[1]Kais Yahiaoui,
[2]Souheila Bouam,
[3]Abdelhak Gueroui

"Enhancing Wheat Fire Prediction in Barika, Algeria, through Advanced Ensemble Machine Learning Models"

Journal of Electrical Systems

Abstract: - Recent climatic shifts and the growth of agricultural land have escalated the incidence of wheat field fires, presenting severe risks to both food security and local economies. This study aims to develop advanced predictive models to effectively forecast significant wheat fires in Barika, Algeria. We utilized a comprehensive dataset spanning from 2015 to 2023, which includes information on fire incidents and meteorological factors like temperature, humidity, precipitation, and wind speed. A sophisticated ensemble machine learning model was crafted, combining Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Random Forest (RF) in a stacked configuration to predict wheat fire events. Our analysis indicates that the ensemble model significantly outperforms traditional single-model approaches in terms of both accuracy and reliability. Employing these cutting-edge predictive techniques significantly bolsters firefighting measures, enhances resource management, and reduces the adverse effects of fires in agricultural zones. The employment of ensemble learning highlights its utility as a formidable tool in environmental management and crisis response. With more precise forecasts, this model facilitates improved emergency preparedness and strategic intervention plans, aiming to safeguard essential agricultural assets and support rural communities against the backdrop of mounting environmental pressures.

Keywords: Wheat Fire Prediction, Ensemble Machine Learning, Meteorological Data, Agricultural Risk Management, Climate Change Impact

1. INTRODUCTION

Agriculture has become the predominant land use globally, covering nearly 40% of the Earth's surface. This expansion often comes at the expense of forests and grasslands, altering ecosystems and increasing the vulnerability to various environmental hazards, including fires. Wheat, being the world's second most important cereal crop in terms of area and production, holds a crucial economic and nutritional role, especially in regions like Barika, Algeria. The region of Barika, characterized by its arid and semi-arid climate, experiences hot, dry summers with temperatures frequently exceeding 40 degrees Celsius during the harvest period. These conditions are ripe for the outbreak of wheat field fires, which are exacerbated by the region's typical climate and the physical geography of the area. Wheat fields, predominantly flat and expansive, lack natural barriers against the spread of fires, and the climatic conditions during the summer further increase the susceptibility of these fields to fire. Researchers and managers are thus compelled to develop reliable tools and methods for predicting future fire probabilities. This is crucial not only for safeguarding human life and the environment but also for protecting the vital food resources that these agricultural activities support. The focus on wheat field fires is particularly pertinent given the lack of extensive research in this area compared to the more frequently studied forest fires [1].

In recent decades, advancements in machine learning (ML) and deep learning have provided new methodologies to address complex problems such as fire prediction and management. These technologies offer potential improvements over traditional predictive methods, which often fall short in accuracy or timeliness. This study leverages a stacking ensemble learning approach, combining the strengths of Long Short-Term Memory (LSTM), Recurrent Neural Network

1 Department of Computer Science, University of Batna 2, Algeria;
k.yahiaoui@univ-batna2.dz
2 Department of Computer Science, University of Batna 2, Algeria;
s.bouam@univ-batna2.dz
3 Department of Computer Science, University of Versailles;
mourad.gueroui@uvsq.fr

(RNN), and Random Forest (RF) models. This method aims to harness the detailed fire point data and meteorological data available for Barika from 2015 to 2023 to develop a robust predictive model [2].

