A Covid-19 Positive Case Prediction and People Movement Restriction Classification

I Made Artha Agastya

Informatics
Universitas Amikom Yogyakarta
Yogyakarta, Indonesia

Abstract—The world experienced a pandemic that changed people's daily life due to Coronavirus Disease 2019 (covid-19). In Jakarta, the covid-19 cases were discovered on March 18, 2020, and the case increased uncontrollably until the government conducted a movement restriction called pembatasan sosial berskala besar (PSBB). The effectivity of movement restriction was not evaluated in detail. Therefore, we investigated the covid-19 cases in the PSBB period to understand the contribution of movement restriction. Moreover, a prediction model is proposed to computerize the decision of movement restriction. The models are divided into regression and classification models. The regression model is developed to forecast the number of infected cases. At the same time, the classification model is used to identify the best movement restriction type. We utilize data transformation named Principal Component Analysis (PCA) to reduce the number of features. In our case, the best regression method is Multiple Linear Regression (MLP). Then, the best classification method is the Support Vector Machine (SVM). The MLP results are 148.38, 37036.37, and 0.250336 for Mean Absolute Error (MAE), Mean Square Error (MSE), and R^2, respectively. In contrast, the SVM achieved an accuracy of 84.81%. Moreover, the prediction system on the website were successfully deployed.

Keywords—Covid-19; movement restriction; machine learning; positive case; infected prediction

I. INTRODUCTION

The Coronavirus Disease 2019 (COVID-19) was first detected in China at the end of 2019. The virus causes breathing symptoms such as fever, coughing, pneumonia, and diarrhea in patients [1]. The World Health Organization (WHO) received a report from the Chinese government on December 31, 2019, regarding this disease with reports on the case of pneumonia or wet lungs in Wuhan, in Hubei Province, China. A week later, the Chinese government confirmed that Covid-19 had been identified as the cause of pneumonia [2]. Then the media reported that many new cases were recorded in other countries because international travel and trade were operating as usual.

As the center of government and trade in Indonesia, Jakarta has the highest population density compared to other places [3]. Consequently, an increased number of positive confirmed Covid-19 in Jakarta. Based on data on March 18, 2020, there were 171 cases. In Jakarta, the number of positive cases of Covid 19 drastically expanded since there were ten times increasing positive cases on April 9, 2020, compared to March 18, 2020. Jakarta suffered 1776 cases at that time. While the average increase per day for the period March 18 until April 9, 2020, is 70 cases. However, at that time, the government had not taken any corrective actions. In contrast, Malaysia held a national scale MCO (Movement Control Order), which began on March 18, 2020 [4]. After three weeks without corrective action, then on April 10, 2020, the DKI Jakarta Regional Government imposed a pembatasan sosial berskala besar (PSBB) consisting of two categories, namely a strict PSBB and a transition PSBB. The Strict PSBB was valid from April 10 - June 4, 2020, and September 14 - October 11, 2020. And the Transition PSBB took place from June 5 - September 10, 2020, and October 12 - October 25, 2020.

Based on [5], [6], the positive confirmation trend per day when Strict PSBB is carried out tends to be stable when compared to No-PSBB. This shows the effect of restrictions on community movements. To reduce the economic burden, the government relaxed community activities by imposing Transition PSBB. However, there is a significant increase in the number of positive cases per day. Comparing the Transition PSBB and No-PSBB, the condition of the Transition PSBB is more worrying as the rate, and cumulative cases are much higher. A comprehensive data collection can be an excellent choice to gain more insight. Moreover, an artificial intelligence system can help the government to decide the type of movement restriction. Therefore, the research question can be determined as follow:

1) What factors or features affect the number of daily positive cases?

2) How to predict the number of daily positive cases?

3) How to determine the type of movement restriction in certain conditions?

This paper is organized into five sections which are the introduction, related works, methods, result and discussion, and conclusion and future work. The introduction section consists of a brief explanation of the problem and the purpose of investigation. The related work section contains the summary of other researcher works that solved a covid-related prediction problem. The methods section shows the data acquisition, preparation, exploration, transformation, and frameworks. After that, the result and discussion section report the experimental findings. The conclusion and future work section contain a final comment on the current work and the potential research in the future.

II. RELATED WORK

The infectious disease outbreaks such as food contamination, severe acute respirations syndrome (SARS), dengue, malaria, and influenza [7]–[9] have been analyzed using the artificial intelligence (AI) approach. The most common research is to forecast the number of infected individuals using the machine learning method. As the data is mostly time series, the model can be following a linear or nonlinear model [10]–[12]. The AI system can be used to prevent the outbreak become uncontrollable as the estimated number of diseased people is calculated. The government can prepare the medical facilities and their staff to match the number of patients. The anticipation of the worst-case scenario makes us more ready to confront the outbreak.

The machine learning-based model [13] has been deployed to predict the positive case of covid-19. The result is satisfying as the R^2 of 0.9998, 0.9996, and 0.9999 for Gompertz, Logistic, and Artificial Neural Networks models, respectively. Based on the results, trends can be forecasted and extended until the termination of the pandemic in Mexico. However, their models were only able to predict new COVID-19-positive cases. Velásquez and Lara [14] utilized a reduced-space Gaussian process regression model to forecast 82 days of positive, dead, and recover cases in the USA. Compared to the actual values, it is discovered that the model can generate the expected case value with a significant correlation coefficient. Buckingham-Jeffer [15] performs prediction using stochastic SIR (susceptible infectious removed) and SEIR (susceptible-exposed infectious removed) models. Both models can immediately forecast the number of positive cases using maximum likelihood inference in a time frame period.

A typical time-series experiment uses the Autoregressive Integrated Moving Average (ARIMA) model to produce a prediction model to estimate the amount of COVID-19 confirmed cases, fatalities, and recoveries [16]. If the current trend in Pakistan continues, the number of actual instances might triple by May. The investigation result concludes that the statistics are accurate and that the trends will continue to rise in the next month. Fortunately, the ARIMA model cannot comprehend the provided historical data pattern, so their findings are questionable.

A comparison of ARIMA with machine learning methods has been conducted by Kamarudin et al. [17] for positive, dead, and recovery cases in Malaysia. They compare support vector regression (SVR), Gaussian process (GP), linear regression (LR), neural network (NN), and ARIMA. They found that the NN outperforms most of the methods in positive and recovery case prediction. The ARIMA can predict the number of dead cases accurately. The most unreliable model is LR because it overestimated the case number so much. Meanwhile, the GP and SVR have shown a promising result as the Root Mean Square Error (RMSE) value is around the ARIMA value.

The previous literature shows the machine learning method can forecast the outbreak damage level by predicting the number of infected people. In this study, the number of positive cases is not only expected but the type of movement restriction can be predicted using the machine learning method. The move restriction has been proved by many countries can reduce the number of positive and dead cases [4], [18]–[23]. The comparison of Multiple Linear Regression (MLR), Support Vector Regression (SVR), Random Forest Regressor, and Decision Tree Regressor are conducted to know the best regression method [24]–[26]. In contrast, the classification methods are implemented to determine the movement restriction status. The Logistic Regression, Support Vector Machine (SVM), Decision Tree, Naïve Bayes, KNN, Adaboost, and XGboost methods are compared to know the most successful approach to predict the correct movement restriction. Moreover, the Particle Component Analysis (PCA) and grid search algorithm are implemented to understand the features reduction and hyperparameter optimization effect in accuracy, Mean Absolute Error (MAE), and Mean Square Error (MSE) [27]–[29].

