International Journal of Multidisciplinary Research and Publications
ISSN (Online): 2581-6187

Implementation of Machine Learning Algorithm Using K-Nearest Neighbors Technique to Predict Indonesian State Budget Deficit

Irma Elita[1], Lulu Chaerani Munggaran[2]
[1,2]Department of Technology and Engineering, Gunadarma University, Jakarta, DKI Jakarta, Indonesia-129590
Email address: [1]irmaelitha @ gmail.com, [2]lulu.chaerani @ gmail.com

***Abstract*—** *The state budget of Indonesia is an instrument used by the government to manage the budget to archive development goals in the economy. Since 1984 to 2019 state budget of Indonesia experiences a budget deficit where state expenditure is bigger than its revenue which has an impact on the decline of the trade balance, a decrease in the level of gross domestic product that indicates the ability of the country's economic resources to weaken, and the increase in government debt to finance the budget deficit. This study applies the machine learning algorithm using the k-Nearest Neighbors classification technique to predict Indonesia's State budget deficit by using the nearest optimum distance of the k-Fold Cross-Validation algorithm. The results showed that the application of the budget deficit prediction can predict a decrease/increase in the budget deficit with an accuracy level of 63%. This level of accuracy is obtained by using the top 9 nearest neighbors distance that is most appropriate for this study.*

Keywords*—Budget Deficit, K-Fold Cross-Validation, K-Nearest Neighbors, Machine Learning*.

I. INTRODUCTION

According to Law No. 20 of 2019, the state budget is an annual government financial plan approved by the House of Representatives. The state budget consists of the state revenue budget, the state expenditure budget, and budget financing.

Based on data from Badan Pusat Statistik (BPS) regarding the rate of growth of Gross Domestic Product (GDP) and data from Kementerian Keuangan Republik Indonesia about the state budget and financial note of 1984-2019 shows that state revenues have fluctuated each year and often do not reach figures that have been targeted in the state budget.

In 2019 the realization of state revenues grew by 1.6% from 2018, amounting to Rp1,957 trillion with an economic growth rate of 5.2%, Rupiah exchange rate of Rp14,250.00 per USD, inflation of 3.1%, oil price of US$63 per barrel, SPN interest rate of 5.6%, oil lifting of 754,000 per day and gas lifting of 1,072,000 barrels of oil equivalent per day. This increase in state revenue was followed by the realization of state expenditures of Rp2,310 trillion [2,3].

The condition where the state expenditure exceeds the amount of revenue is called the budget deficit [4]. This budget deficit can be financed by the government by accumulating loans in the past and making loans in the bond market [5]. The trauma experienced by the government in financing the budget deficit through the printing of money which resulted in high inflation in the 1960s made the government start using foreign debt to be able to finance the budget deficit [6].

The budget deficit, in the long run, has a negative impact, which can be a decline in the trade balance [7]. GDP will also experience a negative impact on increasing budget deficits because this indicates that the ability of economic resources produced by the country is weakening [5].

From this impact, the predicted budget deficit appears necessary in the process of determining the budget deficit target in the draft of the state budget. The budget deficit prediction is used to monitor the economy of the country situation and determine whether there will be a swelling in the realization of the budget deficit from the initial target specified.

According to Cao, Fang, Ottoson, Näslund, and Stendberg in their research said that several algorithms in machine learning had been proven to make accurate predictions. The machine learning algorithm is designed to automatically retrieve data from several sources, create logical rules and display the patterns found from those rules [8].

The process of finding a model used in predicting the class of an unknown object is called classification. One algorithm that can be used to classify is k-Nearest Neighbors (k-NN) [9]. k-NN is a method used to classify objects based on learning data that is the closest distance to the object [10].

The k-Fold Cross-Validation algorithm is one method to find out the average success of a system repeatedly using a randomized attribute [11]. K-Fold Cross-Validation functions to assess the performance of an algorithm by randomizing data and dividing it into groups as many as k-Fold values. Each group will be used as test data and the rest as training data [12]

Based on the description above, this study will apply the k-NN of the machine learning algorithm to be able to predict the realization of the Indonesian state budget deficit by using the most optimal closest neighbors distance value obtained from the k-Fold Cross-Validation algorithm.