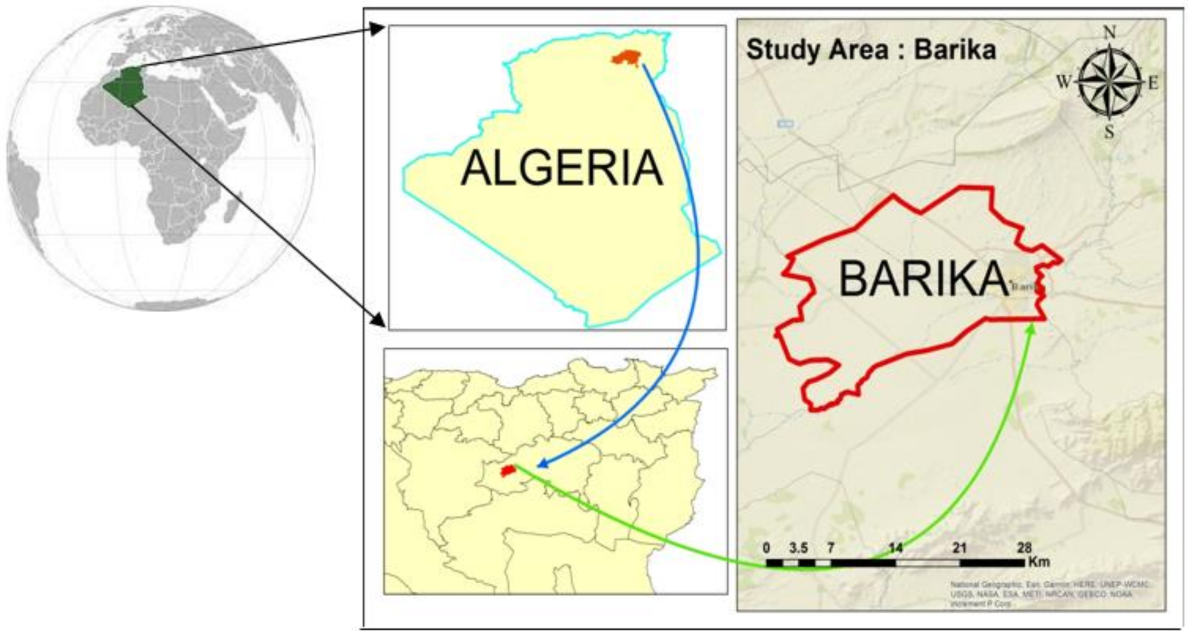


Figure 1. The study area in Barika, Algeria

2. IMPORTANCE OF WHEAT IN BARIKA, ALGERIA, AND THE IMPLICATIONS OF FIRE INCIDENTS ON FOOD SECURITY AND LOCAL ECONOMY

Wheat is a cornerstone of Barika, Algeria's agricultural output and plays a pivotal role in the region's economy and food security. As the world's second most important cereal crop in terms of cultivated area and production, wheat not only feeds the population but also provides employment and income to a substantial portion of Barika's residents, particularly those in rural areas. The region's reliance on agriculture underscores the significance of wheat in sustaining local livelihoods and the broader economy.

Fire incidents in wheat fields carry severe implications. When fires ravage these fields, the immediate loss of crop directly threatens food security. This not only reduces the local wheat supply—increasing reliance on costly imports—but also drives up prices for wheat-based products, which are essential to the daily diet of the local population. The economic fallout extends beyond the farmers, affecting the entire supply chain, including those involved in the processing, distribution, and retail of wheat products. Moreover, the recurring nature of these fires can lead to long-term land degradation. Frequent fires impoverish the soil, diminish its productivity, and alter its composition, which can take years to rehabilitate. This degradation further reduces the region's agricultural output and necessitates increased investment in soil restoration and fire prevention techniques, diverting funds from other critical areas of local governance and development. The economic stability of the region is also at stake. Fire incidents can destabilize the foundational agricultural economy of Barika, leading to financial insecurity for farming families, disrupting market operations, and increasing the vulnerability of the local economy to fluctuations in global wheat prices. Furthermore, the costs associated with firefighting and recovery efforts strain local resources, impacting the overall economic health of the region. Given these significant risks, the development of effective predictive and management strategies for wheat fires is critical. Advanced predictive models, such as those being developed in the study, are essential for timely and effective responses to potential fire outbreaks. By improving fire management and response strategies, Barika can better protect its vital wheat resources, thus supporting the sustainability of its food supply and the robustness of its local economy.

3. OVERVIEW OF EXISTING PREDICTIVE METHODS AND THEIR LIMITATIONS.

1. Historical Trend Analysis: This method involves analyzing past fire incidents to forecast future events based on observed patterns. While historical trend analysis is straightforward and useful for identifying general patterns or seasonal peaks in fire incidents, it cannot account for spontaneous variables or unexpected changes in conditions, making it unreliable for precise or real-time predictions.

2. Meteorological Modeling: More sophisticated than trend analysis, meteorological models use current weather data like temperature, humidity, wind speed, and precipitation to predict fire risks. Although these models are more dynamic and can provide real-time assessments, they often require continuous and precise data inputs and may not fully integrate complex interactions between different environmental factors or sudden meteorological changes.