The number of positive cases can be predicted by using a regression model, and the number of expected cases is combined with factual information (e.g., number of public transport passengers, wind speed, and congestion level) at that time to predict the type of the restriction. The number of positive cases needs to be predicted because the data of positive cases is unknown at that specific time. Therefore, a prototype of the website application is developed and deployed to show the proposed system in a real environment.

III. METHOD

We collect all data such as COVID-19 patient data, weather, and climate data as well as air transportation user data. Then the data is cleaned using imputation method to recover the missing data and scaling method if the data is not on the same scale. Then the clean data is processed with statistical and visual analysis. Only then after getting a deeper understanding, the data can be added or reduced depending on the findings with visualized data. After the data is ready, the data can be trained with a regression algorithm to obtain predictions of the number of patients with Covid 19 in daily basis. The results of the prediction are entered into the classification system and the type of restrictions can be determined. The mean absolute error (MAE), the mean square error (MSE), and accuracy are used as measuring instrument.

A. Data Accuisation

Based on the problems formulation, we collect data that might affect the number of positive cases. To get complete COVID-19 data in Jakarta we access the following website: https://corona.jakarta.go.id. The display form of the website above is shown in Fig. 1. We can see various information related to Covid-19 in Jakarta such as the number of specimens being tested, the number of people who are vaccinated, the amount that recovered, etc.
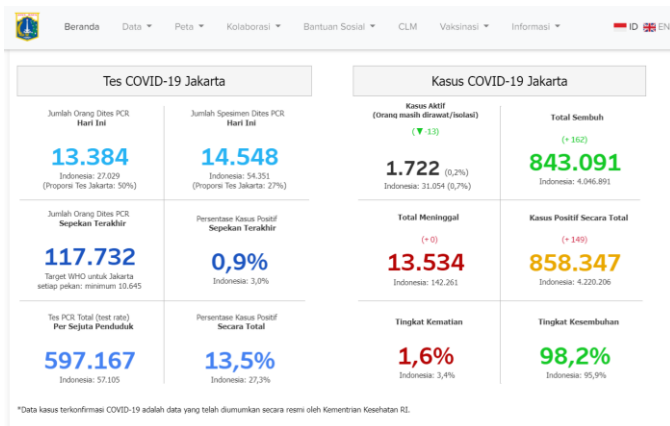
Fig. 1.   The Main Menu of https://corona.jakarta.go.id.

Because we cannot access the data from the website, we traced the data source and found that the dashboard was made using the tableau application. Then we found the tableau address as follows:

https://public.tableau.com/app/profile/jsc.data/viz/dashboardcovid-19jakarta_15837354399300/dashboard22

From the website, we can download data in the form of a PDF File, as shown in Fig. 2. Then, from the PDF data, we change it to Excel File with the help of the PDF Editor application (Nitro Pro).
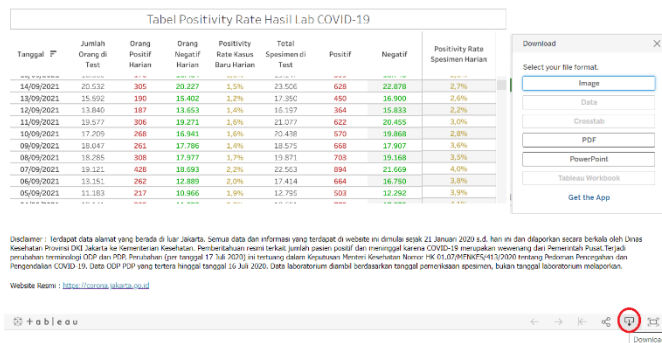
Fig. 2.   Covid-19 Data in Jakarta at Tableau Site.

Then we searched the movement data of Jakarta people by looking at the number of users of public transportation modes. For train user data, we take from the following site: https://www.bps.go.id/indicator/17/72/2/sum-penumpang-kereta-api.html. From that website, we can download the data in an excel file, as shown in Fig. 3.
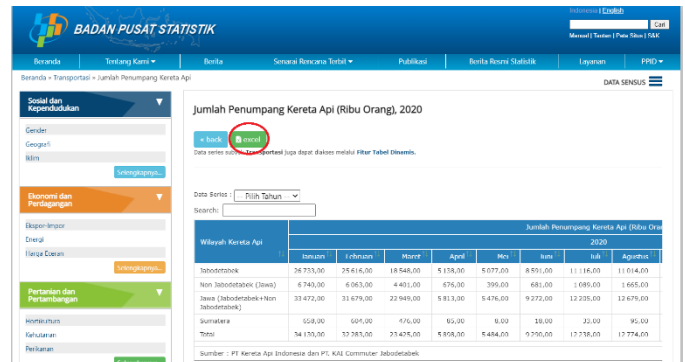

Fig. 3.   The Data of Train Users in Jakarta.

Then we took the data of users of Transjakarta, Mass Rapid Transport (MRT), and LRT (Integrated Rail Cross) via https://data.jakarta.go.id/dataset. In addition to the three-transportation data, we can also obtain other data, such as data on the number of residents moving to Jakarta, the number of family planning participants, and others, as shown in Fig. 4.
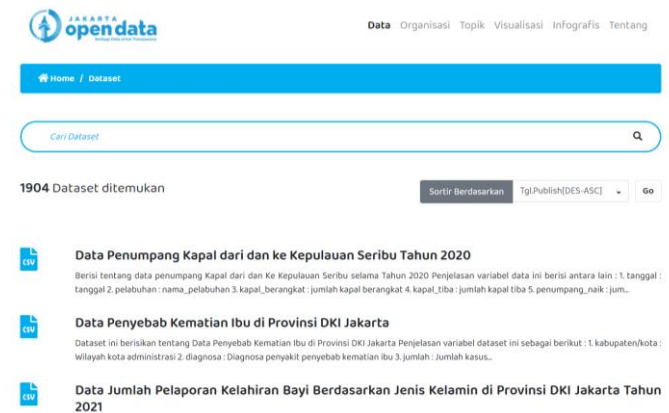

Fig. 4.   Jakarta Open Data Site.

We can perform a search, as shown in Fig. 4, and then select the data we want to access. For MRT data, we can download the data by clicking Download Data, as shown in Fig. 5. Then, we get a comma-separated value (CSV) from it. For the LRT and Transjakarta data, we used the same method as previously described to obtain the data. The addresses of the datasets are as follows:
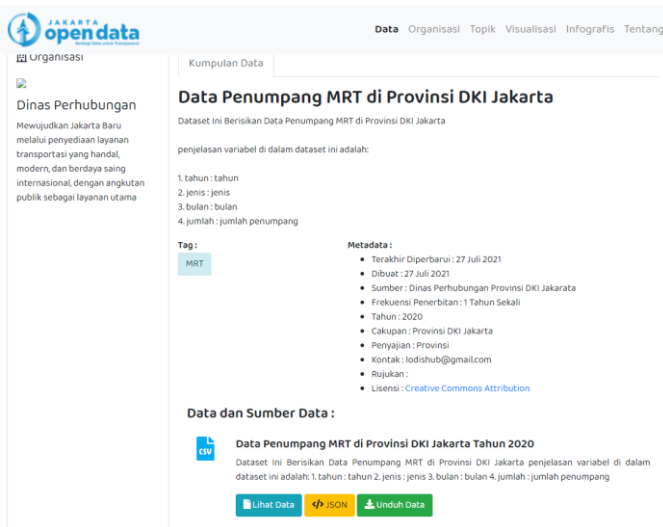
https://data.jakarta.go.id/dataset/data-penumpang-transjakarta-2020/

https://data.jakarta.go.id/dataset/data-penumpang-mrt-di-provinsi-dki-jakarta

https://data.jakarta.go.id/dataset/data-penumpang-lrt-di-provinsi-dki-jakarta

Fig. 5. Download Section in MRT, Jakarta Open Data.

