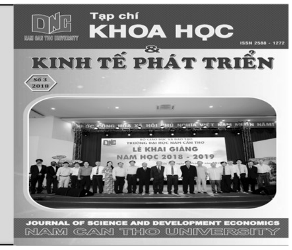
Journal of Science and Development Economics
Nam Can Tho University

Website: jsde.nctu.edu.vn



Application of artificial intelligence in smoke detection with streaming updated data

Tran Ho Xuan Mai[1], Ngo Ho Anh Khoi[1], Trinh The Luc[2], Tran Huynh Khang[1]
[1]*Faculty of Information Technology, Nam Can Tho University*
[2]*Hanoi University of Mining and Geology*

Received: November 10, 2023
Revised: November 30, 2023
Accepted: December 5, 2023

Keywords: *artificial intelligence, decision tree, smoke detection*

Từ khóa: *cây quyết định, phát hiện khói, trí tuệ nhân tạo*

ABSTRACT

At present, artificial intelligence (AI) is one of the most rapidly developing fields in science and technology. In the modern context, AI technologies have become a highly researched area globally, leading to breakthrough technologies that enhance efficiency and effectiveness across various sectors, including environmental security. In today's landscape, where fire and explosion incidents can occur anywhere, there is growing concern about environmental pollution caused by hazardous substances released from fires. These subtainces can come into contact with soil, water, and air, posing a threat to the environment. This technology has the potential to greatly impact human an animal health, as well as the environtment, by reducing the occurrence of diseases and adverse effects caused by fires. To investigate this topic, the "Smoke Detection Dataset" provided by Deep Contractor was selected as the primary dataset. The Decision Tree Classifier algorithm, which has shown significant advancement in various fields of artificial intelligence, was chosen to build classification models for this study. Its proven effectiveness and versatility make it suitable foundation for the system being developed.

TÓM TẮT

Hiện nay, trí tuệ nhân tạo (AI) là một trong những lĩnh vực khoa học công nghệ phát triển nhanh nhất. Trong bối cảnh hiện đại, công nghệ AI đã trở thành một lĩnh vực được nghiên cứu kỹ lưỡng trên toàn cầu, dẫn đến những công nghệ mang tính đột phá giúp nâng cao hiệu suất và hiệu suất trên nhiều lĩnh vực, bao gồm *cả an ninh môi trường. Trong bối cảnh ngày nay, nơi sự cố cháy nổ có thể xảy ra ở bất cứ đâu, mối lo ngại ngày càng tăng về ô nhiễm môi trường do các chất độc hại thoát ra từ hỏa hoạn. Những chất phụ này có thể tiếp xúc với đất, nước và không khí, gây ra mối đe dọa cho môi trường. Công nghệ này có khả năng tác động lớn đến sức khỏe con người và động vật cũng như môi trường bằng cách giảm sự xuất hiện của bệnh tật và tác động bất lợi do hỏa hoạn gây ra. Để điều tra chủ đề này, "Bộ dữ liệu phát hiện khói" do Deep Contractor cung cấp đã được chọn làm bộ dữ liệu chính. Thuật toán Phân loại cây quyết định, đã cho thấy sự tiến bộ đáng kể trong các lĩnh vực trí tuệ nhân tạo khác nhau, đã được chọn để xây dựng các mô hình phân loại cho nghiên cứu này. Hiệu quả và tính linh hoạt đã được chứng minh của nó làm cho nó trở thành nền tảng phù hợp cho hệ thống đang được phát triển.*

1. INTRODUCTION

Fire incidents are one of the leading causes of property damage, loss of life, and sometimes even hinder the development of an area. However, in certain cases, detecting smoke and fire alarms can become challenging, especially in hard-to-reach locations or in dusty and hazy environments. Vietnam is one of the countries with the highest number of fire incidents in the Southeast Asia region. In the article titled "The Haunting Fires of 2021," journalist Van Ngan writes, "Throughout the country, thousands of small and large fires occurred, some of which had extremely serious consequences" (Van Ngan, 2021) [1]. According to statistics from the Ministry of Public Security (Ministry of Public Security, 2023), from 2014 to 2019, there were approximately 11,000 fire incidents, including over 1,000 major fires, causing severe damage to property and human lives. However, the annual report on fire prevention and firefighting by the Ministry of Public Security reveals that only around 20% of buildings in Vietnam are equipped with modern fire alarm and firefighting systems. This underscores the necessity of researching and developing smoke detection and fire alarm solutions. On September 13, 2023, the Ministry of Public Security issued a directive to the Firefighting and Rescue Police Department and local police departments to enhance the effectiveness of state management in fire prevention, firefighting, and rescue operations. The directive aims to address fire and explosion situations and minimize casualties and property damage, particularly in densely populated residential areas like mini apartments and high-density rental service businesses. Detecting smoke and fire alarms is crucial in Vietnam to minimize fire-related losses and ensure the safety of the population. Currently, there is a growing demand for modern smoke detection and fire alarm systems, especially in industrial zones and densely populated residential areas. Detecting smoke and fire alarms is a critical issue in ensuring community safety.

