Implementation of the Neural Network (NN) Algorithm in Analysis of Student Class Increment Data Based on Report Card Value.

Alimuddin*, and Muhammad Saiful

Universitas Hamzanwadi, Lombok, Indonesia

* alimuddin76@gmail.com

Abstract. Formal education is education in schools that takes place regularly and gradually follows clear and strict conditions, the purpose of which is to add insight or knowledge in a person and enrich character, and prepare someone to be able and skilled in a particular field. 12 years formal education that must be obtained by everyone starting from elementary, middle and middle school, in general in the process of class improvement, several stages must be passed by students to be able to proceed to higher grade levels, including good behavior, level attendance of at least 70% and the value must be above the KKM (Minimum Settlement Criteria). KKM is the level of achievement of basic competencies that must be achieved by students. So that in the analysis of determining student achievement classes, data mining techniques are used by applying the Neural Network (NN) algorithm to facilitate the parties involved in analyzing the level of increase in student class and can find out the results of how high the accuracy of the algorithm applied. In this experiment, testing was carried out 3 times with different K-Fold Validations on the cross-validation operator. K-Fold Validation functions to divide the amount of training data and test the data on the tested data, the result of the accuracy that has been tested is 99.26%.

Key words: Analysis of grade increase, Using Neural Network

1. Introduction.

Formal education is education in schools that takes place regularly and stratified following clear and strict conditions, the aim is to add insight or knowledge in a person and enrich character, and prepare someone to be able and skilled in a particular field[1][2]. 12 years of formal education that must be obtained by all groups ranging from elementary, junior high and high school, in general in the process of grade promotion, several stages must be passed by students to progress to higher grade levels, including good behavior, a minimum attendance rate of 70% and the value must be above the KKM (Minimum Completeness Criteria). KKM is the level of achievement of basic competencies that must be achieved by students and is made as a teacher's reference in assessing the competencies of their students. Before entering the report card, the value of each student will be recapitulated, to check whether there are students who need remedial, and after remedial, the value taken is the highest. In this case, in determining students who are worthy and not worthy of going up class are analyzed appropriately and following the criteria for determining grade advancement, the author tries to apply data processing by utilizing data mining techniques, where data mining is a process to obtain useful information from a large database warehouse[3]. One data mining algorithm used by the author in this application is the Neural Network (NN) algorithm. Neural Network (NN) is an information processing system that is designed to imitate the workings of the human brain in solving problems[4]. One of the
advantages of this Neural Network Algorithm is the ability to compute in parallel so that the process is shorter[5]. The weakness is the length of the training process that may occur in a very long time for a very large amount of data[6][7].

2. Methode.

For the completion of Data Mining in this study, the method used is the CRISP-DM framework. The following stages[8]:

![CRISP-DM Framework](image)

**Figure 1. CRISP-DM Framework**

2.1. Neural Network (NN)

Neural network (NN) is an information processing system that tries to mimic the performance of the human brain and is inspired by the neural network of living things[9][10]. Neuron is an information processing unit in a neural network consisting of:

- Set synapses or connecting links that are marked with weights.
- Adders to add weighted input signals called linear combinations.
- Activation function, to limit the amount of output from a neural.

2.2. Neural Network (NN) Model

Like the human brain, nerve tissue also consists of several neurons, and there are connections between these neurons. Figure 24 shows the structure of neurons in which neurons will transform the information received through the output connection to other neurons. In the neural network, this relationship is known as weight. The information is stored at a certain value in the weight. both or maybe more to get data redundancy. This is processed by a propagation function that will add up the values of all future weights. The results of this addition are then compared with an information called input sent to neurons with a certain arrival weight. Enter a certain threshold through the activation function of each neuron. Activation function Is a function of processing the sum of input data into output data[11].

3. Discussion.

3.1. The rate of increase in student class.

Determination of class increments is calculated based on the achievement of odd and even semester learning outcomes in one school year, with the following conditions:

- If the learning outcomes in the odd and even semester are complete, then the subject is declared complete.
- If the learning outcomes in the odd and even semester are not completed, then the subject is declared incomplete.
• If the learning outcomes of a subject in one of the odd and even semester are not completed, then the completeness of the subject must be calculated in that subject.

3.2. Data Required.

3.3. Initial Data Processing

The total amount of initial data in class X of SMAN 1 Suela obtained was 131 records, with a total of 24 attributes, including No, Name, Address, KKM, Attendance, Behavior, Average Value, Information, and all general subjects. consists of 15 subjects[12]. But not all attributes can be used because they have to go through several stages of data processing preliminary (preparation of data) to get quality data. The following explanation of the attributes used can be seen from the table 1

| No | Nama Atribut                  | Information                          | Type Data |
|----|--------------------------------|--------------------------------------|-----------|
| 1  | No                             | Number of student data               | Integer   |
| 2  | Name                           | Names of students                    | Id        |
| 3  | Subjects                       | Grades of all subjects               | Numeric   |
| 4  | Presence                       | Student Attendance Value             | Numeric   |
| 5  | Behavior                       | Student behavior score               | Numeric   |
| 6  | Average value                  | Total number of subjects             | Numeric   |
| 7  | Information                    | Information on student grade promotion | Label     |

