International Journal on Advanced Science Engineering Information Technology

Vol.13 (2023) No. 1
ISSN: 2088-5334

Prediction of Drug Demand Based on Deep Learning Approach and Classification Model

Musli Yanto [a,*], Syafri Arlis [b], Muhammad Reza Putra [a], Hadi Syahputra [b], Vicky Ariandi [a]

[a] *Department of Informatics Engineering, Universitas Putra Indonesia YPTK Padang, Indonesia*
[b] *Department of Informatics System, Faculty of Computer Science, Universitas Putra Indonesia YPTK Padang, Indonesia*
Corresponding author: [*]musli_yanto@upiyptk.ac.id

***Abstract*— The high demand for drugs in the last period has caused problems with drug shortages in several pharmacies. Almost all pharmacies experience the same problem, causing many people who do not to get their drug needs during the current pandemic. To overcome this, analyzing the process of predicting drug demand in the next period is necessary. The prediction process can be used as an initial solution in solving problems to see the number of drug demand numbers that will occur. This study aims to develop a predictive analysis model for drug demand using a deep learning approach and a classification model. Deep learning is an approach that does well in the case of prediction. The classification model also includes the right concept for solving the problem. The prediction and classification analysis methods include K-Means clustering, Multiple Linear Regression (MRL), Artificial Neural Network (ANN), and Decision Tree algorithms C.45. This method can provide better performance results in the prediction process to get precise and accurate output. Prediction results obtained from the learning process provide an accuracy rate of 99.99%. The output of the classification model also provides an overview of the knowledge base in the form of a decision tree. The level of classification model testing carried out gives the accuracy of the classification pattern of 97.05% so that the analytical model developed can predict future drug demand.**

***Keywords*— Drug; pharmacies; demand; prediction; classification.**

Manuscript received 2 Feb. 2022; revised 3 Jun. 2022; accepted 8 Aug. 2022. Date of publication 28 Feb. 2023.

I. INTRODUCTION

In the last two years, it can be seen that the Covid-19 pandemic has not experienced a significant decrease in cases [1]. The conditions in several countries are still struggling to suppress the rate of positive cases [2]. Several efforts have been made, such as lockdowns in several countries, providing education to improve health protocols, vaccine programs, and administering drugs and multivitamin supplements to maintain body resistance to avoid exposure to exposure Covid-19 [3]. This condition is in line with the increasing demand for drugs in pharmacies, causing new problems in drug availability [4], [5].

To overcome this problem, the initial solution needed is the existence of a predictive analysis process in the case of drug demand that will occur in the next period. The predictive analysis process can be used as a solution to the problems encountered [6]. Various analytical models have been developed to present precise and accurate prediction results [7]. Prediction models using Machine Learning can be used as a solution in decision-making to increase the productivity of drug procurement [8]. Other studies also explain that mathematical interpretation algorithms in forecasting cases are used to deal with drug demand problems [9]. The prediction process developed based on a model from sales data can be used to predict drug demand [10].

In previous studies, the resulting prediction process has not significantly contributed to the problem. This is based on the prediction model used that still looks traditional, so the output given still has a fairly high error rate. For this reason, this study will develop an existing model by proposing a classification model for the prediction process. The methods used include K-Means clustering, Multiple Linear Regression (MRL), Artificial Neural Network (ANN) with Deep Learning, and Decision Tree. The performance of this method is expected to provide better results in predicting drug demand.

K-means is a concept applied to group data with an essential similarity [11]. In another explanation, K-Means is an approach that can solve problems by dividing the sample into data sets [12]. The clustering method can be applied to generate patterns used in the prediction process [13]. MRL analysis is a method that can measure the level of influence of the relationship between predictors and outcomes. MRL is an analytical method that can improve the prediction process [14].

ANN is the basis of deep learning, which is applied to carry out learning with large numbers of data [15]. Implementing Deep Learning can produce predictive outputs that have a strong relationship with the accuracy value. The results obtained also have a high confidence level so that they can be used as a reference in decision-making [16]. The Decision Tree method in classification results in a pattern of rules [17]. The classification model in the prediction case can present the output as a decision tree [18].

This study aims to develop a predictive process analysis model in the case of drug demand. The proposed update is to present predictive analysis with a deep learning approach and a classification model. Moreover, the implementation of MRL analysis is also used to measure the accuracy of the variables. In general, the previous prediction process has not yet measured the variables, and thus, the results provided have not been able to prove the accuracy. In addition, this study will also provide output in the form of rules in a decision tree that can be used as a reference in decision-making. In the end, the results focus on the output of the predictive value of drug demand and provide a fairly good analytical model.

II. Materials and Method

Drug demand research using the concept of deep learning provides predictive results with an average accuracy of 92.4% based on the data set of the number of drugs. The results provided can be used as information that is quite valuable in carrying out the drug availability management process [19]. In this case, deep learning is a learning concept that is used to solve certain problems with better accuracy output [20]. The development of deep learning models can be applied to solving forecasting problems. The Multilayer network model will be a learning architecture to get better output [21]. The classification model developed in deep learning is able to carry out the classification process [22]. The results given by the classification model in deep learning get much better results based on the tests carried out [23].

This research method is used to explain the stages of activities that will be carried out in the process of predicting drug demand. The dataset used in the prediction process is sourced from historical real-time sales data that occurred in Kimia Farma Pharmacy, Padang, West Sumatra Province, Indonesia, within the last one year. The stages of this research can be seen in Fig. 1.

