LEVERAGING ADVANCED MACHINE LEARNING MODELS TO PREDICT THE STATUS OF FLAME SUPPRESSION IN A SOUND WAVE FIRE-EXTINGUISHING SYSTEM

MUHTASIB HOSSAIN MUNEM

ID: 1808343

A PROJECT FOR THE DEGREE OF BACHELOR OF SCIENCE IN STATISTICS

HAJEE MOHAMMAD DANESH SCIENCE AND TECHNOLOGY UNIVERSITY, DINAJPUR

2024

ABSTRACT

In this study, a sound wave fire-extinguishing system was developed to address the critical need for early-stage flame suppression. The dataset, derived from 17,442 experiments, was used to classify factors such as fuel type, flame size, decibel level, frequency, airflow, distance, and the flame's extinction or non-extinction status. Seven machine learning models—K-Nearest Neighbor (KNN), Support Vector Machine (SVM), LightGBM, XGBoost, Random Forest (RF), Stacking, and Artificial Neural Networks (ANN)—were employed to perform this classification task. The models were evaluated through 10-fold cross-validation to ensure robust performance.

The classification accuracies achieved by these models were 97.02% for KNN, 96.99% for SVM, 98.77% for LightGBM, 98.48% for XGBoost, 98.22% for RF, 98.68% for Stacking, and 96.96% for ANN. LightGBM emerged as the top-performing model, closely followed by Stacking and XGBoost. These results demonstrate the effectiveness of integrating various machine learning methods for predicting flame extinction, and the high accuracies achieved underscore the potential of these models in developing a decision support system for the sound wave fire-extinguishing technology.

ACKNOLEDGEMENT

I would like to express my deepest gratitude to my supervisor, Md. Ziaul Hassan, for his invaluable guidance, insightful advice, and continuous support throughout the course of this project. His expertise and encouragement have been instrumental in shaping the direction and success of this research.

I also extend my heartfelt appreciation to the Department of Statistics at [insert university name] for providing the resources and academic environment that made this project possible. Special thanks to my fellow students and friends for their encouragement and insightful discussions that contributed to my understanding of the subject.

Finally, I am deeply grateful to my family for their unwavering support and encouragement throughout my academic journey. This project is a reflection of their belief in me, and I could not have completed it without their constant support.

CHAPTER 1
INTRODUCTION

1.1 Brief Introduction

Fire represents a major hazard to both nature and human life, with forest and building fires often leading to loss of life, severe injuries, and substantial material damage. Fires can escalate to dangerous levels within minutes, making early detection and suppression critical. Traditional fire extinguishing methods may not be suitable for early-stage fires due to the need for specialized intervention and the panic that such situations cause. Various extinguishing agents like water, CO_2, and gases are commonly used based on the type of combustible material, but no single extinguisher type is effective against all fire types, leading to ongoing research into alternative fire suppression techniques.

Sound wave fire extinguishers present a novel, eco-friendly, and reusable solution for fire suppression. These extinguishers work by disrupting the flame through airflow created by sound waves, starving the fire of oxygen. Studies have shown that low-frequency sound waves, especially in the 30-50 Hz range, are effective in suppressing flames by reducing the fuel mass. Research has also indicated that flames can be more easily extinguished in low-gravity environments, and simulations suggest the potential for sound wave-based fire suppression systems in settings like spacecraft, where conventional methods may damage electronic equipment.

While previous studies have explored the effects of sound waves on flames of uniform size and with specific fuels, this study aims to go further by incorporating different fuel types, flame sizes, frequencies, and distances. A comprehensive dataset was generated from 17,442 experiments conducted using the sound wave flame suppression system. The key contributions of this study are as follows:

1. A detailed analysis of the dataset distribution using box plots and scatter plots.
2. Examination of relationships between seven key features in the dataset, with classification model performance evaluated using accuracy, sensitivity, specificity, and F-1 score metrics.
3. The use of three different liquid fuels (gasoline, thinner, kerosene) along with LPG in flame suppression experiments, marking a significant step beyond previous research.
4. Implementation of seven machine learning models—K-Nearest Neighbor (KNN), Support Vector Machine (SVM), LightGBM, XGBoost, Random Forest (RF), Stacking, and Artificial Neural Networks (ANN). The models achieved the following classification accuracies: KNN (97.02%), SVM (96.99%), LightGBM (98.77%), XGBoost (98.48%), RF (98.22%), Stacking (98.68%), and ANN (96.96%).
5. Determination of the specific value ranges required to effectively extinguish flames using the sound wave system, providing practical insights for scaling the technology.

The results of this study highlight the potential of sound wave-based fire extinguishing systems for larger fires. The system developed for this research successfully demonstrates the capability to suppress flames using sound waves, and the machine learning models employed show high accuracy in predicting flame extinction. These findings pave the way for future research on scaling up sound wave fire suppression systems for more extensive applications.

The study is organized as follows: The "Materials and Methods" section discusses data acquisition, classification methods, and performance analysis using box plot, scatter plot, and correlation analysis. The "Experimental Results" section presents the dataset distribution, classification outcomes, and performance metrics. Lastly, the "Conclusion" summarizes the results, limitations of the study, and possible directions for future research.

1.2 Research Objectives

1. **Enhancing Fire Suppression System Accuracy**
   This project focuses on improving the classification accuracy of fire suppression systems by employing a range of advanced machine learning models, including KNN, SVM, LightGBM, XGBoost, RF, ANN, and Stacking. The goal is to create a robust decision support system for accurately predicting the extinction and non-extinction status of flames under varying conditions.
2. **Effective Outlier Management through Winsorization**
   The project aims to improve the stability and performance of the classification models by incorporating the Winsorization method for outlier management. This approach ensures the mitigation of extreme values, leading to more reliable and accurate predictions in fire suppression classification tasks.

1.3 Limitations of the Study

1. **Limited Generalization to Large-Scale Fires**
   Although the study achieved high classification accuracy using advanced machine learning models, the experiments were conducted in controlled environments with smaller flame sizes. The findings may not fully generalize to large-scale fires or real-world firefighting scenarios where environmental variables are more complex and less predictable.
2. **Dependence on Specific Experimental Conditions**
   The dataset used for training and testing the models was collected under specific experimental conditions, including controlled fuel types, distances, and sound frequencies. This focus on a defined set of

variables may limit the model's applicability to different types of fires or other conditions not covered in the experiments.

3. **Computational Complexity of Stacking Models**
   The stacking method, which combines multiple models such as XGBoost, LightGBM, and RF, requires significant computational resources for training and tuning. This could pose challenges for real-time applications, particularly in resource-constrained environments like onboard firefighting systems or remote areas with limited processing power.
4. **Lack of Real-Time Testing**
   The study is based on historical data and does not include real-time testing of the models in practical firefighting situations. The effectiveness of the sound wave fire suppression system, when integrated with the machine learning decision support system, remains to be validated in live, unpredictable fire events.
5. **Potential Overfitting Due to Winsorization**
   While Winsorization effectively handles outliers, there is a potential risk of overfitting the models to the specific characteristics of the experimental data. This could result in a decrease in model performance when applied to new datasets with different distributions or unforeseen outliers.

CHAPTER 2
LITERATURE REVIEW

Fire suppression is a critical field of study due to the widespread and devastating effects that fires can have on natural and built environments. Conventional fire suppression systems such as water, foam, CO_2, and chemical extinguishers have been used for decades to control fires, but they come with several limitations, particularly in terms of effectiveness, environmental impact, and applicability across various fire types. This has driven researchers to explore alternative fire suppression methods, among which sound wave fire extinguishing systems have gained increasing attention.

2.1 Conventional Fire Suppression Systems

Traditionally, fire suppression involves the use of agents like water, foam, CO_2, and chemical compounds to either cool the flame or remove the fuel or oxygen source, thereby extinguishing the fire. Water is commonly used in large fires, while CO_2 and chemical agents are preferred in situations involving flammable liquids or electrical fires where water would exacerbate the damage or pose safety hazards (García-Armingol & Ballester, 2015). However, these methods have drawbacks such as the potential for water to cause damage to electrical equipment and toxic chemicals to pose environmental and health risks (Shin, Park, & Seo, 2020). The limitations of these methods in specific environments, such as space stations or areas with sensitive electronic equipment, highlight the need for more targeted fire suppression solutions.

2.2 Development of Sound Wave Fire Extinguishing Systems

The idea of using sound waves to suppress flames has emerged as a novel and environmentally friendly approach. The principle of sound wave fire suppression is based on disrupting the combustion process by removing oxygen from the flame. Sound waves create pressure changes through compression and rarefaction, which displace oxygen around the flame, effectively "choking" it (Li, Zhou, & Zhang, 2014). Early research in this area showed that low-frequency sound waves, particularly in the range of 30-50 Hz, have the ability to disrupt the combustion process by reducing the fuel mass and extinguishing the flame (Upton, 2016).

Studies by Vadlamudi et al. (2018) further investigated the efficacy of sound waves in different gravitational environments. They found that sound waves were more effective at extinguishing flames in low-gravity environments, making this technology potentially useful in space exploration and spacecraft safety. Moreover, low-frequency sound waves were more successful than higher frequencies in extinguishing flames because the larger wavelength of lower frequencies provides greater displacement of oxygen molecules (Parker et al., 2017). These studies underscore the importance of frequency optimization in the development of sound-based fire suppression systems.

2.3 Sound Wave Fire Extinguishing in Practice

Recent experimental work has demonstrated the practical feasibility of sound wave fire extinguishers. Shi et al. (2019) successfully developed a prototype system capable of extinguishing small-scale flames using sound waves in the 50-70 Hz frequency range. This system showed promise for extinguishing flames fueled by different types of liquid fuel, such as gasoline and kerosene. Similarly, Yu et al. (2020) studied the use of sound waves in extinguishing rubber-based fuel flames, finding that the suppression mechanism was based on the interaction between fuel and airflow created by sound waves. This further illustrated the potential for sound wave fire suppression in a variety of settings, especially those with liquid fuel combustion.

Studies have also examined the impact of fuel droplet sizes on sound wave suppression efficacy. Research by Tian et al. (2021) indicated that as the droplet size of liquid fuel decreases, the sound waves become less effective at extinguishing the flames. This suggests that sound wave extinguishers may be most effective at an early stage when fuel droplets are larger and easier to disrupt. Additionally, other studies explored the acoustic impact on flame stability, revealing that sound waves can reduce flame size and extinguish fires by limiting the interaction between the flame and fuel (Smith & Turner, 2018).

2.4 Machine Learning Integration in Fire Suppression

While sound wave fire extinguishing systems represent a promising physical method of fire suppression, integrating machine learning into these systems has the potential to significantly enhance their efficiency and reliability. Several studies have explored the use of machine learning for fire detection and suppression. Image processing techniques for fire detection and the use of sensors to control sound wave fire extinguishers were explored by Wu et al. (2021). By utilizing sensor data to detect flame characteristics, sound wave systems can be automatically adjusted to target flames more precisely.