For data on weather conditions and air quality in Jakarta, we take from two sources, namely:

1) https://dataonline.bmkg.go.id/data_iklim
2) https://aqicn.org/data-platform/covid19/

For Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) online data, we need to create an account first, and we can only download it for a period of 30 days, as shown in Fig. 6. While on the Air Quality Open Data Platform we can download air condition data around the world from 2015 to 2021 as shown in Fig. 7. From BMKG we can obtain data on rainfall, duration of exposure, wind speed, humidity, etc. Meanwhile, on the Air Quality Open Data Platform, we collected PM10 and PM2.5 data, which are international standard air condition indexes.
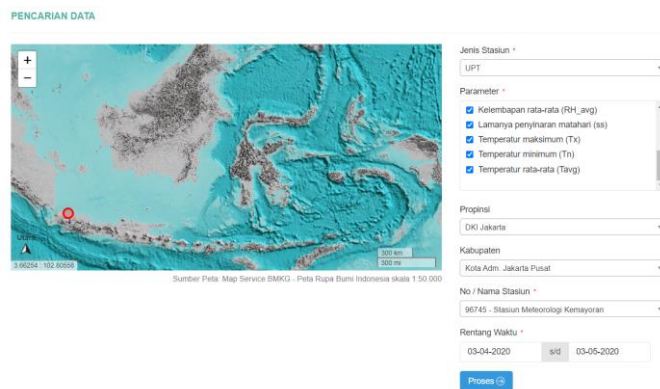

Fig. 6. Badan Meteorologi, Klimatologi, Dan Geofisika (BMKG) Data.

For congestion data, we take from https://www.tomtom.com/en_gb/traffic-index/jakarta-traffic/. We can only take data every month as an indicator of congestion, as shown in Fig. 8. The index is between 0 – 1, where 0 is the value without congestion and 1 is a great traffic jam.

For data on the status of people's movement restrictions, we take the following news:

https://news.detik.com/berita/d-5167032/timeline-psbb-jakarta-to-tarik-rem-darurat

From this news, we can determine the period of strict PSBB, transitional PSBB, and no PSBB. So that the data we need is complete but still not well organized and validated.
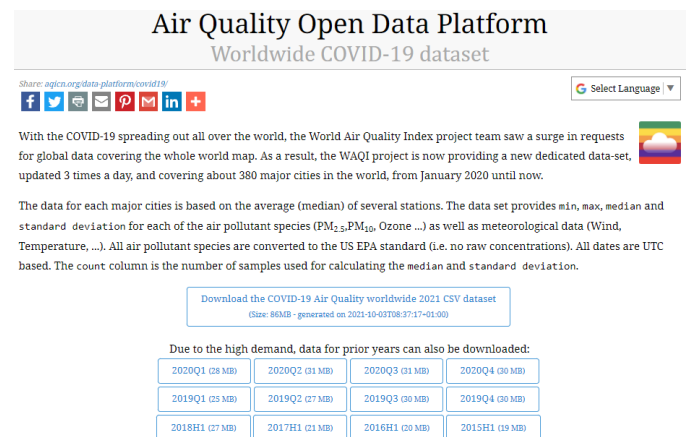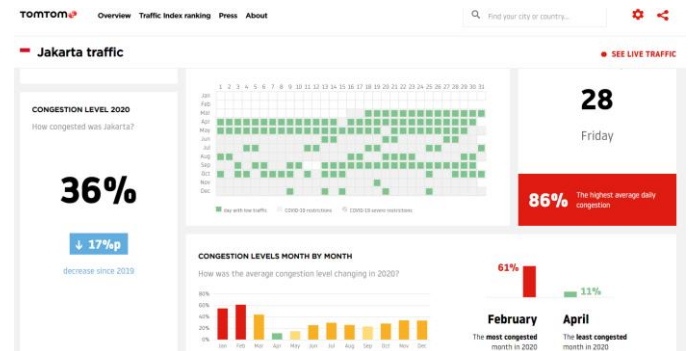

Fig. 7. Air Quality Open Data Platform

Fig. 8. Jakarta Congestion Data

B. Data Preparation

Once we get all the required data, we look at the data. The first criterion used is the availability of data during the PSBB period, which is March 10, 2020 – November 25, 2020. After investigating the dataset, the downloaded data meets these criteria. Then we discard the unnecessary information, such as the highest wind direction, maximum wind speed, maximum/minimum temperature, etc. We delete the data because we only used average values such as average wind speed and average temperature. In addition, we also discard data on the number of suspects, the number of negative covid-19, and the daily percentage of cases because this has been explained in the daily positive number and daily test number. After the needless data is removed, we then combine the data into a comprehensive table.

The overview of this research problems are the daily positive case prediction and the movement restriction type classification problems. We expect that the predicted positive case will be used as pseudo data to indicate the type of people's movement restrictions. Table I contains information

as follows, KRL Passenger, MRT Passenger, LRT Passenger, Bus Passenger, Congestion level, Rain drop rate, Day time, Humidity, Pm10, Pm2.5, Temperature, Wind, Restriction Status, Tested sample, and Positive Case.

TABLE I. COMBINATION OF ALL DATA

Date	KRL	..	Status	Tested Sample	Positive Case
10/03/2020	618266	..	NO_PSBB	52	2
11/03/2020	618266	..	NO_PSBB	48	26
12/03/2020	618266	..	NO_PSBB	56	10
13/03/2020	618266	..	NO_PSBB	30	7
14/03/2020	618266	..	NO_PSBB	44	16
15/03/2020	618266	..	NO_PSBB	41	2
..	..	..	..	..	..
25/11/2020	387400	..	NO_PSBB	15,381	1,124

C. Data Exploration

Descriptive statistical analysis can be seen in Table II below. From Table II, we can find out the average, maximum value, and minimum value of each attribute or feature. From Table II, there is nothing strange about these statistical values, so we can say that the combined data is valid.

TABLE II. STATISTICS OF ALL DATA

	KRL	MRT	LRT	Bus	..	Positive Case
count	261	261	261	261	..	261
mean	326717.6	14614.85	727.3257	251539.8	..	6080.831
std	117980.5	11273.04	461.2927	131444.5	..	4425.058
min	169233	1451	198	80396	..	30
25%	286366	11351	613	149401	..	2211
50%	337600	12991	632	285011	..	5295
75%	370533	17393	872	307195	..	10021
max	618266	46787	2024	559231	..	17871

Fig. 9 shows that the daily positive number changes along with the number of people being tested (blue) for COVID-19 per day. This indicates that the daily positive number (red) is strongly influenced by the number of people who test for Covid.
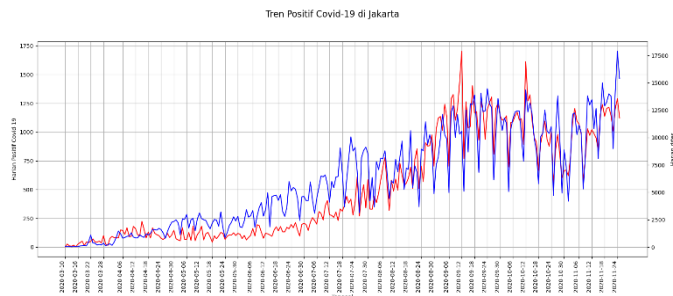

Fig. 9. The Positive Case vs the Sample Tested.

