Journal of Engineering, Technology & Applied Science
p-ISSN: 2721-7949, e-ISSN: 2721-8090

Original Research Paper

Covid Classification System for Covid Detection

Muhammad Akmal Husaini Haris[1], Ahmad Fadli Saad[1]

[1] *Department of Computer Science, Faculty of Computer and Mathematical Science, University Teknologi Mara. Perak, Malaysia.*

Article History
Received:
28.12.2023

Revised:
20.01.2024

Accepted:
04.02.2024

***Corresponding Author:**
Ahmad Fadli Saad
Email:
afadlis@uitm.edu.my

Abstract: The ongoing global pandemic, which has now become an endemic has had a significant impact on the educational sector. Despite advancements in technology, there is no real- time prevention of COVID-19 transmission especially UiTM Tapah students that must go through crowd to reach health unit and have high possibility of spreading the disease. This research aims to develop a mobile application by utilizing machine learning for UiTM Tapah students to classify their COVID-19 status. To ensure a systematic and efficient development process, the research adopted the Mobile Application Development Life Cycle (MADLC) methodology. Within the application, the core of the machine learning functionality lies in the implementation of an Artificial Neural Network (ANN). By using 5,434 samples of data that had previously been classified by previous studies that analyzed student data such as symptoms to identify potential cases of the virus. The ANN approach performed greatly with the accuracy of 98%. Feedback from 32 respondents helped identify students' difficulties during the pandemic, which results in majority of the respondents agreeing that MCO affecting them adapting to online learning, and access to healthcare facilities. Furthermore, the usability testing conducted using the System Usability Scale (SUS) provided valuable insights into the system's user-friendliness and effectiveness. The results indicated a high level of usability, with a SUS score of 88.3 from 30 UiTM Tapah students while 76.25 from UiTM Tapah lecturers. This system was usable to classify the risk of infection among students who had not yet been diagnosed, while having some room for improvements. The system's ability to identify infected individuals promptly and accurately aided in controlling the spread of the virus, thereby protecting both students and staff. Thus, the classification system by using ANN was a valuable tool for public health organizations in their efforts to combat the COVID-19 pandemic.

Keywords: Artificial Neural Network, Classification, Detection.

Journal of Engineering, Technology & Applied Science, vol. 6, no. 1, pp. 9-18, April 2024. **DOI: 10.36079/lamintang.jetas-0601.610**

1. introduction

Technology has made our life even more convenient when compared to traditional ways of handling things in our daily life and has given profound impact in on cultivating our lives in emerging world. These can be seen when the advancement of technology in multiple areas such as web-based sites has made people keen with current technology and are able to get all information within their reach in fast manners. Some of the conveniences of technology include improving health as the spread of COVID-19 can be prevented early as the technology is advanced. Social media and other means of disseminating health messages revealed that it contributed to the betterment of psych behavioral responses to COVID-19 (Lin et al., 2020) and thus, other people can know anyone that have been infected with COVID-19 and could use prevention measure to avoid contact.

However, despite the many advancements in technology, there is no real time preventive measure from people that have been infected with COVID-19 and other people. Each student must physically go to the medical center and go through a crowd before reaching 'Unit Kesihatan'. COVID-19 is spread from one person to another person through close contact with an approximately three feet or around one meter when a person released droplets into the air by talking, coughing, breathing, or sneezing while in some cases spread in enclosed spaces with poor ventilation system (Liu et al., 2020). As a person walking outside, fresh air that constantly moving keeps on spreading this virus further. Moreover, the current traditional means of detecting COVID-19 such as self-reported symptoms and manual data collection can be time consuming and prone to create errors.