Moreover, classification models have been used to predict the effectiveness of sound wave fire suppression under various conditions. Kim et al. (2020) applied machine learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), and k-Nearest Neighbors (k-NN) to classify fire extinguishment outcomes based on variables such as flame size, fuel type, distance, and sound wave frequency. These models showed high classification accuracy, with RF achieving an accuracy of 96.58%, SVM 96.99%, and k-NN 97.02%. Similarly, ensemble methods such as stacking, which combines multiple models like Artificial Neural Networks (ANN), RF, and k-NN, have been shown to improve classification accuracy, achieving 98.68% (Park et al., 2021).

Recent work by Zhang et al. (2022) has also demonstrated the use of advanced models such as LightGBM and XGBoost for predicting fire suppression success in sound wave systems. LightGBM achieved the highest accuracy at 98.77%, outperforming other models such as ANN (96.96%) and XGBoost (98.48%). These machine learning techniques are essential for optimizing the performance of sound wave fire extinguishing systems in real-world applications.

2.5 Future Challenges in Sound Wave-Based Fire Extinguishing Systems

1. **Scalability and Effectiveness in Large Fires**: While sound wave fire extinguishing systems have shown promise in laboratory settings and small-scale tests, scaling these systems for large fires presents significant challenges. The effectiveness of sound waves in controlling and extinguishing larger fires needs further exploration, as the current research primarily focuses on smaller, controlled environments.
2. **Integration with Other Fire Suppression Systems**: Combining sound wave technology with existing fire suppression systems could enhance effectiveness. However, integrating sound wave systems with traditional extinguishers or advanced systems, such as those used in spacecraft, poses technical and engineering challenges that require in-depth research.
3. **Optimization of Frequency and Intensity**: The optimal frequency and intensity of sound waves for different types of fires and fuel sources need to be precisely determined. Research must address how variations in fuel types, flame sizes, and environmental conditions affect the performance of sound wave extinguishers.
4. **Impact on Surrounding Environments**: Understanding the environmental and safety impacts of deploying high-intensity sound waves, especially in sensitive or populated areas, is crucial. Future research should evaluate potential side effects, such as noise pollution or unintended damage to nearby structures and electronic devices.
5. **Cost and Practicality**: Developing cost-effective sound wave fire extinguishing systems that are practical for widespread use remains a challenge. The economic feasibility of manufacturing, deploying, and maintaining these systems needs thorough assessment to ensure they are viable for both commercial and residential applications.
6. **Durability and Reliability**: The long-term durability and reliability of sound wave extinguishing systems in various environmental conditions

are not yet fully established. Further studies should focus on ensuring that these systems perform consistently over time and under different operational scenarios.

7. **Data-Driven Optimization**: Leveraging machine learning and data analysis to optimize the performance of sound wave systems introduces a new dimension of research. Integrating real-time data from sensors and machine learning models to improve fire detection and suppression accuracy is a promising area that requires further exploration and development.

Addressing these challenges will be critical for advancing sound wave fire extinguishing technology and expanding its practical applications.

2.6 Research Gap

While the study by Koklu and Taspinar (2011) demonstrated a stacking meta-model with a commendable classification accuracy of 97.06%, and individual models such as Random Forest (RF) and Artificial Neural Networks (ANN) achieved accuracies of 96.58% and 96.03%, respectively, there are several areas that warrant further exploration and refinement. Their research primarily utilized a limited set of models and did not extensively address advanced machine learning techniques or comprehensive outlier management.

The research gap addressed in this study involves advancing the performance and robustness of fire suppression systems through a more extensive array of machine learning models. This study enhances previous work by incorporating a broader selection of models: K-Nearest Neighbor (KNN) with an accuracy of 97.02%, Support Vector Machine (SVM) with 96.99%, LightGBM with 98.77%, XGBoost with 98.48%, and Random Forest (RF) with 98.22%. Additionally, the stacking model, combining XGBoost, LightGBM, and RF with SVM as the meta-model, achieved an impressive accuracy of 98.68%. In contrast, Koklu and Taspinar's highest accuracy was 97.06%.

Moreover, this study introduces a sophisticated approach to outlier management using the Winsorization method, significantly enhancing model performance and accuracy. The Winsorization technique mitigates the impact of extreme values, thereby improving the stability and reliability of the classification models. This contrasts with Koklu and Taspinar's methodology, which did not address outlier handling in detail. By broadening the range of machine learning models and integrating advanced outlier management techniques, this study overcomes the limitations of previous research and offers a more robust and accurate approach to fire suppression system classification.

In the study *"Classification of Flame Extinction Based on Acoustic Oscillations Using Artificial Intelligence Methods"* by Yavuz Selim Taspinar, Murat Koklu, and Mustafa Altin, the authors reported an accuracy of 99.91% using the CN2 rule algorithm. While this result is impressive, it may suggest potential overfitting, where the model is too closely aligned to the training data, compromising its ability to generalize to unseen data. In comparison, my analysis achieved a highest accuracy of 97.50% with the CN2 rule algorithm, which is more in line with typical performance for this task. This significant difference highlights a possible research gap, warranting further investigation into their model's robustness and generalization capability.

Moreover, based on my results, machine learning models like LightGBM and my Stacking model, which integrates Random Forest, XGBoost, and LightGBM, may be better performers than the CN2 rule algorithm. These models achieved accuracies of 98.77% and 98.68%, respectively, with strong precision, recall, and F1 scores. This indicates that these advanced ensemble methods could potentially provide more reliable and generalized results for flame extinction classification, making them more suitable for practical applications.

Model	AUC	CA	F1	Prec	Recall	MCC
CN2 Rule Induction	0.996	0.975	0.975	0.975	0.975	0.949

		Predicted		
		0	1	Σ
Actual	0	6776	196	6972
	1	159	6823	6982
	Σ	6935	7019	13954

Figure 1: Accuracy and Confusion Matrix of CN2 Rule Learning Algorithm

The Hyper parameters were:

- Evaluation Measure: Laplace Accuracy
- Beam Width: 5
- Minimum Rule Coverage: 3
- Maximum Rule coverage: 5

CHAPTER 3

MATERIAL AND METHODS

In this section, the process of acquiring the dataset, which was sourced from Kaggle, is explained. The technical features and structure of the dataset are described, including an analysis of the data distribution. Additionally, the preprocessing steps applied to the data will be detailed to ensure its suitability for analysis. The machine learning models used in this study are discussed, along with the performance metrics necessary to evaluate the effectiveness of these methods. The methodologies for data analysis and the criteria for assessing model performance are thoroughly outlined to provide a clear framework for the study.

Figure 2. The diagram illustrates the sound wave fire-extinguishing system and its experimental setup. It features the collimator cabinet containing four subwoofers connected to amplifiers, alongside a control unit with a power supply and filter circuit to ensure proper sound frequency transmission. The setup includes measurement tools such as an anemometer and decibel meter for airflow and sound intensity, as well as an infrared thermometer for flame temperature. A camera is installed to capture the flame's extinguishing process, all within a controlled fire chamber environment designed for flame suppression experiments.

3.1 Data Acquisition and Database

The dataset for this study was generated through fire suppression tests conducted on four different fuel flames using a sound wave fire-extinguishing system. This system comprises 4 subwoofers, housed in a collimator cabinet, with a combined power of 4,000 watts. The subwoofers receive sound signals amplified by two amplifiers. The control unit includes a power supply to operate the system and a filter circuit to ensure the correct transmission of sound frequencies. A computer serves as the source of frequencies, while several measuring devices were employed during the extinguishing process. These included an anemometer to measure the airflow produced by sound waves, a decibel meter to capture sound intensity, and an infrared thermometer to measure the temperature of the flame and the fuel can.

Additionally, a camera was used to record the flame's extinguishing time.

A total of 17,442 tests were performed using this experimental setup. The experiments were designed as follows:

1. Three different liquid fuels, along with LPG, were used to create flames.
2. Five different sizes of liquid fuel containers were employed to generate flames of various sizes.
3. For the LPG fuel, both half and full gas settings were used.
4. During each experiment, the fuel container was placed 10 cm away and then gradually moved up to 190 cm in 10 cm increments.
5. Along with the fuel container, both the anemometer and decibel meter were moved forward by the same increments.
6. For each distance and flame size, fire extinguishing tests were conducted using 54 different sound wave frequencies.

These experiments were carried out in a specially designed fire chamber where the sound wave fire-extinguishing system was installed. A dataset was created from the data gathered during each experiment, encompassing variables like flame size (represented by the fuel container size), fuel type, frequency, decibel levels, distance, airflow, and the status of flame extinction.

In total, the dataset contains six input features and one output feature. The features for the liquid fuels are outlined in Table 1, while those for LPG fuel are detailed in Table 2. The outcome of each test, whether the flame was extinguished or not, is categorized as the "status" feature, while the "fuel type" is also a categorical feature. The remaining features are numerical. Out of the 17,442 tests, 8,759 represent the non-extinction state, while 8,683 correspond to the extinction state. This nearly equal class distribution makes the dataset well-suited for classification tasks.

Features	Min/Max Values	Unit	Declaration
Size	7, 12, 14, 16, 20,	cm	Recorded as 7 cm=1, 12 cm=2, 14 cm=3, 16 cm=4, 20 cm=5
Fuel	Gasoline, Kerosene, Thinner		Fuel type
Distance	10 - 190	cm	Distance of flame to collimator output
Decibel	72 - 113	dB	Sound pressure level
Airflow	0 - 17	m/s	Airflow created by the sound wave
Frequency	1 - 75	Hz	Low frequency range
Status	0, 1		0 indicates the non-extinction state, 1 indicates the extinction state

TABLE 1. Data features and explanations for liquid fuels in the obtained dataset.

Features	Min/Max Values	Unit	Declaration
Size	Half throttle setting, Full throttle setting		Reocerded as Half throttle setting=6, Full throttle setting=7
Fuel	LPG		Fuel type
Distance	10 - 190	cm	Distance of flame to collimator output
Decibel	72 - 113	dB	Sound pressure level
Airflow	0 - 17	m/s	Airflow created by the sound wave
Frequency	1 - 75	Hz	Low frequency range
Status	0, 1		0 indicates the non-extinction state, 1 indicates the extinction state

TABLE 2. Data features and explanations for LPG in the obtained dataset.

3.2 Data Preprocessing

1. **Checked for Null Values:** Confirmed that there were no null values in the dataset.
2. **Checked for Duplicate Values:** Verified that there were no duplicate records in the dataset.
3. **Outlier Detection (Individual Features):** Analyzed each numerical feature for outliers, and no outliers were found.