II. LITERATURE REVIEW

A. State Budget

The state budget consists of the state revenue budget, state expenditure budget, and budget financing. According to Law No. 20 of 2019, it is stated that what is meant by the state budget is the annual financial plan of the government of the country approved by the House of Representatives.

Irma Elita and Lulu Chaerani Munggaran, "Implementation of Machine Learning Algorithm Using K-Nearest Neighbors Technique to Predict Indonesian State Budget Deficit," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 2, Issue 9, pp. 17-22, 2020.

State opinion consists of tax revenues, non-tax state revenues, and grant receipts which are the right of the central government to be recognized as an addition to net wealth whereas state expenditures consist of central government spending and transfers to regions and village funds which are the obligations of the central government to be recognized as a deduction from the net worth.

Basic macroeconomic assumptions consisting of Gross Domestic Product (GDP), economic growth, inflation y-o-y, Rupiah exchange rates, 3-month SPN interest rates, oil prices, oil lifting, gas lifting are used as a reference in compiling various components state budget posture.

B. K-Nearest Neighbors

K-Nearest Neighbors (k-NN) is a machine learning algorithm that is used to classify objects based on training data which is the closest distance to the object. The results of query instances are classified by the majority so this technique is included in the supervised learning algorithm. Class classification results are determined based on the class that appears the most. This algorithm aims to classify new objects using attributes and training data samples [9].

Since this algorithm only uses very few training data for learning purposes, the k-NN algorithm is often called a lazy learning algorithm or memory-based learning or instance-based learning. The hypothesis in k-NN depends on the training data that has been provided so that the more data that is added, the more complicated the process to achieve that hypothesis. The results obtained from k-NN depend on the amount of data of the nearest neighbors chosen (k). To avoid the appearance of the same amount of distance, in general, the k value chosen is odd [10].

C. K-Fold Cross-Validation

In the k-NN algorithm, new data is classified based on the distance of the data with its nearest neighbor. The k-Fold Cross-Validation algorithm is a method to find out the average success of a system. This method shuffles the attributes entered and tests repeatedly as many k-Fold values as specified to divide the training data into groups. Once each group chosen becomes the test data, the other groups will become the training data. The error rate that occurs in the system can usually be measured using 5-Fold or 10-Fold Cross-Validation.

Each test data is predicted using the k-NN algorithm for each k value. The results of the k-NN prediction classification are determined by comparing the actual data with the correct number of predictions from all the actual data. The highest level of accuracy that will be selected to be the best k [11].

III. Method and Materials

A. Identification

Based on Indonesian state budget data from 1984-2019, Indonesia always experiences a budget deficit. In APBN 2019 (Direktorat Penyusunan APBN, 2020) there was a swelling budget deficit of around Rp353 trillion or 2.2% of GDP. The realization of this budget deficit is greater than the initial target, which is Rp296 trillion or 1.84% of GDP.

The budget deficit was caused by declining revenue (Rp1,957 trillion from the initial target of Rp2,165 trillion) and stable spending (Rp2,310 trillion from the initial target of Rp2,461 trillion) increasing the primary balance (Rp77.5 trillion from the initial target of Rp. 20.1 trillion) which means that funds are not available to pay debt interest and the government must issue new debt to pay the principal and interest of the debt.

According to research conducted by Darma Anita [7] in his research showed that the budget deficit in the long term and short term can cause the trade balance to decline. The inflation rate has an indirect effect on increasing the budget deficit. With the price increase due to inflation, the country is forced to spend more than what has been budgeted. The increased expenditure due to inflation which can cause an increased budget deficit [1].

Yahya [26] said that there are five impacts of the budget deficit policy on the macroeconomy: (1) Demand for money increases and the interest rate rises due to reduced tax revenue so the country needs additional capital; (2) The higher the interest rate, the lower the domestic investment which means there will be a swelling in the trade balance and the Rupiah weakens; (3) An increase in prices caused by state spending that exceeds its revenue can result in a decline in real income of the community and a reduction in consumption and savings so that it affects the level of investment; (4) The enthusiasm of investors to invest will decrease due to high-interest rates so that the impact on project termination because it can no longer be continued. With the termination of the project there will be a dismissal of workers which can cause an increase in the unemployment rate; (5) Economic growth will weaken if investment decreases and interest rates rise.