Fig. 10 shows a strong correlation between the number of daily tests and daily positives. There is a weak correlation between the number of Bus passengers and the daily test. There

is also a weak correlation between congestion levels and daily tests.

Fig. 10. The Correlation Matrix.

Based on Fig. 11 the distribution shows a positive correlation between the number of daily tests and the number of daily positives.
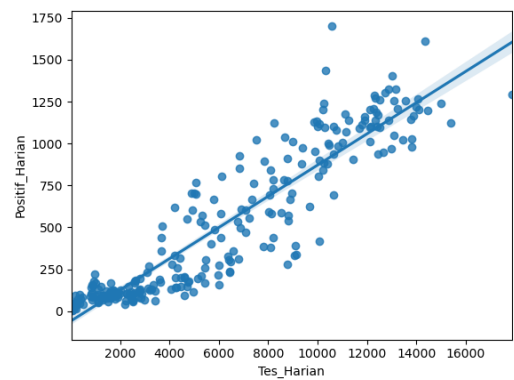

Fig. 11. Tested Sample vs. Positive Case.

Since we don't have daily data on congestion levels, we use monthly data, creating a scatter plot graph as shown in Fig. 12. There is a weak correlation between congestion with a positive total.
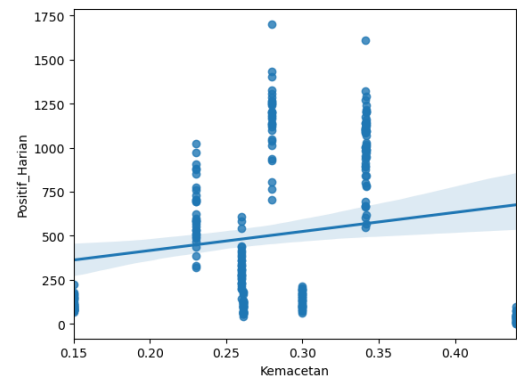

Fig. 12. Congestion vs. Positive Case.

Because data on the number of bus passengers per day is not available, we use the average per day of the number of passengers per month as shown in Fig. 13. We get a weak positive correlation between Bus vs. Total positive.
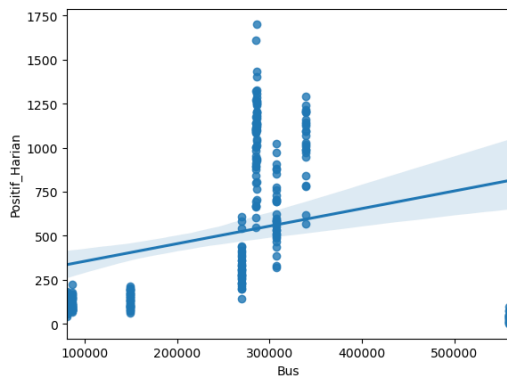
Fig. 13. Bus Passenger vs. Positive Case.

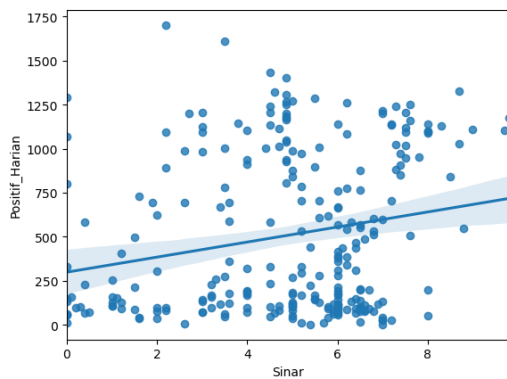There is a very weak correlation between solar irradiance and daily positive, as shown in Fig. 14.

Fig. 14. Day Time vs. Positive Case.

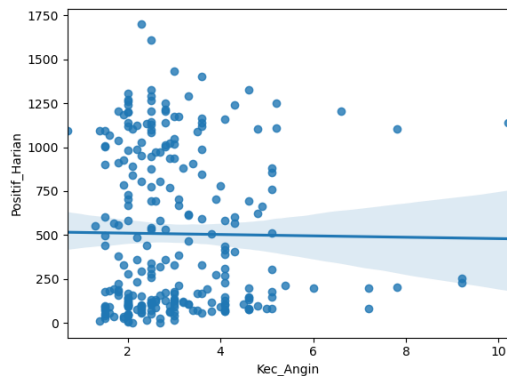There is a very weak negative correlation between wind speed and positive total, as shown in Fig. 15.

Fig. 15. Wind Speed vs. Positive Case.

Based on Fig. 16, there is a difference in the median value between no PSBB and strict PSBB. However, there was no significant difference between the three PSBB statuses in the number of daily positive cases.
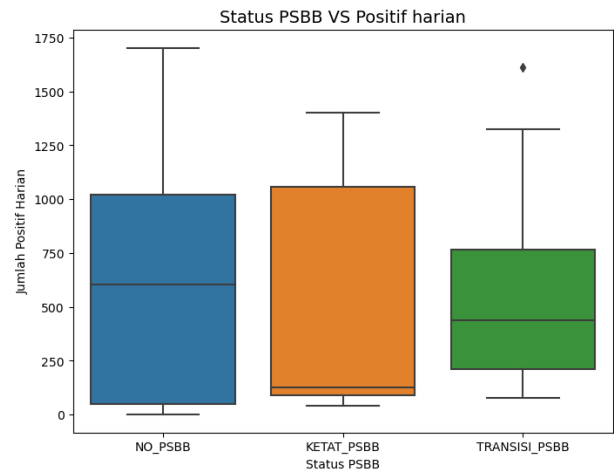

Fig. 16. Block Plot of People Restriction Status.

D. Data Transformation

Since we do not have daily data on the number of KRL passengers, we use the mean of total monthly passengers. However, giving the same value for one month is very unrealistic, so we need to transform the data to be more realistic by generating random data that meets the criteria for the mean of total passengers and standard deviation = 0.2 * mean, as shown in Fig.17.
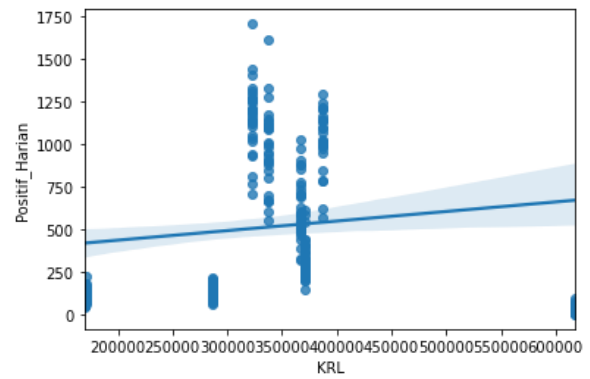
(a)
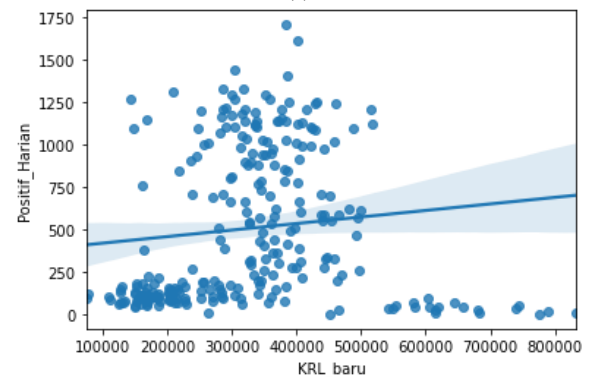
(b)

Fig. 17. KRL Passenger Data Transformation.