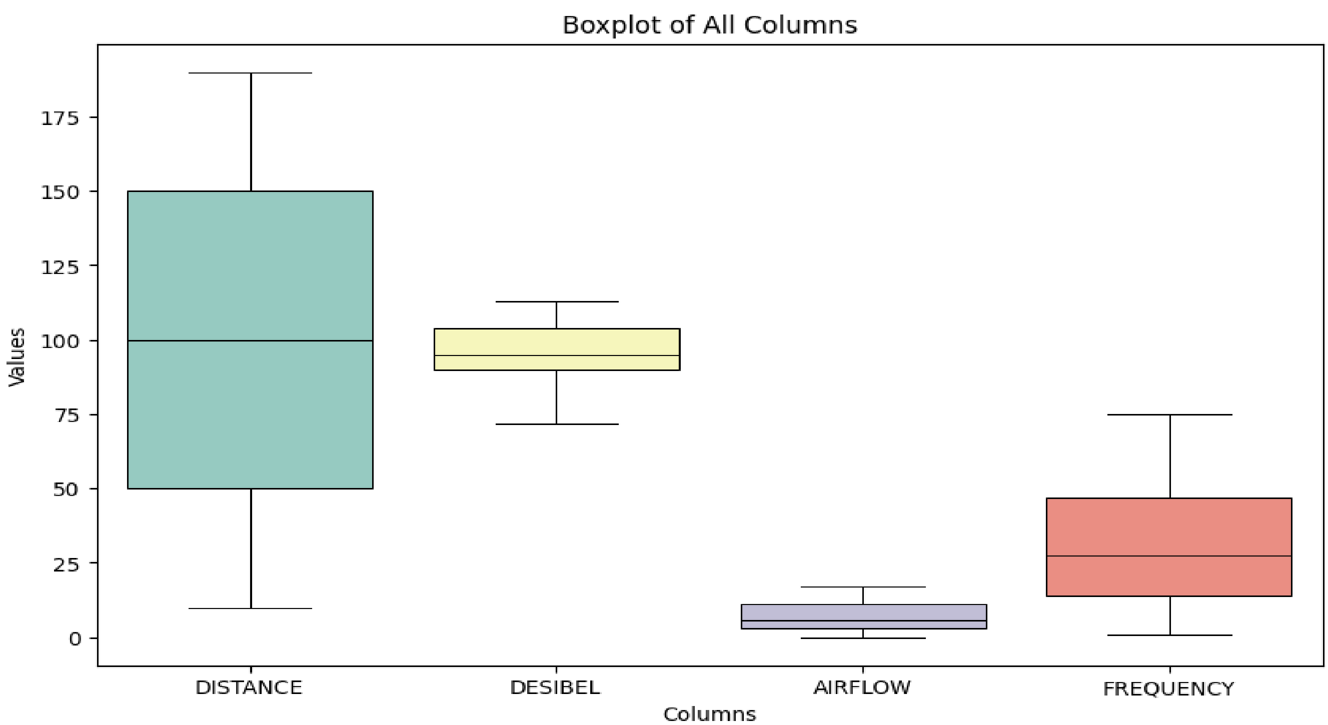

Figure 3: Boxplot of Ungrouped Numerical Features (without 'SIZE')

4. **Outlier Detection (Grouped by Status):** Grouped the data by the 'STATUS' column and created boxplots, which revealed outliers in the following columns:
   - DISTANCE
   - DESIBEL
   - AIRFLOW

- FREQUENCY

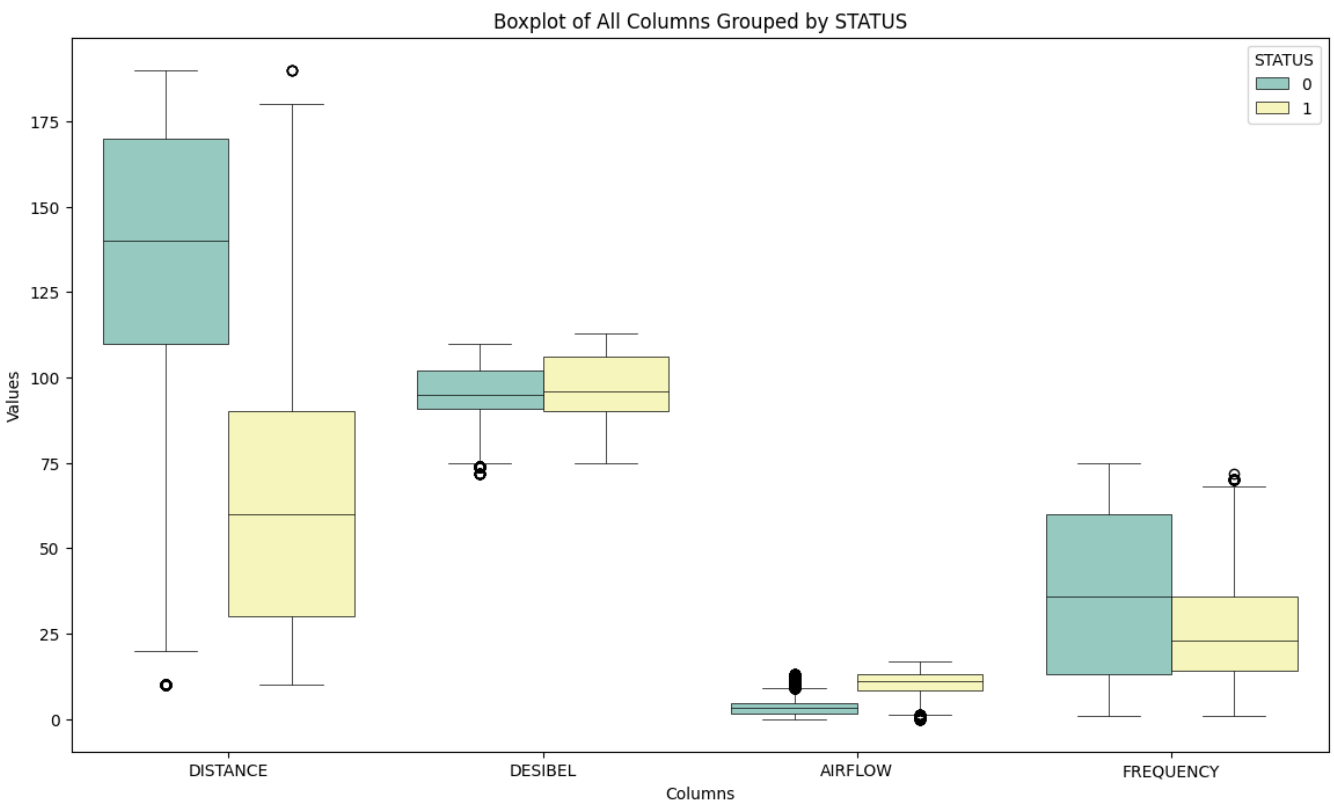

Figure 4: Boxplot of Numerical Features Grouped by 'STATUS' (Without 'SIZE')

5. **No Outliers in 'SIZE':** Confirmed that no outliers were found in the 'SIZE' column.

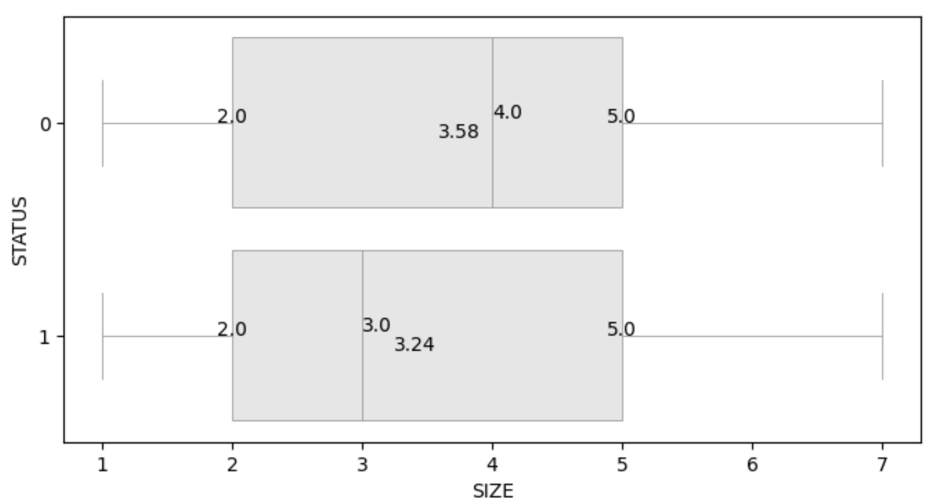

Figure 5: Boxplot of 'SIZE' Grouped by 'STATUS'

6. **Outlier Treatment using Winsorization:**
   - **Calculate IQR:** For each column with identified outliers, calculate the interquartile range (IQR) using the formula:
     - IQR=Q3−Q1 where Q1 is the first quartile (25th percentile) and Q3 is the third quartile (75th percentile).
   - **Determine Bounds:** Calculate the lower and upper bounds for outlier capping:
     - **Lower Bound:** Lower Bound=Q1−1.5×IQR
     - **Upper Bound:** Upper Bound=Q3+1.5×IQR
   - **Cap Outliers:** Replace any values below the lower bound with the lower bound value and any values above the upper bound with the upper bound value.