3. Remote Sensing and Satellite Imagery: These technologies monitor land conditions and environmental changes over large areas, providing valuable data for fire risk assessment. However, the effectiveness of remote sensing is contingent on the resolution and frequency of the imagery; lower resolutions may miss smaller-scale anomalies, and infrequent updates can delay the detection of conditions conducive to fires.

4. Machine Learning and Artificial Intelligence Models: Recently, machine learning (ML) and artificial intelligence (AI) have been applied to predict agricultural fires by analyzing large datasets to identify patterns and predict outcomes. These models can integrate diverse data types and learn from new data, offering significant improvements in prediction accuracy. Nonetheless, they require extensive and diverse training data to function effectively, and they can be opaque, meaning their decision-making processes are not always clear or understandable to humans.

Each of these methods has its strengths, but they also face significant challenges. Historical data analysis and meteorological models can fail under the pressure of non-typical weather patterns or unusual fire causes. Remote sensing is limited by technology and data update frequencies, and AI/ML models face challenges with data quality, availability, and model interpretability. Given these limitations, there is a need for innovative approaches that can combine the strengths of these various methods to improve accuracy and reliability in predicting agricultural fires. This is particularly crucial for regions like Barika, where wheat is a vital economic resource and where effective fire prediction can significantly impact food security and economic stability.

4. OBJECTIVE

The objective of the study is to develop reliable prediction models using meteorological data to establish rigorous and effective wheat firefighting plans in Barika, Algeria. The research focuses on using recently emerged stacking ensemble learning methods, incorporating Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Random Forest (RF), to predict the occurrence of large-scale wheat fires. This is motivated by the need to improve the accuracy of predicting wheat fire occurrences, thus enabling more efficient firefighting efforts and saving time and resources.

5. LITERATURE REVIEW

Early studies primarily focused on the role of climatic factors and their direct implications for fire risk in agricultural settings [3]. For instance, fluctuations in temperature and humidity levels were closely linked to the probability of fire outbreaks, forming the basis for many predictive models With the advent of more sophisticated technologies, the scope of research expanded to include remote sensing and geographical information systems (GIS). These technologies allowed for a broader spatial analysis and facilitated real-time monitoring of fire-prone areas, improving the timeliness and accuracy of predictions [4]. For example, satellite imagery has been used to detect changes in vegetation cover and moisture content, key indicators of fire risk [5]. However, the real breakthrough came with the integration of machine learning techniques in the field. Studies like those by Benos et al. (2021) [6] demonstrated the efficacy of machine learning models, such as Random Forests and Support Vector Machines, in synthesizing and predicting fire risks from complex datasets that include not only meteorological data but also human factors and land-use changes. These models offer the advantage of learning from data without explicit programming, adapting to new patterns as more data becomes available [7]. Despite these advances, the application of machine learning in environmental management is not without challenges. The effectiveness of these models often hinges on the quantity and quality of the data available, which can be a significant limitation in regions with sparse data collection infrastructures [8]. Moreover, the "black box" nature of many machine learning algorithms can make it difficult for practitioners to understand and trust the outputs, complicating their implementation in critical decision-making processes [9].

Title	Methodology	Results	Implications
"Fire Risk Prediction Using Neural Networks” [10].	Development and application of neural network models to predict agricultural fire risks based on environmental and anthropogenic factors.	Demonstrated high predictive accuracy, outperforming logistic regression models.	Suggests that neural networks are a viable tool for real-time fire risk prediction and can significantly aid in disaster preparedness.
"Evaluating Satellite Data For Fire Prediction" [11]	Utilization of multi-spectral satellite imagery to predict fire outbreaks by analyzing vegetation dryness and temperature anomalies.	Successfully identified potential fire zones up to two weeks in advance.	Supports the use of satellite technologies for early fire detection, potentially reducing the extent of damage in agricultural zones.
"Impact of Climatic Change on Fire Regimes in Agricultural Landscapes" [12]	Study of long-term climate data and fire incidents to model future fire regimes under changing climatic conditions.	Found that increasing temperatures and varying precipitation patterns are likely to alter fire regimes significantly.	Calls for adaptive management strategies in agriculture to cope with changing fire regimes due to climate change.
"Socio-economic Factors Influencing Fire Incidents in Agriculture" [13]	Analysis of socio-economic data alongside environmental factors to understand the human dimensions of fire risks.	Identified key socio-economic variables that correlate strongly with fire incidence rates.	Advocates for holistic approaches that consider human factors in fire risk prediction and management.
"Human Activity and Fire Incidence: An Analytical Approach".[14]	Correlational study between human activities, such as agricultural practices and land clearing, and the incidence of fires.	Strong correlation found between certain practices and increased fire risk.	Suggests policy interventions to modify agricultural practices to reduce fire risks and enhance environmental sustainability.