The World Health Organization (WHO, 2018) [2] reported that in 2018, approximately 180,000 people died globally due to fires, and this number continues to rise over time. In 2019, in Europe, wildfires in Portugal and Spain devastated thousands of hectares of forests, causing significant environmental and economic damage. Additionally, the presence of smoke from these fires also impacted air quality and the health of the population. The article by Pham and colleagues proposes a method to assess fire risk and design firefighting systems for universities. This method utilizes multi-state assessment techniques and evaluation criteria such as independence, dispersion, functionality, and efficiency. The results indicate that this approach can enhance the fire safety level for tall building structures. In the past, fire alarms were often given through the use of bells or sirens to alert everyone in a potentially dangerous area. However, this method may not guarantee effectiveness in certain situations, especially in large buildings or when there are many people present. Additionally, manually activating fire alarms can also pose challenges, particularly when people are not familiar with how to use them or when there are not enough alarms distributed throughout the entire area. Furthermore, the use of manual fire alarms can lead to errors as users may not recognize the hazardous situation or may not use the alarms correctly. This highlights the importance of modernizing and improving fire detection and alarm systems to ensure quicker and more accurate responses in case of fire emergencies. The article by Rongbin Xu investigates the impact of the large wildfires that occurred in Australia during the 2019-2020 summer on human health and visibility. The research findings indicate that smoke from the wildfires had a significant adverse effect on the health of people in cities located hundreds of kilometers away. Health risks increased, especially for individuals with pre-existing respiratory conditions. Additionally, the smoke reduced visibility in major cities, affecting daily life and transportation. In the article of Khan Muhammad and colleagues, the authors propose a cost-effective CNN architecture for fire detection in surveillance videos. The model draws inspiration from the GoogleNet architecture and is specifically fine-tuned to focus on computational complexity and detection accuracy. Through experiments, it has been demonstrated that the proposed architecture outperforms existing fire detection methods based on manual features as well as those based on the AlexNet architecture. The article of Panagiotis Barmpoutis and colleagues (Panagiotis, 2014) [3] introduces a novel method for smoke detection in videos. This method aims to distinguish smoke from moving objects by applying spatio-temporal analysis, smoke motion modeling, and dynamic texture recognition. It initially identifies candidate smoke regions in a frame using background subtraction and color analysis based on the HSV model. Subsequently, a spatio-temporal smoke model, including spatial energy analysis and spatio-temporal energy analysis, is applied in candidate regions. Gradient orientation histograms and optical flow (Hoghofs) are computed to capture appearance and motion information, while dynamic texture recognition is applied within each candidate region using linear dynamic systems and a bag-of-features approach. Finally, a combined dynamic saliency score is used to determine the presence of smoke in each candidate image region. Experimental results presented in the article show the significant potential of the proposed method.

In the article titled Fire Safety Search (2020) [4], the authors introduce fundamental elements, actions, and best practices related to effective early warning systems. It supports the development and implementation of early warning fire detection systems that prioritize human-centric, timely, and understandable alerts for at-risk individuals, including guidance on how to act upon receiving an alert. In the article by Penghui Dong, a fuzzy comprehensive evaluation and Bayesian network-based method is proposed for building fire risk assessment. The advantages of Bayesian networks in handling uncertain problems have been used to model and analyze fuzzy issues that can be well addressed in practical scenarios. The Bayesian network model is analyzed using the example of a high-rise residential building fire accident on Yuyao Road in Shanghai. Prior probabilities come from the statistical history of fire incidents in high-rise buildings in recent years, and fuzzy probabilities are provided by experts. In the article by Myoung-Young Choi and Sunghae Jun (Choi and Jun, 2020) [5], the authors propose to build a fire risk model using statistical machine learning and optimized risk indexing. Data related to fire risk factors, including explanatory variables (X) influencing the occurrence of fire incidents and responses (Y) indicating the frequency of fire incidents, were collected. The research shows that the proposed model can predict fire risk with high accuracy and could be useful for fire safety management in high-rise buildings. The article by Souad Kamel and colleagues in South Korea focuses on developing an efficient system for early fire management to prevent material and human losses. This system utilizes Internet of Things (IoT) technology, making it feasible with the advancement of IoT. It combines low-cost IoT sensors to collect real-time data (such as temperature, the number of people at the scene of a fire, etc.) and presents all sensor readings on a single web-based dashboard. When values collected exceed specific thresholds, the system sends alerts to the building manager, allowing them to notify authorities or dispatch firefighters in real-time. A crucial aspect of this system is its ability to monitor the number of people at the scene of a fire, simplifying evacuation processes and enabling civil defense agencies to manage resources effectively. The system has been successfully tested in various scenarios within an educational building (Al-Faisaliah Women's Campus, Jeddah University, Saudi Arabia). In the article by Bá Tuấn and Văn Ngọc (Bá Tuấn - Văn Ngọc, 2021) [6] in Vietnam, the authors propose a fire risk prediction model for high-rise buildings using deep learning techniques. To build the model, they utilized a dataset of fire incidents in high-rise buildings collected from various sources. The experimental results demonstrate that the proposed model has the potential to accurately predict fire risk and could be valuable for fire safety management in high-rise buildings in Vietnam. An IoT-based system is applied for monitoring and automatic fire warnings: the system includes a central fire alarm cabinet, fire alarm devices, control devices, IoT Wi-Fi devices, and remote monitoring software for users. The system uses

LPWAN (Low Power Wide Area Network) IoT technology to connect wireless sensors for real-time data transmission to the cloud, helping to pinpoint the location of a fire outbreak (Science and Development, 2021) [7].

Continuous learning is currently a crucial approach in the field of artificial intelligence, particularly when dealing with data in constantly changing environments. In this project, artificial intelligence will be used in conjunction with a sliding window technique along with the Smoke Detection Dataset collected by Stefan Blattmann in the "Real-time Smoke Detection with AI-based Sensor Fusion" project and updated by Deep Contractor (comprising 62,630 rows with 15 data fields, measured by IoT sensors, including parameters such as temperature, humidity, brightness, and various other metrics) to develop models for classifying smoke and non-smoke from sensor signals collected by IoT devices.

To contribute to addressing the issue of fire and explosions, utilizing the collected and analyzed database to develop the main algorithms for a smoke detection and fire alerting system is a promising approach. By applying artificial intelligence and machine learning technologies using the mined database, organizations can have additional tools, research insights, and deployment results for addressing this problem through artificial intelligence systems. The research project aims to find a suitable method that can serve as the core algorithm for developing a sensor-based recognition system. The purpose of this research is to leverage artificial intelligence systems to provide the most accurate fire alert predictions, benefiting the social life in Vietnam. This project holds the potential to significantly contribute to minimizing the damage caused by fires and explosions, especially in emergency situations.