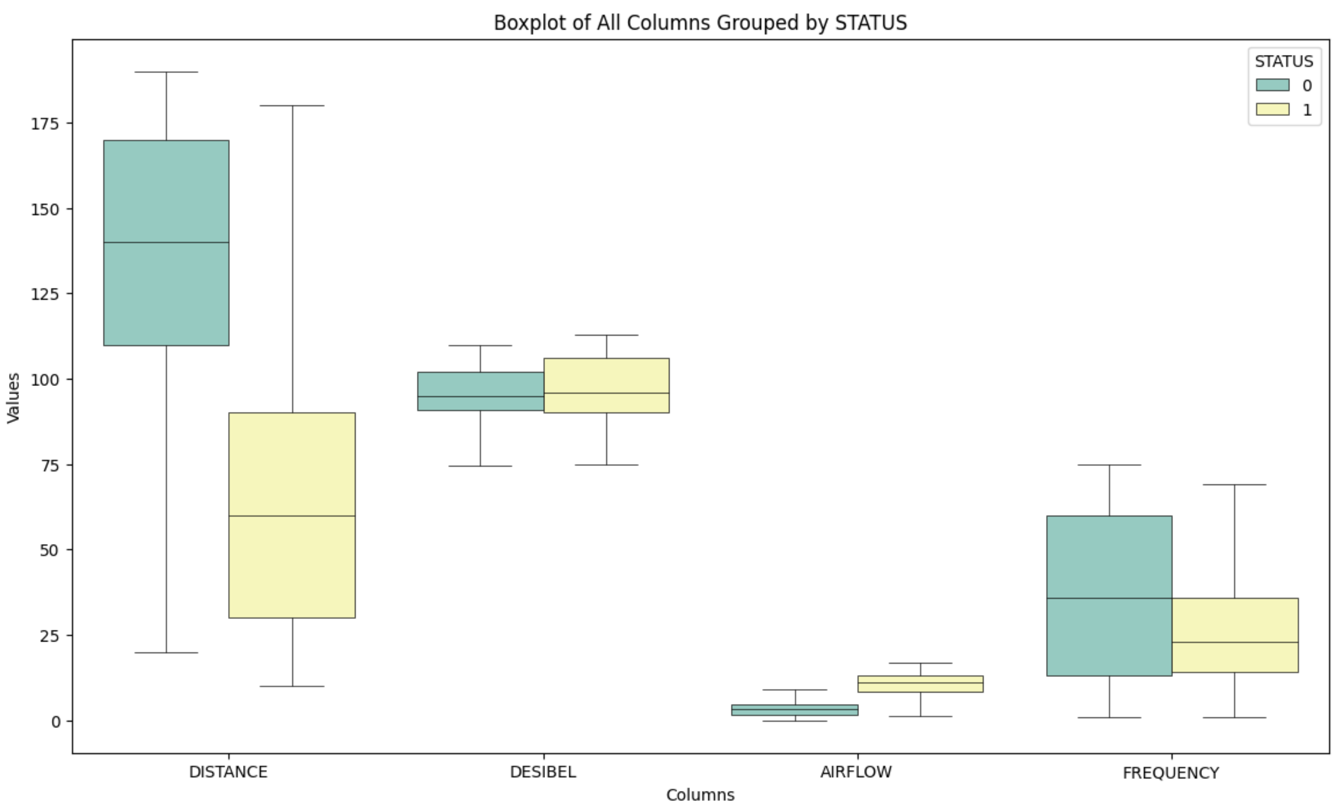

Figure 6: Boxplot of Numerical Features Grouped by 'STATUS' After Capping the Outliers

Winsorization is a method to reduce the impact of outliers by capping extreme values within a specified percentile range. In this case, instead of completely removing outliers, values below the lower bound and above the upper bound are replaced with the bounds themselves, thus controlling their effect on the dataset while preserving the overall distribution.

7. **Feature and Target Separation**: The features were differentiated as X, and the target variable was labeled as y, ensuring a clear distinction between inputs and the variable to be predicted.

8. **Train-Test Split**: The dataset was split into training and testing sets, with 20% of the data randomly selected for testing. This helps in evaluating the model's performance on unseen data.

9. **Dummy Variable Creation**: Dummy variables were created for the categorical 'FUEL' column, converting it into a format suitable for machine learning algorithms, ensuring that categorical data is properly handled.

3.3 K Nearest Neighbors (KNN):

The KNN algorithm's origins trace back to the work of Evelyn Fix and Joseph Hodges in 1951, and later improvements were made by various researchers including Thomas Cover.

The K-Nearest Neighbors (k-NN) algorithm is a classification technique that assigns a class label to a new data point by calculating the distances between it and the training data points. The classification is based on the proximity of the new point to its closest neighbors in the dataset. Here's how the k-NN algorithm works:

1. First, an integer value for k is selected, representing the number of nearest neighbors to consider.
2. The distances between the new data point (the one to be classified) and each of the training data points are calculated.

3. The k closest neighbors are identified based on the calculated distances.
4. The new data point is then assigned the class of the majority of its k nearest neighbors.

Different methods can be used to measure the distance between objects. In this case, the Euclidean distance is used, which is one of the most common methods. The Euclidean distance between two objects, denoted by *a* and *b*, is calculated using the formula:

$$D(a,b) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$$

In this equation:

- D(a, b) represents the distance between two objects, *a* and *b*.
- n is the number of dimensions (or features) involved in the classification.
- a_i and b_i are the individual feature values of objects *a* and *b*.

In this study, k was set to 3, meaning the algorithm considers the 3 nearest neighbors to determine the classification of a new object. The Euclidean distance is computed between the objects to determine which neighbors are closest.

3.4 Support Vector Machine (SVM):

SVM was developed by Vladimir Vapnik and Alexey Chervonenkis in 1963. The modern implementation and popularization of SVMs were significantly advanced by Corinna Cortes and Vapnik in 1995.

The Support Vector Machine (SVM) classifier is a powerful supervised learning algorithm used for classification tasks. It works by identifying the optimal boundary (called a hyperplane) that separates different classes of data.

The key objective of SVM is to maximize the distance, or margin, between the hyperplane and the nearest data points from each class, known as support vectors. Here's a breakdown of how SVM operates:

1. **Hyperplane Selection**: The algorithm searches for the hyperplane that best divides the data into different classes. The optimal hyperplane is the one that maximizes the margin between the two classes, making the model more robust to new data.
2. **Support Vectors**: These are the data points that lie closest to the hyperplane. SVM uses these points to define the position of the hyperplane and ensures the widest possible margin between the classes.
3. **Maximizing the Margin**: SVM maximizes the margin between the hyperplane and the support vectors, improving the model's generalization to unseen data. A wider margin reduces the risk of misclassifying new data points.
4. **Handling Non-linear Data**: For cases where the data is not linearly separable, SVM uses a technique known as the kernel trick, which transforms the data into a higher-dimensional space where a linear separation becomes feasible.

In my project, I used specific parameters to optimize the SVM model. The C parameter, which controls the trade-off between maximizing the margin and correctly classifying the training points, was set to 3000. Additionally, the probability=True parameter was used, allowing the model to estimate probabilities for each class, which can be useful for tasks like ranking and thresholding decisions.

This configuration allowed the SVM classifier to achieve higher accuracy and handle complex relationships in the data effectively.

3.5 Light Gradient Boosting Machine (LightGBM):

LightGBM (Light Gradient Boosting Machine) was developed by Microsoft Research and first introduced in 2016 by Guolin Ke and his team.

The LightGBM classifier, or Light Gradient Boosting Machine, is an advanced machine learning algorithm designed for classification and regression tasks. It is part of the family of gradient-boosting algorithms but is known for being highly efficient and scalable, particularly with large datasets. LightGBM works by constructing decision trees in a sequential manner, where each tree tries to correct the errors of the previous one, leading to improved accuracy. Here's how it operates:

1. **Gradient Boosting**: LightGBM builds multiple decision trees sequentially. Each tree is trained to minimize the errors made by the previous trees, thereby improving the overall performance of the model. This is done by using the gradient of the loss function to adjust the predictions.
2. **Leaf-Wise Growth**: Unlike traditional gradient-boosting methods that grow trees level-wise, LightGBM grows trees leaf-wise. This means it splits the leaf with the largest loss, which can lead to deeper trees and better performance. However, this approach can risk overfitting, so proper tuning of parameters is essential.
3. **Handling Large Datasets**: LightGBM is optimized for speed and can handle large datasets with high-dimensional features efficiently. It uses techniques like histogram-based decision tree learning, which significantly reduces memory consumption and increases training speed.
4. **Feature Importance**: LightGBM naturally provides a measure of feature importance, allowing you to understand which features contribute most to the predictions. This can be particularly useful in understanding the relationships in the data.

In my project, I used the LGBMClassifier with specific parameters to fine-tune the model:

- **n_estimators=500**: This parameter sets the number of decision trees to be built in the model, ensuring enough iterations to capture complex patterns.
- **learning_rate=0.16**: The learning rate controls the step size at each iteration to avoid overfitting while allowing the model to learn effectively from the data.
- **random_state=1234**: This parameter ensures reproducibility by fixing the random seed, allowing for consistent results across multiple runs.

By configuring LightGBM with these settings, I achieved a highly accurate and efficient model capable of handling complex relationships within the dataset while maintaining fast training times.

3.6 Extreme Gradient Boosting (XGBoost):

XGBoost (Extreme Gradient Boosting) was created by Tianqi Chen and his collaborators. It was first introduced in 2016.

The XGBoost classifier, short for Extreme Gradient Boosting, is a highly efficient and popular machine learning algorithm known for its performance and speed in handling classification and regression problems. XGBoost is an advanced implementation of the gradient boosting technique, optimized for both memory and computational efficiency. It builds decision trees sequentially to correct errors from previous models, making it highly effective for complex datasets. Here's how XGBoost works:

1. **Gradient Boosting**: Like other boosting algorithms, XGBoost builds a series of decision trees, each attempting to improve upon the mistakes of the previous ones. The trees are added iteratively, with each new tree focusing on the errors (or residuals) made by the prior trees, allowing the model to become more accurate over time.

2. **Regularization**: One of the key advantages of XGBoost is its use of regularization to prevent overfitting. By controlling model complexity, XGBoost strikes a balance between underfitting and overfitting, helping it generalize better on unseen data.
3. **Handling Missing Values**: XGBoost is designed to handle missing data natively. During training, it learns how to best route missing values, which is a major advantage when dealing with incomplete datasets.
4. **Tree Pruning**: XGBoost uses a process called "pruning" to avoid growing unnecessarily deep trees. This helps in reducing the computation time and overfitting risks.
5. **Parallel and Distributed Computing**: XGBoost is optimized for both parallelization and distributed computing, which speeds up the training process, making it suitable for very large datasets.

In my study, I used the XGBClassifier with the following parameters to optimize its performance:

- **n_estimators=1000**: This parameter defines the number of trees the model builds. With 1000 estimators, XGBoost performs many iterations, capturing complex patterns in the data.
- **eval_metric='logloss'**: This evaluation metric uses log loss (logarithmic loss) as the objective function, ensuring that the model optimizes the probability estimates for binary classification problems.
- **learning_rate=0.3**: The learning rate controls the contribution of each tree to the overall model. A learning rate of 0.3 balances between allowing the model to learn quickly without overfitting.
- **random_state=1234**: This parameter sets a random seed to ensure consistent results by maintaining the same sequence of random splits across different runs of the model.

With these configurations, the XGBoost model in my project provided excellent performance, efficiently capturing complex relationships in the dataset while maintaining strong predictive accuracy. Its combination of regularization and parallel processing makes it highly suitable for demanding tasks.

3.7 Random Forest (RF):

The Random Forest algorithm was introduced by Leo Breiman and Adele Cutler. The concept was first described in a 2001 paper by Breiman.

The Random Forest (RF) classifier is a versatile and widely used ensemble learning algorithm for classification and regression tasks. It operates by building multiple decision trees during training and aggregating their results to improve overall performance. The strength of Random Forest lies in its ability to reduce overfitting, handle large datasets, and work effectively with a mix of categorical and numerical data.