Based on the impact of the budget deficit, a system is needed to be able to predict whether the realization of the budget deficit with the basic macroeconomic assumptions given to the state budget in the future will increase or decrease. With this prediction, it is expected that the impact of the budget deficit can be anticipated by the government by making policies that can overcome these conditions.

B. Analysis

In this step, an analysis of the k-NN technique is implemented to predict budget deficits.

a) Collection of dataset

The dataset used for this research is data on the amount of realization of the state revenue and state expenditure, along with basic macroeconomic assumption data consisting of Gross Domestic Product (GDP), percentage of economic growth, percentage of inflation y-o-y, Rupiah exchange rate against US$, percentage of 3-month SPN interest rate, oil price in units of US$ per barrel, the amount of oil lifting in units of thousand barrels per day, and the amount of lifting gas in units of thousand barrels per day. This dataset was obtained from Indonesian state budget data for the 1984-2019 period and grouped according to the fiscal year.

The data used in this study are secondary data obtained from the official BPS website on the rate of growth of gross domestic products and the website of Kementerian Keuangan

Irma Elita and Lulu Chaerani Munggaran, "Implementation of Machine Learning Algorithm Using K-Nearest Neighbors Technique to Predict Indonesian State Budget Deficit," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 2, Issue 9, pp. 17-22, 2020.

Republik Indonesia on the law of state budget and financial note.

b) Normalization dataset

The data that has been obtained is then normalized before being processed using the k-NN algorithm. This data normalization is done using the feature scaling formula. Refer to (1), data that has to be normalized is expressed with variable X. Whereas X_{min} is a variable for the minimum data set in the state budget dataset and X_{max} is for the state budget maximum data. X' contains the results of the normalization. The equation of the feature scaling formula is as follows.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

Data that are not normalized will be difficult to apply to machine learning, for example, data for state revenues and state expenditures have units of trillions of Rupiah, while for GDP data have percent units. The difference in the range will make the calculation of the distance between one data with other data produces a very long range so that the normalization of data is needed.

c) Determine the amount of distance to the nearest neighbor

The next step is to determine the value of k, which is the number of nearest neighbors to be used as a comparison. The k value that will be used in the application of the k-NN algorithm is the k value with an odd number to avoid the same amount of distance from occurring. The k values chosen in this study were all odd numbers between 1-30, namely: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23.