3.4. Testing

In this experiment, researchers conducted a test using K-Fold Validation 2 to 10, but as an example here the author only presents tests that use K-Fold Validation 3, 5 and 7, for other validation test results can be seen in table 4.7. Following is the testing process:

• Testing using K-Fold Validation 3

Testing using K-Fold Validation 3 will divide the data into three parts, namely 2 training data and 1 testing data, the amount of data processed is 131 records. The confusion matrix model will form a matrix consisting of true positive or positive tuples and true negative or negative tuples, then input the testing data that is ready to be processed into the confusion matrix so that it gets the desired results.

\[ \text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn} = \frac{120 + 10}{120 + 10 + 1 + 0} = 99.22\% \]

\[ \text{Precision} = \frac{tp}{tp + fp} = \frac{120}{120 + 1} = 100.00\% \]

From Figure 4.8, it explains the accuracy results obtained by using K-Fold Validation 3, which is 99.22%.

**Figure 2.** Accuracy Results Using K-Fold Validation 3

**Figure 3.** Precision results using K-Fold Validation 3

**Figure 4.** Classification Performance

**Table 1.** Data Attribute Types Used
Figure 4. Recall results using K-Fold Validation 3

\[
\text{Recall (Use)} = \frac{tp}{tp+fn} = \frac{120}{120+1} = 88.89\%
\]

AUC obtained from testing using K-Fold Validation 3 is equal to 0.944%.

- Testing using K-Fold Validation 5
  Testing using K-Fold Validation 5 will divide the data into five parts, namely 4 training data and 1 testing data. The confusion matrix model will form a matrix consisting of true positive or positive tuples and true negative or negative tuples, then input the testing data that is ready to be processed into the confusion matrix so that it gets the desired results.

Figure 5. Accuracy Results Using K-Fold Validation 5

Following calculations to find the value of accuracy, precision, and recall with the equation below:

\[
\text{Accuracy} = \frac{tp+tn}{tp+tn+fp+fn} = \frac{120+10}{120+10+1+0} = 99.26\%.
\]

Figure 6. Precision results using K-Fold Validation 5

\[
\text{Precision (Use)} = \frac{tp}{tp+fp} = \frac{120}{120+1} = 100.00\%
\]

AUC obtained from testing using K-Fold Validation 5 is equal to 0.975%.

Figure 7. Recall results using K-Fold Validation 5

\[
\text{Recall (Use)} = \frac{tp}{tp+fn} = \frac{120}{120+1} = 93.33\%
\]

AUC obtained from testing using K-Fold Validation 5 is equal to 0.975%.

- Testing using K-Fold Validation 7
  In testing using K-Fold Validation 7 data will be divided into 7 parts, namely 6 training data and 1 testing data from the total data to be processed as many as 131 records. Confusion matrix model will form a matrix consisting of true positive or positive tuples and true negative or negative tuples, then input the testing data that is ready to be processed into the confusion matrix so that it gets the desired results.
Following calculations to find the value of accuracy, precision, and recall with the equation below:

\[
\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn} = \frac{120 + 10}{120 + 10 + 1 + 0} = 99.25\%.
\]

\[
\text{Precision (Reliability)} = \frac{tp}{tp + fp} = \frac{120}{120 + 1} = 100.00\%.
\]

\[
\text{Recall (Use)} = \frac{tp}{tp + fn} = \frac{140}{140 + 28} = 92.86\%.
\]

AUC of the Neural Network (NN) algorithm obtained from test II using K-Fold Validation 7 is equal to 0.962%.

### 3.5. Discussion result

In predicting the value of the grade increase of students in SMAN1 Suela, based on experiments conducted where tests using K-Fold Validation 5 get higher accuracy results than tests using other K-Fold Validation, which is as much as 99.26% tested that use, this study aims to find out how much the level of accuracy resulting from the use of the Confusion Matrix K-Fold Validation by applying Neural Network Algorithms (NN) in predicting the value of the grade increase of students based on the value of rapport on SMAN1 Suela[13][12][14][15]. From these tests, we obtained the results of Accuracy, Precision, Recall and AUC in the equation below.

| Table. 2 Accuracy, Precision, Recall and AUC results using Confusion Matrix |
|----------------|----------------|----------------|----------------|
| Validation     | Accuracy       | Precision      | Recall         | AUC           |
| 2              | 99.2 %         | 100.00%        | 90.00%         | 0.960         |
| 3              | 99.22%         | 100.00%        | 88.89%         | 0.944         |
| 4              | 99.24%         | 100.00%        | 91.67%         | 0.958         |
| 5              | 99.26%         | 100.00%        | 93.33%         | 0.975         |
| 6              | 99.24%         | 100.00%        | 91.67%         | 0.967         |
| 7              | 99.25%         | 100.00%        | 92.86%         | 0.962         |
| 8              | 99.22%         | 100.00%        | 87.50%         | 0.958         |
| 9              | 99.26%         | 100.00%        | 90.91%         | -             |
| 10             | 99.23%         | 100.00%        | 90.00%         | 0.958         |