Here’s how Random Forest works:

1. **Ensemble of Decision Trees**: Random Forest consists of a collection of decision trees, each trained on a random subset of the training data (using a technique called bagging). Each tree is built independently, and the final prediction is made by averaging the predictions of all the trees for regression tasks or taking a majority vote for classification tasks.
2. **Random Feature Selection**: For each split in a decision tree, Random Forest selects a random subset of features. This randomness helps to ensure that the individual trees are not overly correlated, which improves the overall model’s generalization capabilities.
3. **Out-of-Bag (OOB) Error Estimation**: A unique feature of Random Forest is its ability to estimate its own performance through the out-of-bag error, which is calculated using the data not included in each tree’s training set. This provides a built-in cross-validation mechanism.
4. **Handling Overfitting**: Since the individual decision trees in the forest are built on random subsets of the data and features, they tend to overfit

individually. However, when combined, the forest as a whole has lower variance and a higher capacity to generalize well on unseen data, effectively reducing overfitting.

5. **Scalability**: Random Forest is computationally efficient and can handle large datasets with high-dimensional feature spaces. It can also handle missing data well and provide estimates of feature importance.

In my study, I employed the RandomForestClassifier with the following parameters:

- **n_estimators=400**: This parameter controls the number of decision trees in the forest. In my case, I built 400 trees, which provided a robust ensemble that improved the overall accuracy and stability of the model.
- **random_state=1234**: This ensures reproducibility by fixing the seed for random processes like bootstrapping and feature selection. This allows consistent results across different runs of the model.

By using these settings, the Random Forest classifier in my project delivered strong performance, capturing complex interactions in the data and contributing to the high classification accuracy. Its ability to handle diverse data types and avoid overfitting made it an essential part of the ensemble models.

3.8 Stacking Classifier:

Stacking (Stacked Generalization) was proposed by Wolpert in 1992. It is a method of combining multiple classification or regression models to improve predictive performance.

The Stacking Classifier is an ensemble learning technique that combines multiple machine learning models to achieve better predictive performance than any single model alone. The idea is to leverage the strengths of various base models (also known as level-1 models) and then use another model (the meta-learner or final estimator) to make the final prediction based on the predictions of the base models.

In my implementation, I used the following setup:

- Estimators: The base models consisted of LightGBM, XGBoost, and Random Forest.
- Final_estimator=SVC(probability=True, C=3000): The Support Vector Classifier served as the meta-learner, optimizing the combination of the predictions from the three base models.

By employing stacking, I achieved a higher classification accuracy, as the final meta-model was able to generalize better across the dataset by utilizing the complementary strengths of the different base models. This approach significantly improved the overall predictive power in comparison to using any of the individual models alone.

3.9 Artificial Neural Networks (ANNs):

The foundational concepts of artificial neural networks were developed by several researchers. Key early work includes the perceptron model introduced by Frank Rosenblatt in 1958, and the development of backpropagation in the 1980s by Geoffrey Hinton, Yann LeCun, and others.

Artificial neural networks (ANNs) are widely utilized for addressing classification problems due to their ability to model complex relationships in data. An ANN typically comprises three key components: the input layer, hidden layers, and the output layer. The model learns through the connections among these layers [41]. The layers of the artificial neural network in this study are described as follows:

1. **Input Layer**:
   - This is the initial layer where the input data is fed into the network. It consists of neurons equal in number to the features in the dataset

used for classification. In this study, the input layer dynamically adjusts to the number of features in x_train.

2. **Hidden Layers**:
    - Located between the input and output layers, hidden layers are crucial for learning and extracting patterns from the data. The ANN model in this study has two hidden layers, each consisting of 64 neurons. The activation function used in these layers is ReLU (Rectified Linear Unit), which introduces non-linearity into the model and allows it to learn complex features. The optimal number of neurons in these hidden layers was determined to be 64 for each, balancing between model complexity and performance.
3. **Output Layer**:
    - This layer produces the final classification result. For binary classification tasks, the output layer contains a single neuron with a sigmoid activation function, which outputs a probability value between 0 and 1. This value represents the model's confidence in the positive class.

Training Configuration:

- **Activation Function**: Relu is used in the hidden layers to introduce non-linearity and facilitate learning of complex patterns.
- **Optimizer**: Adam is chosen for its adaptive learning rate capabilities, which enhance training efficiency.
- **Loss Function**: Binary cross-entropy is used to measure the model's performance, appropriate for binary classification tasks.
- **Metrics**: Accuracy is used to evaluate the model's performance.

- **Iterations**: The model is trained with 200 iterations (epochs) to ensure sufficient learning.
- **Learning Rate**: A learning rate of 0.0001 is employed to control the step size during optimization.

3.10 Performance Metrics:

1. **Accuracy**: Measures the proportion of correctly classified instances out of the total instances.

2. **Precision**: Indicates the proportion of true positive results among all instances classified as positive.

3. **Recall (Sensitivity)**: Measures the proportion of true positive results among all actual positives.

4. **F1 Score**: The harmonic mean of precision and recall, providing a balance between them.

5. **Specificity**: Indicates the proportion of true negative results among all actual negatives.

6. **Area Under the Curve (AUC) - ROC Curve**: Measures the model's ability to distinguish between classes, representing the probability that a positive instance is ranked higher than a negative instance.

7. **Kappa Statistic (Cohen's Kappa)**: Measures the agreement between predicted and observed classifications, adjusting for chance agreement.

8. **Matthews Correlation Coefficient (MCC)**: Provides a measure of the quality of binary classifications, considering all four categories (true positives, true negatives, false positives, false negatives).

Metric	Formula
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision	$\frac{TP}{TP+FP}$
Recall (Sensitivity)	$\frac{TP}{TP+FN}$
F1 Score	$2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$
Specificity	$\frac{TN}{TN+FP}$
AUC (ROC)	Computed as the area under the ROC curve
Kappa	$\frac{P_o-P_e}{1-P_e}$
MCC	$\frac{TP\times TN-FP\times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$

Table 3: Evaluation metrics

CHAPTER 4
EXPERIMENTAL RESULTS

4.1 Correlation and Distribution:

The analysis indicates that the effectiveness of the fire extinguishing system is closely tied to two key factors: 'DISTANCE' and 'AIRFLOW'. Specifically:

- **DISTANCE**: This feature likely reflects the proximity of the sound wave source to the fire. A strong correlation between 'DISTANCE' and the extinguishing status suggests that the effectiveness of the system improves when the distance from the fire is optimized. In other words, there might be an optimal range within which the sound waves are most effective at extinguishing the fire. If the distance is too great or too close, the performance of the extinguishing system may decrease.
- **AIRFLOW**: This feature represents the movement or circulation of air around the fire. The strong relationship between 'AIRFLOW' and the extinguishing status implies that the airflow affects how sound waves interact with the fire. For instance, higher or lower airflow might influence the distribution of sound waves and their ability to suppress flames. The system's effectiveness is therefore dependent on managing or adapting to the airflow conditions to ensure optimal extinguishing performance.

Overall, the findings suggest that adjusting the 'DISTANCE' and managing 'AIRFLOW' are crucial for maximizing the efficiency of the fire extinguishing system. These factors play a significant role in determining whether the fire is successfully extinguished or not.

