Probabilistic Risk Assessment of COVID-19 Patients at COVID-19 Assessment Centre

Choo-Yee Ting¹, Helmi Zakariah², Yasmin Zulaikha Mohd Yusri¹

¹Faculty of Computing and Informatics, Multimedia University, 63000 Cyberjaya, Selangor, Malaysia
²AIME Healthcare Sdn Bhd, Mid Valley City, 59200 Kuala Lumpur, Malaysia

Abstract. COVID-19 started impacting Malaysia in early 2020, and the cases have reached 4.4 million as of April 27, 2022, with 35507 deaths. Since then, federal and state governments have set up COVID-19 Assessment Centres (CACs) to monitor, manage and assess the risk of COVID-19-positive patients. However, a large number of patients within a day has caused the CACs to experience a shortage in medical officers and subsequently resort to overwhelming administrative work. A misassignment of a patient to either home quarantine or COVID-19 Quarantine and Treatment Center or immediate hospital admission (PKRC) could potentially increase the Brought-In-Dead (BID) cases. Therefore, this study aimed to overcome the challenges by achieving the following two main objectives: (i) to identify the optimal feature sets for adult and child patients when they require hospital admission, (ii) to construct predictive models that perform preliminary assessment of a patient, which a medical officer later confirms. In this study, the predictive models developed were Naive Bayes, Random Forest, K-Nearest Neighbors, Logistic Regression and Decision Tree. The datasets were obtained from one of the CACs in Malaysia and were imbalanced in nature. The empirical findings showed that Logistic Regression outperformed the rest with a slight difference. The findings suggested that while there are shared symptoms among adult and child patients, such as runny nose and cough, the child patients exhibited extra symptoms such as vomiting, lung disease, and persistent fever.

Keywords: Covid-19 assessment centre; Home quarantine; Hospitalization; Machine learning

1. Introduction

A pneumonia pandemic was discovered in Wuhan, China, by the Chinese health authorities in late December 2019 (Ciotti et al., 2019; Berawi et al., 2020) and was later recognized as “COVID-19” by the World Health Organisation (WHO). The impact of COVID-19 pandemic has influenced various economic sectors of developing countries (Al-Doori et al., 2021). In Malaysia, the earliest incidence of COVID-19 was detected on January 25, 2020. The virus was transported by three Chinese nationals that had close contact with the infected individual in Singapore (Elengoe, 2020). Since the major outbreak via a religious event named Tabligh at Sri Petaling, Malaysia, the Ministry of Health (MOH) developed the standard operating procedure for the treatment of COVID-19 (Othman & Babulal, 2020). Thirty-four government hospitals that had established screening centers for COVID-19 patients in every state in Malaysia (Ravindran, 2020).

¹Corresponding author’s email: cyting@mmu.edu.my, Tel.: +603-83125217, Fax.: +603-83125264
doi: 10.14716/ijtech.v13i6.5882
Malaysia has gone through several waves of COVID-19. The first COVID-19 wave happened from January 25 to February 15, 2020, with 22 cases, while the second wave lasted from February 27 to June 30, 2020 (Rampal & Liew, 2021). The third wave occurred on September 20, 2020 (Rampal & Liew, 2021). Since then, the demand for more beds in critical care units has increased, not only for COVID-19 patients, but also for a number of other severe non-coronavirus cases (Zoey, 2021). In March 23, 2021, Malaysia declared that the third wave of COVID-19 infections is progressively subsiding since the increase in vaccine distribution (Anand, 2021). Then, the fourth wave arrived as few as four weeks after the third, in May 2021 (Singh, 2021). The Ministry of Health has mandated that individuals who test positive for COVID-19 and have mild or no symptoms must undergo home self-quarantine for ten days (MOH, 2021). The home quarantine applies to COVID-19 Category 1 and Category 2 patients, i.e., patients who have tested positive but are experiencing minimal or minor symptoms, as existing health facilities are unable to cope with the increase in daily cases (The Straits Times, 2021). With the increase in COVID-19 patients over time, the government established the COVID-19 assessment center (CAC) as the one-stop er that serves as a referral center for the identification, monitoring, managing, and assessment COVID-19 positive patients. It was established to assess these infected people undergoing home monitoring treatment (MOH, 2021). The treatment plan for COVID-19 patients in Malaysia is divided into five clinical categories of patients (MOH, 2020), namely (i) Category 1. asymptomatic, (ii) Category 2. symptomatic and no pneumonia, (iii) Category 3. symptomatic and pneumonia, (iv) Category 4. symptomatic, pneumonia and supplemental oxygen required, (v) Category 5. critically ill with multiple organ failure.

