Hospital Length of Stay Prediction based on Patient Examination Using General features

Rabiatul Adawiyah, Tessy Badriyah, Iwan Syarif

Informatics Engineering Department, Politeknik Elektronika Negeri Surabaya
Email: wilyanyatul@gmail.com, *tessy@pens.ac.id, iwanarif@pens.ac.id

Abstract

As of the year 2020, Indonesia has the fourth most populous country in the world. With Indonesia's population expected to continuously grow, the increase in provision of healthcare needs to match its steady population growth. Hospitals are central in providing healthcare to the general masses, especially for patients requiring medical attention for an extended period of time. Length of Stay (LOS), or inpatient treatment, covers various treatments that are offered by hospitals, such as medical examination, diagnosis, treatment, and rehabilitation. Generally, hospitals determine the LOS by calculating the difference between the number of admissions and the number of discharges. However, this procedure is shown to be unproductive for some hospitals. A cost-effective way to improve the productivity of hospital is to utilize Information Technology (IT). In this paper, we create a system for predicting LOS using Neural Network (NN) using a sample of 3055 subjects, consisting of 30 input attributes and 1 output attribute. The NN default parameter experiment and parameter optimization with grid search as well as random search were carried out. Our results show that parameter optimization using the grid search technique give the highest performance results with an accuracy of 94.7403% on parameters with a value of Epoch 50, hidden unit 52, batch size 4000, Adam optimizer, and linear activation. Our designated system can be utilised by hospitals in improving their effectiveness and efficiency, owing to better prediction of LOS and better visualization of LOS done by web visualization.

Keywords: Predict, length of stay, Neural Network, Data, Hospital.

1. INTRODUCTION

The world population is expected to reach 8.5 billion by 2030 [1]. With a population of 270.203.917 people in 2020, Indonesia has the fourth largest population in the world [2]. This is influenced by several factors, including the increase in the birth rate, the uneven geographical distribution of the population, and a steady fall in the mortality rate. These factors together contribute to an increase in the number of hospital patients, including outpatients, inpatients and emergency rooms. Therefore, hospitals or health

centers must provide adequate facilities and services [3], especially for inpatients. LOS or length of stay is inpatient services provided by the hospital to patients, including medical examination, diagnosis, treatment, and rehabilitation services.

LOS has a significant potential in terms of revenue value. On the other hand, it entails a significant maintenance costs [4]. LOS is determined by summing the days from when a patient is admitted to the hospital until the day the same patient leaves (either a patient is discharged from a hospital or died at the hospital). The number of days is obtained from medical records kept by the hospital's or local clinic's administration team.

This method of keeping records is certainly unproductive for hospitals and local clinics. Applying the latest information technology is one of the ways to manage the productivity of a hospital or a local clinic.

Previous research has shown that it is important to predict LOS without sacrificing the quality of existing health services [5] using a statistical or predictive method approach based on artificial intelligence (AI), which has proliferated in recent years. The advantage of AI is that it provides cost-effective health services and helps hospitals to manage inpatients effectively and efficiently [6]. Based on this research, we aim to build a LOS prediction system using one of the types of algorithms. The algorithm we use is a machine learning algorithm, which is a neural network algorithm also known as the neural network.

The use of Neural Networks in the health sector has been expanding rapidly. There are several studies that combine neural networks with machine learning in identification. It has been shown that the neural network algorithm is an algorithm that gives accurate prediction [7]. Other studies have also shown that the predicted neural network results are not significantly different from the actual average LOS [8]. Hence, the NN algorithm is chosen to build a prediction system in this study, which can assist hospitals in planning and allocating their limited resources [9].

This study aims to build a *length of stay* prediction system using the *neural network*. The data used in this study are field data from the UPT Puskesmas Arjasa Kangean, Sumenep Regency, East Java, Indonesia. The data contains 22 attributes or features with a total sample of 3056.

2. RELATED WORK

E. Kutafina et al. [10] This study examines the efficient planning of hospital beds and how to manage the conditions needed to minimize hospital costs, by predicting the number of available hospital beds in a few months. The data used come from administrative records and school holidays, totaling around 60 days. Based on existing data and the data generated by prediction, the authors obtain the mean absolute error percentage (MAPE) of 6.24% calculated from 8 validation sets.