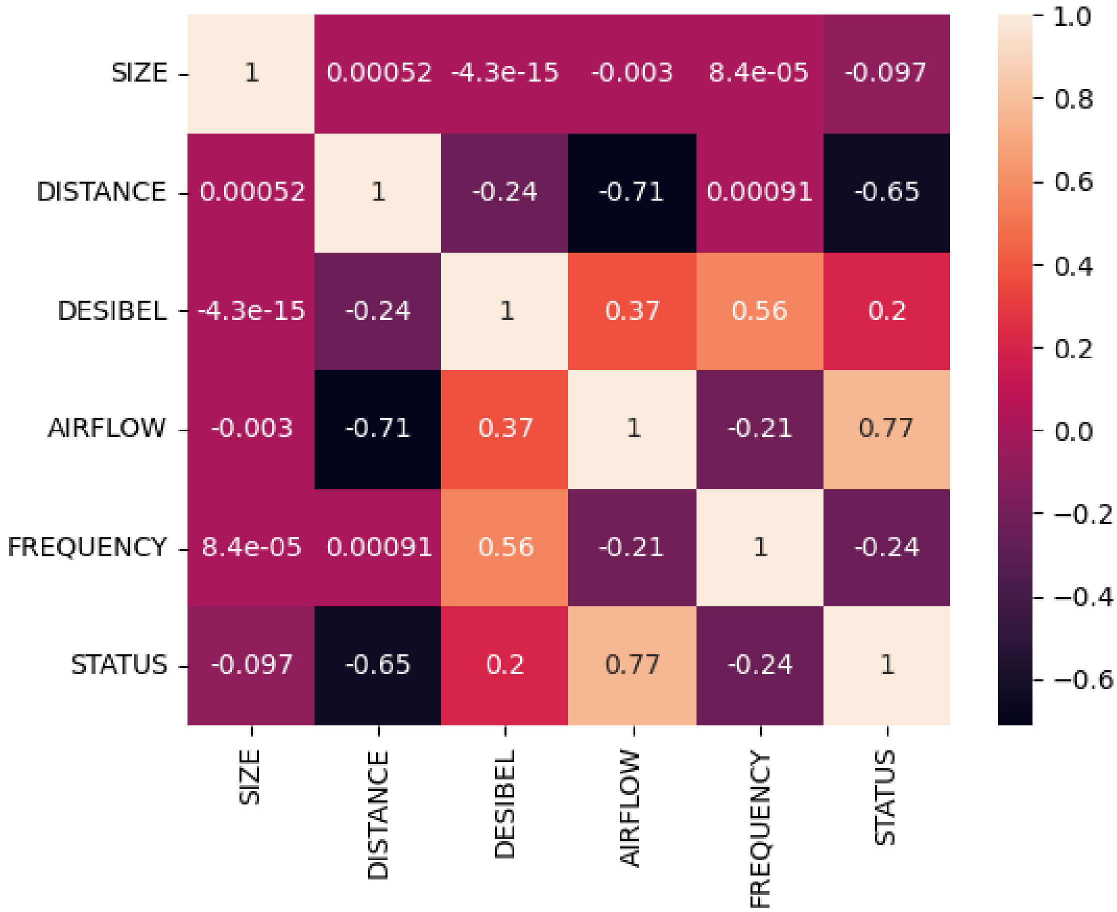

Figure 6: Correlation Heatmap

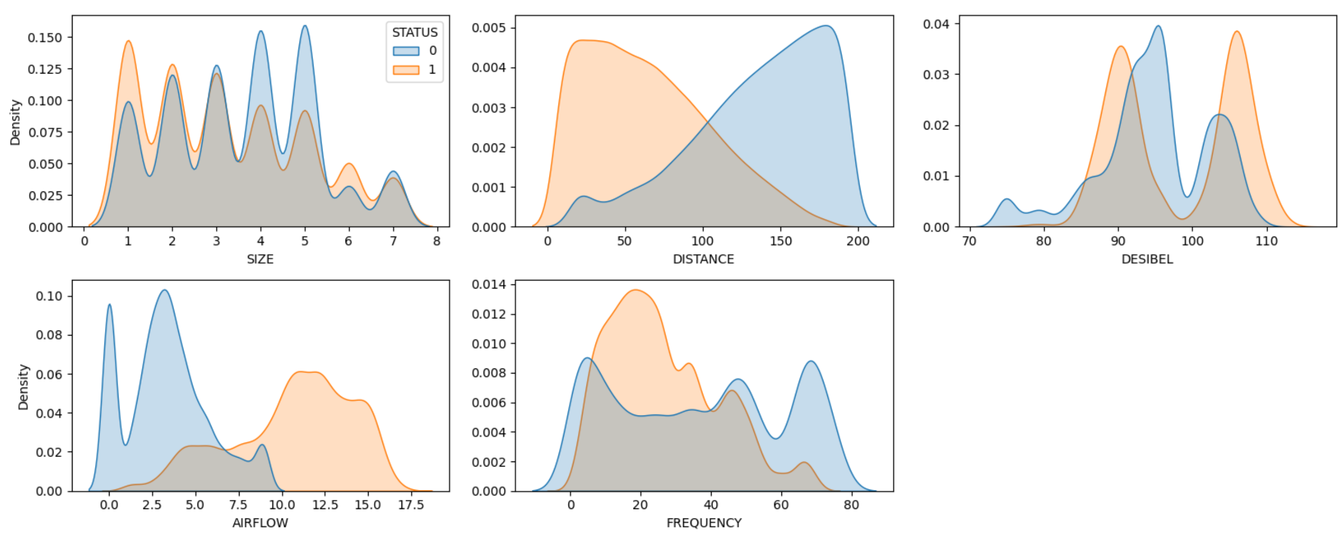

Figure 7: Distribution of Numerical Features grouped by 'STATUS'

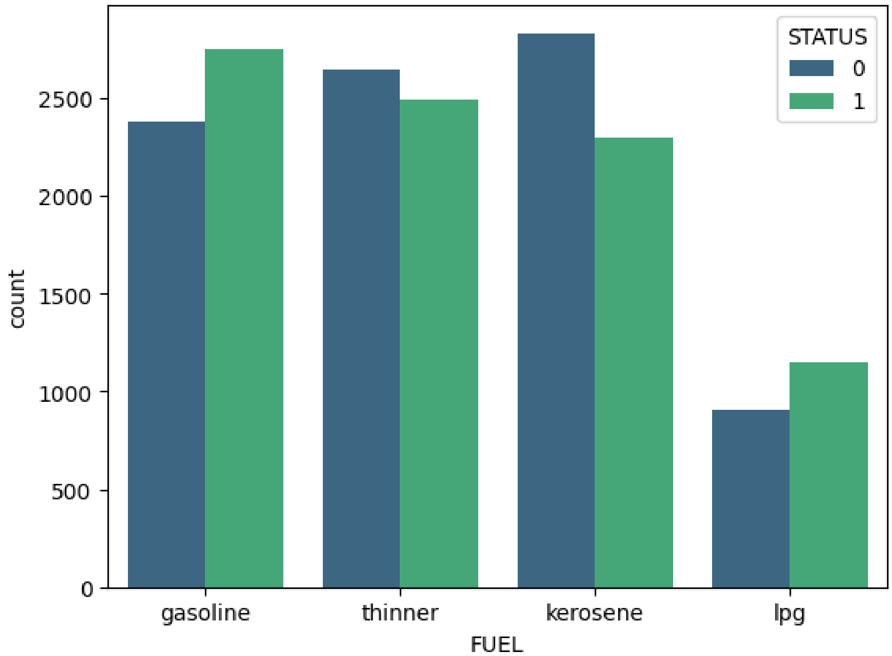

Figure 8: ‘STATUS’ counts by ‘FUEL’s

4.2 Model Fitting and Results:

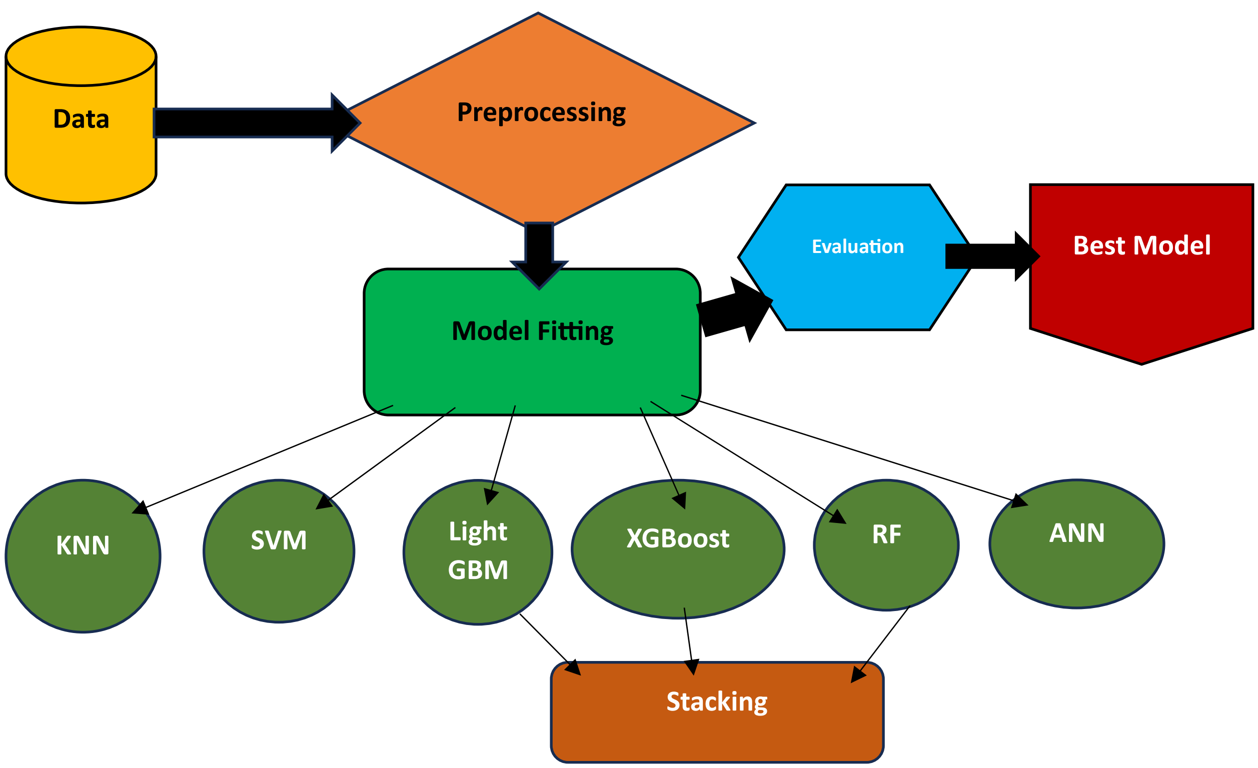

Figure 9: Conceptual Diagram

In this study, several machine learning models were employed to classify the flame extinction status based on sound wave data. The performance of these models was assessed using various evaluation metrics such as Accuracy, Precision, Recall, F1 Score, Specificity, AUC (Area Under the Curve), Cohen's Kappa, and Matthews Correlation Coefficient (MCC). Below is a detailed description of the models' performance:

The K-Nearest Neighbors (k-NN) model achieved a high accuracy of 97.02%, demonstrating strong predictive power. Its precision, recall, and F1 score were all similarly high (97.02%, 97.13%, and 97.07% respectively), indicating that the model is effective in both identifying true positives and minimizing false positives. The AUC score of 0.9910 further confirms the model's ability to distinguish between flame extinguishing and non-extinguishing states. The Cohen's Kappa and MCC scores, both 0.9404, signify a strong agreement between predicted and actual labels, highlighting the model's overall robustness.

The Support Vector Machine (SVM) model performed similarly, with an accuracy of 96.99%. While its precision, recall, and F1 score were marginally lower than those of k-NN, they still reflect excellent performance (96.96%, 97.13%, and 97.04%, respectively). The model's AUC of 0.9968 is notably high, indicating strong classification performance. Cohen's Kappa and MCC values of 0.9398 point to a high level of agreement, suggesting reliable predictions across different instances.

The LightGBM model outperformed all other models, achieving the highest accuracy of 98.77%. Its precision (98.65%), recall (98.93%), and F1 score (98.79%) indicate its exceptional capability to correctly classify both positive and negative cases. The model's AUC score of 0.9991 underscores its near-perfect ability to differentiate between classes. Cohen's Kappa and MCC, both 0.9753, indicate excellent model stability and consistency in its predictions, making LightGBM a top performer in this classification task.

XGBoost also demonstrated strong performance with an accuracy of 98.48%, and similarly high precision (98.48%), recall (98.54%), and F1 score

(98.51%). The model's AUC of 0.9988 suggests excellent separability between flame extinction and non-extinction cases. Cohen's Kappa (0.9696) and MCC (0.9696) reinforce XGBoost's consistent and reliable classification.

The Random Forest model achieved an accuracy of 98.22%, slightly trailing behind XGBoost and LightGBM. Its precision (98.53%), recall (97.97%), and F1 score (98.25%) remained strong, indicating a well-rounded classification model. The AUC score of 0.9981, along with Cohen's Kappa (0.9645) and MCC (0.9645), show that Random Forest is another reliable model in this study, although it slightly underperformed compared to LightGBM and XGBoost.

The Stacking model, which combines multiple models (LightGBM, XGBoost, and Random Forest), showed remarkable performance with an accuracy of 98.68%. Its precision (98.59%), recall (98.82%), and F1 score (98.71%) are close to LightGBM's values, reflecting its potential to generalize well across varying datasets. With an AUC of 0.9919 and Cohen's Kappa (0.9736) and MCC (0.9736), the stacking model demonstrates that the combination of base learners enhances predictive performance, making it one of the best-performing models in this task.

Finally, the Artificial Neural Network (ANN) performed solidly, with an accuracy of 96.96%. Although slightly behind some of the other models, its precision (97.49%), recall (97.58%), and F1 score (97.03%) indicate strong performance. The AUC score of 0.9965 further supports its good classification capabilities, while Cohen's Kappa (0.9392) and MCC (0.9393) show that the ANN model is reliable, though not as high-performing as models like LightGBM and Stacking.

Overall, while all models demonstrated strong classification capabilities, LightGBM and Stacking stood out as the top performers. Their higher accuracy, precision, recall, and AUC scores suggest that these models are highly suitable for the task of classifying flame extinction, possibly offering better generalization than traditional models like k-NN and SVM.

Additionally, the consistent performance of the Stacking model highlights the potential of model ensemble techniques to boost prediction accuracy.