To be able use total monthly passengers, number of LRT, number of Bus passengers, and daily data of congestion in realistic way, data transformations must be done. The random data are generated using mean and standard deviation (it is calculated by 0.2 * mean) of the respective datasets. The transformations of the dataset have shown in Fig. 18, Fig. 19, Fig. 20, and Fig. 21 respectively.
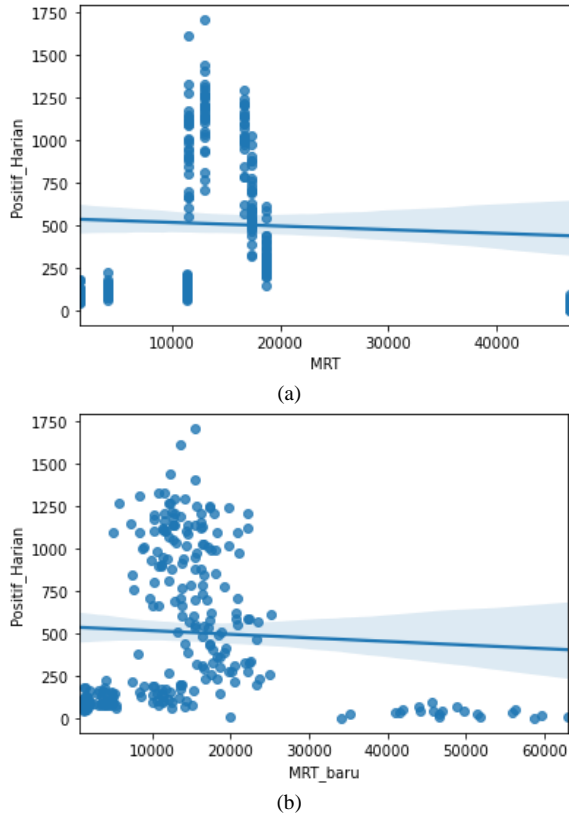

(a)

(b)

Fig. 18. MRT Passenger Data Transformation. (a) Before Transformation, (b) After Transformation.
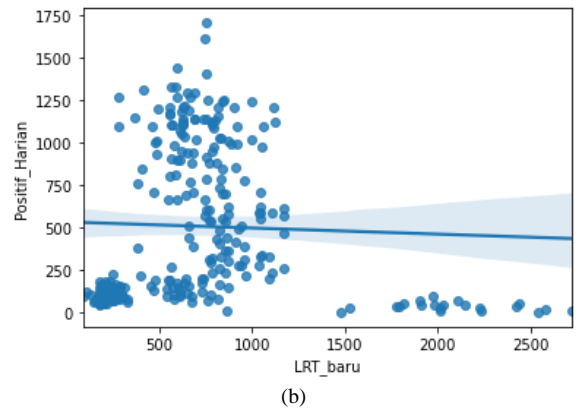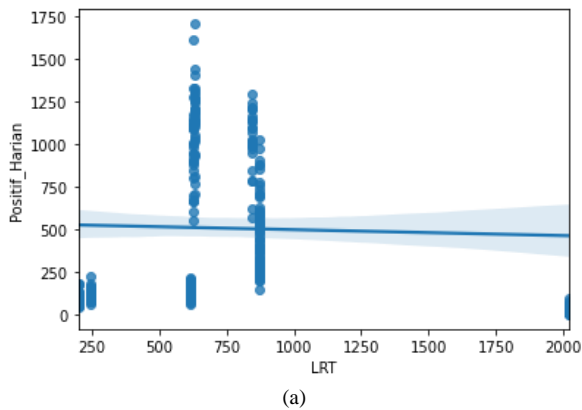

(a)

(b)

Fig. 19. LRT Passenger Data Transformation. (a) Before Transformation, (b) After Transformation.
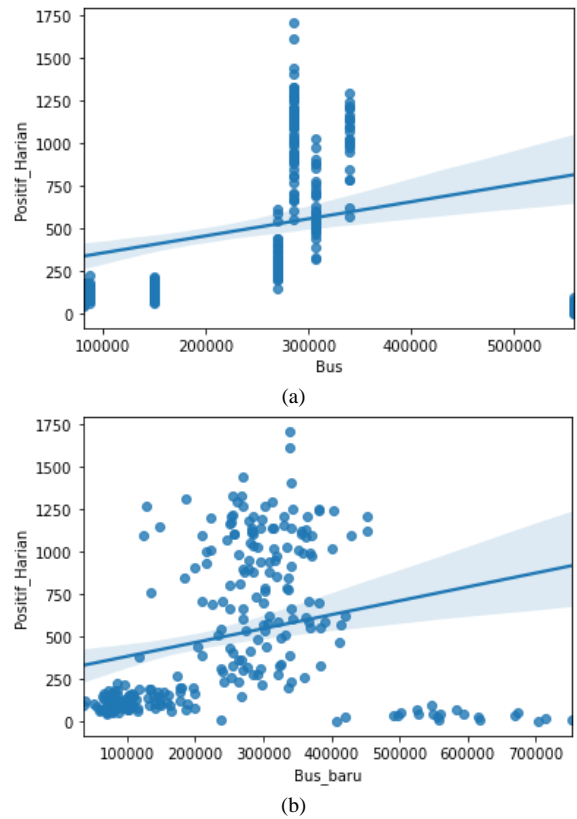

(a)

(b)

Fig. 20. Bus Passenger Data Transformation. (a) Before Transformation, (b) After Transformation.
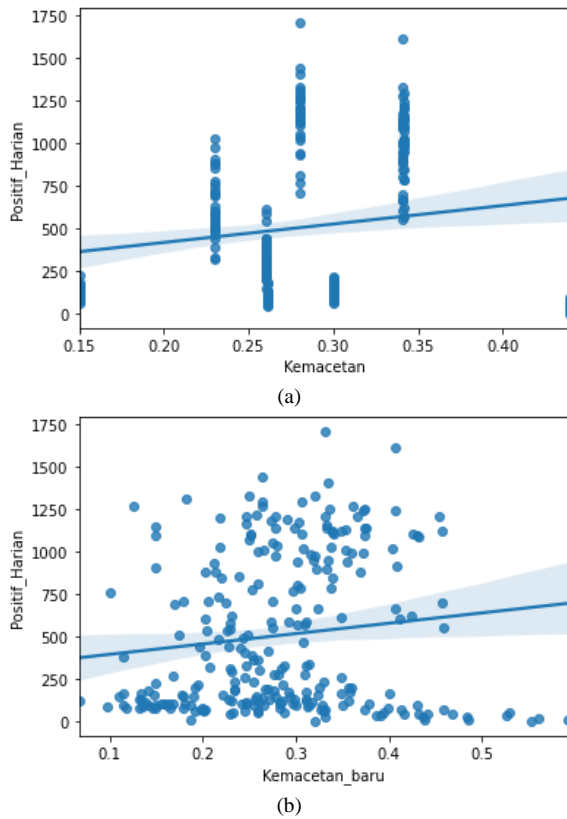
(a)

(b)

Fig. 21. Congestion Data Transformation. (a) Before Transformation, (b) After Transformation.