2. MATERIALS AND METHODS

During the data search process for the topic, numerous datasets were found (approximately 232 datasets related to smoke detection). However, there are 8 datasets specifically focused on smoke recognition for fire detection and alerting, which possess comprehensive parameters and the highest level of availability.

The "Smoke Detection using Classification Models" dataset by Sandesh Singh comprises two folders containing different types of data. The first folder contains 3423 images related to smoke and non-smoke, while the second folder contains 316 videos related to the images in the first folder. This dataset is used to build classification models for smoke and non-smoke based on machine learning algorithms. "Firesense by Chris Gorgolewski" is a dataset that consists of 14 data fields and over 200,000 rows. It provides information about forest fires, climate factors, geography, and the environment. The purpose of this dataset is to assist in predicting and managing forest fires more effectively.

The "Sensor-Fusion Smoke Detection Classification" dataset by Gaurav Dutta was updated and published in 2020. This dataset contains 1074 rows with 8 data fields. The data fields include parameters such as acceleration, angular velocity, and altitude measured using sensors. The data was collected during fire simulation to simulate fire alarm situations and smoke detection. This dataset is used to train classification models for smoke and non-smoke based on sensor signals. The "Smoke Detection Dataset" by Deep Contractor was collected by

Stefan Blattmann in the project "Real-time Smoke Detection with AI-based Sensor Fusion" and updated by Deep Contractor (Contractor, 2020). This dataset comprises 62,630 rows with 15 data fields measured using IoT sensors, including parameters such as temperature, humidity, brightness, and several others. The purpose of this dataset is to develop classification models for smoke and non-smoke based on sensor signals. This data is used for smoke detection and triggering fire alarms.

The "Fire and Smoke dataset" was collected by DataCluster Labs (DataCluster, 2023) [8]. It contains over 7,000 images of fire and smoke. Each image in this dataset has been manually reviewed and verified by computer vision experts at DataCluster Labs. This dataset is large, high-resolution (98% of images are HD or higher), and collected from over 400 urban and rural areas. It can be used for early fire and smoke detection, smart camera systems, fire and smoke alarm systems, and more. The "WildFire-Smoke-Dataset-Tensorflow" was created by Aluru V N M Hemateja, a student at Vellore Institute of Technology in Chennai, Tamil Nadu, India (Hemateja, 2021) [9]. This dataset contains images of smoke from wildfires collected from various sources. It consists of a total of 1,500 images with a size of 256x256 pixels and is divided into two folders, "smoke" and "non-smoke." The images in the "smoke" folder depict smoke from wildfires, while the images in the "non-smoke" folder do not have any relation to wildfires. The article "Forest Fire prediction using Machine Learning" on the Analytics Vidhya website was written by Aman Preet Gulati and published in the paper "Forest Fire prediction using Machine Learning." (Gulati, 2021) [10]. In this article, it is mentioned that the dataset contains 36,011 rows and 15 columns. The article introduces the use of Machine Learning to predict the likelihood of forest fires based on certain attributes. It also discusses the importance of predicting forest fires and the benefits of using Machine Learning for this purpose.

The datasets mentioned above are relatively good, but they were not used for classification purposes and some of them are outdated, making them unsuitable for the current research. To be usable for this research, the dataset must be numerical, well-classified, and have multiple fields to yield the most objective results. The Smoke Detection Dataset is the only one that meets these requirements, which is why it was chosen for this project's system. The dataset used in this study must fulfill specific criteria, including being numerical, well-classified, and having multiple fields to ensure the most objective results. Among the available options, the Smoke Detection Dataset is the most suitable, making it the primary dataset for this research's smoke detection and recognition system. The dataset's distinctive features are sourced from the aforementioned data and have been verified by domain experts in this field. Each feature will have the following metrics:

UTC (Coordinated Universal Time): is a time representation measured in seconds according to the UTC time standard. UTC is an international time standard used to synchronize time globally.

Temperature: Air temperature is a measurement of the degree of hotness or coldness of the air in a specific area. It is typically measured using temperature units such as degrees Celsius (°C), degrees Fahrenheit (°F), or Kelvin (K). Air temperature has an impact on weather, climate, and the activities of humans and animals.

Humidity: Air humidity, also known as humidity or relative humidity, is a measurement of the amount of water vapor present in the air relative to the maximum amount the air could hold at the same temperature and pressure. Humidity is expressed as a percentage and can range from 0% (no water vapor in the air) to 100% (air is saturated, unable to hold any more water vapor).

TVOC: Volatile Organic Compounds (VOC) which are a group of organic substances that have a high tendency to evaporate into the air. TVOC is usually measured as a percentage. These compounds are commonly utilized to evaluate the air quality in various settings, including residential, workplace, and public environments. They are of particular concern in buildings, offices, and residential areas.

eCO2: Equivalent CO2 concentration is often calculated based on various values such as TVCO (Total Volatile Organic Compounds).

Raw H2: Raw molecular hydrogen, uncompensated (Deviation, temperature, etc.).

Raw Ethanol: Crude ethanol gas, in the form of ethanol (C2H5OH), that has not undergone processing or refinement. Ethanol is an organic compound, a type of alcohol, and a common solvent used in various industries.

Pressure: Air pressure is the force exerted on a specific area. In the case of air pressure, it is measured in units called Pascals (Pa), or other pressure units such as bar, psi (pounds per square inch), atm (atmosphere), mmHg (millimeters of mercury), etc. Air pressure plays a crucial role in various fields, such as aerospace, metrology, and industrial manufacturing processes.