d) Determine the number of groups

YEAR	SR	SE	GDP	EG	I	RER	SPN	OP	ML	GL
2013	0.7380	0.7354	0.5240	0.9116	0.0688	0.7135	0.0326	0.9307	0.0940	0.8568
1992	0.0134	0.0151	0.0692	0.9163	0.0384	0.0759	0.1902	0.1996	0.9016	0.4465
2011	0.5931	0.5499	0.4119	0.9209	0.0238	0.5856	0.0290	0.9886	0.1574	0.9278
2006	0.3196	0.2935	0.1356	0.8884	0.0794	0.6150	0.1594	0.4995	0.2885	0.9230
1986	0.0025	0.0026	0.0057	0.8884	0.0899	0.0194	0.1938	0.2658	0.7377	0.0878
1985	0.0000	0.0000	0.0057	0.7302	0.0304	0.0064	0.2772	0.4736	0.6426	0.0614
2000	0.0685	0.0775	0.0798	0.8419	0.0966	0.5592	0.2047	0.0414	0.7672	0.8279
2007	0.3434	0.3184	0.1769	0.9116	0.0608	0.6137	0.0978	0.5822	0.2142	0.8664
1995	0.0253	0.0245	0.1412	1.0000	0.0873	0.0925	0.1721	0.0207	0.9005	0.6450
1987	0.0010	0.0022	0.0000	0.8465	0.0913	0.0467	0.1504	0.2182	0.7279	0.1348
1996	0.0288	0.0267	0.1702	0.9814	0.0595	0.0996	0.1902	0.0724	0.9027	0.7124
1999	0.0566	0.0839	0.0701	0.6512	0.0000	0.5164	0.3460	0.0765	0.7148	0.8363
2002	0.1404	0.1312	0.1115	0.8186	0.1058	0.6265	0.2174	0.0776	0.5847	0.9266
2012	0.6566	0.6343	0.4668	0.9070	0.0304	0.6322	0.0000	1.0000	0.1148	0.9170
2009	0.4134	0.3959	0.2887	0.8326	0.0106	0.7081	0.0779	0.4716	0.2077	0.8387
1997	0.0342	0.0331	0.1569	0.8372	0.1204	0.1425	0.1449	0.0321	0.8776	0.7473
2003	0.1618	0.1545	0.1518	0.8279	0.0410	0.5710	0.1268	0.1313	0.4612	1.0000
2008	0.4794	0.4165	0.2444	0.8977	0.1204	0.6558	0.1105	0.8376	0.1279	0.7762
2001	0.1345	0.1390	0.0752	0.7814	0.1389	0.6983	0.2609	0.0889	0.6918	0.7858
1993	0.0164	0.0185	0.0907	0.9209	0.1019	0.0803	0.1105	0.0920	0.9115	0.4874
1991	0.0134	0.0129	0.0563	0.9395	0.0992	0.0699	0.3533	0.2027	1.0000	0.3875
1984	0.0005	0.0004	0.0114	0.9395	0.0886	0.0000	0.2138	0.5026	0.8208	0.0000
1998	0.0362	0.0361	0.0191	0.0000	1.0000	0.6796	1.0000	0.0124	0.8372	0.7244
1988	0.0030	0.0043	0.0114	0.8837	0.0450	0.0500	0.2192	0.2141	0.6765	0.2034
1989	0.0040	0.0069	0.0260	0.9628	0.0516	0.0563	0.1522	0.2937	0.7945	0.2455
2010	0.4864	0.4410	0.3441	0.9070	0.0661	0.6098	0.0616	0.6556	0.2175	0.9952
2004	0.1945	0.1807	0.0974	0.8000	0.0661	0.5983	0.0779	0.2068	0.4109	0.9290
2005	0.2382	0.2108	0.2044	0.8791	0.1997	0.6563	0.1069	0.3868	0.3738	0.9410
1994	0.0203	0.0219	0.1124	0.9674	0.0952	0.0858	0.1504	0.0000	0.9126	0.6390
1990	0.0079	0.0095	0.0408	0.9535	0.0992	0.0618	0.2645	0.3992	0.8579	0.2984

Legend:
Group-1
Group -2
Group -3
Group -4
Group -5

Fig. 1. Normalized and Grouped State Budget Data

In research conducted by Mutiara, Irwan & Andi [11] it is said that with 5-Fold or 10-Fold Cross-Validation can be used to measure the level of errors that occur. Therefore in this study used 5-Fold Cross-Validation to find the most optimal k value. The distribution of data is determined based on the percentage that will be used for training data and testing data.

Since this research uses 5-Fold, the data to be used in this search only uses 30 data. The data is divided by the amount of 80:20 which means that the training data is 24 data and testing data are 6 data. Each training data already has its class. The classes used in the training data are increasing or decreasing. The most optimal k search experiment is carried out by running the k-NN algorithm 5 times and dividing 30 randomized data into 5 groups with the same amount of data, which is 6 data. Each experiment will predict 6 different data.

Each experiment will use a different group to be tested data. Referring to Fig. 1, it can be seen in the experiment, the k-NN algorithm will be run using group-1 as test data. That is, data in groups 2, 3, 4 and 5 will be used as training data. When running the second experiment, the k-NN algorithm is run using data testing in group-2 and data in groups 1, 3, 4, and 5 will be the training data. It is certain that for each experiment will use different test data.

e) Calculate the nearest distance

The k-NN algorithm is very dependent on the nearest neighbors to the test data. Usually, the distance or proximity between neighbors is calculated using the Euclidean formula. The smaller the distance between the test data and the training data, the greater the distance between the two. Equation (2) is a Euclidean formula.

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \qquad (2)$$

Variable d contains the results of distance calculation based on the year of test data (x) with the current training data year (y). This formula takes the sum of the square root of the running column test data (x_i) reduced by the running column training data (y_i).

f) Classification of testing data

At this stage, the classification of predicted data is determined. The classification of this data depends on the value of k, which means the results of the classification for each k value can be different. In Tab. I is an example of predictions of the data in 2013 data using the value $k = 1$ which means the classification is done by taking one smallest data neighbor to determine predictions.