To be able to see the correlation between PSBB status and other attributes, it is necessary to transform the data from PSBB status with one hot encoding. Then by using the Pearson correlation method, a correlation table is obtained as shown in Fig. 22.
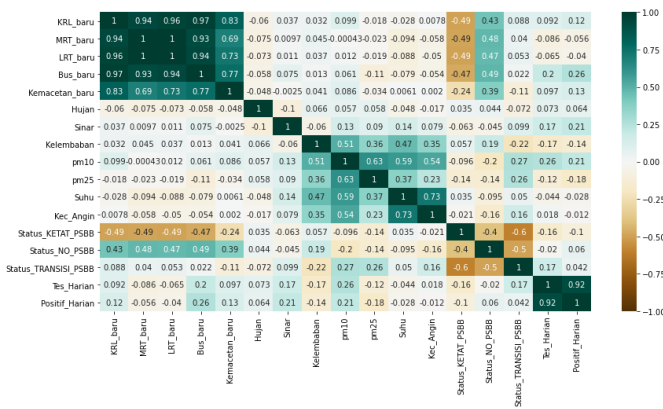

Fig. 22. Correlation Analysis after Data Transformation.

Based on Fig. 22, the daily test has the most significant correlation, which is 0.92. Then Bus_baru, which is the number of bus passengers per day, has a weak correlation of 0.26. Then the light, which is the length of irradiation in hours, has a correlation of 0.21. Then pm10, the air pollution index, has a weak correlation of 0.21. And humidity has a weak correlation of -0.14.
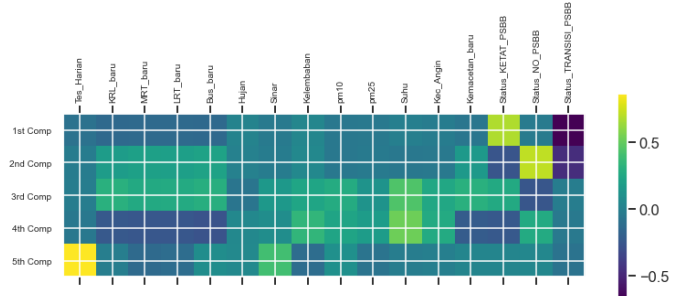

Fig. 23. The Contribution of Each Feature to PCA Value.

Based on Fig. 23, the strict PSBB status gave a positive contribution to the 1st component. In addition, the No PSBB status made a positive contribution to the 2nd component. The transitional PSBB status made a negative contribution to 1st component. Then the daily test made a very large contribution to the 5th component. The top five of Principle Component Analysis (PCA) value for each component can be found in Table III.

TABLE III. PCA VALUE

1st Comp	2nd Comp	3rd Comp	4th Comp	5th Comp
-0.30117	1.454034	0.571142	-0.52336	-0.3233
-0.18254	1.251233	0.288714	-0.26551	-0.23284
-0.22774	1.363212	0.258748	-0.54626	-0.30674
-0.33262	1.537455	0.586097	-0.69514	-0.45323
-0.33639	1.458526	0.629993	-0.52833	-0.28187

E. Framework

Based on the aim of this project is to create a system that can predict the number of people infected with Covid-19 and determine what restrictions should be made. Therefore, the PSBB status and the number of daily positive confirmations are labeled to be predicted. We created two models, namely the regression model and the classification model. The Regression Model is used to indicate the daily confirmed number of Covid-19 under certain conditions, and the classification model is used to determine what restrictions should be placed on the state of the daily positive count, as shown in Fig. 24.
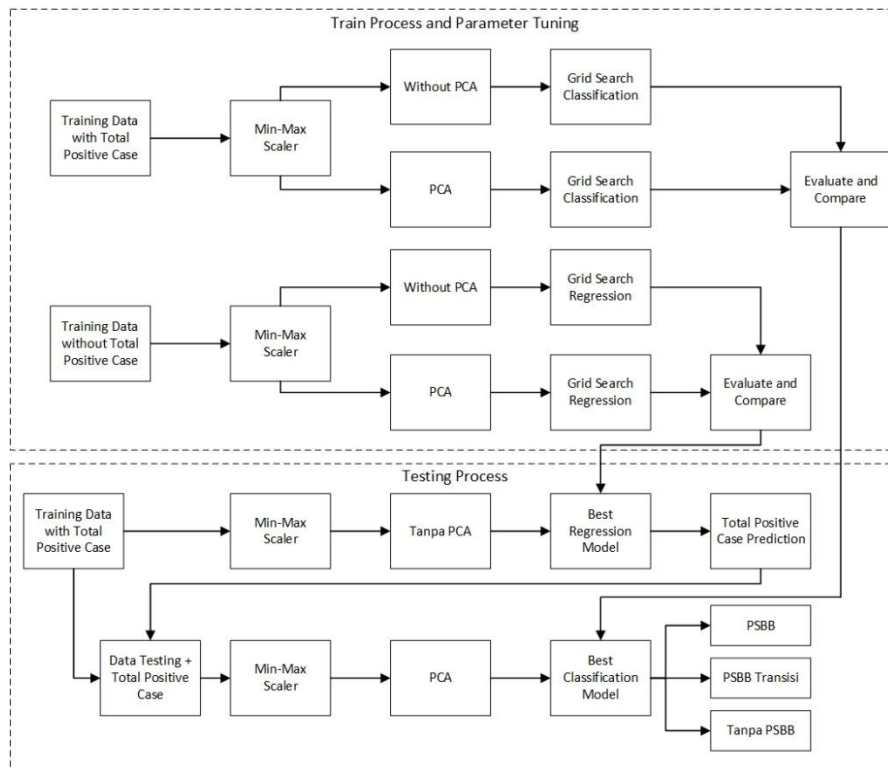
Fig. 24. Covid-19 Case Prediction and Movement Restriction Classification System Framework.
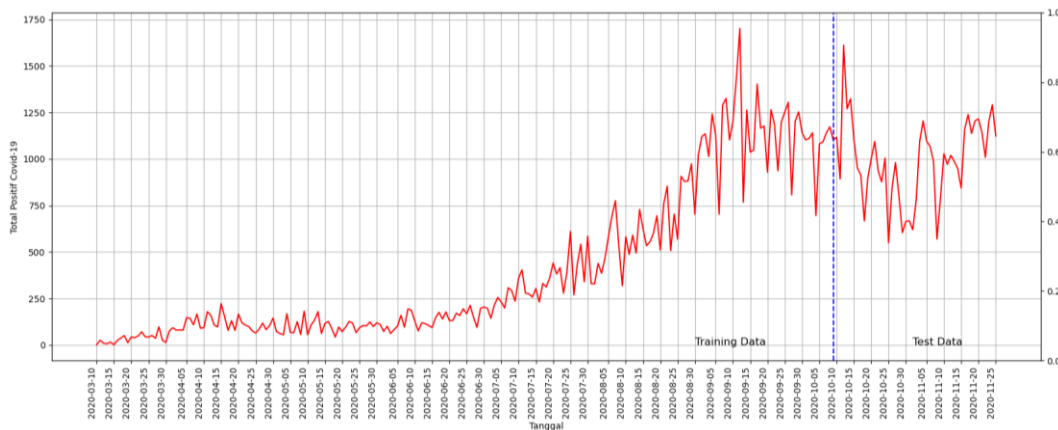

Fig. 25. Data Training and Testing Separation.