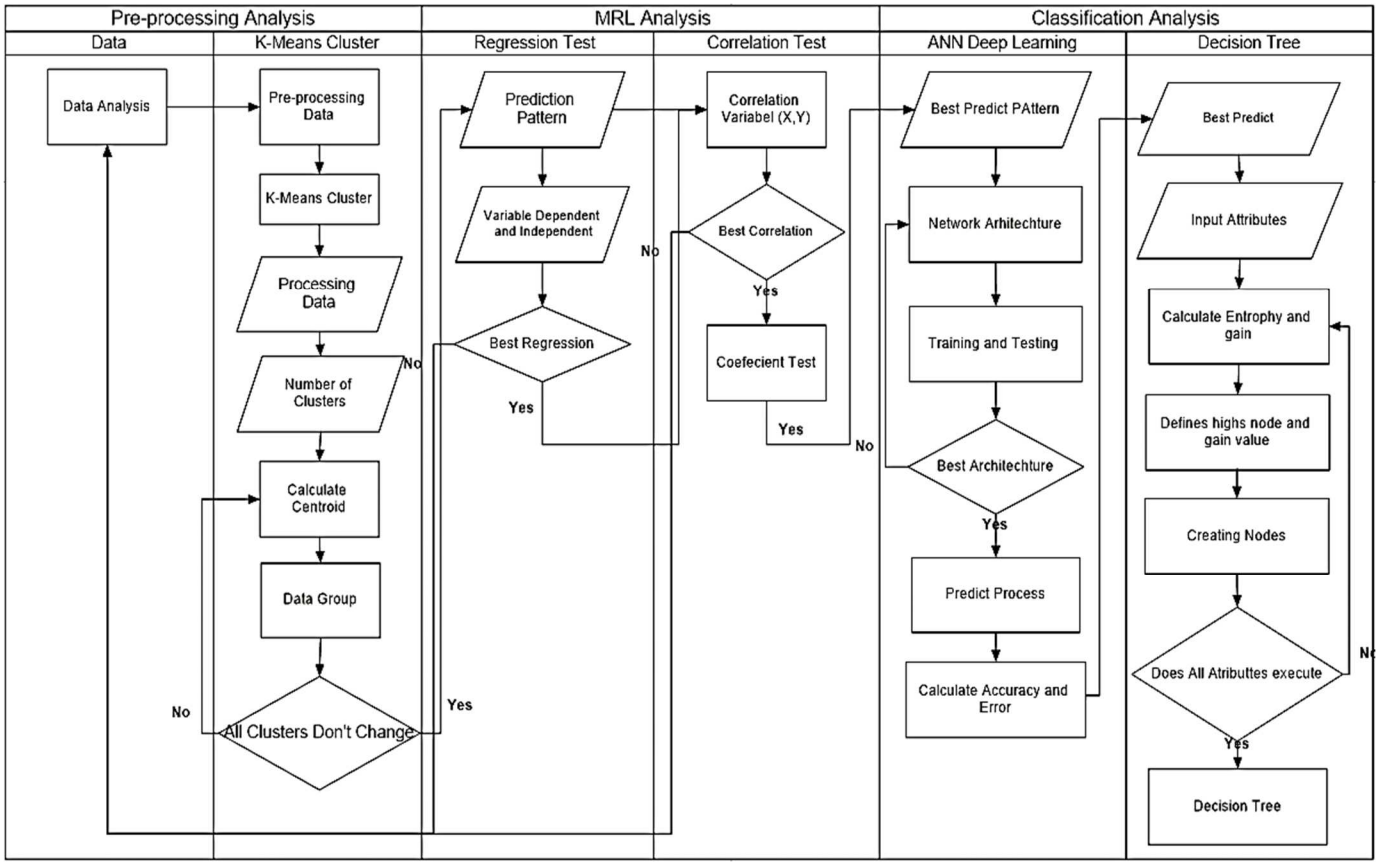

Fig. 1 Research Framework

Figure. 1 explains that the research framework can describe the stages of research consisting of several activities. This process starts with pre-processing the data and then continues with the prediction and classification process of drug demand cases that will occur in the next period. The output results are presented in the form of information of predictive output and a knowledge base in the form of a decision tree.

A. Data Pre-processing

The pre-processing process aims to analyze early in the prediction, and this is needed to ensure the prediction process provides precise and accurate output. The stages in Figure.1 show that the pre-processing process consists of 2 stages, including the clustering process to form a predictive rule pattern and the MRL analysis process to prove the accuracy

of the variables by performing regression and correlation tests. The results of this data pre-processing will be used to carry out the prediction and classification process.

B. K-Means Clustering

K-Means is a concept used to cluster data. K-Means is an algorithm for grouping and is often used in learning [24]. This algorithm is very simple and easy to present the data collection cluster process [25]. The clustering process is carried out using Formula 1 [26]:

$$\sum_{j=1}^{k}\sum_{xi\epsilon}||Xi - \mu||_2^2 \quad (1)$$

This formula explains that the mathematical calculation process in finding the optimal centroid can be determined based on the number of clusters used. To produce optimal cluster results, it is necessary to re-analyze to get results that do not change.

C. Multiple Regression Linear (MRL)

The MRL method is a mathematical technique that can carry out the process of analyzing the relationship that occurs between variables and the output [27]. The results presented can be seen based on the R. Square test in the analysis carried out [28]. The analysis output presented can see a significant relationship in the data used [29]. The MRL analysis used is seen in this Formula 2&3 [30],[31]:

$$Yi = \beta 0 + \beta 1xi1 + \beta 2\ xi\ 2 + \beta 3xi\ 3 ..... \beta p.xip \quad (2)$$

$$R^2 = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2(\sum x^2) - \sqrt{n(\sum y^2(\sum y^2)}} \quad (3)$$

The formula above is a mathematical model in linear regression with (Y) as the output and the variable (Xi) as the criteria used. The results are presented as a correlation level measured by R. Square so that the strength between each variable will be a benchmark in carrying out the next analysis process.