TABLE I. Classification of Testing Data in 2013 with $k = 1$.

d(x,y)	Distance	Classification
(2013,2012)	0.1953	Increase
(2013,2008)	0.5248	Increase
..	..	..
(2013,1991)	1.8320	Increase
(2013,1998)	2.4730	Decrease

Since the selected data is only one neighbor, the 2013 budget deficit is predicted to increase because only one data is

Irma Elita and Lulu Chaerani Munggaran, "Implementation of Machine Learning Algorithm Using K-Nearest Neighbors Technique to Predict Indonesian State Budget Deficit," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 2, Issue 9, pp. 17-22, 2020.

used as a comparison. Whereas for $k = 9$ in Tab. II it is predicted that the realization of the budget deficit in 2013 will decrease because the majority of the classes in the 9 smallest data neighbor are the declining class with a total of 5 out of 9 data.

TABLE II. Classification of Testing Data in 2013 with $k = 9$.

d(x,y)	Distance	Classification
(2013,2012)	0.1953	Increase
(2013,2008)	0.5248	Increase
(2013,2010)	0.5514	Decrease
(2013,2009)	0.7197	Decrease
(2013,2007)	0.7747	Increase
(2013,2005)	1.0180	Increase
(2013,2004)	1.2062	Decrease
(2013,2003)	1.2830	Decrease
(2013,2002)	1.3878	Decrease
(2013,2001)	1.4434	Increase
..	..	..
(2013,1991)	1.8320	Increase
(2013,1998)	2.4730	Decrease

g) Test the accuracy of the predicted results

After getting the results of predictions on all values of k, then the next step is to test the level of accuracy of the prediction results provided. This level of accuracy is obtained by calculating the total number of correct prediction results (C) at the value of k divided by the total number of data tested (X) and multiplied by 100%.

$$Accuracy = \frac{C}{X} \times 100\% \qquad (3)$$

h) Determine the most optimal nearest neighbor distance

After running all five k-NN algorithm experiments and getting all the results of the k-level accuracy value in each experiment, the next step is to determine the most optimal k value. Based on the total average obtained from all levels of accuracy in the first to fifth experiments for each k value, it can be concluded that there are two k values that have the same level of accuracy, namely the value of $k = 9$ and the value of $k = 23$ by 63%. According to research conducted by Hana, Fhira and Danang [27] the greater the k value used in the k-NN algorithm, the slower the performance provided by the system. Therefore, this study will use the value of $k = 9$ as the most optimal k value.

C. Design

After analyzing the most optimal k value for use in the application of the k-NN algorithm in this study, the next step is designing the application to be created using UML diagrams. This design phase consists of designing the application workflow and designing the appearance of the application. The design of the application display is made to display predictive information simply accompanied by a chart of the state budget data for the 1984-2019 period.

D. Implementation

At this stage, the application design that has been created is then implemented using the R and Shiny programming languages to create an application interface that allows users to enter data that they want to predict into the application through a user-friendly display.

IV. Result and Discussion

A. Prediction Page

Fig. 2 is a display of the prediction page results. The design of the prediction page is created by displaying a navigation bar that lists the menus in the application. Furthermore, on the main body of the prediction page, there are three parts, namely a section to enter prediction data, a section to display prediction results, and a section to display a graph of accuracy test levels and a summary of the results of normalization of all training data and testing data.

The form for inputting prediction data is created with 12 input columns according to the number of variables that will be used for the k-NN algorithm and one button to execute the command to run the k-NN algorithm. In the prediction results section displayed in a tabular form that can be equipped with several features, namely the predictive data search feature, the feature to display the amount of data to display, and the pagination feature to move to the next table page if the table contains more than the amount of data displayed every page of the table.