P. F. J. Tsai *et al.* [7]. This study predicts the earliest possible length of stay to assist in monitoring the quality of inpatient services. This helps in

managing hospital administration by using 3 main diagnoses in data collection including, Coronary atherosclerosis (CAS), heart failure (HF), and acute myocardial infarction (AMI). Based on the experiments conducted, the accuracy results obtained were 88.07%-89.95% for CAS patients at the discharge state88.31%-91.53% for the admission stage. Meanwhile, for AMI or HF patients, an accuracy of 64.12%-66.78% was obtained at the discharge stage and 63.69%-67.47% at the admission stage.

A. Yin [11]. This study carries out the length of stay prediction to alleviate the strain on hospital's limited resources, help optimize hospital operations and reduce treatment costs by using MLP in optimizing parameters for LOS prediction. The steps taken by the authors were as follows: preprocessing the data, training and testing the MLP (Multilayer perceptron), then adjusting the neural network parameters to reach the desired level of accuracy. The feature selection stage involves selecting features that have better performance in obtaining accuracy. Based on the steps carried out, the best accuracy value was 100%.

3. ORIGINALITY

The originality of this research lies in the application of LOS prediction in the context of hospitals and local clinics in Indonesia. The use of LOS prediction and the web visualisation that comes with it could help hospitals and local clinics in managing their limited resources better, using the data that they have on hand.

Before pre-processing, the data a total of 3055 subjects, with 22 attributes or features with 1 target attribute used as the outcome variable. The features used were taken from demographic data, obtained from the hospital's or local clinic's administrative department, which keep records of patients' detailed characteristics. This data includes age, gender, address, profession, source of funds, inpatient class, and LOS.

Clinical data are data obtained from patients' general examination. The data includes blood pressure, respiration rate, pulse, temperature.

Laboratory data are data obtained from the results of patients' laboratory examinations. The data includes hemoglobin, leukocytes, albumin, uric acid, blood sugar, thyroid, creatinine, urea, SGOT (Serum Glutamic Oxaloacetic Transaminase), SGPT (Serum Glutamic Pyruvate Transaminase).

It is known that there are 6 nominal data attributes and 16 numeric data attributes. Based on these data, it is necessary to transform data from nominal into binary using a one-hot-encoding approach. Encoding approach was chosen because it adds attribute columns or features according to the given data. From this approach, 1724 columns are generated.

Value 1 indicates that the row contains the data in question, while value 0 indicates otherwise. Based on this, a feature selection process is required in order to obtain the best features (these are features with a high degree of accuracy). The feature selection process uses the random forest algorithm. This process will take 30 best feature data selected from other features. The

use of random forest in feature selection is carried out to divide each node (vertices above) to produce a better error rate [12], The system built can be used to help hospitals, and more importantly encourage improvements in hospital's efficiency. Moreover, the web visualization system allows user to see the LOS visual duration with ease.

4. SYSTEM DESIGN

Our LOS prediction system based on general features using neural networks consists of the following stages: Data Pre-processing, Classification algorithm, Classification model, Length of Stay Analysis, Visualization of web app predictions, and Duration of LOS.
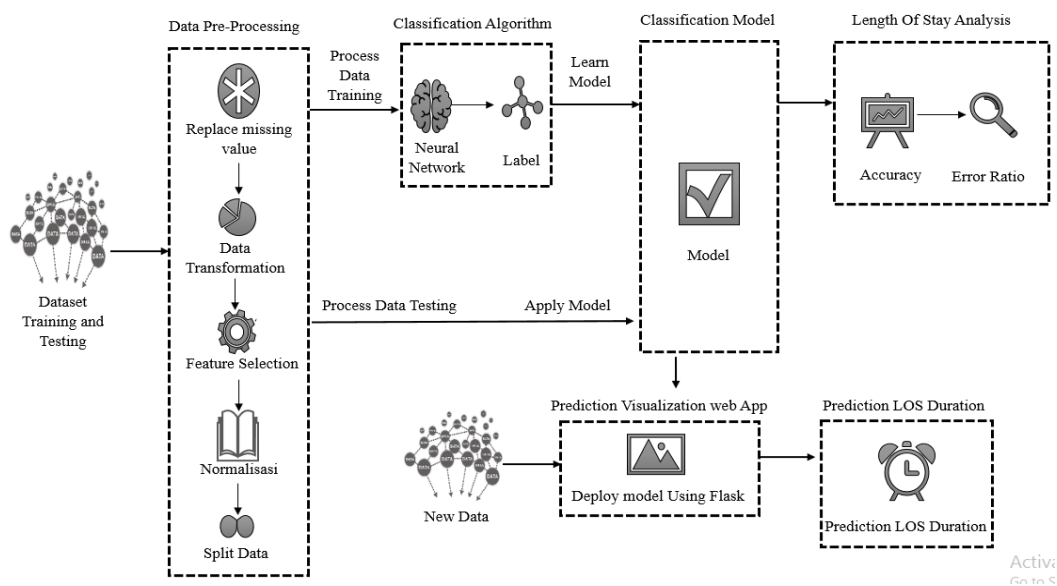
**Figure 1:** The system design of our LOS prediction system.