D. Deep Learning

Deep learning is one of the concepts of machine learning developed for processing big data [32]. Learning architecture can be designed to perform several tasks, such as classification, identification, and prediction [33]. Deep learning enables predictive serving using a neural network architecture consisting of several layers [34].

E. Artificial Neural Network (ANN)

ANN is a computational concept based on human neural networks to solve problems. The process is presented as a training process to produce a solution [35]. ANN is also a technique derived from the mechanism of the human brain system [36]. Conceptually, ANN can solve complex problems for the best solution [37], [38].

F. Classification Decision Tree

Classification Decision Tree is a method used to produce decisions from various problem-solving [39]. In the same explanation, a decision tree is a technique widely implemented in classifying data into information in the form of a decision tree [40]. The results described can be interpreted as a graph consisting of several nodes in each attribute. Each set of decision tree results has an output in the form of a cover containing one or more classes [41].

G. Pre-processing Analysis

The discussion of this research starts from the pre-processing analysis stage. The initial process is the clustering process of drug demand data. This clustering process aims to group data based on dataset variables used in the prediction process [42]. The results of the clustering process used some data can be seen in Table 1.

TABLE I
RESULT OF DRUG DEMAND CLUSTER PROCESS

Group	Promotion	Packaging	Brand	Stock	Price	Sold Out	Demand	Cluster
Limited	Not Promotion	BTL	Branded	10	57000	2	2	3
Potent Drug	Promotion	TAB	Not Branded	360	486111	360	422	1
Standard Herbal	Not Promotion	STR	Branded	50	161052	19	19	3
Potent Drug	Promotion	TAB	Not Branded	317	430599	317	322	1
Potent Drug	Not Promotion	TAB	Not Branded	1190	750420	1190	1252	1
Potent Drug	Not Promotion	TAB	Branded	1060	500	1060	1200	1
Potent Drug	Not Promotion	TAB	Branded	100	400	40	40	1
Standard Herbal	Not Promotion	PCS	Branded	300	2926111	108	108	1
Limited	Not Promotion	TAB	Branded	370	483783	370	398	3
Potent Drug	Not Promotion	TAB	Not Branded	995	482914	995	1102	1
Potent Drug	Not Promotion	TAB	Not Branded	400	246915	130	130	1
Limited	Not Promotion	BTL	Branded	10	26000	4	4	3
Potent Drug	Not Promotion	TAB	Not Branded	1000	314818	745	745	1
Potent Drug	Not Promotion	TAB	Not Branded	400	11047	420	488	1
Potent Drug	Promotion	TAB	Not Branded	1086	295395	1086	1231	1
Standard Herbal	Not Promotion	BTL	Not Branded	10	23000	4	4	2
Potent Drug	Promotion	TAB	Not Branded	1028	247276	1028	1134	1
Potent Drug	Not Promotion	BTL	Not Branded	100	5500	50	50	1
Potent Drug	Promotion	BTL	Not Branded	10	35000	6	6	1
Phytopharma Drugs	Not Promotion	STR	Not Branded	10	15000	3	3	2
Potent Drug	Not Promotion	TAB	Not Branded	400	426315	190	190	1
Limited	Not Promotion	STR	Branded	100	166304	46	46	3
Potent Drug	Not Promotion	TAB	Not Branded	300	110344	290	290	1
Standard Herbal	Not Promotion	STR	Not Branded	10	13000	6	6	2
Potent Drug	Not Promotion	TAB	Branded	500	994444	360	360	1

Group	Promotion	Packaging	Brand	Stock	Price	Sold Out	Demand	Cluster
Potent Drug	Promotion	TAB	Branded	400	781481	270	270	1
Limited	Not Promotion	PCS	Branded	10	21000	2	2	3
Limited	Not Promotion	TAB	Not Branded	200	1500	80	880	3
Limited	Not Promotion	TAB	Not Branded	615	104878	615	700	3
Potent Drug	Not Promotion	BTL	Not Branded	10	32000	10	10	1
Potent Drug	Not Promotion	TAB	Not Branded	250	400	250	311	1
Potent Drug	Not Promotion	STR	Branded	200	187619	105	105	1
Phytopharma Drugs	Promotion	PCS	Branded	100	266666	30	30	2
Phytopharma Drugs	Promotion	BTL	Branded	10	28000	1	1	2

Table. 1 explains that the results of the cluster process describe that the high cluster (C1) has 21 items, the medium cluster (C2) has 5 items, and the low cluster (C3) has 8 items. The results of this cluster will be a pattern in the prediction process [43]. Before the prediction process is carried out, the pre-processing process continues to test the resulting pattern using the MRL method. The MRL work process is able to measure the level of correlation between predictor variables (X) and prediction output (Y). The results of the correlation test process with MRL can be seen in Table 2. Table. 2 shows that the results of the R. Square test produced have a significant level of relationship between all predictor variables (X) and the output (Y) [44]. Based on the output obtained, the value of the R. Square test results obtained is 90.7%. To prove which variables can affect the prediction output, the correlation test process for each variable will be carried out [45]. The results of the correlation test can be seen in Table 3.