PM1.0, PM2.5: Particle size < 1.0 µm: PM1.0 (Particulate Matter 1.0). Particle size between 1.0 µm and 2.5 µm: PM2.5 (Particulate Matter 2.5).

NC0.5, NC1.0, NC2.5: Particle size < 0.5 µm: NC0.5 (Number Concentration 0.5). Particle size between 0.5 µm and 1.0 µm: NC1.0 (Number Concentration 1.0). Particle size between 1.0 µm and 2.5 µm: NC2.5 (Number Concentration 2.5)

Aspect Ration: Aspect Ratio refers to the ratio between the width and height of an image or video. It is represented as a numerical value or a percentage, helping determine the shape and format of the content.

Label: These are critical and particularly important values. The label takes one of two values: "0" or "1". If the class is "smoke present," the label is 1, if the class is "no smoke," the label is 0. The predictions include a total of 17,873 data points of type 0 and 44,757 data points labeled as 1.

The experiments will be conducted through a nonlinear experimental model (Batch Learning, batch size = 100), meaning the system will have to perform 100 steps (Batch 100), with each step containing approximately 438 data points. In this project, there are two outputs: Fire Detection and No Fire Detection. Therefore, positive can be considered as Fire Detection, and negative as No Fire Detection. The indicators TP, TN, FP, FN have the following meanings respectively:

- TP (True Positive): Correctly detecting the presence of fire.

- TN (True Negative): Correctly detecting the absence of fire.

- FP (False Positive): Incorrectly detecting the presence of fire when there is none.

- FN (False Negative): Incorrectly detecting the absence of fire when it's present.

It's important to have:

These indicators are important for assessing the performance and accuracy of the model in fire detection

Incorrect Ground Truth Labels (0)	**True Ground Truth Labels (1)**	
FP	TP	**True positive predictions. (1)**
TN	FN	**False positive predictions. (0)**

Correct, TPR (True Positive Rate) is also known as the sensitivity or recall, and it represents the ratio of true positive predictions (correctly detecting the positive class) to the actual positive instances in the dataset. It's calculated using the formula:

$$\text{TPR} = \frac{TP}{TP+FN}$$

Correct, TNR (True Negative Rate) is also known as specificity and represents the ratio of true negative predictions (correctly detecting the negative class) to the total actual instances of the negative class in the dataset. It's calculated using the formula:

$$\text{TNR} = \frac{TN}{TN+FP}$$

After obtaining these two metrics, Balanced Accuracy is calculated using the formula:

$$Balance\ Accuracy = \frac{TPR + TNR}{2}$$

3. RESULTS AND DISCUSSION

The model used for conducting the scientific experiment has been specifically discussed in the previous section. Therefore, this section focuses on analyzing and comparing the results among the algorithms. The bar chart (both stepwise and averaged) below will provide a comprehensive evaluation of the algorithms' effectiveness. The age of the algorithm model used for comparison is set to 15 (15 data groups). To enhance objectivity, the average ratio is calculated from the results of 10 runs, i.e., the total result of 10 runs divided by 10. The dataset used in the experiment consists of two parts: one for training and the other for testing. The training dataset comprises 43,841 data rows, and the testing dataset contains 18,789 data rows. Although this dataset will vary in each experiment, the data remains unchanged (data conservation) and is reshuffled. This is achieved through the application of artificial intelligence methods, specifically algorithms combined with the sliding window approach. By utilizing charts to compare the average results of the algorithms, this method ensures the fairest representation of data authenticity during result comparison. The average experimental results of the algorithms are represented in the chart below (Figure 1).