As a result, these UiTM Tapah students have a higher risk of infecting other people around them as they go to 'Unit Kesihatan' due to their proximity and interactions with one another. Meanwhile, this traditional method is time-consuming. This self- reported method requires students to fill out forms, answer questions and have their temperature checked. It takes a significant amount of time, leading to a long line and delays especially during the surge of COVID-19 cases. Another consequence is that it leads to potential errors in the screening process. Manual data collection also leaves room for human error, such as mistyping or misreporting of information. These errors can lead to false positives or negatives, resulting in students who are infected but not identified and students who are not infected but are incorrectly identified as such. This inefficient and inaccurate traditional method could also risk the disease spreading to other classmates and further to other classes. This leads to detrimental impact on the smooth functioning of the class and academic progress of the students. Therefore, the classification system for COVID-19 status is crucial in preventing the transmission of the disease. The system allows for easy and efficient screening of students, reducing the risk of the virus spreading among the students.

2. Literature Review

2.1. COVID-19

On December 31, 2019, pneumonia of unknown origin detected in Wuhan City, Hubei Province of China in which World health Organization (WHO) China Country Office was informed. Since then, the total number of pneumonia cases in China reported rapidly increasing and Chinese authorities identified this new type of coronavirus on January 7, 2020. As the authorities trying to figure out the source of this new disease and finding cure, Thailand reported its first coronavirus case making WHO declared an outbreak of COVID-19 a public health emergency of international concern at the end of January 2020. As the virus spread across the globe, Malaysia is also unable to avoid this tragic.

The COVID-19 pandemic is Malaysia's worst transmitted disease outbreak since the 1918 Spanish Flu, which killed 34,644 persons, or 1% of the then-British Malaya's population. The Nipah virus epidemic in 1999 killed 105 Malaysians, but the SARS outbreak in 2003 killed only two. The ongoing COVID-19 pandemic has claimed over 36,753 lives so far. In vulnerable group among the age of 55 to 64 years, the incidence of COVID-19 cases was the highest. In terms of economic, Malaysia lost approximately 2.4 billion a day during Movement Control Order (MCO) period and accumulated up to RM63 billion up to the end of April 2021 (Hashim et al.,2021).

Journal of Engineering, Technology & Applied Science, vol. 6, no. 1, pp. 9-18, April 2024. **DOI: 10.36079/lamintang.jetas-0601.610**

2.2. Data Analytics

Data analytics may be defined as the analysis of comprehensive, dynamic, low- cost, huge, and diverse data sets to provide advanced solutions (Khanra et al., 2020). Data analytics has become increasingly popular as science and technology have progressed, particularly to find trends and draw conclusions about the information they contain. Formerly, data were mostly used to test hypotheses, but this has changed with the introduction of neural networks and data mining. Datasets are employed in the new regime to develop hypotheses and find relationships that logical approaches cannot detect. Most big data researchs use open datasets, which are released in the public domain. Big data has enabled healthcare organizations to have access to massive datasets that link systems utilizing evidence-based information to demonstrate how pharmaceuticals interact and uncover underlying safety risks and adverse effects. The beauty of such systems is that they can create enormous amounts of relevant data in a short period of time (Batko & Ślęzak, 2022), which is important in health economics and pharmacoepidemiology. The systems also enable the measurement, analysis, and storage of complicated structured and unstructured information, which was previously impossible with archaic information systems that could not interact with one another.

Biomonitoring is one of the areas where big data has provided major improvements in health economics. As individuals continue to utilize passive bio- monitoring devices like smart watches and heart rate monitors, such real-time data may be pooled to provide health indicators that can be used to make clinical and policy choices. Big data has also given health economists the ability to create real-world evidence and advocate customized care. For example, pharmaceutical corporations use various big data approaches to trawl clinical data systems to analyze the efficacy of their drugs before releasing them to the market (McPadden Jacoband Durant, 2019).

2.2 Predictive Data Analytics

While healthcare data firms build more complex analytics solutions, customized healthcare organizations are shifting away from traditional analytics and toward predictive health insights to better comprehend present difficulties and outcomes. Rather than just presenting an end user with knowledge from prior events, healthcare predictive analytics estimate the likelihood of a conclusion based on important findings in historical data - a huge step forward in performance for many customized health firms (Van Calster et al., 2019). This enables medics, financial analysts, and administrative workers to be "ahead of the curve" on potential conditions and make initiative-taking decisions about how to proceed. The importance of predictive modelling in healthcare may be seen in emergency care, surgery, and intensive care, where a patient's outcome is directly tied to the care provider's rapid reaction and acute decision making when or if the situation takes an unanticipated turn for the worse (Waghmare & Hemalatha, 2017).