TABLE II
R. SQUARE TEST RESULTS

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error	Change R Square
1	.952[a]	0,907	0,882	144,86117	2,239

a. Predictors: (Constant), Sold Out, Packaging, Promotion, Price, Brand, Group,
b. Dependent Variable: Demand

TABLE III
CORRELATION TEST RESULTS

Correlations

		Group (X1)	Promotion (X2)	Packaging (X3)	Brand (X4)	Stock (X5)	Price (X6)	Sold Out (X7)	Demand (Y)
Group (X1)	Pearson Correlation	1	0,279	0,244	-0,017	-0,172	0,137	-0,177	-0,249
	Sig. (2-tailed)		0,110	0,164	0,923	0,331	0,440	0,317	0,155
	N	34	34	34	34	34	34	34	34
Promotion (X2)	Pearson Correlation	0,279	1	-0,078	-0,041	0,096	0,025	0,132	0,105
	Sig. (2-tailed)	0,110		0,659	0,816	0,590	0,890	0,458	0,556
	N	34	34	34	34	34	34	34	34
Packaging (X3)	Pearson Correlation	0,244	-0,078	1	0,313	-0,036	.377*	-0,078	-0,086
	Sig. (2-tailed)	0,164	0,659		0,072	0,841	0,028	0,660	0,627
	N	34	34	34	34	34	34	34	34
Brand (X4)	Pearson Correlation	-0,017	-0,041	0,313	1	-0,274	0,221	-0,288	-0,331
	Sig. (2-tailed)	0,923	0,816	0,072		0,117	0,210	0,098	0,056
	N	34	34	34	34	34	34	34	34
Stock (X5)	Pearson Correlation	-0,172	0,096	-0,036	-0,274	1	0,208	.977**	.927**
	Sig. (2-tailed)	0,331	0,590	0,841	0,117		0,239	0,000	0,000
	N	34	34	34	34	34	34	34	34
Price (X6)	Pearson Correlation	0,137	0,025	.377*	0,221	0,208	1	0,118	0,068
	Sig. (2-tailed)	0,440	0,890	0,028	0,210	0,239		0,505	0,001
	N	34	34	34	34	34	34	34	34
Sold Out (X7)	Pearson Correlation	-0,177	0,132	-0,078	-0,288	.977**	0,118	1	.945**
	Sig. (2-tailed)	0,317	0,458	0,660	0,098	0,000	0,505		0,000
	N	34	34	34	34	34	34	34	34
Demand (Y)	Pearson Correlation	-0,249	0,105	-0,086	-0,331	.927**	0,068	.945**	1
	Sig. (2-tailed)	0,155	0,556	0,627	0,056	0,000	0,701	0,000	
	N	34	34	34	34	34	34	34	34

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 3 shows that the correlation test results for each variable can present the level of relationship to the output (Y) [46]. Based on the results, it can be seen that a significant relationship occurs in the Stock variable (X5) of 0.000 < 0.05, the Price variable (X6) of 0.001 < 0.5, and the Sold Out variable (X7) of 0.000 < 0.05. This result proves that the variable with a significant value < 0.05 indicates that the variable is very influential on the output [47]. To re-test the results of the correlation test that has been carried out, the test process will continue to measure the coefficients of each variable on the output. The results of these tests can be presented in Table 4.

TABLE IV
COEFFICIENTS TEST RESULTS

Coefficients[a]

Model		Unstandard		Standard	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	168,514	137,540		1,225	0,231					
	Group	-50,654	34,433	-0,100	-1,471	0,153	-0,249	-0,277	-0,088	0,778	1,285
	Promotion	18,255	63,815	0,019	0,286	0,777	0,105	0,056	0,017	0,842	1,187
	Packaging	21,623	34,753	0,043	0,622	0,539	-0,086	0,121	0,037	0,738	1,354
	Brand	-63,547	57,473	-0,075	-1,106	0,279	-0,331	-0,212	-0,066	0,771	1,296
	Stock	0,168	0,364	0,149	0,461	0,000	0,927	0,090	0,028	0,034	29,074
	Price	0,0031	0,000	-0,040	-0,541	0,001	0,068	-0,105	-0,032	0,658	1,519
	Sold Out	0,861	0,358	0,765	2,407	0,000	0,945	0,427	0,144	0,035	28,198

a. Dependent Variable: Demand

Table. 4 can also prove that the coefficient value of each variable correlation with output still shows that the Stock (X5), Price (X6), and Sold Out (X7) variables can affect the output of drug demand. With the results of this analysis, the previous network pattern will be updated again to carry out the prediction process.

III. RESULTS AND DISCUSSION

Prediction of the number of drug requests that will be used a deep learning concept approach. Deep learning is suitable for use in prediction cases with large numbers of data [48]. The deep learning approach can also be used in the case of inventory stock prediction [49]. Supervised learning is one form of learning that will be applied in the case of prediction [50]. The learning process adopts the pattern formed on the results of the previous pre-processing process. The pattern will be described in a network to carry out the training and testing process. The Artificial Neural Network method with the Multi Back propagation algorithm will be used in the analysis process to produce the right prediction results in the case of drug demand. From the learning process that has been carried out, the results obtained are initially in the form of the best network for making predictions. The best prediction network can be seen in Table 5.