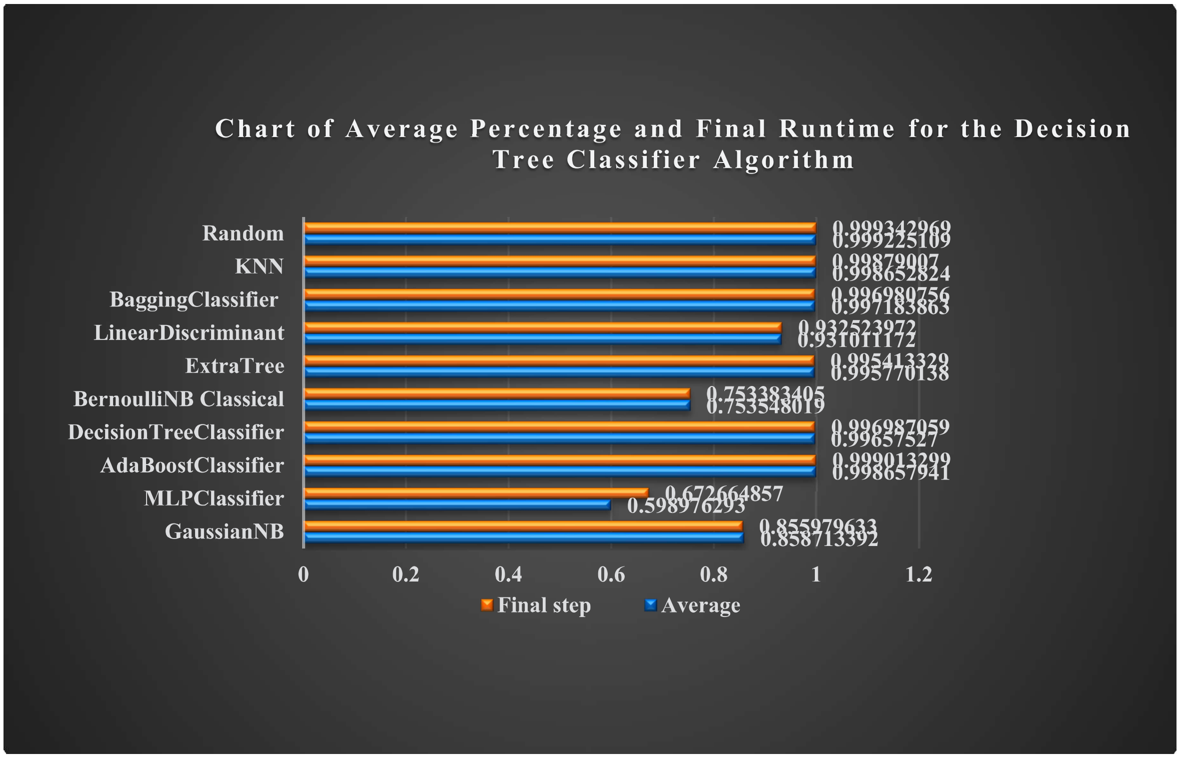


Figure 1. Average Experimental Algorithm Chart Based on Age (DecisionClassifier)

Based on the data in the chart, we can analyze the performance of the DecisionClassifier as follows:

Observing the chart, we can see that the DecisionClassifier maintains a stable performance above the 95% threshold. The average accuracy is achieved above 95% (specifically 99.70%), and the accuracy at the final step is also noteworthy, exceeding 95% (specifically 99.70%). This demonstrates the stability and ability to maintain a high-performance level of the model. This capability ensures safety and protection in situations where high sensitivity to smoke detection is needed. In addition to calculating the algorithm's average results, another approach to analyzing the effectiveness of the experimental model is based on age-wise analysis. This analysis provides a comprehensive and detailed view, helping us understand the model's capabilities for each specific age group. This approach allows us to evaluate and draw more accurate conclusions about the experimental model's performance. To illustrate this analysis, the chart below shows the results of the experimental model based on age groups (Figure 2):

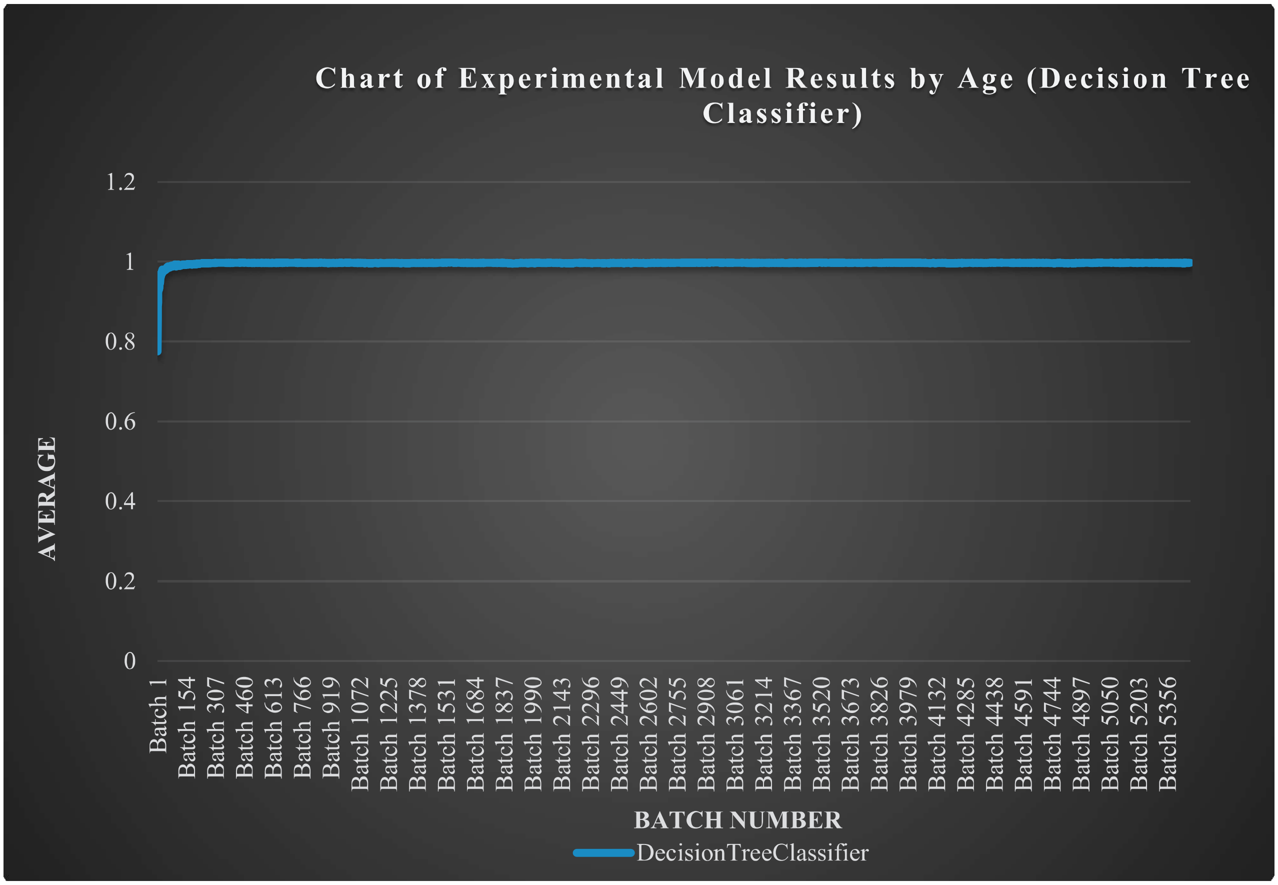


Figure 2. Experimental Model Results Chart by Age Groups (Data Groups)

Analyzing the chart, we observe that the DecisionClassifier algorithm has a relatively moderate start, achieving quite good results in the initial steps ranging from 79% to 98% over the first 84 steps, and then it starts to stabilize. Looking at the chart, we can see that the DecisionClassifier exhibits relative stability. However, it's important to note that the highest accuracy of the DecisionClassifier can reach up to 99.9%, which is a considerably high value that can be compared to many other algorithms. Specifically, during the phase from step 3505 to 3514, as shown in the chart below (Figure 3).