In this study, we build a system specifically catered for predicting LOS in hospitals and local clinics. We start from the attributes that are relevant for the hospital environment as a healthcare provider. The attributes collected in the prediction process were patient records spanning 3 years, starting from 2018 to 2020. The patients have been admitted to the hospital as inpatients due to various causes. Therefore, we propose the use of predictive attributes to be selected based on common diseases that are mostly present with inpatients. The system will then provide an overview of LOS for each disease and treatments.

Data analysis is needed to utilize the existing data collected by hospitals and local clinics to better predict future LOS. This system is coupled with a web visualization that will enable healthcare staff to easily understand the information generated by our OLS predicted system.

4.1 Data Pre-Processing

In the pre-processing stage, we manipulate the dataset before inputting the data into the model, so that the data will be compatible with the library used. If there exist missing values in the column attributes, the function 'ffill' in Python, short for "forward fill" will be used to fill in the columns with either forward or top index.

Next, we proceed by transforming the data from nominal to binary with the aid of one-hot-encoding approach. This transformation step is crucial as the neural network does not function well with nominal data. The one-hot-encoding approach will add a feature according to the given data. The value 1 will show that the data is present in a given row, and zero if otherwise. Then, the feature selection process will be carried out using Random Forest (RF) algorithm. This process will choose the 30 best feature data from the existing features. Feature selection will be used to obtain features that are relevant for data accuracy, as seen in Figure 2.
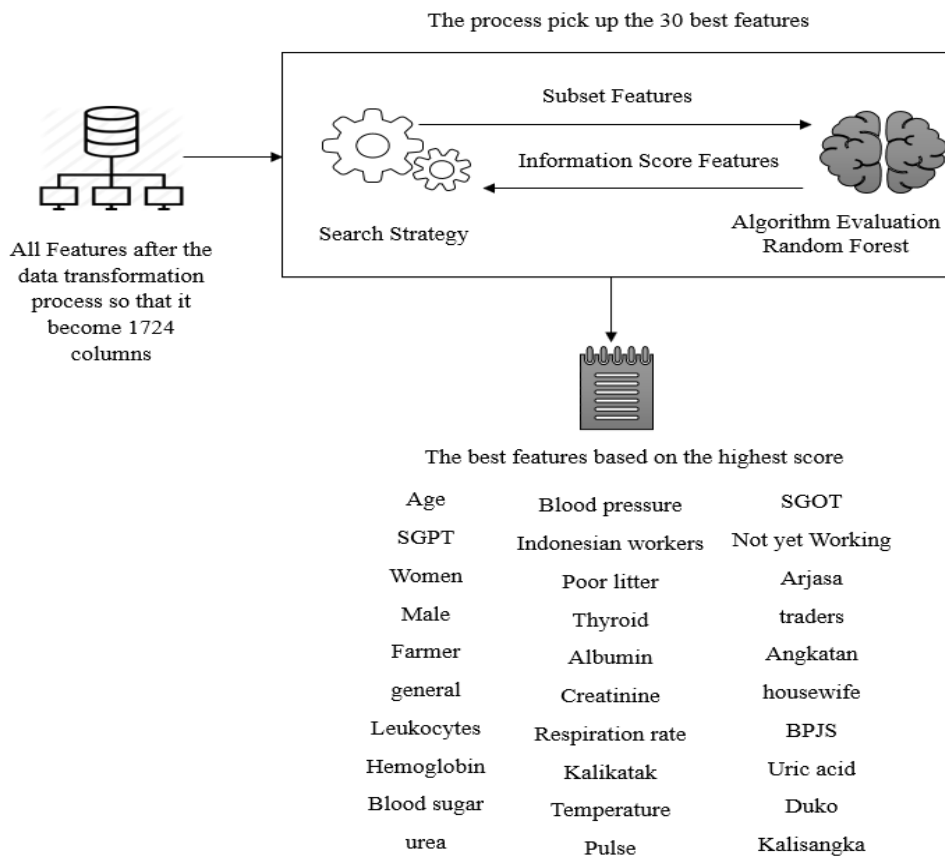
Figure 2: Feature selection process