TABLE V
BEST PREDICTION NETWORK COMPARISON RESULTS

Architecture	Single Hidden Layer				Architecture	Single Hidden Layer			
	Accuracy	MSE	Sensivity	Validation		Accuracy	MSE	Sensivity	Validation
(3-3-1)	99,9959	0,0041	0.95697	0.9689	(3-18-1)	99,9941	0,0059	0.99591	0.9886
(3-5-1)	99,9938	0,0062	0.95338	0.98509	(3-20-1)	99,997	0,003	0.99251	0.98067
(3-8-1)	99,9959	0,0041	0.98838	0.97611	(3-23-1)	99,9879	0,0121	0.9956	0.96442
(3-10-1)	99,9968	0,0032	0.96457	0.98822	(3-25-1)	99,9987	0,0013	0.99098	0.97366
(3-13-1)	99,9976	0,0024	0.99765	0.97591	(3-28-1)	99,9885	0,0115	0.99804	0.95732
(3-15-1)	99,9971	0,0029	0.9728	0.94809	(3-30-1)	99,9906	0,0094	0.99805	0.50601
(3-3-3-1)	99,9965	0,0035	0.99576	0.99926	(3-5-3-3-1	99,9986	0,0014	0.94822	0.99611
(3-3-5-1)	99,9962	0,0038	0.99014	0.9987	(3-5-5-3-1)	99,9939	0,0061	0.98905	0.9796
(3-3-8-1)	99,9978	0,0022	0.95108	0.97556	(3-5-8-3-1)	99,9989	0,0011	0.98802	0.89879
(3-3-10-1)	99,9993	0,0007	0.91751	0.98944	(3-5-10-3-1)	99,998	0,002	0.99672	0.99495
(3-3-13-1)	99,9921	0,0079	0.98901	0.84824	(3-5-13-3-1)	99,9995	0,0005	0.9948	0.9979
(3-3-15-1)	99,9965	0,0035	0.99816	0.94516	(3-5-15-3-1)	99,9978	0,0022	0.98899	0.99386
(3-5-3-1)	99,9992	0,0008	0.89181	0.97303	(3-5-5-3-3-1)	99,9359	0,0641	0.99064	0.9407
(3-5-5-1)	99,9963	0,0037	0.95397	0.9986	(3-5-5-5-3-1)	99,9971	0,0029	0.99477	0.98566
(3-5-8-1)	99,998	0,002	0.99541	0.99096	(3-5-5-8-3-1)	99,989	0,011	0.97539	0.73608
(3-5-10-1)	99,9985	0,0015	0.9916	0.99854	(3-5-5-10-3-1)	99,9936	0,0064	0.98691	0.99674
(3-5-13-1)	99,9994	0,0006	0.98642	0.78818	(3-5-5-13-3-1)	99,998	0,002	0.99251	0.87313
(3-5-15-1)	99,9763	0,0237	0.988894	0.97953	(3-5-5-15-3-1)	99,9971	0,0029	0.96555	0.74007

Table 5 shows that the best predictive network can be seen in the network architecture 5-5-13-3-1. The architecture consists of 1 input layer (X) with 3 neurons, 3 hidden (Z) layers with 5-13-3 neurons in each layer, and 1 output layer with 1 output neuron (Y). After the best network architecture is obtained, the prediction process is carried out to obtain the output of drug demand that will occur in the next period. The results prediction process can be seen in the output graph of Fig. 2.