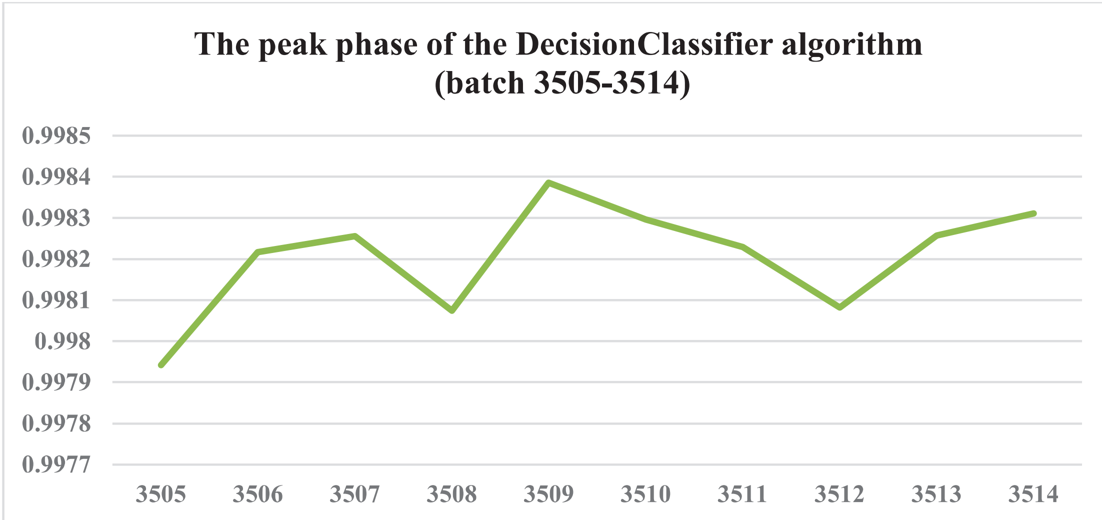


Figure 3. Highest Phase of the DecisionClassifier Algorithm (Batch 3505-3514)

The DecisionClassifier not only maintains a performance level above 99% but also excels and demonstrates greater stability compared to other algorithms. This is attributed to its ability to handle non-continuous and missing data. This capability reduces the preprocessing workload while increasing the flexibility and effectiveness of the model across various types of data. The DecisionClassifier consistently demonstrates reliability, even in the initial stages before reaching its best performance during the run, although it remains at a relatively high level. Notably, this model quickly recovers and maintains stability, a trend clearly evident in the chart below (Figure 4). The model's rapid recovery and sustained high performance following the initial phases underscore the robustness and dependability of the DecisionClassifier in data prediction and classification tasks.

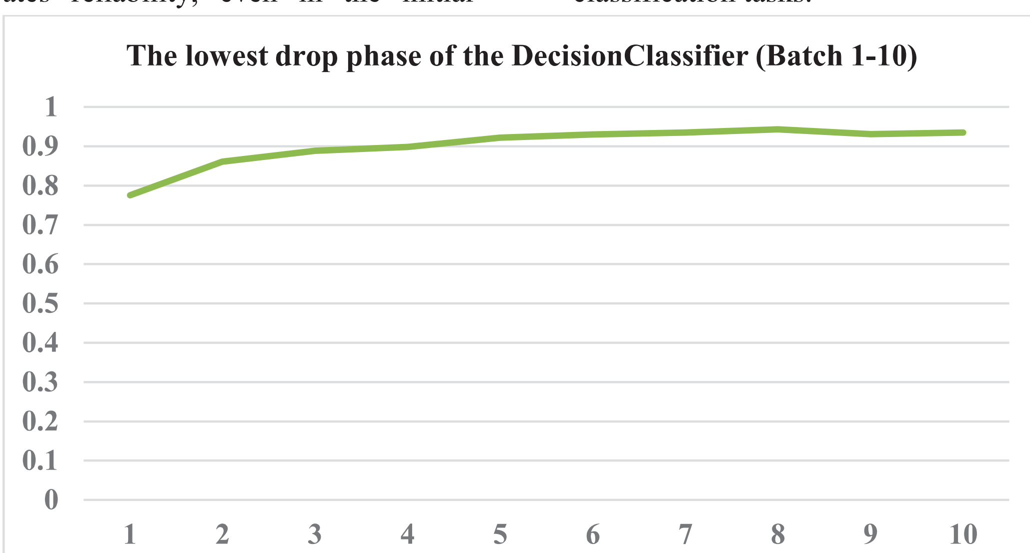


Figure 4. Lowest Drop Phase of the DecisionClassifier (Batch 1-10)

When examining the data from Batch 1 to Batch 10, we can clearly observe the variations in the achieved accuracy of the Decision Tree model. Initially, the accuracy increased from 0.793 (79.30%) to 0.880 (88.02%), creating significant fluctuations during the initial stabilization phase. However, from Batch 5 onwards, the model started to exhibit more stability and rapidly elevated the accuracy from 0.930 (93.04%) to 0.955 (95.56%) in a short time span. This demonstrates the model's ability to quickly recover and maintain stability after reaching a consistent level. In summary, within the range from Batch 1 to Batch 10, the Decision Tree model shows significant initial variations followed by maintaining a stable level and a quick recovery capability during the runtime. The model's performance continues to gradually improve and maintains a high level of accuracy after reaching a good accuracy rate. This evaluation underscores the impressive potential of the Decision Tree in prediction and classification tasks.