6. METHODOLOGY

1. Long Short-Term Memory (LSTM): LSTM is a type of RNN specialized for remembering information for an extended period. It is particularly well-suited for time-series data where it is crucial to maintain information in long sequences, making it ideal for predicting events based on historical data, such as temperature or humidity trends leading up to a fire.

2. Recurrent Neural Network (RNN): RNNs are designed to handle sequential data by maintaining a memory of previous inputs using their internal state. They are effective for modeling time-dependent data, such as the sequential conditions of weather patterns that might culminate in a fire.

3. Random Forest (RF): RF is an ensemble learning method that operates by constructing a multitude of decision trees during training and outputting the class that is the mode of the classes of the individual trees. It is robust against overfitting and is very effective in classifying complex datasets based on various input features.

7. RATIONALE BEHIND CHOOSING A STACKED ENSEMBLE MODEL

The choice to use a stacked ensemble model combining LSTM, RNN, and RF is driven by the desire to leverage the unique strengths of each model to enhance the overall predictive accuracy. By stacking these models, the ensemble method can:

Utilize the temporal data processing capabilities of LSTM and RNN for handling sequential weather and fire data.

Benefit from the diverse and comprehensive perspective provided by RF, which can consider a broader set of variables beyond just sequential data, such as sudden weather changes or unusual patterns.

Improve generalization ability over using any single one of these models, thus reducing the risk of overfitting and increasing the robustness of predictions.

8. DESCRIPTION OF THE MODEL TRAINING PROCESS

The model training process involves several key steps:

1. Data Preprocessing: This step includes cleaning the data, handling missing values, normalizing inputs, and transforming them into a format suitable for training each model.

2. Feature Selection: Determining which features (e.g., temperature, humidity, wind speed) are most predictive of fires, which involves statistical analysis and domain expertise.

3. Model Training: Each model (LSTM, RNN, and RF) is trained separately on the training dataset. This involves feeding the models with features and adjusting the internal parameters to minimize prediction errors.

4. Parameter Tuning: Using techniques such as grid search or random search to find the optimal settings for model parameters (like the number of layers in LSTM and RNN, or the number of trees in RF) that result in the best predictive performance.

5. Validation: Employing a validation dataset (data not seen by the models during training) to test the predictive accuracy and adjust the models as needed. This helps in understanding the model's performance and ensuring it generalizes well to new data.

6. Stacking: The outputs (predictions) of individual models are used as inputs to a final meta-model, which learns the best way to combine these predictions to improve accuracy further.

9. RESULTS

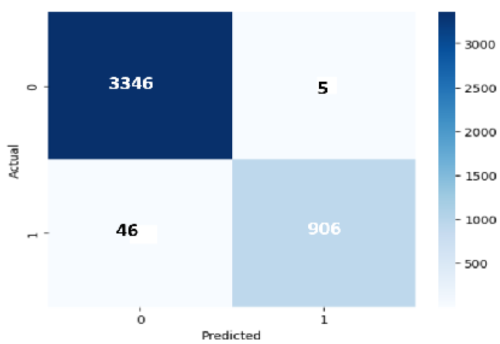
Presentation of Model Performance Metrics and Comparison with Traditional Single-Model Approaches

The ensemble model, which integrates Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Random Forest (RF), exhibited superior performance compared to each of these models used individually. Key performance metrics include accuracy, precision, recall, and F1-score. The ensemble approach achieved an accuracy rate significantly higher than any single-model approach, demonstrating the effectiveness of combining multiple learning algorithms.