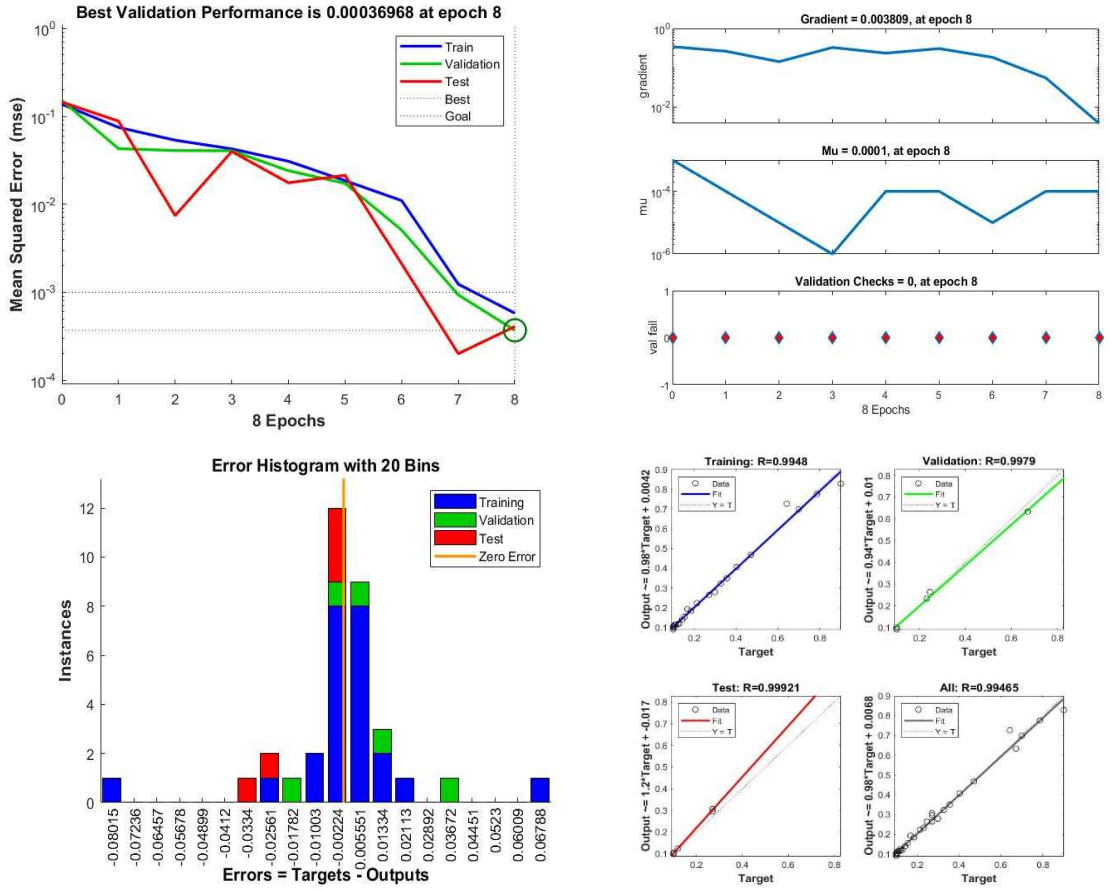

Fig. 2 Prediction Result Graph

Figure.3 shows that the error rate (MSE) resulting from the prediction process is 0.003809. This result is relatively small to represent the level of error obtained. The resulting accuracy value is very good, with a value of 99.99%. Based on these outputs, it can be stated that ANN in deep learning provides fairly accurate prediction results [51]. After the prediction process is obtained, the final process will classify the prediction pattern that has been generated using the decision tree algorithm C.45 method. The results given are in the form of a decision tree that can be used as a rule in the case of prediction [52]. The classification results can be seen in the picture of the decision tree formed in Fig. 3.

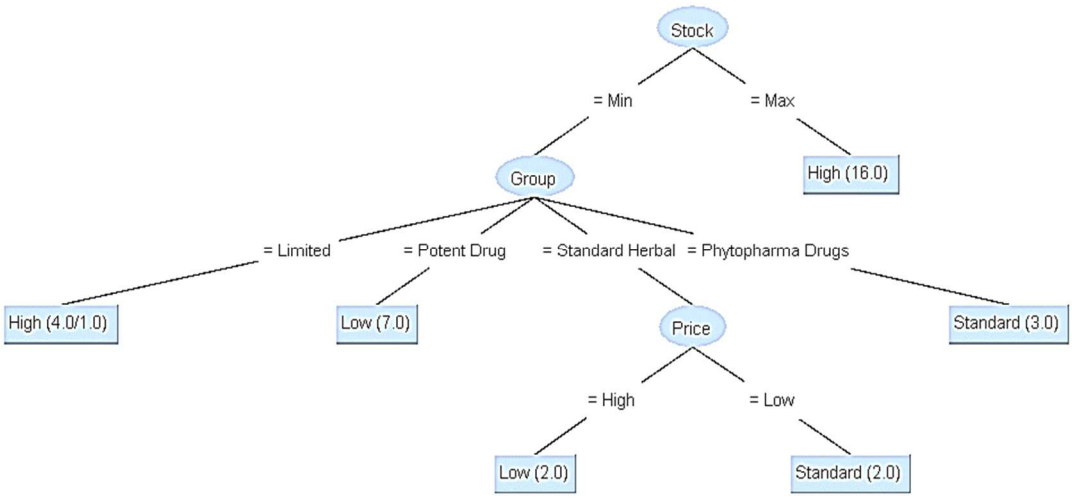

Fig. 3 Drug Predictor Decision Tree Classification Results

IV. CONCLUSION

In conclusion, the demand for drugs in the next period will experience a fairly high trend of increasing stock, price, and number of drugs sold in the previous time period. This is based on the predictive pattern that has been tested using MRL. The results presented provide a correlation rate of 90.7%. The network learning process with the concept of deep learning is also able to provide fairly good prediction results with an accuracy of 99.99%. These results indicate that the prediction analysis presented is accurate enough to see the number of drug requests that will occur. Decision tree analysis in this prediction process is able to describe a decision tree in the form of classification results on the level of drug demand. The results show that the decision tree method is very good at classifying 94.11%. This research can also provide a model developed for predicting the amount of drug demand. Thus, the prediction output can be used as an initial control to prepare for the availability of drugs in the next period.

ACKNOWLEDGMENT

The authors thank Mrs. Dr. Hj. Zerni Melmusi, SE, AK, CA, as the chairman Yayasan Perguruan Tinggi Komputer YPTK Padang, who has encouraged this research. Then Prof. Dr. Sarjon Defit, S.Kom, M.Sc as the Rector Universitas Putra Indonesia YPTK Padang that supports the completion of this research.

REFERENCES

[1] R. A. Zayed, D. Omran, and A. A. Zayed, "COVID-19 clinical and laboratory diagnosis overview," *J. Egypt. Public Health Assoc.*, vol. 96, no. 1, pp. 1–5, 2021.

[2] M. Khan, M. T. Mehran, Z. U. Haq, Z. Ullah, and S. R. Naqvi, "Applications of artificial intelligence in COVID-19 pandemic: A comprehensive review," *Expert Syst. Appl.*, p. 115695, 2021.

[3] V. Reinstadler *et al.*, "Monitoring drug consumption in Innsbruck during coronavirus disease 2019 (COVID-19) lockdown by wastewater analysis," *Sci. Total Environ.*, vol. 757, p. 144006, 2021.

[4] H. A. Badreldin and B. Atallah, "Global drug shortages due to COVID-19: impact on patient care and mitigation strategies," *Res. Soc. Adm. Pharm.*, vol. 17, no. 1, pp. 1946–1949, 2021.

[5] E. A. Gehrie, S. M. Frank, and S. M. Goobie, "Balancing supply and demand for blood during the COVID-19 pandemic." Lippincott Williams & Wilkins, 2020.