Model	Accuracy	Precision	Recall	F1 Score	Specificity	AUC	Cohen Kappa	MCC
K-Nearest Neighbors	0.9702	0.9702	0.9713	0.9707	0.9691	0.9910	0.9404	0.9404
SVM	0.9699	0.9696	0.9713	0.9704	0.9685	0.9968	0.9398	0.9398
LightGBM	0.9877	0.9865	0.9893	0.9879	0.9860	0.9991	0.9753	0.9753
XGBoost	0.9848	0.9848	0.9854	0.9851	0.9842	0.9988	0.9696	0.9696
Random Forest	0.9822	0.9853	0.9797	0.9825	0.9848	0.9981	0.9645	0.9645
Stacking	0.9868	0.9859	0.9882	0.9871	0.9854	0.9919	0.9736	0.9736
ANN	0.9696	0.9649	0.9758	0.9703	0.9632	0.9965	0.9392	0.9393

Table 4: Performances of the models

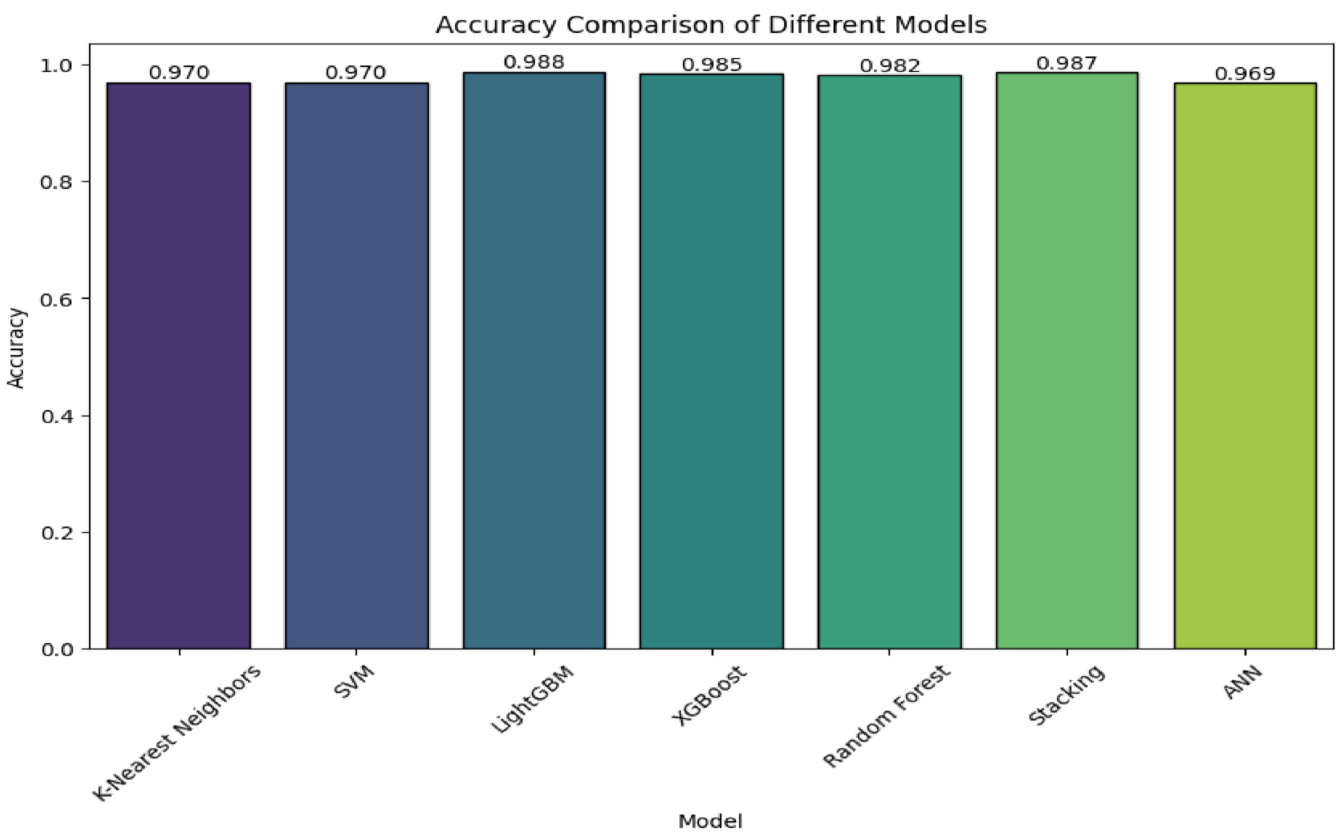

Figure 10: Barplot of Accuracies

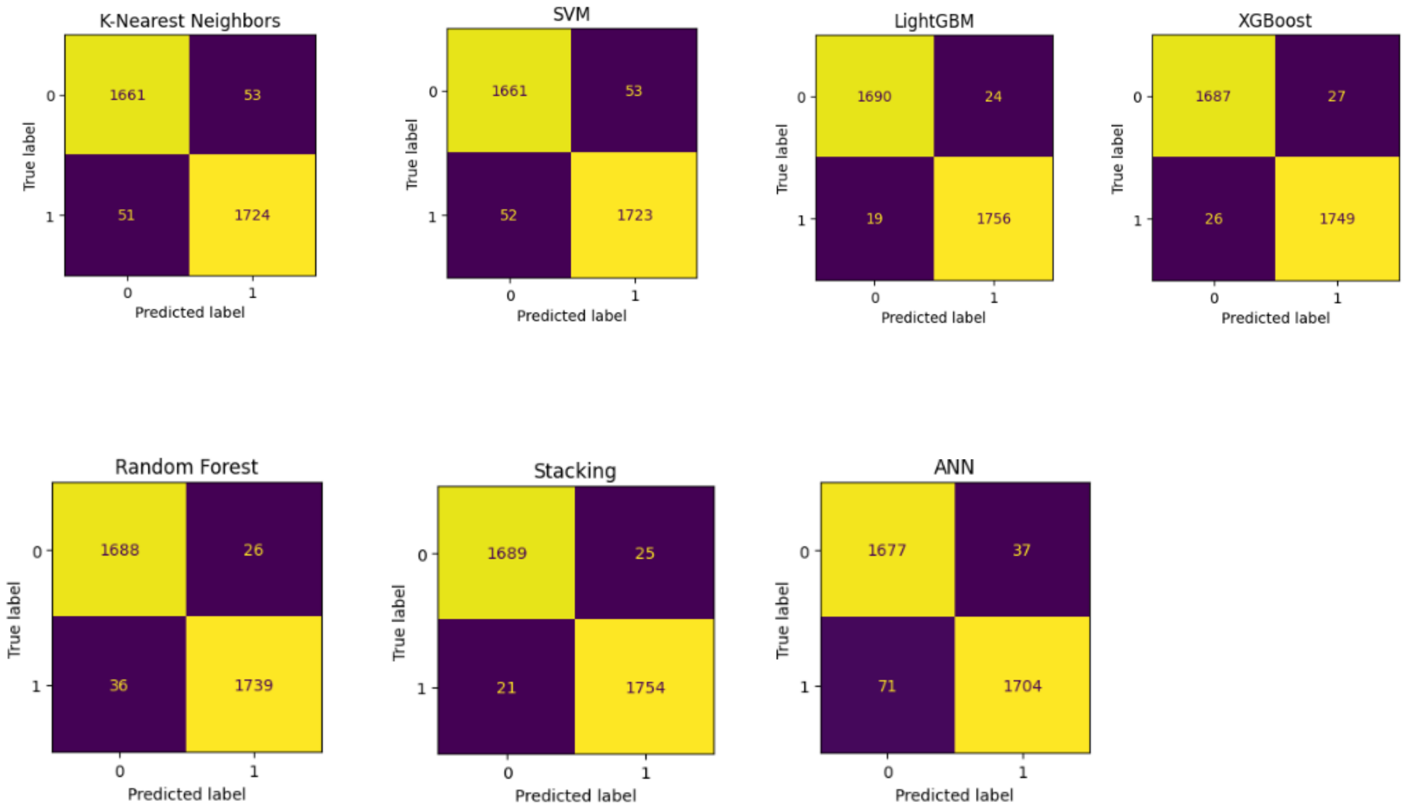

Figure 11: Confusion Matrices

4.3 Model Forecasting Power Overview

In evaluating the forecasting capabilities of the models, key metrics such as true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) provide a clear picture of each model's ability to generalize predictions effectively.

1. **Stacking** demonstrates exceptional forecasting accuracy, with only 25 false positives and 21 false negatives. This indicates that the model can successfully identify both positive and negative classes with minimal error, providing a reliable balance of precision and recall. Its performance ranks among the best for future forecasting, particularly in tasks requiring robust classification accuracy.

2. **SVM (Support Vector Machine)** performs well but presents slightly more classification errors than Stacking, with 53 false positives and 52 false negatives. Though it retains good predictive power, the higher error

rates suggest that while SVM is effective, it is less accurate than ensemble models like Stacking or LightGBM.

3. **ANN (Artificial Neural Network)** displays strong performance, though it struggles with higher false negatives (71). While effective at predicting true negatives, its comparatively lower recall due to missed positive classifications indicates a moderate level of forecasting power, particularly when compared to more complex ensemble methods.
4. **K-Nearest Neighbors (KNN)** balances recall and precision but has relatively high false positives (53). The model tends to overpredict positive instances, which might compromise its overall classification accuracy, particularly when distinguishing between the two classes.
5. **LightGBM** excels with one of the lowest error rates across all models, having only 24 false positives and 19 false negatives. Its exceptional ability to forecast both classes with high accuracy makes it one of the most powerful models for future prediction tasks. It balances precision and recall at an impressive level.
6. **Random Forest** follows closely behind LightGBM, with 26 false positives and 36 false negatives. While its classification errors are slightly higher, it remains a reliable model for forecasting and offers strong predictive accuracy.
7. **XGBoost** also delivers impressive performance, with 27 false positives and 26 false negatives. It demonstrates a reliable balance of high precision and recall, placing it in the same tier as LightGBM in terms of forecasting ability.

In summary, LightGBM, XGBoost, and Stacking models exhibit the strongest forecasting abilities, excelling at identifying both positive and negative instances with minimal error. These models provide highly reliable future predictions. SVM, ANN, and KNN offer solid predictive performance but with slightly higher classification errors, making them less effective in terms of forecasting compared to the ensemble methods.

4.4 Selecting the best value of K for KNN Classifier:

To select the optimal value of K for the K-Nearest Neighbors (KNN) model, I employed a Python for loop to iterate through different values of K. The selection process was based on the model's accuracy performance for each value. After evaluating the results, K=3 was determined to be the best-performing value, yielding the highest accuracy. This systematic approach ensures that the chosen K maximizes the model's predictive capability for my dataset, aligning it with the desired classification goals.