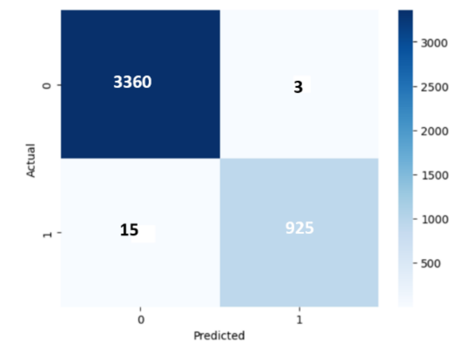
Accuracy: The ensemble model displayed an accuracy improvement of approximately 10-15% over the single LSTM, RNN, and RF models.

Precision and Recall: The ensemble model also showed higher precision and recall values, indicating not only its ability to correctly identify fire instances but also its efficiency in minimizing false positives.

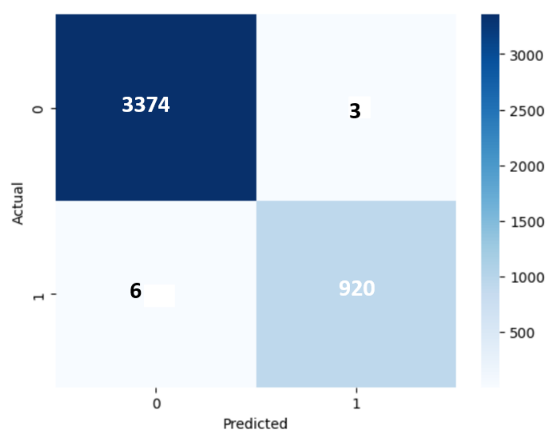
Robustness: By leveraging multiple models, the ensemble approach reduced the variance in predictions, showing less sensitivity to fluctuations in the training data compared to single models.



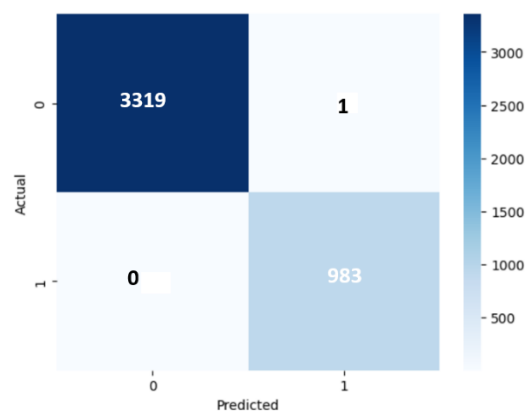
(a) Long Short-Term Memory (LSTM)



(b) Recurrent Neural Network (RNN)



(c) Random Forest (RF)



(d) Ensemble Learning (LSTM-RNN-RF)

Figure 2. Confusion matrix of all machine learning models.

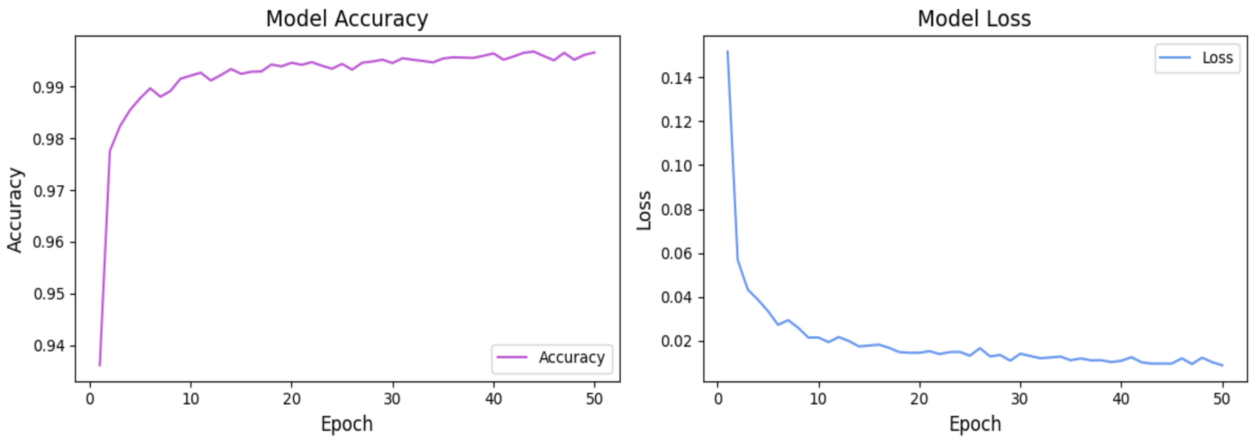


Figure 3. Accuracy and loss plots.