However, not all predictive analytics in healthcare necessitate the deployment of an experienced team. Predictive analytics is the practice of studying data to find patterns and trends that can predict future occurrences. Predictive analytics may be used in healthcare to estimate the likelihood of specific health issues or the chance that a patient will respond to a specific therapy (Leung et al., 2020).

Table 1. Types of Predictive Data Analytics

Types of Predictive Data Analytics	Description
Neural Network	A neural network is a set of algorithms that can recognizepatterns in data and make predictions based on those patterns.
Regression	studies how a dependent variable (the target) and independent factors are related (predictor). Forecasting, time series modelling, and determining the causal relationship between the variables are all done using this method.
Decision Tree	Based on how one set of questions was answered, supervised machine learning is used to classify orforecast.

Predictive analytics in healthcare can help doctors make better judgments about which therapies to provide patients and how to personalize those treatments to their specific requirements. Predictive

Journal of Engineering, Technology & Applied Science, vol. 6, no. 1, pp. 9-18, April 2024. **DOI: 10.36079/lamintang.jetas-0601.610**

healthcare analytics can also assist identify people at risk of complications or relapse and intervene before issues arise (Brennan et al., 2019). Overall, predictive analytics has the potential to improve healthcare delivery quality and efficiency. There are elements of predictive data analytics that can be used which are neural network, regression, and decision tree. Table 1 shows the types of predictive data analytics.

2.3. Convolutional Neural Network

A convolutional neural network (CNN or ConvNet) is a deep learning network design that learns directly from data. CNNs are extremely helpful for detecting patterns in pictures that may be used to distinguish items, classifications, and categories. They can also be used for categorizing audio, time series, and signal data. When compared to other classification methods, CNNs requires less pre-processing. While filters in primitive approaches are hand-engineered, CNNs can learn these filters/characteristics with adequate training (Li et al., 2021). CNN's design is like the connecting pattern of neurons in the human brain and was inspired by the arrangement of the Visual Cortex. Individual neurons only respond to stimuli in a small area of the visual field known as the Receptive Field.

Based on their capacity to learn and extract characteristics from input, Convolutional Neural Networks (CNNs) have become a great choice for image and video processing applications. This is very efficient and has been shown to perform well in image and video classification tasks. However, CNNs, like every other algorithm, have their own set of benefits and drawbacks. The effectiveness of CNNs is one of its primary advantages. They perform well in image and video processing applications, making them a popular choice for tasks like object detection and facial recognition. Another benefit is that they are translation invariant. CNNs can detect patterns and features in images that have been translated, rotated, or scaled. This makes them suitable for jobs like image search, where a picture may appear in a variety of orientations.

CNNs have certain limits, despite their great advantages. One of the most significant downsides is their intricacy. CNNs may be difficult to build and train, necessitating a substantial amount of computer power. Furthermore, they are highly specialized for image and video processing jobs and may struggle with other forms of data such as text or audio. The potential of overfitting is another shortcoming of CNNs. Overfitting happens when a model becomes overly complicated and memorizes the training data rather than generalizing to new data. This might result in poor performance on unknown data. Furthermore, CNNs are considered black-box models, which means that it is difficult to comprehend how the model makes its predictions, which might limit interpretability and explainability.