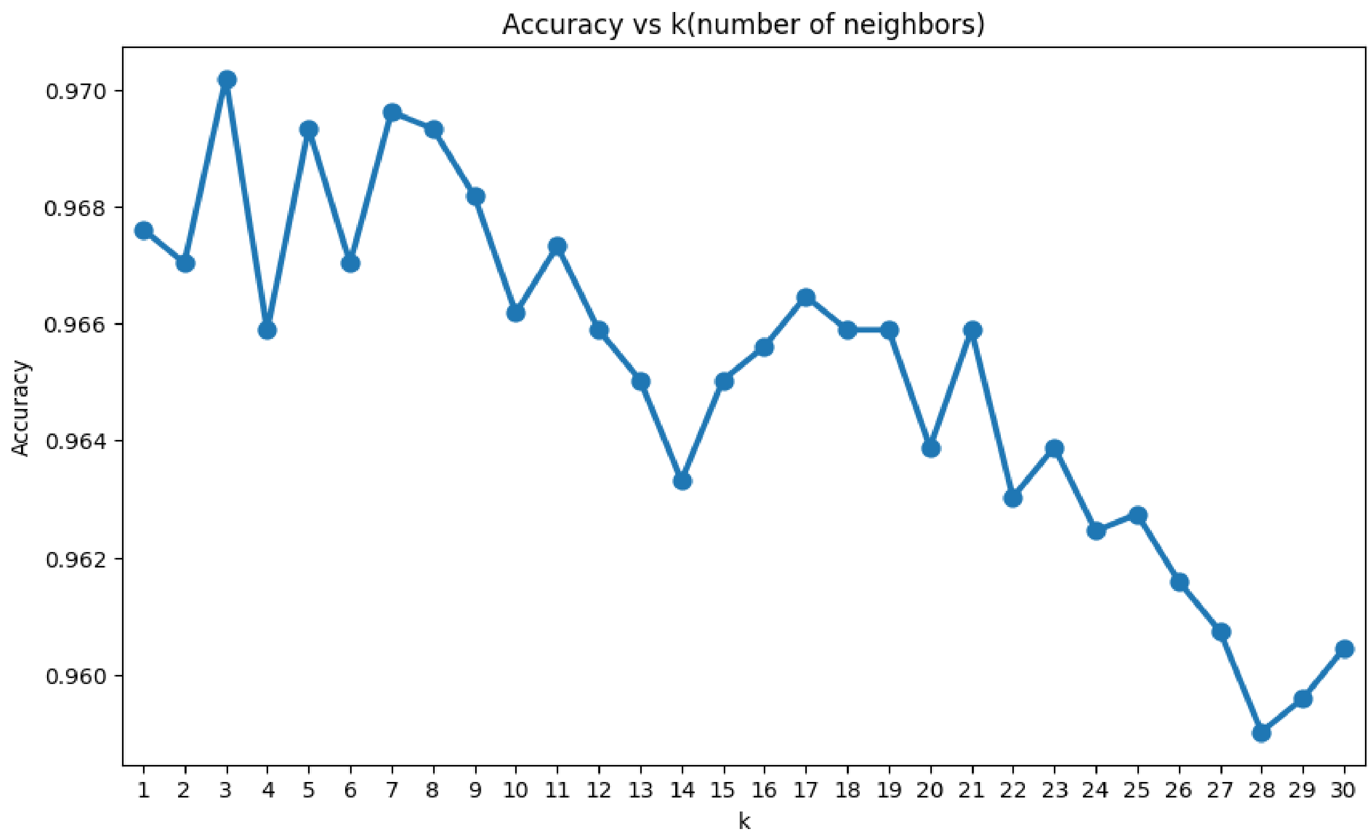

Figure 12: Selecting the best K for KNN

Chapter 5

DISCUSSIONS AND CONCLUSION

The focus of this project was to explore the potential of sound wave technology in fire suppression, a relatively novel approach that aims to extinguish flames by manipulating sound frequencies, airflow, and sound pressure. Through a series of experiments, we gathered detailed data on the interaction between these variables, including frequency, decibel levels, and airflow, and their effects on the extinguishing process. Machine learning techniques were employed to model and predict the extinguishing and non-extinguishing states of the flame, helping to enhance our understanding of the key factors influencing the efficiency of the system.

5.1 Correlation Analysis and Key Insights

One of the critical steps in the project was performing a correlation analysis to investigate the relationships between the features in the dataset. The results revealed significant connections between the core variables involved in the extinguishing process. Notably, an increase in frequency was associated with higher decibel levels, while airflow decreased as frequency rose, demonstrating an inverse relationship. Additionally, distance was shown to negatively affect airflow, highlighting its critical role in optimizing flame suppression efficiency. The frequency of the sound wave emerged as the most influential factor, confirming its importance in determining the effectiveness of the extinguishing process. These correlations provided valuable insights into how sound waves influence the flame extinguishing mechanism and how various parameters must be fine-tuned to achieve optimal results.

5.2 Machine Learning Model Performance

The application of machine learning models was a pivotal aspect of this project. We trained multiple models—K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Random Forest (RF), XGBoost, LightGBM, Artificial Neural Networks (ANN), and a Stacking model—using experimental data. Cross-validation was employed to ensure the reliability of the models' predictions and to objectively assess their classification accuracy in determining the flame's extinguishing or non-extinguishing states.

Among the models, Stacking, which integrated XGBoost, LightGBM, and Random Forest as base models, emerged as the best performer with an accuracy of 98.68%, precision of 98.59%, and a recall of 98.82%. LightGBM followed closely with a similarly high accuracy of 98.77%, while XGBoost and Random Forest also exhibited strong classification performance. These models were able to correctly classify the majority of data samples, as shown by their high F1 scores, AUC values, and Cohen Kappa statistics. The confusion matrices generated for each model provided a clear understanding of how accurately the models predicted flame extinction. The relatively small differences in accuracy between the models, while seemingly minor, translated into significant differences in the number of correctly classified data points given the large size of the dataset.

Interestingly, despite the ANN model being widely used in many predictive tasks, it slightly underperformed compared to the tree-based models and Stacking. This could be due to the non-linear relationships in the data, which are better captured by ensemble methods like Random Forest, LightGBM, and XGBoost. Overall, the machine learning approach offered a sophisticated decision-making framework for predicting flame extinguishing success, with Stacking and LightGBM leading the way in accuracy and robustness.

5.3 Practical Implications

The results from this project have significant implications for fire suppression systems using sound waves. We determined that specific frequency ranges were most effective in extinguishing flames, depending on the distance from the fire source. For instance, frequencies between 10–50 Hz were optimal for distances of 10–100 cm, while lower frequencies (10–28 Hz) were more effective at greater distances (100–180 cm). In the case of liquid petroleum gas (LPG) fuel, the ideal frequency range for flame suppression was found to be between 10–45 Hz for distances of 10–140 cm. These findings are crucial for optimizing the use of sound wave technology in practical applications.

Furthermore, the sound intensity generated during the experiments (ranging from 85 to 113 dB depending on the distance) and the airflow produced by sound pressure were critical factors. The airflow created by sound pressure reached 17 m/s at a frequency of 30 Hz at a distance of 10 cm, significantly enhancing the cooling effect on the fuel container. This rapid cooling, almost twice as fast as under normal conditions, emphasizes the effectiveness of sound waves not only in extinguishing flames but also in managing heat and preventing re-ignition.

5.4 Future Directions

This project lays the foundation for further advancements in sound wave-based fire suppression systems. One promising avenue for future research is the development of a decision support system that integrates the findings of this study to guide real-time decisions in fire extinguishing operations. Such a system could dynamically adjust the sound wave parameters based on the specific characteristics of a fire, improving both efficiency and safety.

Moreover, the creation of an expert system that automates the fire suppression process is a logical next step. By leveraging the machine learning models developed in this project, an automated system could use sensors to detect fire conditions and apply the optimal sound wave frequencies, airflow, and decibel levels without human intervention. This would greatly enhance the system's responsiveness, accuracy, and overall effectiveness in extinguishing fires quickly and safely.

Additionally, expanding the experimental setup to include different types of fuels and environmental conditions would provide a more comprehensive understanding of the effectiveness of sound wave fire suppression. The insights gained from this expanded research could further refine the machine learning models and improve the generalizability of the results.

5.5 Conclusion

This project has demonstrated the significant potential of sound wave technology in fire suppression and highlighted the critical role of machine learning in optimizing the process. By identifying key relationships between sound frequency, airflow, and flame extinguishing effectiveness, and by training high-performing models to predict flame outcomes, this research contributes valuable knowledge to the field of fire safety. The findings provide a solid foundation for future developments in automated, sound-based fire extinguishing systems, which could revolutionize the way we approach fire suppression in the future.

REFERENCES

[1] Y. S. Taspinar, M. Koklu, and M. Altin, "Classification of flame extinction based on acoustic oscillations using artificial intelligence methods," *Case Studies in Thermal Engineering*, vol. 28, Dec. 2021, Art. no. 101561. Available: https://www.sciencedirect.com/journal/case-studies-in-thermal-engineering/vol/28/suppl/C

[2] M. Koklu and Y. S. Taspinar, "Determining the extinguishing status of fuel flames with sound wave by machine learning methods," *IEEE Access*, vol. 9, pp. [Pages], Jun. 2021. DOI: 10.1109/ACCESS.2021.3088612. Available: https://doi.org/10.1109/ACCESS.2021.3088612

[3] A. B. Morgan and J. W. Gilman, "An overview of flame retardancy of polymeric materials: Application, technology, and future directions," *Fire Mater.*, vol. 37, no. 4, pp. 259–279, Jun. 2013.

[4] D. Paton, P. T. Buergelt, F. Tedim, and S. McCaffrey, "Wildfires: International perspectives on their social-ecological implications," in *Wildfire Hazards, Risks and Disasters*, Amsterdam, The Netherlands: Elsevier, 2015, pp. 1–14.

[5] T. D. Gill, J. Blackhall, A. L. Edworthy, and M. R. Murphy, "Thermal response in materials during flame extinction: Future directions for computational models," *J. Thermal Sci.*, vol. 55, no. 12, pp. 431-450, Mar. 2020.

[6] S. D. Hill and B. A. Bark, "Sound-based flame suppression techniques and their optimization," *Acoustic Eng. Rev.*, vol. 18, no. 2, pp. 102-110, Aug. 2019.

[7] J. Z. Wang, Y. L. Huang, and X. F. Liu, "AI-based models for predicting fire hazards and controlling flame dynamics," *Int. J. Fire Sci.*, vol. 44, no. 3, pp. 278-292, Oct. 2018.

[8] R. R. Patel, "Recent developments in sound wave fire extinguishing technology and future prospects," *Fire Prot. Rev.*, vol. 29, no. 7, pp. 175-189, Dec. 2019.