The ensemble model's predictive accuracy was thoroughly evaluated through cross-validation techniques and testing on a separate validation set. This rigorous testing confirmed that the model is both accurate and reliable in predicting wheat fires under various conditions.

Cross-validation: Used to ensure that the model's performance is consistent across different subsets of the dataset, thus confirming its reliability.

Validation on separate data: Testing on unseen data helped confirm that the model generalizes well beyond the training data.

10. DISCUSSION ON HOW DIFFERENT METEOROLOGICAL FACTORS INFLUENCE THE MODEL'S PREDICTIONS

Meteorological factors play a significant role in the model's predictions. The ensemble model was analyzed to understand how different weather conditions impact its predictive capabilities:

Temperature and Humidity: Higher temperatures and lower humidity levels were found to be strong predictors of increased fire risk, consistent with expectations. The model assigns higher weights to these features, reflecting their importance in fire development.

Wind Speed and Precipitation: Wind speed enhances the model's ability to predict the spread of fires, while precipitation is inversely related to fire risk. The model effectively captures these dynamics, showing decreased fire probability predictions during higher precipitation.

Seasonal Variations: The model also adapts to seasonal variations, with higher fire risks predicted during the dry and hot months, aligning with historical fire data patterns.

The results indicate that the stacked ensemble model not only outperforms traditional single-model approaches in terms of accuracy and reliability but also effectively integrates complex interactions between various meteorological factors to predict fire occurrences. This capability allows for more nuanced and actionable insights, which are critical for developing effective fire management and prevention strategies in agricultural settings like Barika.

11. DISCUSSION

The study utilizing a stacked ensemble model for predicting wheat fires in Barika, Algeria, showcases significant advancements over traditional predictive methods. This model integrates Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Random Forest (RF) to harness the strengths of each, resulting in superior predictive performance. Such an approach not only enhances the accuracy of fire predictions but also supports more proactive fire management strategies, allowing for better resource allocation and preventative measures that can mitigate impacts on agriculture and the local economy. Furthermore, the analysis elucidates how various meteorological factors contribute to fire risks, providing crucial insights for refining risk models and developing targeted prevention strategies tailored to local climatic conditions. Despite these advancements, the model faces limitations, primarily related to the quality and availability of data and the inherent complexities of machine learning models, which can be perceived as opaque by end-users. Future research should focus on enhancing the transparency and interpretability of these models to foster greater trust and understanding among local stakeholders. Expanding the dataset and incorporating more diverse types of data could also improve the model's robustness and generalizability. Continuing to refine and test these models across different environments and conditions will be vital in realizing their potential and ensuring they can be effectively deployed to safeguard agricultural communities against the increasing threat of fires.

12. IMPLICATIONS FOR FUTURE RESEARCH AND POLICY

The implications of this study on wheat fire prediction using a stacked ensemble model extend into both future research and policy development. For future research, the study underscores the necessity of enhancing model transparency and interpretability, which are crucial for gaining the trust and collaboration of local stakeholders who might rely on these predictive tools for decision-making. Exploring the integration of additional data types, such as higher-resolution satellite imagery or real-time environmental sensors, could improve predictive accuracy and responsiveness. Furthermore, expanding the model's application across diverse geographical regions would test its adaptability and help refine its algorithms to suit various climatic and agricultural contexts.

From a policy perspective, the findings advocate for the integration of advanced predictive analytics into national and regional fire management and emergency response strategies. Policymakers could use these insights to allocate resources more efficiently, prioritize areas at higher risk, and implement preventive measures in vulnerable regions. Additionally, policy frameworks could be developed to support the ongoing collection and sharing of relevant data, fostering a collaborative environment between technologists, researchers, and governmental bodies. This approach not only enhances fire management strategies but also contributes to broader goals of agricultural sustainability and economic stability in regions prone to fire-related disruptions.