The data is then normalized by transforming several variables into having similar range of values, making statistical analysis easier. Feature selection is a method for determining features are relevant for increasing the

accuracy of data. The selection of data features is done using the Random Forest (RF) algorithm. This process will retrieve 30 best feature data selected from the existing features. The use of RF on random feature selection is done to divide each node (vertices above) in order to produce a better error rate [12].The number of trees formed affect the accuracy of the classification results. The more trees there are, the more accurate the selection will be. In addition, RF can handle large variable input, balancing errors in unbalanced datasets [13].

4.2 Classification Algorithm

The classification stage involves using a specific algorithm in the predictive classification process called neural network.
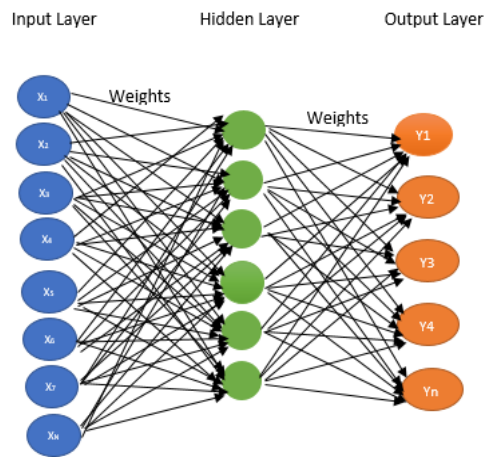

Figure 3: Neural network design

Figure 3 is an architecture or structure of the ANN. It lays out the arrangement of the component layers and neurons in input, hidden and output. These are connected to weights, activation functions and learning functions. The ANN adopts the thinking mechanism resembling the human brain, and it could either accept, process, or tolerate errors. With the details of the network architecture, there are 30 input neurons connected with weights that are initialized randomly into 1 hidden layer containing 52 neurons. The input neurons that are connected to the weights are then processed using activation. The activations used are Relu, tanh and linear. Furthermore, the data from the hidden layer results are connected with the weights contained in the hidden layer, resulting in 25 output neurons done by the prediction classification. Then, we compare the results obtained with the class or target data. The target data has a range of values from day 0 to 25 days of care in the hospital. We also obtain the error rate from this process. This process applies to all parameter setting experiments performed on default parameters, grid search parameters and random search parameters.

4.3 Classification Model

The model stages are used to recognize, study, and determine the relationship between attributes and attribute labels and provide instructions in evaluating the model. The detailed Model Training process can be seen in Figure 3. A training model is formed, which can then examine new data and perform LOS prediction classification. This results in a LOS predicted value e.g. 4 days, 5 days, and so on.
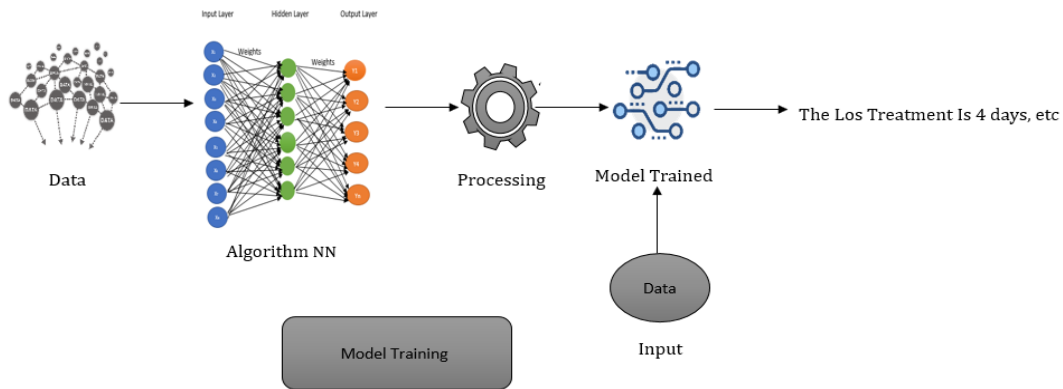
Figure 4: Model Training Process

4.4 Length of Stay Analysis

This stage evaluates models that have succeeded in obtaining the prediction accuracy and error ratio. Prediction accuracy is calculated based on the evaluation results from the Confusion Matrix. These criteria are presented below in Table 1.