In general, COVID-19 patients in category do not require special treatment. Patients in Category 2 and 3 will receive symptomatic treatment, such as fever, cough, and flu medicines. Patients will also be looked after to ensure that their nutritional status is optimal and that their blood circulation is maintained. Antivirals, immunomodulators (to suppress inflammatory response), and anticoagulants (to prevent clotting) are reserved for COVID-19 patients that are in more serious conditions, such as those in Category 4 and 5 (MOH, 2020). All COVID-19-positive patients in Malaysia will be treated at authorized treatment centers, including those asymptomatic, either in hospitals or PKRC. This means that all COVID-19 patients will be separated from the rest of the community. This approach is effective in preventing the COVID-19 virus from spreading in society. The clinical progress of COVID-19 patients was monitored regularly at treatment centers for vital signs and blood tests, as needed. These aid in early detection and treatment to prevent patients from deteriorating, particularly individuals from Category 3, 4, and 5 (MOH, 2020). A good estimation of patients in each category could help in subsequent medical intervention, including the design of modular isolation units as reported by (Yatmo et al., 2021).

This study aimed to identify whether COVID-19 patients who reported themselves at CACs required hospital admission. That is, to determine the clinical symptoms (features) that were strongly related to the needs of hospital admission for both the adult and child patients. The obtained features were then used to construct machine learning models that can analyse a patient’s condition and subsequently provide tailored recommendations.

2. Related Work

2.1. Feature Selection

Machine learning research has been utilizing feature selection techniques to reduce the data dimension so that quality of the developed model can be sustained. Feature selection has been applied in machine learning work related to COVID-19. Examples of work are
reported in (Chassagnon et al., 2021) and (Alotaibi et al., 2021) where features are ranked based on their importance. Another work by (Saeed et al., 2021) proposed a multilevel feature selection consisting of Mutual Information (MI) and Relief to analyze the features affecting teachers’ conditions during the COVID-19. Researcher has used one of the common feature selection algorithm named “BORUTA” A study by (Heldt et al., 2021) used machine learning to predict the clinical outcome measures using data from 879 patients who suffer from COVID-19 from two London urban tertiary hospitals. The COVID-19 risk assessment proved useful, with models that achieved high predictive accuracy. Age of the patient, oxygenation status, and certain laboratory tests were significant predictors of outcome. Three machine learning algorithms were used to predict patient outcomes: XGBoost, Random Forest, and Logistic Regression. A method for calculating the likelihood of an outcome measured as a linear function is known as logistic regression (Heldt et al., 2021). The trained algorithms were used to extract the most useful clinical features for detecting these patient trajectories. Based on the inclusion criteria, they identified that 15% of patients needed critical care, 7% needed artificial respiration, and 31%died in the hospital. After successfully learning based on preliminary clinical data and accurately predicting clinical outcomes, the most effective model achieved an AUC of 0.76 to 0.87 and an F1 score of 0.42 to 0.60.

Xu et al. (2021) examined the scientific features of COVID-19 ARDS patients and used artificial intelligence approaches to identify the presence of ARDS in COVID-19 individuals. The study reported that ARDS is the leading reason for death in COVID-19 individuals with chronic illness affecting 67 percent of them. Data from 659 patients with COVID-19 from 11 different China’s regions were used for this research. The scientific aspects of COVID-19 patients with ARDS and those who do not have ARDS were thoroughly investigated, then prediction models were built using machine learning as well as deep learning technologies. Referring to the results of the univariate analysis, 19 factors (e.g., admission severity score, BMI, age, temperature, gender, cough, and many more.) were identified as predictors to the model to determine whether COVID-19 patients would cause ARDS (Xu et al., 2021). These factors were predicted using five machine learning. The decision tree outperformed the others with a hit rate of 97%. It was integrated into the electronic medical record system for infectious diseases to help physicians identify COVID-19 at-risk patients.