IV.   RESULT AND DISCUSSION

We first made a regression model by comparing the evaluation results using PCA and without PCA. The Grid Search Method is used to determine the best method and the best parameters. To produce a model that can predict future values, it is necessary first to divide the training and testing data as shown in Fig. 25. The best methods, parameters, and evaluation results are shown in Table IV. Based on Table IV the best method is Multiple Linear Regression (MLR) with MAE = 148.3892 and MSE = 37036.37. If you look at the R2 value, only MLR without PCA gets a positive value. So, the MLR method without PCA was chosen as the regression method used.

Next, we need to determine a suitable method for classification, using PCA and without PCA, and determine the best parameters with grid search. The results of the grid search are shown in Table V. Based on Table V, the classification method used is a support vector machine (SVM) with PCA with a training accuracy of 0.72 and a testing accuracy of 0.8481. Because the testing accuracy exceeds 0.7, the SVM method is good enough to predict what kind of tightening will be done to the people of Jakarta. In Fig. 26, after all, forms are filled in, the system can predict the confirmed number of Covid-19 on that day totaling 871 people. In addition, we are recommended to implement the Transition PSBB by the system.
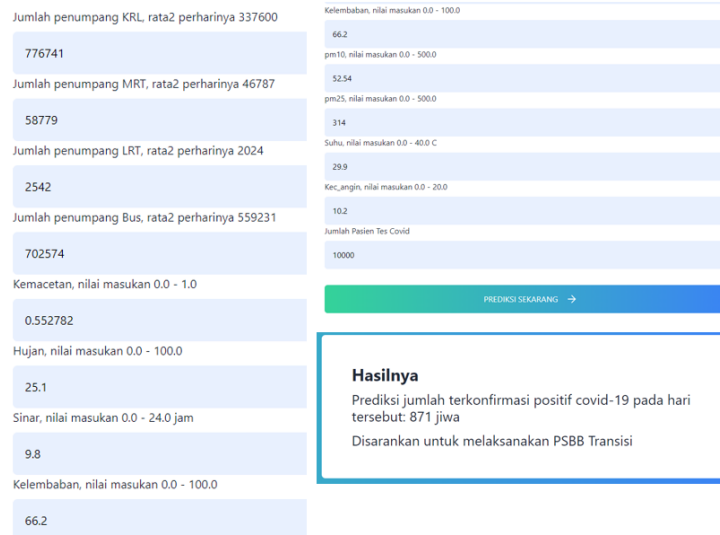
Fig. 26. System Deployment https://prediksi-covid-psbb.herokuapp.com.

TABLE IV. REGRESSION RESULT

Method	Parameter	Transforms	MAE	MSE	R2-score
MLR	**fit_intercept=True, normalize=False**	**Without PCA**	**148.3892**	**37036.37**	**0.250336**
SVR	kernel='poly', degree=3	Without PCA	536.7085	333113.5	-5.74265
Random Forest Regressor	n_estimators=100, max_depth=10	Without PCA	231.3127	77746.98	-0.5737
Decision Tree Regressor	splitter='best', max_depth=10	Without PCA	271.6456	122728.9	-1.48419
MLR	fit_intercept=True, normalize=False	PCA	307.6172	118276.6	-1.39407
SVR	kernel='poly', degree=3	PCA	866.0381	796039.1	-15.1129
Random Forest Regressor	n_estimators=50, max_depth=5	PCA	348.5216	178841.1	-2.61998
Decision Tree Regressor	splitter='best', max_depth=5	PCA	357.3924	181732.3	-2.6785

TABLE V. CLASSIFICATION RESULT

Method	Parameter	Transforms	Acc. Training	Acc. Testing
Logistic Regression	C: 10, penalty: 'l1', solver: 'liblinear'	Without PCA	0.703097	0.797468
SVM	C: 10, degree: 2, gamma: 1, kernel: 'rbf'	Without PCA	0.730237	0.772152
Decision Tree	criterion: 'gini', max_depth: 4, min_samples_leaf: 3	Without PCA	0.702732	0.670886
Naïve Bayes	alpha: 0.001, binarize: 0, fit_prior: False	Without PCA	0.489162	0.443038
KNN	algorithm: 'auto', n_neighbors: 4, p: 1, weights': 'distance'	Without PCA	0.758015	0.810127
Adaboost	algorithm: 'SAMME', learning_rate: 0.1, n_estimators: 200	Without PCA	0.713934	0.772152
XGboost	booster: 'gbtree', eta': 0.1	Without PCA	0.774499	0.810127
Logistic Regression	C: 10, penalty: 'l1', solver: 'liblinear'	PCA	0.642532	0.683544
SVM	**C: 10, degree: 2, gamma: 1, kernel: 'rbf'**	**PCA**	**0.724681**	**0.848101**
Decision Tree	criterion: 'gini', max_depth: 4, min_samples_leaf: 3	PCA	0.708379	0.683544
Naïve Bayes	alpha: 0.001, binarize: 0, fit_prior: False	PCA	0.483698	0.506329
KNN	algorithm: 'auto', n_neighbors: 4, p: 1, weights': 'distance'	PCA	0.724863	0.78481
Adaboost	algorithm: 'SAMME', learning_rate: 0.1, n_estimators: 200	PCA	0.680783	0.708861
XGboost	booster: 'gbtree', eta': 0.1	PCA	0.729964	0.734177

V. Conclusion and Future Work

In this study, the data collection and data preprocessing were explained in detail to show the reader how to collect and prepare the data before it is fed to the classification or regression methods. By detailed explanation of both the tasks, the reader can follow and develop their own data acquisition and preprocessing tasks. From the correlation analysis, it was found that the number who carry out the Covid-19 test affects the number of confirmed Covid-19 significantly. Adding data transformations such as PCA can enhance the accuracy of SVM. However, the MLR did not gain improvement. In addition, based on the test results, we were able to obtain an accuracy of 84%, which is a pretty good result. With this level of accuracy, we can be confident in using the model in actual cases. From the deployment results, we can see that the machine learning model can make predictions as we desire. The model can predict the number of Covid-19 sufferers and provide recommendations for restricting the movement of people.

Even though the accuracy of classification is relatively high, the system is still limited to the Jakarta region. The research expansion to the whole nation (Indonesia) can be more conclusive and comprehensive. The challenge will be the data collection that most of the regions do not have interactive web base covid data. So, collaboration with the government must be established. The provided data transformation is still limited to PCA. The extensive study in data transformation can be exciting as a lot of feature extraction and selection methods are available. Moreover, the technique is still limited to conventional machine learning approaches. The exploration of deep learning algorithms such as long short-term memory (LSTM) and convolutional neural networks (CNN) can be interesting discussions.

Acknowledgment

References

[1] H. A. Rothan and S. N. Byrareddy, "The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak," J. Autoimmun., vol. 109, no. February, pp. 18–21, 2020, doi: 10.1016/j.jaut.2020.102433.

[2] Z. Wu and J. M. McGoogan, "Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases from the Chinese Center for Disease Control and Prevention," JAMA - J. Am. Med. Assoc., vol. 323, no. 13, pp. 1239–1242, 2020, doi: 10.1001/jama.2020.2648.

[3] Z. E. Rasjid, R. Setiawan, and A. Effendi, "A Comparison: Prediction of Death and Infected COVID-19 Cases in Indonesia Using Time Series Smoothing and LSTM Neural Network," Procedia Comput. Sci., vol. 179, no. 2020, pp. 982–988, 2021, doi: 10.1016/j.procs.2021.01.102.