Installation:

The system comprises two main functionalities: algorithm installation (developer) and prediction (end user). The project will encompass functional buttons like prediction, running classic algorithms, a list of processed models, system configuration, and login. It will be implemented in a web-based environment and depicted using the use case diagram below (Figure 5).

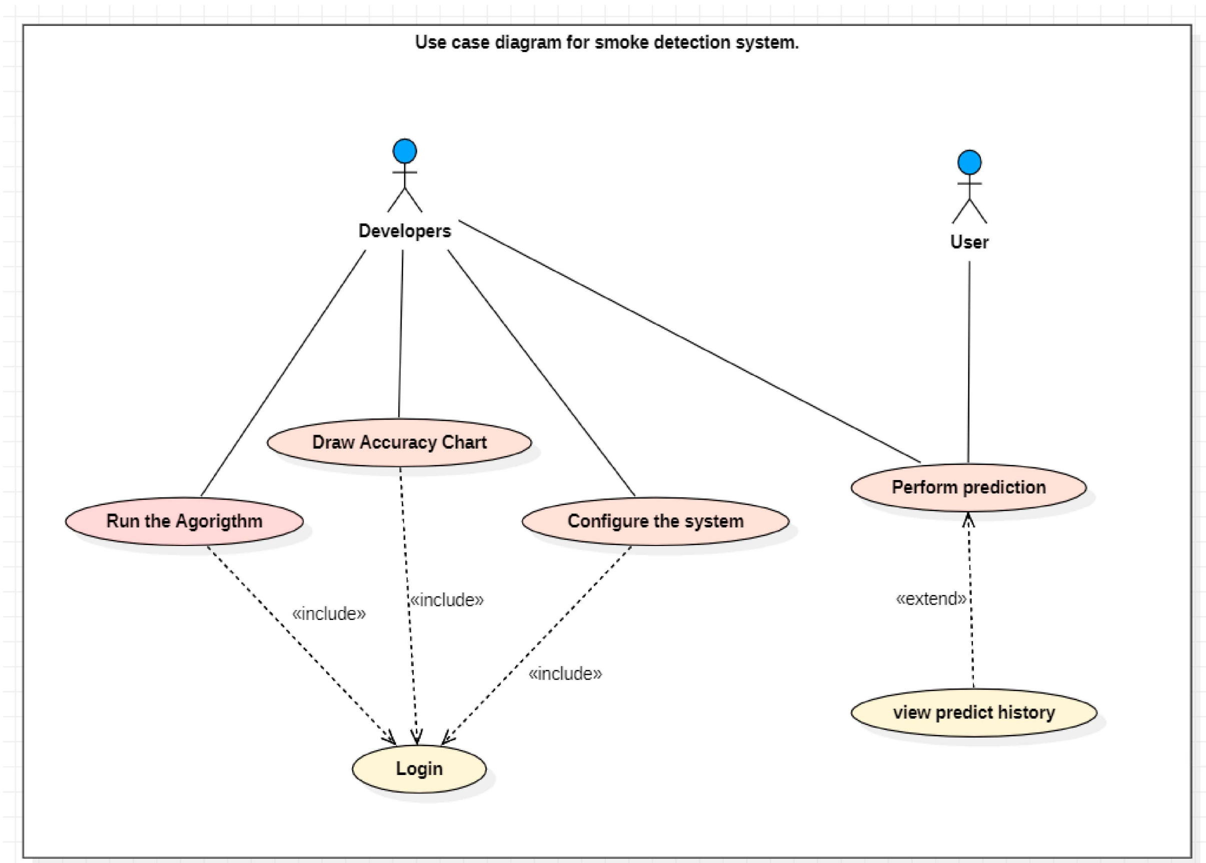


Figure 5. Use-case diagram of the system

System Configuration:

In this project, the system is packaged as a .ZIP file and compressed into a file named "BaoCaoThucTap_main". After users download and extract it, they will find a folder named "BaoCaoThucTap." Inside this folder, there are files for installation and running the program, such as "_SETUP_", requirements.txt, and RunServer.bat. The system requires a computer with a stable internet connection, minimum configuration of Windows 10, 2GB RAM, and 10GB or more of available disk space to ensure smooth and stable performance. To perform the installation, navigate to the "_Setup_" folder. The Python version required for installation is 3.9.9 (python-3.9.9-amd64.exe). Install the relevant libraries by running the CaiThuVien.bat file. The Remove.bat file can be used to delete all program data. The system can change the administrator account in 'BaoCaoThucTap_main/dataUser.csv'.

After completing the environment setup process, there will be a "Remove.bat" file, which is used to delete unnecessary data files, including those used for testing. This file should only be used in two cases: right after extraction and installation, and when users want to delete all previously run data. Run the RunServer.bat file to start the program, then open a web browser and access the address http://127.0.0.1:8000. Below is the main interface page, featuring a button to initiate the prediction process, The Start Prediction and 13 input fields to enter diagnosis data corresponding to the input parameters in the data constraints table.