2.2. Machine Learning for COVID-19

In a study by (Kim et al., 2020), different machine learning algorithms were developed. The models were Decision Trees, Artificial Neural Networks, XGBoost, and Random Forest. They are developed to examine the prognosis of a COVID-19 patient. The goal was to develop a prognostic model based on information that patients can provide, such as demographic data, concomitant diseases, and subjective symptoms. XGBoost performed better than the other 55 machine learning models in relation to discrimination. In addition, Cheng et al. (2020) built a risk prioritization tool based on machine learning that predicts ICU transfers within the next 24 hours. To train a Random Forest model data from a time series were used as variables for input. The result is a predictive model to recognize individuals who may need to be admitted to the intensive care unit within the next 24 hours. This method has the chance to enhance hospital resource management and patient throughput planning, leading to better COVID-19 patient care. To predict the COVID-19 individuals’ mortality, An et al. (2020) used LASSO, Linear SVM, k-Nearest Neighbours, and Random Forest. Age, pre-existing DM, or cancer were the most important factors for LASSO; Age, method of infection, and underlying hypertension were the most important predictors for Random Forest. A study by Izquierdo et al. (2020) has integrated both the traditional epidemiological methods and natural language processing to summarise the COVID-19 patient clinical
characteristics and find the parameters predictive of ICU admission. They created a decision chart to sort COVID-19 patients into categories depending on their likelihood of being admitted to the ICU. As a result, two classes are evaluated in the model; whether the patient has to be admitted to the ICU or does not have to be admitted to the ICU. The tree consists of child nodes depending on the variables of the patient. Each branch of the tree is intended to provide as much distinction as possible between the two classes. A node with entropy equal to one represents the lowest possible classification, while a node with zero entropy represents a large categorization (Izquierdo et al., 2020).

In another study by Assaf et al. (2020), a machine learning-based predicting method to classify the high-risk COVID-19 patients was discussed. The primary outcome was a disease defined as a multi-organ failure, ICU admission, mechanical ventilation, and death. Machine learning was constructed to forecast patient health failure and compare it to existing recommended predictors and the APACHE II risk prediction score. One hundred sixty-two patients with non-severe COVID-19 were hospitalized, and 25 (15.4%) of them worsened to critical COVID-19. The machine learning models exceed other indicators in estimating critical COVID-19 outcomes, including the APACHE II score. The model’s most influential variables were the APACHE II score, the count of white blood cells, the count of blood lymphocytes, and the time from symptom onset to saturation of oxygen and admission. Compared to the most effective available solutions, machine learning algorithms showed excellent efficacy in predicting important COVID-19. As a result, artificial intelligence could be used to accurately detect the risk, enabling better patient triage and hospital assignment, prioritization of medical resources, and overall COVID-19 pandemic management (Assaf et al., 2020; Casiraghi et al., 2020).

3. Method

3.1. Dataset Description

In this study, two COVID-19 patient datasets from 22/2/2021 to 21/8/2021 were collected from one of the COVID-19 Assessment Centres in Malaysia (Table 1). Both datasets recorded information on whether a patient required hospital admission. As shown in Table 2, both datasets are considered imbalanced. Most of the patients did not require hospital admission. The first dataset, D_ad, consists of basic profiles and clinical symptoms for adult patients. It has 36 columns with 95666 records; 91834 did not require hospital admission, while 2832 were admitted. Examples of clinical variables for this dataset are adult active, adult runny nose, adult cough, adult difficulty breathing, adult fever, and adult loss of smell.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Category</th>
<th>Records</th>
<th>Hospital Admission (before SMOTE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_ad</td>
<td>Adult</td>
<td>94666</td>
<td>91834 2832</td>
</tr>
<tr>
<td>D_ch</td>
<td>Children</td>
<td>27575</td>
<td>26707 868</td>
</tr>
</tbody>
</table>