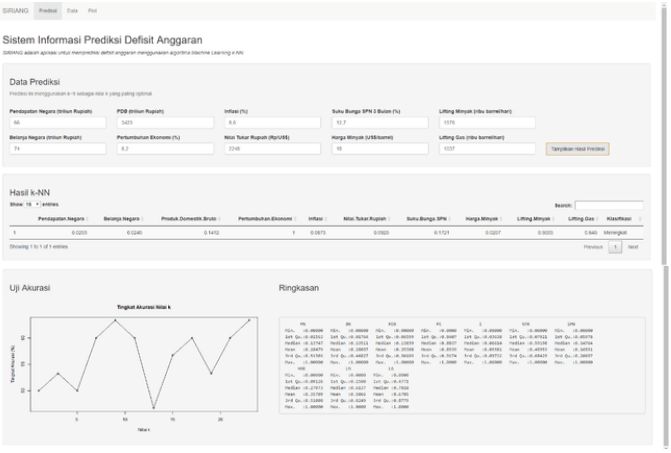

Fig. 2. Prediction Page

B. Data Page

The design of the data page can be seen in Fig. 3. In this design, there is only a table that is equipped with a state budget data search feature making it easier for users to find the data sought.

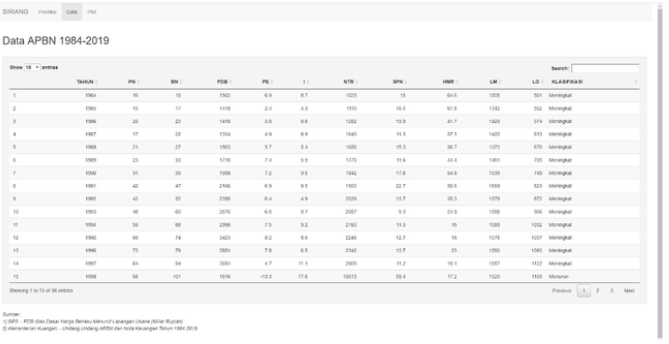

Fig. 3. Data Page

Irma Elita and Lulu Chaerani Munggaran, "Implementation of Machine Learning Algorithm Using K-Nearest Neighbors Technique to Predict Indonesian State Budget Deficit," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, Volume 2, Issue 9, pp. 17-22, 2020.

There is the feature of the amount of data to be displayed on one page, with these features allowing users to display all data on one page or divide it into several pages. By default, the table will be displayed in 15 data rows per page. Then there is the pagination feature that will appear if the user displays the data by dividing it into several pages.

C. Plot Page

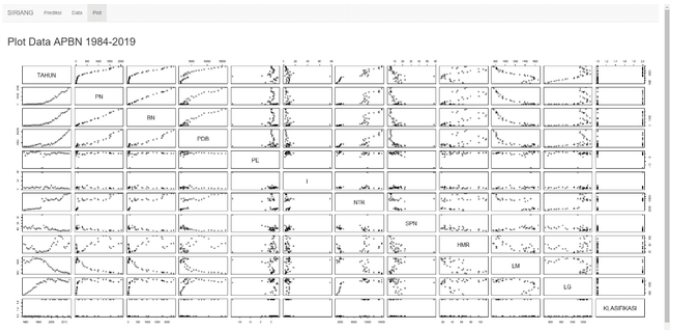

Fig. 4. Plot Page

Figure 4 is an illustration of the plot page design. This page only contains a collection of graphs from the 12 data variables used in this study. This page also features a graphic information feature that will appear if the user presses one of the points on the chart you want to appoint.

D. Unit Testing

After deploying an application project, the next step is to test the application to determine whether the application is made under the expected results. The testing technique used in this study is unit testing. This test is done by making a test code to call every method in the application of the budget deficit prediction application.

This research uses the help of the R language library named *testthat* in making test code. With this library, the results obtained from the method being tested can be compared to whether or not the same as the results that should be. Based on the results in Tab. III it can be shown that the testing of all the methods in this application gives the results as expected so that it can be concluded that the predicted results displayed on the application are obtained from the implementation of the appropriate k-NN algorithm stages.

TABLE III. Unit Testing.

Method	Result
normalization(data, min, max)	Pass
knn_result(test, train, class)	Pass
xy_str(x,y)	Pass
xy_range_str(xmin, xmax, ymin, ymax)	Pass

V. CONCLUSION

The budget deficit prediction application has been successfully made by implementing the k-Nearest Neighbors technique of machine learning algorithm to classify the class of the predicted object, namely the budget deficit. This application can display prediction results by utilizing state budget data for the period 1984-2019 as training data which consists of 10 main indicator variables in the preparation of Indonesian state budget posture: state revenue, state expenditure, Gross Domestic Product (GDP), economic growth, inflation y-o-y, Rupiah exchange rates, 3-month SPN interest rates, oil prices, oil lifting, and gas lifting.