2.4. Artificial Neural Network

Artificial neural networks (ANN) are implemented as a network of linked processing components known as nodes that function in a comparable way as biological neurons. Artificial neural networks are non-parametric classifiers. Artificial neural networks are constructed similarly to the human nervous system (Abiodun et al., 2019). This type of network's fundamental unit is the neuron, which is a unified computing rudiment that acts as its fundamental unit. A neuron's growth is divided into two stages: training and use. During the training phase, the neuron learns to do an activity and then utilizes the training data to estimate the response during the testing phase. In general, these neural networks are used to identify trends or patterns within a set of data. ANN applications for real-world issues vary and are classified into three categories: pattern classification, prediction, and control and optimization. For example, if the network is charged with detecting faces, the first hidden layer will act as a line detector, with the second hidden layer taking these lines as input and combining them to form a nose. The third hidden layer may couple the nose with an eye, and so on until the entire face is produced.

The capacity of ANNs to learn from data is one of its primary advantages. ANNs may learn to identify patterns and make predictions without the need for explicit programming when trained on a big dataset. This makes them suitable for jobs like picture identification and natural language processing, where writing clear rules to explain the data might be challenging. The capacity of ANNs to handle vast and complicated datasets is another advantage. ANNs may learn to identify patterns and characteristics that are not immediately obvious to the human eye and can handle vast volumes of data. This makes them suitable for applications like image recognition, which require a big dataset to train the model.

Journal of Engineering, Technology & Applied Science, vol. 6, no. 1, pp. 9-18, April 2024. **DOI: 10.36079/lamintang.jetas-0601.610**

Despite their numerous benefits, ANNs have several limits. One of the most significant downsides is the possibility of overfitting. Overfitting happens when a model becomes overly complicated and memorizes the training data rather than generalizing to new data. This might result in poor performance on unknown data. Furthermore, ANNs may be computationally costly to train and execute, necessitating a large amount of computer resources. Another disadvantage of ANNs is their inability to be interpreted. ANNs are classified as black-box models, which means that it is difficult to comprehend how the model makes its predictions. When it comes to interpretability and explainability, this might be a problem, making it difficult to comprehend the logic behind the model's predictions.

2.5. Multilayer Perceptron

A multi-layer perceptron (MLP) is a feed forward neural network augmentation. As illustrated in Figure 2.5, it is made up of three layers: the input layer, the output layer, and the hidden layer. The input layer receives the data to be processed, the output layer performs tasks such as prediction and categorization, and the hidden layers are the core computational component. The data flows in a forward direction, like a feed forward network, and the MLP is trained using the back propagation technique. MLPs are commonly used for tasks such as pattern classification, recognition, prediction, and approximation, and are particularly useful for solving problems that cannot be solved linearly.

Some of the Multilayer Perceptron's advantages are:

- Can be applied to complex nonlinear problems.
- Works well with large input data.
- Provides quick predictions after training.
- The same accuracy ratio can be achieved even with smaller data.

Some of the Multilayer Perceptron's disadvantages are:

- It is not known to what extent each independent variable is affected by the dependent variable. Computations are difficult and time consuming.
- The proper functioning of the model depends on the quality of the training.

2.6. Recurrent Neural Network

Recurrent Neural Network (RNN) is a sort of Neural Network in which the previous step's output is given as input to the current step. In typical neural networks, all inputs and outputs are independent of one another; however, when predicting the next word in a sentence, the prior words are necessary, and so the previous words must be remembered. As a result, RNN emerged, which utilized a Hidden Layer to resolve this issue. The Hidden state is the core and most essential aspect of RNN which remembers certain information about a sequence. RNNs have a "memory" that recalls everything that has been computed (Nassif et al., 2019). It employs the same settings for each input since it performs the same work on all inputs or hidden layers to generate the result. Unlike other neural networks, this decreases the complexity of parameters.

RNNs' capacity to handle sequential data is one of its key features. RNNs are built to handle data with a temporal component, such as time series or text. They can learn to detect patterns and forecast outcomes based on the temporal structure of the data, making them excellent for tasks like speech recognition and natural language processing. RNNs also have the capacity to recall information from prior inputs. RNNs contain a hidden state memory component that allows them to remember information from prior inputs. As a result, they are suitable for jobs like language translation, where knowing the context of the statement is critical.