[6] R. V. G. Peddapatla *et al.*, "Process Model Approach to Predict Tablet Weight Variability for Direct Compression Formulations at Pilot and Production Scale," *Pharmaceutics*, vol. 13, no. 7, p. 1033, 2021.

[7] Y. Ma and M. Gan, "DeepAssociate: A deep learning model exploring sequential influence and history-candidate association for sequence recommendation," *Expert Syst. Appl.*, vol. 185, p. 115587, 2021.

[8] M. Gurnani, Y. Korke, P. Shah, S. Udmale, V. Sambhe, and S. Bhirud, "Forecasting of sales by using fusion of machine learning techniques," in *2017 International Conference on Data Management, Analytics and Innovation (ICDMAI)*, 2017, pp. 93–101.

[9] A. R. B. Oglu and I. I. T. Kizi, "A method for forecasting the demand for pharmaceutical products in a distributed pharmacy network based on an integrated approach using fuzzy logic and neural networks," in *International Conference on Intelligent and Fuzzy Systems*, 2020, pp. 998–1007.

[10] S. Miller, A. El-Bahrawy, M. Dittus, M. Graham, and J. Wright, "Predicting Drug Demand with Wikipedia Views: Evidence from Darknet Markets.," in *Proceedings of the web conference 2020*, 2020, pp. 2669–2675.

[11] V. S. S. Somisetti and S. H. Palla, "Efficient Clustering of Water Distribution Network Using Affinity Propagation.," *Ingénierie des Systèmes d'Information*, vol. 25, no. 4, 2020.

[12] Y. W. Lee, J. W. Choi, and E.-H. Shin, "Machine learning model for diagnostic method prediction in parasitic disease using clinical information," *Expert Syst. Appl.*, p. 115658, 2021.

[13] Y. Chang and T. Li, "Prediction Algorithm for Drugstore Consumption Members," in *2021 IEEE 6th International Conference on Big Data Analytics (ICBDA)*, 2021, pp. 214–218.

[14] A. Jozaghi *et al.*, "Multi-model streamflow prediction using conditional bias-penalized multiple linear regression," *Stoch. Environ. Res. Risk Assess.*, pp. 1–19, 2021.

[15] H. Wang, R. Czerminski, and A. C. Jamieson, "Neural networks and deep learning," in *The Machine Age of Customer Insight*, Emerald Publishing Limited, 2021.

[16] K. Abbas *et al.*, "Application of network link prediction in drug discovery," *BMC Bioinformatics*, vol. 22, no. 1, pp. 1–21, 2021.

[17] T. S. Kumar, "Data mining based marketing decision support system using hybrid machine learning algorithm," *J. Artif. Intell.*, vol. 2, no. 03, pp. 185–193, 2020.

[18] H. Wang, F. Yang, and S. Shen, "Supply Fraud Forecasting using Decision Tree Algorithm," in *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, 2021, pp. 344–347.

[19] J. Y. Ryu, H. U. Kim, and S. Y. Lee, "Deep learning improves prediction of drug–drug and drug–food interactions," *Proc. Natl. Acad. Sci.*, vol. 115, no. 18, pp. E4304–E4311, 2018.

[20] W. Deelder *et al.*, "Using deep learning to identify recent positive selection in malaria parasite sequence data," *Malar. J.*, vol. 20, no. 1, pp. 1–9, 2021.

[21] A. Jha, G. Verma, Y. Khan, Q. Mehmood, D. Rebholz-Schuhmann, and R. Sahay, "Deep convolution neural network model to predict relapse in breast cancer," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2018, pp. 351–358.

[22] Y. Pathak, P. K. Shukla, A. Tiwari, S. Stalin, and S. Singh, "Deep transfer learning based classification model for COVID-19 disease," *Irbm*, 2020.

[23] G. Chen and Z. Xu, "Usage of intelligent medical aided diagnosis system under the deep convolutional neural network in lumbar disc herniation," *Appl. Soft Comput.*, vol. 111, p. 107674, 2021.

[24] T. A. Sipkens and S. N. Rogak, "Using k-means to identify soot aggregates in transmission electron microscopy images," *J. Aerosol Sci.*, vol. 152, p. 105699, 2021.

[25] M. S. Alam *et al.*, "Automatic human brain tumor detection in MRI image using template-based K means and improved fuzzy C means clustering algorithm," *Big Data Cogn. Comput.*, vol. 3, no. 2, p. 27, 2019.

[26] Y. Hozumi, R. Wang, C. Yin, and G.-W. Wei, "UMAP-assisted K-means clustering of large-scale SARS-CoV-2 mutation datasets," *Comput. Biol. Med.*, vol. 131, p. 104264, 2021.

[27] M. Yanto, S. Sanjaya, Y. Yulasmi, D. Guswandi, and S. Arlis, "Implementation multiple linear regresion in neural network predict gold price," Indonesian Journal of Electrical Engineering and Computer Science, vol. 22, no. 3, p. 1635, Jun. 2021, doi: 10.11591/ijeecs.v22.i3.pp1635-1642.

[28] C. Zhou, "House price prediction using polynomial regression with Particle Swarm Optimization," in *Journal of Physics: Conference Series*, 2021, vol. 1802, no. 3, p. 32034.

[29] A. Rinaldi and Y. Devi, "Mathematical modeling for awareness, knowledge, and perception that influence willpower to pay tax using multiple regression," in *Journal of Physics: Conference Series*, 2021, vol. 1796, no. 1, p. 12062.