From table.2 we can see the results of Accuracy, Precision, Recall and AUC obtained from Neural Network (NN) testing using K-Fold Validation 2 to 10 confusion matrix where tests using K-Fold Validation 5 get Accuracy, Precision, Recall results and AUC which is higher than testing using other validations, with an accuracy value of 99.26%, a recall value of 93.33%, a precision value of 100.00% and an AUC value of 0.975%.
4. Conclusions

From the discussion and research that has been done, it can be concluded that the application of the Neural Network (NN) algorithm in data processing of students' grade rise results in very good accuracy. Based on tests conducted using k-fold validation 2 to 10, obtained the results of Accuracy, and higher AUC using k-fold validation 5 when compared with other tests using k-fold validation, namely with an accuracy value of 99.26% and AUC value of 0.975%. In this case, the Neural Network (NN) method is very accurate in processing student grade-level data. So that SMAN1 Suela can apply this method in predicting grades for students' grades increase based on student rapport grades.

References

[1] F. Saltan and Ö. Arslan, “The use of augmented reality in formal education: A scoping review,” Eurasia J. Math. Sci. Technol. Educ., vol. 13, no. 2, pp. 503–520, 2017.
[2] D. Elkind, “Formal education and early childhood education: An essential difference,” Phi Delta Kappan, vol. 67, no. 9, pp. 631–636, 1986.
[3] A. DeCapua, “Reaching students with limited or interrupted formal education through culturally responsive teaching,” Lang. Linguist. Compass, vol. 10, no. 5, pp. 225–237, 2016.
[4] D. Costarelli, “Neural network operators: constructive interpolation of multivariate functions,” Neural Networks, vol. 67, pp. 28–36, 2015.
[5] A. Rohman, “Model Algoritma K-Nearest Neighbor (K-Nn) Untuk Prediksi Kelulusan Mahasiswa,” Neo Tek., vol. 1, no. 1, 2015.
[6] J. McNulty, “No Title עלון הנוטע: תמונת מצב ענף הקיווי, vol. 66, no. 1997, pp. 37–39, 2013.
[7] S. Solikhun, M. Safii, and A. Trisno, “Jaringan Saraf Tiruan Untuk Memprediksi Tingkat Pemahaman Siswa Terhadap Matapelajaran Dengan Menggunakan Algoritma Backpropagation,” J-SAKTI (Jurnal Sains Komput. dan Inform.), vol. 1, no. 1, pp. 24–36, 2017.
[8] A. P. Fadillah, “Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ),” J. Tek. Inform. dan Sist. Inf., vol. 1, no. 3, 2015.
[9] T. Triyuni, “The influence of science learning set using scientific approach and problem solving model on learning outcomes of junior high school students in the subject of heat and temperature,” J. Pendidik. IPA Indones., vol. 5, no. 2, pp. 177–185, 2016.
[10] E. S. Noviando et al., “Studi Penerapan ANN (Artificial Neural Network) Untuk Menghilangkan Harmonisa Pada Gedung Pusat Komputer Abstrack Gambar 2 Bentuk Gelombang Harmonisa (Dugan dan Mc Granaghan, 1996 dalam Sungkowo H, 2013) Sedangkan Untuk Batas distorsi tegangan yang,” vol. 3, no. 2, pp. 1–6, 2016.
[11] J. A. Naser, “Neural networks - a brief introduction,” Proc. Am. Power Conf., vol. 53, no. pt 2, pp. 943–945, 1991.
[12] D. Fauziah, I. Irzani, and R. Ripai, “APLIKASI ADAPTIVE NEURAN FUZZY INFEERENCE SYSTEM (ANFIS) SEBAGAI MODEL DIAGNOSIS KONSENTRASI JURUSAN PADA SISWA SMA/MA,” in Prosiding Seminar Nasional Matematika dan Pendidikan Matematika, 2016.
[13] E. Irfiani and F. Indriyani, “Data Mining Untuk Sistem Pengambilan Keputusan Menentukan Kenaikan Kelas Berbasis Web,” Informatics Educ. Prof., vol. 2, no. 1, pp. 19–28, 2017.
[14] S. Solikhun, M. Safii, and A. Trisno, “Jaringan Saraf Tiruan Untuk Memprediksi Tingkat Pemahaman Siswa Terhadap Matapelajaran Dengan Menggunakan Algoritma Backpropagation,” J-SAKTI (Jurnal Sains Komput. dan Inform.), vol. 1, no. 1, p. 24, 2017, doi: 10.30645/j-sakti.v1i1.26.
[15] V. Canals, A. Morro, A. Oliver, M. L. Alomar, and J. L. Rosselló, “A new stochastic computing methodology for efficient neural network implementation,” IEEE Trans. neural networks Learn. Syst., vol. 27, no. 3, pp. 551–564, 2015.