[4] N. T. P. Pang, K. Assis, M. A. Mohd Kassim, and C. M. Ho, "Analyses of the effectiveness of movement control order (Mco) in reducing the covid-19 confirmed cases in Malaysia," J. Heal. Transl. Med., vol. 24, no. Special Issue Covid-19, pp. 16–27, 2021.

[5] R. S. Pontoh et al., "Jakarta Pandemic to Endemic Transition: Forecasting COVID-19 Using NNAR and LSTM," Appl. Sci., vol. 12, no. 12, p. 5771, 2022, doi: 10.3390/app12125771.

[6] R. O. Nanda et al., "Community Mobility and COVID-19 Dynamics in Jakarta, Indonesia," Int. J. Environ. Res. Public Health, vol. 19, no. 11, p. 6671, 2022, doi: 10.3390/ijerph19116671.

[7] S. Liu et al., "Predicting the outbreak of hand, foot, and mouth disease in Nanjing, China: a time-series model based on weather variability," Int. J. Biometeorol., vol. 62, no. 4, pp. 565–574, 2018, doi: 10.1007/s00484-017-1465-3.

[8] B. Modu, N. Polovina, Y. Lan, S. Konur, A. Taufiq Asyhari, and Y. Peng, "Towards a predictive analytics-based intelligent malaria outbreakwarning system," Appl. Sci., vol. 7, no. 8, pp. 1–20, 2017, doi: 10.3390/app7080836.

[9] R. Jain, S. Sontisirikit, S. Iamsirithaworn, and H. Prendinger, "Prediction of dengue outbreaks based on disease surveillance, meteorological and socio-economic data," BMC Infect. Dis., vol. 19, no. 1, pp. 1–16, 2019, doi: 10.1186/s12879-019-3874-x.

[10] E. O. Nsoesie, J. S. Brownstein, N. Ramakrishnan, and M. V. Marathe, "A systematic review of studies on forecasting the dynamics of influenza outbreaks," Influenza Other Respi. Viruses, vol. 8, no. 3, pp. 309–316, 2014, doi: 10.1111/irv.12226.

[11] S. Lahmiri and S. Bekiros, "Intelligent forecasting with machine learning trading systems in chaotic intraday Bitcoin market," Chaos, Solitons and Fractals, vol. 133, 2020, doi: 10.1016/j.chaos.2020.109641.

[12] Y. Hu, C. Hu, S. Fu, P. Shi, and B. Ning, "Predicting the popularity of viral topics based on time series forecasting," Neurocomputing, vol. 210, pp. 55–65, 2016, doi: 10.1016/j.neucom.2015.10.143.

[13] O. Torrealba-Rodriguez, R. A. Conde-Gutiérrez, and A. L. Hernández-Javier, "Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models," Chaos, Solitons and Fractals, vol. 138, 2020, doi: 10.1016/j.chaos.2020.109946.

[14] R. M. Arias Velásquez and J. V. Mejía Lara, "Forecast and evaluation of COVID-19 spreading in USA with reduced-space Gaussian process regression," Chaos, Solitons and Fractals, vol. 136, 2020, doi: 10.1016/j.chaos.2020.109924.

[15] E. Buckingham-Jeffery, V. Isham, and T. House, "Gaussian process approximations for fast inference from infectious disease data," Math. Biosci., vol. 301, no. February, pp. 111–120, 2018, doi: 10.1016/j.mbs.2018.02.003.

[16] M. Yousaf, S. Zahir, M. Riaz, S. M. Hussain, and K. Shah, "Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan," Chaos, Solitons and Fractals, vol. 138, 2020, doi: 10.1016/j.chaos.2020.109926.

[17] A. N. A. Kamarudin, Z. Zainol, N. F. A. Kassim, and R. Sharif, "Prediction of COVID-19 cases in Malaysia by using machine learning: A preliminary testing," 2021 Int. Conf. Women Data Sci. Taif Univ. WiDSTaif 2021, 2021, doi: 10.1109/WIDSTAIF52235.2021.9430222.

[18] L. Di Domenico, G. Pullano, C. E. Sabbatini, P. Y. Boëlle, and V. Colizza, "Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies," BMC Med., vol. 18, no. 1, pp. 1–13, 2020, doi: 10.1186/s12916-020-01698-4.

[19] S. Panneer et al., "The Great Lockdown in the Wake of COVID-19 and Its Implications: Lessons for Low and Middle-Income Countries," Int. J. Environ. Res. Public Health, vol. 19, no. 1, 2022, doi: 10.3390/ijerph19010610.

[20] T. Sardar, S. S. Nadim, S. Rana, and J. Chattopadhyay, "Assessment of lockdown effect in some states and overall India: A predictive mathematical study on COVID-19 outbreak," Chaos, Solitons and Fractals, vol. 139, p. 110078, 2020, doi: 10.1016/j.chaos.2020.110078.

[21] N. G. Davies et al., "Association of tiered restrictions and a second lockdown with COVID-19 deaths and hospital admissions in England: a modelling study," Lancet Infect. Dis., vol. 21, no. 4, pp. 482–492, 2021, doi: 10.1016/S1473-3099(20)30984-1.

[22] B. Paital, K. Das, and S. K. Parida, "Inter nation social lockdown versus medical care against COVID-19, a mild environmental insight with special reference to India," Sci. Total Environ., vol. 728, p. 138914, 2020, doi: 10.1016/j.scitotenv.2020.138914.

[23] M. A. Acuña-Zegarra, M. Santana-Cibrian, and J. X. Velasco-Hernandez, "Modeling behavioral change and COVID-19 containment in Mexico: A trade-off between lockdown and compliance," Math. Biosci., vol. 325, no. May, p. 108370, 2020, doi: 10.1016/j.mbs.2020.108370.

[24] A. Farezi, S. A. Akbar, A. Yudhana, K. H. Ghazali, P. A. Rosyady, and A. Aminuddin, "Glucose content analysis using image processing and machine learning techniques," 2022 5th Int. Conf. Inf. Commun. Technol. ICOIACT 2022, 2022.

[25] S. A. Akbar et al., "Classification of gram-positive and gram-negative bacterial images based on machine learning algorithm," 2022 5th Int. Conf. Inf. Commun. Technol. ICOIACT 2022, 2022.

[26] E. Y. Sari, A. D. Wierfi, and A. Setyanto, "Sentiment Analysis of Customer Satisfaction on Transportation Network Company Using Naive Bayes Classifier," 2019 Int. Conf. Comput. Eng. Network, Intell. Multimedia, CENIM 2019 - Proceeding, vol. 2019-Novem, Nov. 2019, doi: 10.1109/CENIM48368.2019.8973262.

[27] A. Aminuddin and F. Ernawan, "AuSR1: Authentication and self-recovery using a new image inpainting technique with LSB shifting in fragile image watermarking," J. King Saud Univ. - Comput. Inf. Sci., Feb. 2022, doi: 10.1016/J.JKSUCI.2022.02.009.

[28] A. Aminuddin and F. Ernawan, "AuSR2: Image watermarking technique for authentication and self-recovery with image texture preservation," Comput. Electr. Eng., vol. 102, p. 108207, Sep. 2022, doi: 10.1016/J.COMPELECENG.2022.108207.

[29] F. Ernawan, A. Aminuddin, D. Nincarean, M. F. A. Razak, and A. Firdaus, "Three Layer Authentications with a Spiral Block Mapping to Prove Authenticity in Medical Images," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 4, 2022, doi: 10.14569/IJACSA.2022.0130425.