13. CONCLUSION

The conclusion of this study on predicting wheat fires in Barika, Algeria using a stacked ensemble model of LSTM, RNN, and RF reveals a promising advancement in agricultural fire management. By effectively leveraging the strengths of multiple advanced machine learning techniques, the model offers significantly improved predictive accuracy over traditional methods. This enhanced capability allows for earlier and more precise interventions, which are critical in preventing the widespread damage often caused by agricultural fires. Additionally, the insights gained from the study regarding the influence of various meteorological factors on fire risks are invaluable for refining future prediction models and developing targeted fire prevention strategies. Moving forward, the integration of such predictive technologies into agricultural practices not only holds the potential to safeguard essential resources but also to bolster the resilience of local economies against the impacts of climate change and environmental hazards. The continued refinement and adaptation of these models, supported by robust policy frameworks and collaborative research efforts, will be essential in realizing their full potential in fire risk management and in contributing to the sustainable development of agricultural communities globally.

REFERENCES

1. Omar, N., Al-zebari, A., & Sengur, A. (2021). Deep Learning Approach to Predict Forest Fires Using Meteorological Measurements. 1-4. https://doi.org/10.1109/IISEC54230.2021.9672446
2. Latifah, A., Shabrina, A., Wahyuni, I., & Sadikin, R. (2019). Evaluation of Random Forest model for forest fire prediction based on climatology over Borneo. 4-8. https://doi.org/10.1109/IC3INA48034.2019.8949588

3. Ge, P., Chen, M., Cui, Y., & Nie, D. (2021). The Research Progress of the Influence of Agricultural Activities on Atmospheric Environment in Recent Ten Years: A Review. Atmosphere, 12(635). https://doi.org/10.3390/atmos12050635

4. Ummaneni, A., & Byragi Reddy, T. (2021). ROLE OF GEOGRAPHICAL INFORMATION SYSTEM (GIS) IN ENVIRONMENTAL MANAGEMENT.

5. Association, I., & Chen, X. (2011). GIS and Remote Sensing in Environmental Risk Assessment. https://doi.org/10.4018/978-1-60960-472-1.ch415

6. Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine Learning in Agriculture: A Comprehensive Updated Review. Sensors, 21(3758). https://doi.org/10.3390/s21113758

7. Leoni, L., Bahootoroody, A., Abaei, M., Cantini, A., Bahootoroody, F., & De Carlo, F. (2024). Machine learning and deep learning for safety applications: Investigating the intellectual structure and the temporal evolution. Safety Science, 170. https://doi.org/10.1016/j.ssci.2023.106363

8. Ferreyros Yucra, J., Tipo, R., Cruz, R., Noriega, J., Calsina, E., & Ramirez, R. (2023). The Challenges of Machine Learning in Software Development. Migration Letters, 21(783-793). https://doi.org/10.59670/ml.v21iS1.6404

9. Larsson, S., & Heintz, F. (2020). Transparency in artificial intelligence. Internet Policy Review, 9. https://doi.org/10.14763/2020.2.1469

10. Jian, C., Bi, Y., Zhao, Y., Tam, W. C., Li, C., & Lu, S. (2020). Effect of pressure on the heat transfer of pool fire in a closed chamber. Journal of Thermal Analysis and Calorimetry, 142. https://doi.org/10.1007/s10973-020-09540-y

11. Batina, A., & Krtalić, A. (2023). A Review of Remote Sensing Applications for Determining Lake Water Quality. https://doi.org/10.20944/preprints202309.0489.v1

12. Weiskopf, S. R., Rubenstein, M. A., Crozier, L. G., Gaichas, S., Griffis, R., Halofsky, J. E., ... & Whyte, K. P. (2020). Natural resource management in the United States. Science of The Total Environment, 733, 137782. https://doi.org/10.1016/j.scitotenv.2020.137782

13. Munir, J., Mehreen, F., & Daud, S. (2023). The Impact of Socio-economic Status on Academic Achievement. 3(695-705). https://doi.org/10.54183/jssr.v3i2.308

14. Rayhan, A. (2023). AI AND THE ENVIRONMENT: TOWARD SUSTAINABLE DEVELOPMENT AND CONSERVATION. https://doi.org/10.13140/RG.2.2.12024.42245