Similar to the first dataset, the second dataset, D_ch, captured children’s basic profile and clinical symptoms (i.e., those with ages <18). The size of this dataset is relatively smaller; it has 26 columns with an approximation of only 30% of the records in D_ad. Examples of clinical variables are child active, child fever, runny child nose, child cough, child vomit, child seven days, child lethargic, and child loss appetite. While there are differences in the variables, there are common variables such as blood pressure, respiratory rate, height, weight, and BMR.
3.2. Feature Selection

One of the objectives of this study was to rank the feature importance of hospital admission. To achieve the objective, this study employed a feature selection algorithm named BORUTA to perform the feature ranking on both datasets. It was hypothesized that a portion of the features could be different, with some overlapping between the datasets. In this study, only features with scores greater than 50% will be identified for predictive model construction.

3.3. Model Construction

Five machine learning models were developed using the selected clinical variables as predictors and “require hospital admission” as the class variable. The datasets were split into 75% for the training phase while reserving 25% of the data for testing purpose. For prediction, the F1-Score and ROC curve was used as the main instruments to compare the performance of Naive Bayes (NB), Random Forest (RF), Logistic Regression (LR), K-Nearest Neighbors (K-NN), and Decision Tree (DT). This study's accuracy cannot be used as an important metric mainly because of the highly imbalanced dataset (Table 1). Due to such imbalance characteristics inherent within the two datasets, the AUC and F1-score were considered when identifying the suitable model. In this study, none of the datasets was SMOTE-d because the study aimed to investigate the performance of predictive models without any manipulation of the datasets.

4. Results and Discussion

This section begins with a discussion on features influencing the decision to admit a patient (i.e. adult and child patient) to a hospital for further treatment. Subsequently, predictive models were trained using the datasets to investigate their predictive powers under imbalanced characteristics. The findings are presented using AUC, F1-score, and Accuracy to compare the models.

4.1. Features Contributing to Hospital Admission

Table 2 compares top 10 features influencing hospital admission for adults and children. As shown in the above table, runny nose, cough, respiratory rate, pulse rate, and BMI are some crucial factors in recommending whether an adult or a child requires hospital admission. In contrast to adult patients, children do not have hypertension or loss of smell and taste. However, children with lung disease could end up in hospital admission. In addition, fever (no indication of low, medium, or high) and persistent fever can also be common indicators of whether the child requires hospital admission. As shown in the table, BMI is the 3rd highest contributor indicating that whether a child requires hospital admission can be inferred from their BMI. Therefore, maintaining a good BMI can be a determinant of whether hospitalization is required. The overall ranking of the features for child patients is shown in Figure 1, while Figure 2 shows the ranking of features for adult patients.

Table 3 compares five different models using three metrics, namely, AUC, F1Score, and Accuracy. Figure 3 shows the AUC comparison of the five models using a line chart. The dataset provided by the CAC is imbalanced; therefore, AUC and F1-Score should be the determinant to suggest the fitness of a model. Based on the findings, Logistic Regression outperformed the rest with AUC=0.65 and F1-Score=0.10, despite having the lowest accuracy (68%). In Table 4, the highest accuracy, as depicted by K-Nearest Neighbour, could be attributed to its ability to correctly classify cases that did not require hospital admission. However, K-NN suffered from the lowest F1 score, indicating that it should not be used to classify hospital admission. The chart shown in Fig. 4 compares the performance
of the five classifiers. It is noted that Logistic Regression (green line) outperformed other classifiers.