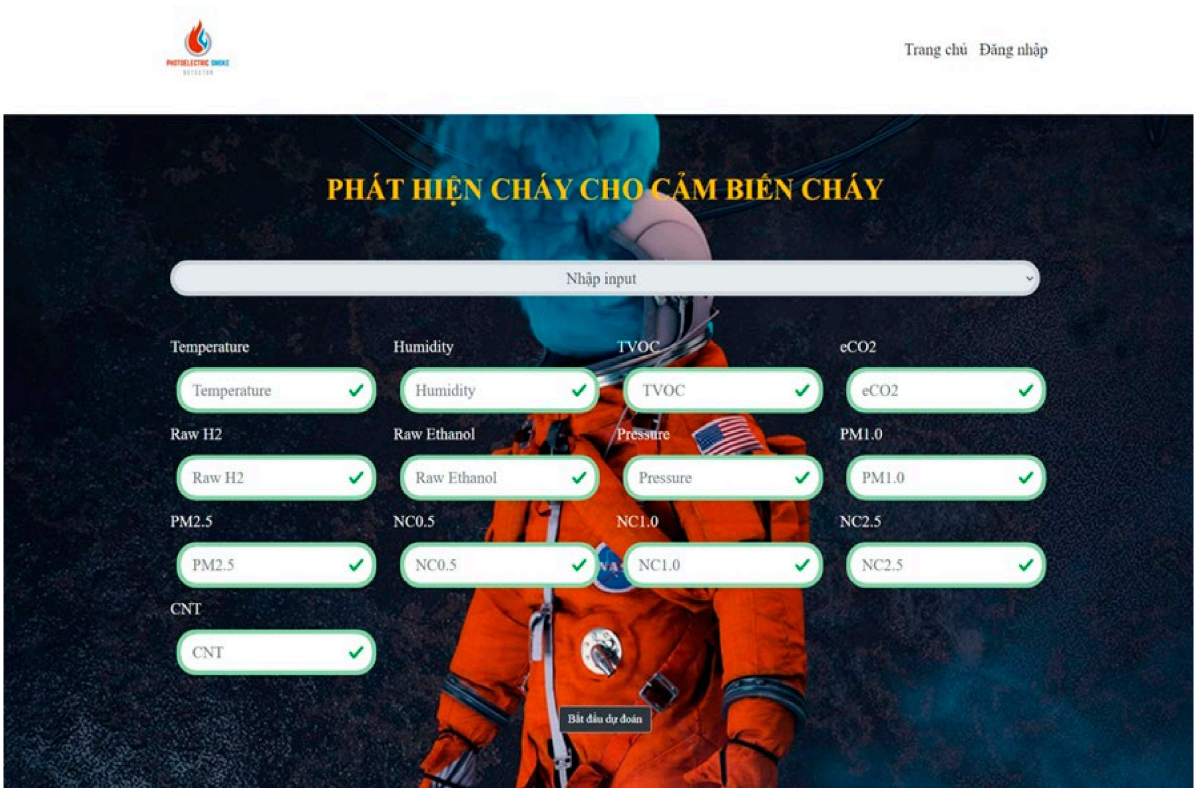


Figure 6. User Interface

4. CONCLUSION AND SUGGESTIONS

Upon concluding the research and report compilation, it can be evaluated that the achieved outcomes are reasonably comprehensive. Concerning the system, the classical DecisionClassifier algorithm has been successfully integrated into the model training and prediction process, addressing the challenge of dynamic data processing—a feat that other algorithms have not yet accomplished. Furthermore, a user-friendly graphical interface has been constructed alongside a command-line system. In terms of the report, crucial components such as data terminology, employed algorithms for training, and detailed descriptions of each chart have been meticulously covered within the content. It's important to note that this project was confined to the research phase, hence it has been developed and deployed to the extent of research. As a result, numerous avenues for expansion remain in the future. Future enhancements could encompass automated raw data processing within the system, optimizing the model training process, improving the system's interface for smoother functionality, and real-world deployment. The latter step aims to facilitate swift and precise smoke detection and fire alerting for users. Additionally, the project is also looking to expand its utility through diversification on mobile platforms, enabling users to access real-time information for recognition on-the-go via mobile devices.

REFERENCES

[1] Văn Ngân (2021). *Việt Nam có hơn hơn 1.200 vụ cháy từ đầu năm đến nay.* Ám ảnh những vụ cháy kinh hoàng trong năm 2021 (vov.vn).

[2] World Health Organization. (2018). *Burns.* https://www.who.int/news-room/fact-sheets/detail/burns.

[3] Barmpoutis, P., Dimitropoulos, K., & Grammalidis, N. (2014). Smoke detection using spatio-temporal analysis, motion modeling and dynamic texture recognition. *2014 22nd European Signal Processing Conference (EUSIPCO)*, INSPEC Accession Number: 14775489 Publisher: IEEE Conference Location: Lisbon, Portugal. https:// ieeexplore.ieee.org/abstract/document/6952375.

[4] Fire Safety Search. (2020). *Very early warning fire detection-fire safety search.* https://www.firesafetysearch.com/very-early-warning-fire-detection/ (2020). Accessed on April 26, 2020.

[5] Choi, M-Y., and Jun. S. (2020). Fire Risk Assessment Models Using Statistical Machine Learning and Optimized Risk Indexing, *Appl.Sci, 10*(12), 4199. https://doi.org/10.3390/app10124199

[6] Bá Tuấn - Văn Ngọc (2021). *Thực trạng và giải pháp hạn chế nguy cơ cháy nổ trong các nhà, chung cư cao tầng. Cảnh sát phòng cháy, chữa cháy và cứu nạn, cứu hộ.* http://canhsatpccc.gov.vn/ArticlesDetail/tabid/193/cateid/1136/id/9668/language/vi-VN/Default.aspx

[7] Khoa học và Phát triển. (2021). *Giám sát và cảnh báo cháy tự động bằng công nghệ IoT.* https://congnghiepcongnghecao.com.vn/.

[8] DataCluster Labs (2023). *Fire and Smoke Dataset.* Bộ dữ liệu phát hiện sớm cháy và khói, Camera thông minh, Hệ thống báo cháy.

https://www.kaggle.com/datasets/dataclusterlabs/fire-and-smoke-dataset.

[9] Hemateja (2021). *Wild Fire-Smoke-Dataset-Tensorflow.* https://www.kaggle.com/datasets/ahemateja19bec1025/wildfiresmokedataset

[10] Gulati, A.P. (2021). *Forest Fire Prediction Using Machine Learning.* https://www.analyticsvidhya.com/blog/2021/10/forest-fire-prediction-using-machine-learning/.