Table 1: Confusion Matrix

	Positive Actual	Negative Actual
Positive Predicted	TP (True Positif)	FP (False Positif)
Negative Predicted	FN (False Negative)	TN (True Negative)

There are four terms contained in the Confusion Matrix are: TP, TN, FP, FN. TP (True Positive) is the number of positive data classified correctly by the system. TN (True Negative) is the amount of negative data classified correctly by the system. Meanwhile, FN (False Negative) is the number of positive data detected as negative data by the system. FP (False Positive) is the amount of negative data detected as positive data by the system [14].

After making a discussion matrix, the evaluation matrix in general can be seen as follows:

Accuracy represents the percentage of applications that are correctly classified as compared to the total number in the system.

$$Accuracy = \frac{TP+FN}{TP+TN+FP+FN} \times 100\% \qquad (1)$$

Error rate measures the errors detected in the system.

$$Err = \frac{FP+FN}{FP+TN+TP+FN} \times 100\% \qquad (2)$$

5. EXPERIEMENT AND ANALYSIS

We use a neural network algorithm to classify LOS into a day range between day 0 and day 25. Neural network is an algorithm that mimics the problem-solving mechanism of the human brain. NN consists of 3 layers of nodes that are interconnected between the input layer, hidden layer, and output layer. This study applies the default NN parameter, where the value of the parameter is pre-determined. Hence, we do not need to carry out an optimization process to find the optimal value combinations. Then, we optimize the NN parameter using Grid Search and random search to obtain the accuracy rates. NN performs a search on all combinations of 1-5000 batch size and select the best combination from 1-5000 and determine how parameters are initialized and updated during optimization. We test for different accuracy levels using the value of epoch, K-fold 10, hidden unit, batch size, optimize, activation for default parameters. Epoch, hidden unit, batch size, optimize, activation for grid search and random search optimization.

Table 2 shows the use of three tests between the default NN parameters and NN optimization. The test includes three activations, namely tanh, linear, relu with epoch 200 for default NN and optimization of NN, batch size 5 for default parameters, 1-5000 for optimization of NN parameters. Hence, Table 2 includes tests with different numbers of trials for the same test parameters. The results in Table 2 are the superior results of the 3 NN default tests and NN optimization. The result obtained for activation linear, epoch 200, K-fold 10, batch size 5 have high accuracy. We record an accuracy of 94.66% for default NN, linear activation, epoch 50, batch size 4000 with 94.7403% accuracy for Grid Search optimization, and linear activation, epoch 1, batch size 2000 with 94.7369% accuracy for Random Search optimization. Hence, it can be seen that the NN algorithm with Grid search optimization has better accuracy compared to the default NN parameter experiment and the Random Search optimization for LOS prediction classification.

The recapitulation of the accuracy test results with default parameters and optimization of the grid search and random search parameters are shown in table 2. Table 2 shows the superior test of the 3 activation tests, namely relu, tanh and linear.

Table 2: Experiment recapitulation NN default parameters, grid search, and random search parameter optimization

Parameter	Testing	Number of trials	Accuracy
Default Parameter	Epoch	200	94.66%
	K-fold	10	
	Hidden Unit	52	
	batch size	5	
	Optimize	Adam	
	Activation	Linear	
Hyperparameter Grid Search	Epoch	50	94.7403%
	Hidden Unit	52	
	batch size	4000	
	Optimize	Adam	
	Activation	Linear	
Hyperparameter Random Search	Epoch	1	94.7369%
	Hidden Unit	52	
	batch size	2000	
	Optimize	Adam	
	Activation	Linear	

Figure 5 shows the results of the block diagram of the superior process based on the experimental stages that have been carried out, namely optimization using grid search with Epoch 50, hidden unit 52, batch size 4000, optimize Adam, linear activation, the results obtained have 94.7403% accuracy. It can be seen that the difference in the block diagram prediction result is lower than the actual LOS. This happens because there is a fairly high error generated.
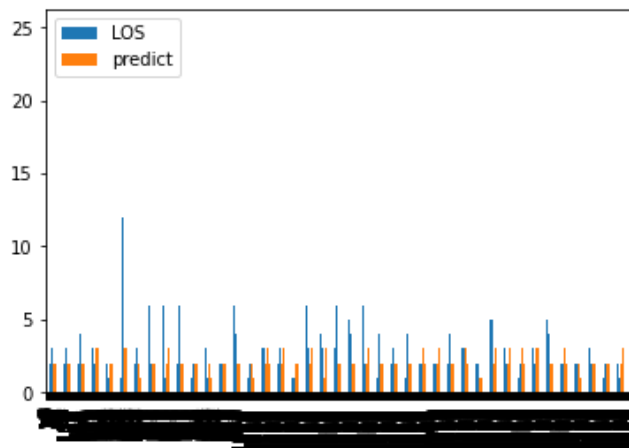

Figure 5: Block diagram actual LOS and predict LOS.