Based on the results of testing the level of accuracy using the k-Fold Cross-Validation parameter optimization method with the value of 5-Fold Cross-Validation which means that this test is repeated 5 times by dividing 30 training data into 5 groups, showing that by taking the nearest neighbor distance (k) as much 9 and 23 neighbors in determining class classification produce the same level of accuracy, which is 63%. Since the greater the distance is taken from the k-NN algorithm, the slower the performance provided by the system therefore that this study uses the nearest distance of 9 neighbors.

REFERENCES

[1] Ratnah S, "Faktor-Faktor Yang Berpengaruh Terhadap Defisit APBN Indonesia," J. Econ., vol. 3, no. 2, pp. 1–11, 2015.

[2] Badan Pusat Statistik, "Laju Pertumbuhan Y on Y Produk Domestik Bruto Menurut Lapangan Usaha (Persen), 2017 - 2019," 2019. [Online]. Available: https://www.bps.go.id/dynamictable/2017/05/05/1255/-seri-2010-laju-pertumbuhan-y-on-y-produk-domestik-bruto-menurut-lapangan-usaha-persen-2017---2019.html. [Accessed: 26-Jan-2020].

[3] Direktorat Penyusunan APBN, Informasi APBN 2020. Jakarta: Kementerian Keuangan Republik Indonesia, 2020.

[4] Maychel, Josep & Dennij, "Analisis Pengaruh Produk Domestik Bruto, Defisit Anggaran, & Tingkat Kurs Terhadap Utang Luar Negeri Indonesia (Periode Tahun 1996-2016)," J. Berk. Ilm. Efisiensi, vol. 18, no. 1, pp. 69–78, 2018.

[5] Dwinanda & Berly, "Pengaruh Defisit Fiskal terhadap Pertumbuhan Ekonomi di Indonesia Pendahuluan Tinjauan Referensi," J. Ekon. & Pembang. Indones., vol. 15, no. 2, pp. 144–157, 2015.

[6] Fatturroyhan & Mufti, "Pembiayaan Defisit APBN Menurut Umer Chapra (Studi Analisa Kritik Terhadap Pembiayaan Defisit APBN Indonesia Periode 2010-2015)," J. Stud. Islam, vol. 12, no. 1, pp. 1–12, 2017, doi: 10.31603/cakrawala.v12i1.1652.

[7] Darma Anita, "Dampak Defisit Anggaran Pendapatan Belanja Negara & Investasi Asing Pada Industri Pengolahan Non Migas Terhadap Neraca Perdagangan," J. Online Mhs. Fak. Ekon. Univ. Riau, vol. 4, no. 1, pp. 365–378, 2016.

[8] I. Matijosaitiene, A. McDowald & V. Juneja, "Predicting safe parking spaces: A machine learning approach to geospatial urban and crime data," Sustain., vol. 11, no. 10, pp. 1–15, 2019, doi: 10.3390/su11102848.

[9] Rifkie Primartha, Belajar Machine Learning Teori & Praktek. Bandung: INFORMATIKA, 2018.

[10] Ika Kurniawati, "Komparasi Metode Machine Learning Pada Prediksi Kelulusan Mahasiswa A Comparative Method of Machine Learning On Predicting Students Graduation," no. May, pp. 37–45, 2019.

[11] Mutiara, Irwan & Andi, "Penerapan K-Optimal Pada Algoritma Knn Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan Ip Sampai Dengan Semester 4," Klik - Kumpul. J. Ilmu Komput., vol. 2, no. 2, pp. 159–173, 2015, doi: 10.20527/KLIK.V2I2.26.

[12] Anjelika, Tanzil & Indriati, "Penerapan Algoritma Modified K-Nearest Neighbour Pada Pengklasifikasian Penyakit Kejiwaan Skizofrenia," J. Pengemb. Teknol. Inf. & Ilmu Komput., vol. 2, no. 10, pp. 3957–3961, 2018.

[13] Dita Meyliana & Ade Sofyan Mulazid, "Pengaruh Produk Domestik Bruto (PDB), Jumlah Bagi Hasil & Jumlah Kantor terhadap Jumlah Deposito Muḍārabah Bank Syariah di Indonesia Periode 2011-2015," J. Ekon. Islam, vol. 8, no. 2, pp. 263–284, 2017, doi: http://dx.doi.org/10.21580/economica.2017.8.2.1442.

