IOT system: The FIREfly Project'. Accessed: Feb. 29, 2024. [Online]. Available: https://apnic.foundation/projects/firefly-project

[4] P. Puttapirat, K. Woradit, H. Hesse and D. Bhatia, "FireFly Project: UAV Development for Distributed Sensing of Forest Fires," *International Conference on Unmanned Aircraft Systems (ICUAS)*, Greece, Jun. 2024, pp. 594601, doi:10.1109/ICUAS60882.2024.10556892.

[5] V. Sevinc, O. Kucuk, and M. Goltas, 'A Bayesian network model for prediction and analysis of possible forest fire causes', *Forest Ecology and Management*, vol. 457, Feb. 2020, doi:10.1016/J.FORECO.2019.117723.

[6] Usha Prashant Kosarkar, Gopal Sakarkar, Mahesh Naik, " A Hybrid Deep Learning Model for Robust Deepfake Detection", International Conference on Advanced Communications and Machine Intelligence(MICA), 30th & 31st October 2023,pp 117-127, https://doi.org/10.1007/978-981-97-6222-4_9

[7] A. Chaube, "ACO-Enhanced Siamese Networks for Robust Feature Matching in Copy-Move Image Forgery Detection," *2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application (ICAIQSA)*, Nagpur, India, 2024, pp. 1-6, doi:10.1109/ICAIQSA64000.2024.10882433

[8] X. T. Cham, M. L. Soh, F. Trujillano, P. C. Y. Yau, O. Choy, X. Cheh, K. Fornace, N. Poh, C. K. Seow, H. Hesse, Q. Cao, G. J. Garay, "AI-Assisted Manual Segmentation Web Application for Geospatial Satellite and Imagery Data," *IEEE World Forum on Internet of Things*, Portugal, 2023, pp. 1-5, doi:10.1109/WFIoT58464.2023.10539446.

[9] S. Al Janabi, I. Al Shourbaji, and M. A. Salman, 'Assessing the suitability of soft computing approaches for forest fires prediction', *Applied Computing and Informatics*, vol. 14, no. 2, pp. 214–224, Jul. 2018, doi:10.1016/J.ACI.2017.09.006.

[10] G. F. Shidik and K. Mustofa, 'Predicting size of forest fire using hybrid model', Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 8407 LNCS, pp. 316–327, 2014, doi:10.1007/978-3-642-55032- 4_31/COVER.

[11] A. Zaidi, 'Predicting wildfires in Algerian forests using machine learning models', *Heliyon*, vol. 9, p. e18064,

2023, doi:10.1016/j.heliyon.2023.e18064.

[12] L. Gigović, H. R. Pourghasemi, S. Drobnjak, and S. Bai, 'Testing a New Ensemble Model Based on SVM and Random Forest in Forest Fire Susceptibility Assessment and Its Mapping in Serbia's Tara National Park', *Forests*, vol. 10, no. 5, p. 408, May 2019, doi:10.3390/F10050408.

[13] W. Ma, Z. Feng, Z. Cheng, S. Chen, and F. Wang, 'Identifying Forest Fire Driving Factors and Related Impacts in China Using Random Forest Algorithm', *Forests*, vol. 11, no. 5, p. 507, May 2020, doi:10.3390/F11050507.

[14] C. Gao, H. Lin, and H. Hu, 'Forest-Fire-Risk Prediction Based on Random Forest and Backpropagation Neural Network of Heihe Area in Heilongjiang Province, China', *Forests*, vol. 14, no. 2, p. 170, Jan. 2023, doi:10.3390/F14020170.

[15] 'Fire Information for Resource Management System (FIRMS) | Earthdata'. Accessed: Feb. 29, 2024. [Online]. Available: https://www.earthdata.nasa.gov/learn/find-data/near-real-time/firms

[16] 'TMD Open Data'. Accessed: Feb. 29, 2024. [Online]. Available: https://data.tmd.go.th/dataset/index.php

[17] V. López, A. Fernández, S. García, V. Palade, and F. Herrera, 'An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics', *Information Sciences*, vol. 250, pp. 113–141, Nov. 2013, doi:10.1016/J.INS.2013.07.007.

[18] J. Srivastava and A. Sharan, 'SMOTEEN Hybrid Sampling Based Improved Phishing Website Detection', *TechRxiv*, Oct. 2023, doi:10.36227/TECHRXIV.2020