Figure 6 shows the movement of the validation training score and cross-validation score curves. It can be seen that when the training process is carried out, there is an imbalance in the error value which is significantly lower than the cross-validation process. However, when it comes to the number of epochs 1250, the error value became stable and is in the direction of the error when the training score process is carried out, with the cross validation score up to the epoch to 2500.
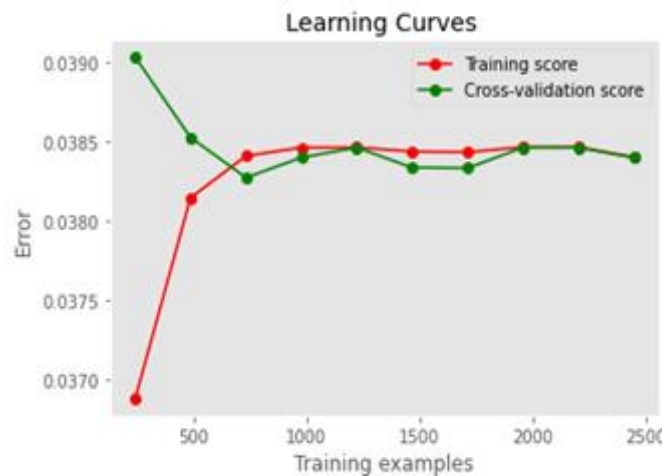
Figure 6: Learning Curves

Figure 7 shows the implementation of the predictions that have been made using the flask framework as a visualization display and a hard library for the data mining process. The front view is the main website page, namely the dashboard which briefly explains the meaning of the length of stay. Besides this explanation, there is an explanation of which features are used in predictive input.
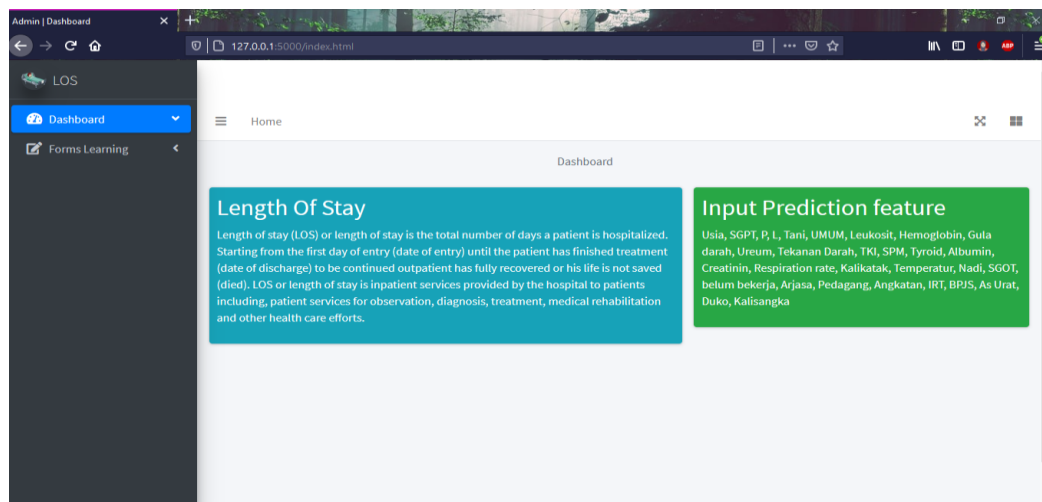
Figure 7: Dashboard