Despite their many benefits, RNNs have certain drawbacks. One of the major drawbacks is the possibility of disappearing slopes. Backpropagation is used to train RNNs, which needs the gradient to be transmitted back over time. However, when the gradient is transmitted back in time, it might become very tiny, making the model difficult to learn. The computational cost of RNNs is another restriction. RNNs can be computationally costly to train and run since they require many computational resources. Furthermore, RNNs may be challenging to tune and train, particularly when the dataset is vast, or the model is sophisticated. Some of the application of Recurrent Neural Network:

- Language Modelling and Generating Text

Journal of Engineering, Technology & Applied Science, vol. 6, no. 1, pp. 9-18, April 2024. **DOI: 10.36079/lamintang.jetas-0601.610**

- Speech Recognition
- Machine Translation
- Image Recognition, Face detection
- Time series Forecasting

3. Methodology

The methodology used in this research is Mobile Application Development Life Cycle (MADLC). The framework that this research works on is based on the method and all the important steps to follow. All the phases in the MADLC explained. The system architecture for this research shows how the system works and how the system affects the people involved. Lastly, the hardware and software requirements for this research started to make sure that this system ran smoothly without having any errors.

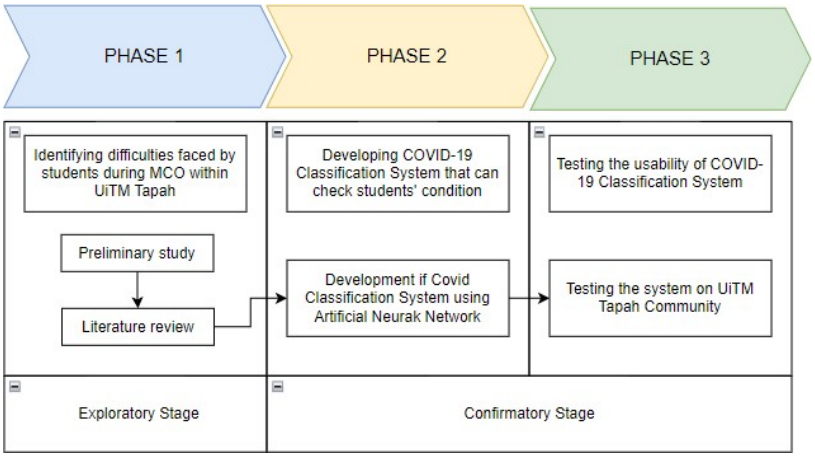


Figure 1. Design and Methodology

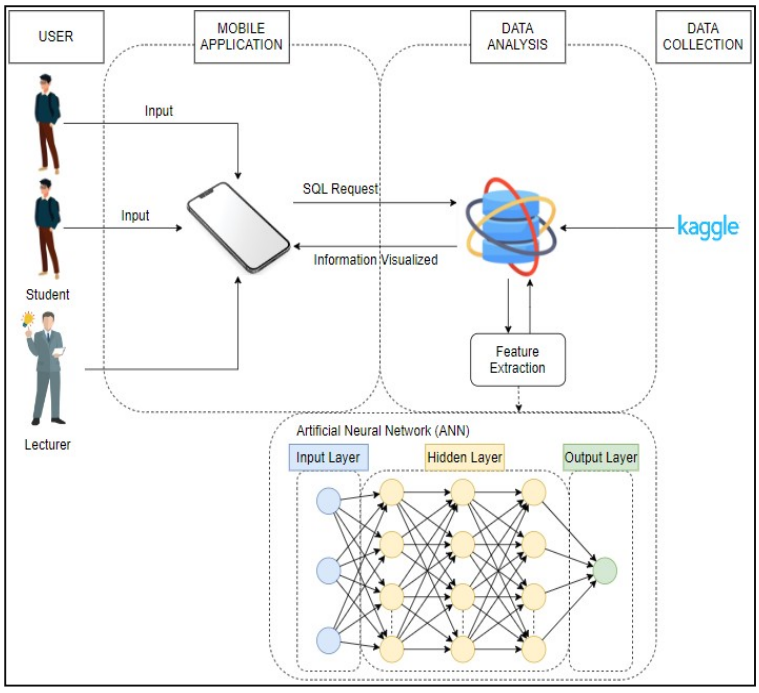


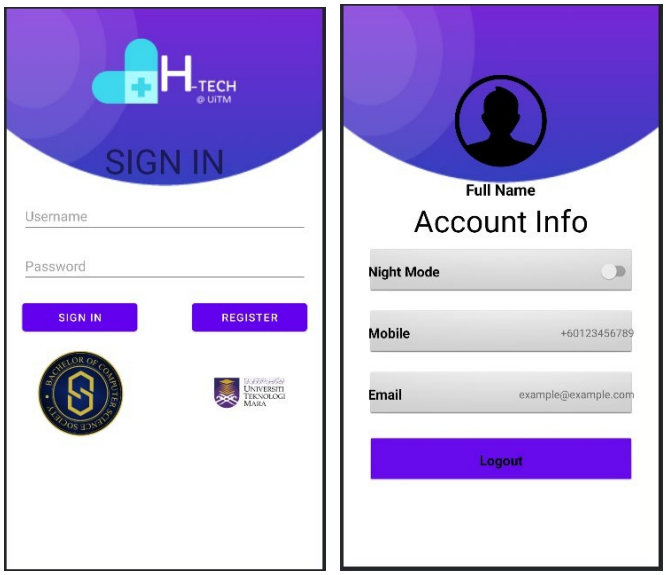
Figure 2. System Architecture

Journal of Engineering, Technology & Applied Science, vol. 6, no. 1, pp. 9-18, April 2024. **DOI: 10.36079/lamintang.jetas-0601.610**

4. Finding and Discussion

4.1. Interface

The login page is an essential component of any system that requires user authentication. It serves as the gateway for users to access the system by providing their credentials. The interface design for the login page prioritizes usability, security, and a seamless user experience. The login page also includes error handling with informative error messages to guide the user. Clear and concise error messages such as invalid credential provided. Login page included a password field where users can securely input their passwords. This is done by employing password masking techniques which display dots instead of characters. A "Sign In" button allows users to submit their login credentials and initiate the authentication process. Figure 3 shows the login page of the system.



(a) (b)

Figure 3. Interface: (a) Login Page of the System, (b) Profile Page

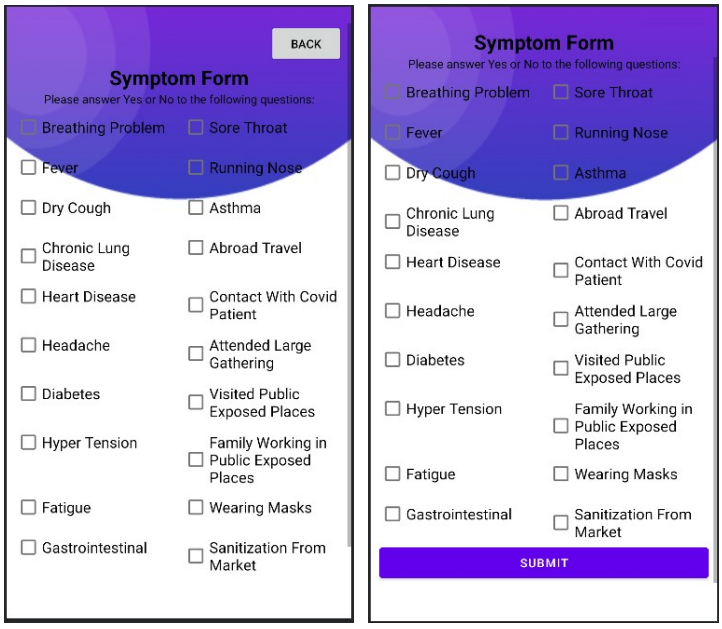


Figure 4. Symptom Fill Form Page