[14] Azwar, "Inflasi di Provinsi Sulawesi Selatan : Analisis Hubungan Dinamis Inflasi Komoditas Barang/Jasa," J. BPPK, vol. 9, no. 1998, pp. 47–65, 2016.

[15] F. Mubarak, "Pemodelan Tingkat Suku Bunga Surat Perbendaharaan Negara 3 Bulan," Kaji. Ekon. & Keuang., vol. 1, no. 3, 2017.
[16] Pusat Kajian Anggaran Badan Keahlian DPR RI, "Buletin APBN," 2020. [Online]. Available: https://berkas.dpr.go.id/puskajianggaran/buletin-apbn/public-file/buletin-apbn-public-55.pdf. [Accessed: 14-Feb-2020].
[17] Y. Firmansyah, "Penerapan Metode SDLC Waterfall Dalam Pembuatan Sistem Informasi Akademik Berbasis Web Studi Kasus Pondok Pesantren Al-Habi Sholeh Penerapan Metode SDLC Waterfall Dalam Pembuatan Sistem Informasi Akademik Berbasis Web Studi Kasus Pondok Pesantren Al-Habi," J. Teknol. Manaj. Inform., vol. 4, no. 1, pp. 184–191, 2018, doi: 10.26905/jtmi.v4i1.1605.
[18] Juniardi & Sari, "Implementasi Model Waterfall Pada Pengembangan Sistem Informasi Perhitungan Nilai Mata Pelajaran Berbasis Web Pada Sekolah Dasar Al-Azhar Syifa Budi Jatibening," Paradigma, vol. 19, no. 2, pp. 142–147, 2017.
[19] M. Tabrani & E. Pudjiarti, "Penerapan Metode Waterfall Pada Sistem Informasi Inventori Pt. Pangan Sehat Sejahtera," J. Inkofar, vol. 1, no. 2, pp. 30–40, 2017.
[20] Suendri, "Implementasi Diagram UML (Unified Modelling Language) Pada Perancangan Sistem Informasi Remunerasi Dosen Dengan Database Oracle (Studi Kasus: UIN Sumatera Utara Medan)," J. Ilmu Komput. & Inform., vol. 3, no. 1, pp. 1–9, 2018.
[21] Ade Hendri, "Pemodelan UML Sistem Informasi Monitoring Penjualan & Stok Barang (Studi Kasus: Distro Zhezha Pontianak)," J. Khatulistiwa Inform., vol. 4, no. 2, pp. 107–116, 2016.
[22] RStudio, "About RStudio," 2020. [Online]. Available: https://rstudio.com/about/. [Accessed: 18-Feb-2020].
[23] Septa, Renny & Irmasari, "Feature Selection pada Dataset Faktor Kesiapan Bencana pada Provinsi di Indonesia Menggunakan Metode PCA (Principal Component Analysis)," J. Tek. ITS, vol. 5, no. 2, pp. 88–92, 2016.
[24] M. Komarudin, "Pengujian perangkat Lunak metode Black box berbasis partitions pada aplikasi sistem informasi di sekolah," J. Mikrotik, vol. o6, no. 3, pp. 02–16, 2016, doi: http://dx.doi.org/10.1016/j.brat.2012.09.004.
[25] Taufik Nur Adi, "Generator Kode Unit Testing Untuk Javascript Berbasis Framework Qunit," J. Rekayasa Sist. Ind., vol. 2, no. 3, pp. 80–88, 2015, doi: 10.25124/jrsi.v2i03.69.
[26] M. Yahya, "Analisis Normatif Kritis Kebijakan Pemanfatan Obligasi Syariah (Sukuk) Dalam Menutup Defisit Apbn," Econ. J. Ekon. Islam, vol. 6, no. 2, p. 37, 2016, doi: 10.21580/economica.2015.6.2.793.
[27] Hana, Fhira & Danang, "Prediksi Penyakit Menggunakan Algoritma K-Nearest Neighbors & Algoritma Genetika Untuk Data Berdimensi Tinggi," e-Proceeding Eng., vol. 3, no. 2, pp. 3771–3777, 2016.