Figure 8 shows the learning form page, namely the prediction menu. This page shows the input form. There exist 30 input attributes, where each input is carried out in value according to the patients symptoms except for attributes P, L, farmer, general, TKI, SPM, Kalikatak, belum bekerja, Arjasa, traders, Angkatan, IRT, BPJS, Duko and Kalisangka. This attribute is done by inputting binary values. The value 1 indicates that the rows which contain the data in question, while the value 0 indicates otherwise. Outside of the input form, there is a button with the command predict. This button will perform the process so that it will display the prediction output.
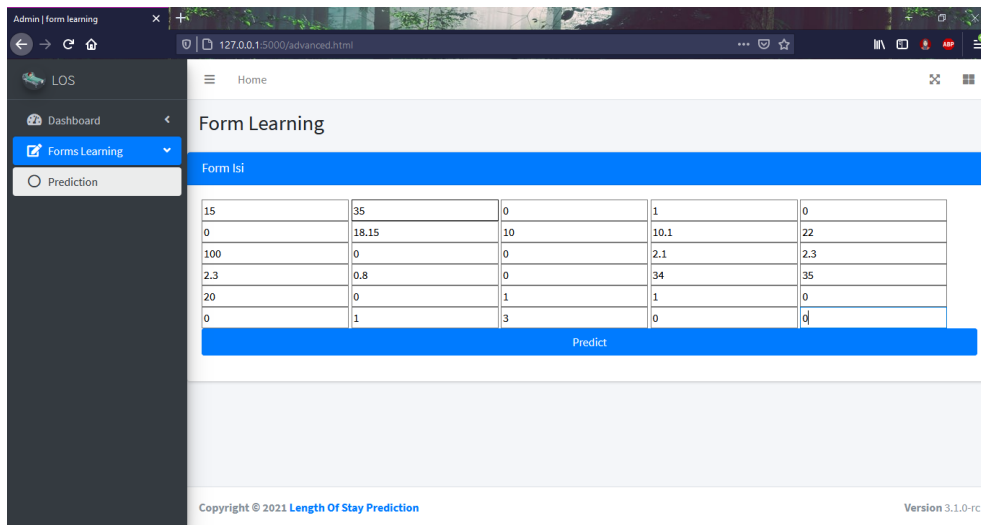

Figure 8: Form Learning Input Predict

Figure 9 shows the prediction results of the length of stay. Based on the input attributes of patients with symptoms, the LOS prediction result is 4, which means that the patient is predicted to undergo treatment for 4 days in the hospital.
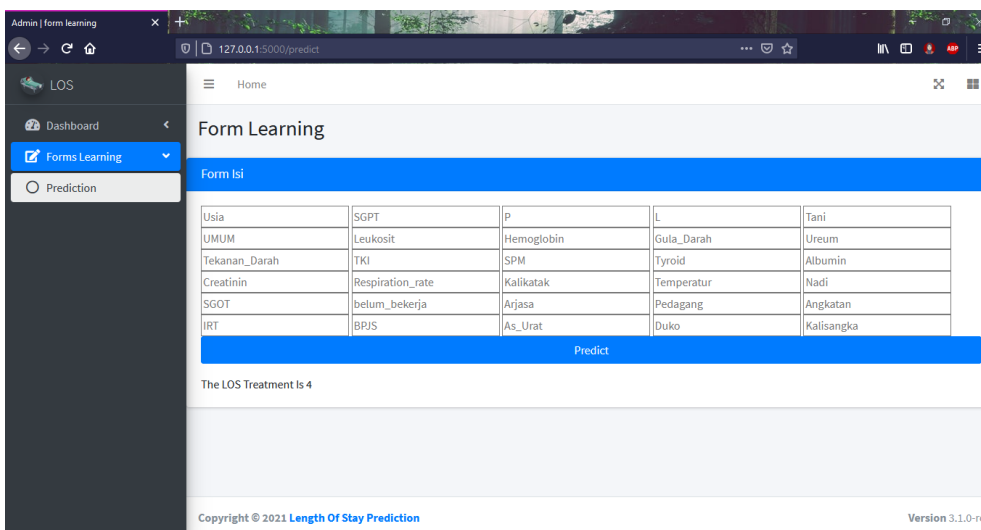

Figure 9: Learning Output Predict

6. CONCLUSION

In this study, we built a system that predicts the length of stay in the hospital. Neural network algorithms are used in predictive classification. Attribute settings are limited to attributes that have a high degree of accuracy, namely features that could help increase the accuracy of data. Based on the feature selection process, we obtained 30 input attributes and 1 target attribute or class as a reference for prediction. Our target data or class has a range of values from day 0 to 25 days of hospitalization. The default NN parameter experiment, parameter optimization with grid search and random search were carried out. Based on the accuracy results obtained, grid search optimization is the superior optimization of the experiment in this study with the value of Epoch 50, hidden unit 52, batch size 4000 and Adam optimizer with linear activation. We obtained an accuracy of 94.7403%. Based on the research that has been done, the neural network algorithm performs classification predictions well, as can be seen from the accuracy value obtained.