Table 2 Top-10 Features for Hospital Admission among Adult and Children

<table>
<thead>
<tr>
<th>Features (Adult)</th>
<th>Score</th>
<th>Features (Children)</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>runny nose</td>
<td>1.00</td>
<td>fever</td>
<td>1.00</td>
</tr>
<tr>
<td>cough</td>
<td>1.00</td>
<td>pulse rate</td>
<td>1.00</td>
</tr>
<tr>
<td>BMI</td>
<td>0.97</td>
<td>BMI</td>
<td>0.96</td>
</tr>
<tr>
<td>respiration rate</td>
<td>0.94</td>
<td>cough</td>
<td>0.91</td>
</tr>
<tr>
<td>pulse rate</td>
<td>0.91</td>
<td>respiration rate</td>
<td>0.83</td>
</tr>
<tr>
<td>loss of smell</td>
<td>0.88</td>
<td>active</td>
<td>0.78</td>
</tr>
<tr>
<td>hypertension</td>
<td>0.85</td>
<td>persistent fever</td>
<td>0.74</td>
</tr>
<tr>
<td>active</td>
<td>0.82</td>
<td>lung disease</td>
<td>0.70</td>
</tr>
<tr>
<td>loss of taste</td>
<td>0.79</td>
<td>vomit</td>
<td>0.65</td>
</tr>
<tr>
<td>diabetes mellitus</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) (b)

Figure 1 Feature Ranking for $D_{ch}$

Figure 2 Feature Ranking for $D_{ad}$

Figure 3 ROC Comparison using $D_{ad}$

Figure 4 ROC Comparison using $D_{ch}$
4.2. Model Comparison using $D_{ad}$ and $D_{ch}$

Table 4 shows the model performance comparison using $D_{ch}$. Figure 4 shows the AUC comparison of the five models using a line chart. The same patterns can be observed for $D_{ch}$ when the same evaluation metrics were used. Logistic regression performed highest in terms of AUC and F1-Score, while scoring the lowest in accuracy. K-Nearest Neighbour, Random Forest, and Decision Tree scored higher than 90% but suffered in AUC and F1-Score. This can explain that no obvious patterns can be observed for those who did not require hospital admission. Another reason could be due to a lack of data points. This can be seen in Table 1. The ROC again demonstrated that Logistic Regression performed better than the other models when treated with $D_{ch}$.

Table 3 Performance Comparison between Models for $D_{ad}$

<table>
<thead>
<tr>
<th>Model</th>
<th>AUC</th>
<th>F1</th>
<th>Acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes</td>
<td>0.64</td>
<td>0.07</td>
<td>0.95</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.53</td>
<td>0.08</td>
<td>0.80</td>
</tr>
<tr>
<td>K-Nearest Neighbour</td>
<td>0.51</td>
<td>0.00</td>
<td>0.97</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>0.65</td>
<td>0.10</td>
<td>0.68</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>0.54</td>
<td>0.08</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Table 4 Model Comparison using $D_{ch}$

<table>
<thead>
<tr>
<th>Model</th>
<th>AUC</th>
<th>F1</th>
<th>Acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes</td>
<td>0.63</td>
<td>0.06</td>
<td>0.80</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.58</td>
<td>0.05</td>
<td>0.94</td>
</tr>
<tr>
<td>K-Nearest Neighbour</td>
<td>0.54</td>
<td>0.02</td>
<td>0.97</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>0.65</td>
<td>0.10</td>
<td>0.76</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>0.51</td>
<td>0.06</td>
<td>0.90</td>
</tr>
</tbody>
</table>

5. Conclusions

This study aimed to investigate the (i) features contributing to hospital admission for both adult and child patients, (ii) the performance of different classifiers when datasets are imbalanced. Empirical findings via the feature selection algorithm suggested that runny nose, cough, BMI, respiration rate, and pulse rate are the top 5 features that contributed to hospital admission. In contrast, variables that commonly appeared among child patients who required hospital admission are fever, pulse rate, BMI, cough, and runny nose. Despite the slight differences, the empirical study observed common features were respiration and pulse rate. The study also found that lung disease and persistent fever can indicate whether hospital admission is needed. When the models were trained and evaluated using imbalanced datasets, all of them suffered from low F1-Score and AUC. The high accuracy was mainly contributed by the model that correctly detected patients that did not require hospital admission. While SMOT-ing the dataset can be a common practice to boost the performance of the models, we do not see that should be the direction forward because there is a tendency to alter the natural characteristics of the datasets. Instead, we propose to use cluster analysis such as DBScan to detect the possible clusters inherent inside the datasets. In addition, we also deploy deep learning models to extract the hidden patterns within the datasets. Future work can also consider investigating whether there is a significant difference in symptoms by different COVID-19 variants.
References