[30] A. A. Pishro, S. Zhang, D. Huang, F. Xiong, W. Li, and Q. Yang, "Application of artificial neural networks and multiple linear regression on local bond stress equation of UHPC and reinforcing steel bars," *Sci. Rep.*, vol. 11, no. 1, pp. 1–20, 2021.

[31] B. W. Seok, K. Wee, J. Park, D. A. Kumar, and N. S. Reddy, "Modeling the teacher job satisfaction by artificial neural networks," *Soft Comput.*, pp. 1–13, 2021.

[32] J. Gröhl, M. Schellenberg, K. Dreher, and L. Maier-Hein, "Deep learning for biomedical photoacoustic imaging: A review," *Photoacoustics*, p. 100241, 2021.

[33] T. M. Shah, D. P. B. Nasika, and R. Otterpohl, "Plant and Weed Identifier Robot as an Agroecological Tool Using Artificial Neural Networks for Image Identification," *Agriculture*, vol. 11, no. 3, p. 222, 2021.

[34] M. Sharma, I. Kandasamy, and V. Kandasamy, "Deep Learning for predicting neutralities in Offensive Language Identification Dataset," *Expert Syst. Appl.*, p. 115458, 2021.

[35] A. Nikolopoulos, C. Samlis, M. Zeneli, N. Nikolopoulos, S. Karellas, and P. Grammelis, "Introducing an artificial neural network energy minimization multi-scale drag scheme for fluidized particles," *Chem. Eng. Sci.*, vol. 229, p. 116013, 2021.

[36] K. Zaorska, P. Zawierucha, M. Świerczewska, D. Ostalska-Nowicka, J. Zachwieja, and M. Nowicki, "Prediction of steroid resistance and steroid dependence in nephrotic syndrome children," *J. Transl. Med.*, vol. 19, no. 1, pp. 1–13, 2021.

[37] M. Rabiej and S. Rabiej, "Application of the artificial neural network for identification of polymers based on their X-ray diffraction curves," *Comput. Mater. Sci.*, vol. 186, p. 110042, 2021.

[38] M. A. Mojid and A. B. M. Z. Hossain, "Comparative performance of multiple linear regression and artificial neural network models in estimating solute-transport parameters," *Sains Tanah-Journal Soil Sci. Agroclimatol.*, vol. 18, no. 1, pp. 27–35.

[39] G. Pappalardo, S. Cafiso, A. Di Graziano, and A. Severino, "Decision tree method to analyze the performance of lane support systems," *Sustainability*, vol. 13, no. 2, p. 846, 2021.

[40] Y. Lian, J. Chen, Z. Guan, and J. Song, "Development of a monitoring system for grain loss of paddy rice based on a decision tree algorithm," *Int. J. Agric. Biol. Eng.*, vol. 14, no. 1, pp. 224–229, 2021.

[41] P. Ducange, F. Marcelloni, and R. Pecori, "Fuzzy hoeffding decision tree for data stream classification," *Int. J. Comput. Intell. Syst.*, vol. 14, no. 1, pp. 946–964, 2021.

[42] V. Malik, Y. Kalakoti, and D. Sundar, "Deep learning assisted multi-omics integration for survival and drug-response prediction in breast cancer," *BMC Genomics*, vol. 22, no. 1, pp. 1–11, 2021.

[43] S. Arlis and S. Defit, "Machine Learning Algorithms for Predicting the Spread of Covid–19 in Indonesia," 2021.

[44] S. L. Makali *et al.*, "Comparative analysis of the health status of the population in six health zones in South Kivu: a cross-sectional population study using the WHODAS," *Confl. Health*, vol. 15, no. 1, pp. 1–11, 2021.

[45] F. S. Neves, "Correlation of the rise and fall in COVID-19 cases with the social isolation index and early outpatient treatment with hydroxychloroquine and chloroquine in the state of Santa Catarina, southern Brazil: A retrospective analysis," *Travel Med. Infect. Dis.*, vol. 41, p. 102005, 2021.

[46] C. Xu, X. Chen, and L. Zhang, "Predicting river dissolved oxygen time series based on stand-alone models and hybrid wavelet-based models," *J. Environ. Manage.*, vol. 295, p. 113085, 2021.

[47] S. Pokharel, P. Sah, and D. Ganta, "Improved Prediction of Total Energy Consumption and Feature Analysis in Electric Vehicles Using Machine Learning and Shapley Additive Explanations Method," *World Electr. Veh. J.*, vol. 12, no. 3, p. 94, 2021.

[48] Z. Yu, S. Ye, Y. Sun, H. Zhao, and X.-Q. Feng, "Deep learning method for predicting the mechanical properties of aluminum alloys with small data sets," *Mater. Today Commun.*, p. 102570, 2021.

[49] P. R. Srivastava, Z. J. Zhang, and P. Eachempati, "Deep Neural Network and Time Series Approach for Finance Systems: Predicting the Movement of the Indian Stock Market," *J. Organ. End User Comput.*, vol. 33, no. 5, pp. 204–226, 2021.

[50] M. Mesgarpour *et al.*, "Prediction of the spread of Corona-virus carrying droplets in a bus-A computational based artificial intelligence approach," *J. Hazard. Mater.*, vol. 413, p. 125358, 2021.

[51] W.-J. Wang, Y. Tang, J. Xiong, and Y.-C. Zhang, "Stock market index prediction based on reservoir computing models," *Expert Syst. Appl.*, vol. 178, p. 115022, 2021.

[52] S. Asvapoositkul and R. Preece, "Decision tree-based prediction model for small signal stability and generation-rescheduling preventive control," *Electr. Power Syst. Res.*, vol. 196, p. 107200, 2021.