Learning from previous research, data has a significant effect on the algorithm, because the neural network cannot perform the classification process properly if nominal value exists in the data, hence data transformation and parameter settings are required. So far, existing research has not fully explained the ways that could be done to improve accuracy and how exactly NN reads data. Our paper attempts to lay out the process and steps taken in NN data processing. The prediction system and web visualization that comes with it can be used to help hospitals manage their limited resources better, ultimately increasing their effectiveness and efficiency in patient care.

7. FUTURE WORK

Unfortunately, this system still produces a fairly high error value in the predicting process, so that the data generated is not all correct. Further research is necessary to improve the accuracy of the model. This can be done by parameter setting, adding the number of attributes and the amount of data that used, and the web display can be improved by using another process.

REFERENCES

[1]   M. A. Umar, "**Bonus Demografi Sebagai Peluang Dan Tantangan Di Era Otonomi Daerah**," *Genta Mulia*, vol. 8, no. 2, pp. 90–99, 2020.
[2]   Badan Pusat Statistik, "**Badan Pusat Statistik.**" pp. 335–58, 2017, doi: 10.1055/s-2008-1040325.
[3]   W. M. Baihaqi, M. Dianingrum, and K. A. N. Ramadhan, "**Regresi Linier Sederhana Untuk Memprediksi Kunjungan Pasien di Rumah Sakit Berdasarkan Jenis Layanan dan Umur Pasien**," *J. SIMETRIS*, vol. 10, no. 2, pp. 671–680, 2019.
[4]   Menteri Kesehatan, **"BAB II TINJAUAN PUSTAKA,"** *Keputusan Menteri Kesehat.*, vol. 18, pp. 19–28, 2003.

[5]   A. Cedars, L. Benjamin, S. V. Burns, E. Novak, and A. Amin, "**Clinical predictors of length of stay in adults with congenital heart disease**," *Heart*, vol. 103, no. 16, pp. 1258–1263, 2017, doi: 10.1136/heartjnl-2016-310841.

[6]   C. S. Yang, C. P. Wei, C. C. Yuan, and J. Y. Schoung, "**Predicting the length of hospital stay of burn patients: Comparisons of prediction accuracy among different clinical stages**," *Decis. Support Syst.*, vol. 50, no. 1, pp. 325–335, 2010, doi: 10.1016/j.dss.2010.09.001.

[7]   P. F. J. Tsai *et al.*, "**Length of Hospital Stay Prediction at the Admission Stage for Cardiology Patients Using Artificial Neural Network**," *J. Healthc. Eng.*, vol. 2016, 2016, doi: 10.1155/2016/7035463.

[8]   C. Gholipour, F. Rahim, A. Fakhree, and B. Ziapour, "**Using an artificial neural networks (ANNS) model for prediction of intensive care unit (ICU) outcome and length of stay at hospital in traumatic patients**," *J. Clin. Diagnostic Res.*, vol. 9, no. 4, pp. 19–23, 2015, doi: 10.7860/JCDR/2015/9467.5828.

[9]   N. Caetano, P. Cortez, and R. M. S. Laureano, "**Using data mining for prediction of hospital length of stay: An application of the CRISP-DM methodology**," *Lect. Notes Bus. Inf. Process.*, vol. 227, pp. 149–166, 2015, doi: 10.1007/978-3-319-22348-3_9.

[10]   E. Kutafina, I. Bechtold, K. Kabino, and S. M. Jonas, "**Recursive neural networks in hospital bed occupancy forecasting**," vol. 1, pp. 1–10, 2019.

[11]   A. Yin, "**Predicting Hospital Length of Stay Using Multilayer Perceptron Neural Networks Abstract**," no. August, 2019.

[12]   Y. Freund and R. E. Schapire, "**Experiments with a New Boosting Algorithm,**" *Proc. 13th Int. Conf. Mach. Learn.*, pp. 148–156, 1996, doi: 10.1.1.133.1040.

[13]   D. Fransiska Amalia Kurniawan, "**Analisis dan Implementasi Random Forest dan Regression Tree (CART) Untuk Klasifikasi pada Misuse Intrussion Detection System**," *Fak. Tek. Inform.*, no. Data Mining, pp. 1–7, 2011.

[14]   V. Kotu and B. Deshpande, "**Model Evaluation**," *Predict. Anal. Data Min.*, pp. 257–273, 2015, doi: 10.1016/b978-0-12-801460-8.00008-2.