An Automated Essay Scoring Based on Neural Networks to Predict and Classify Competence of Examinees in Community Academy

I Gusti Putu Asto Buditjahjanto, Mohammad Idhom, Munoto Munoto, Muchlas Samani

Universitas Negeri Surabaya, Doctoral Program of Vocational Education, Surabaya, Indonesia

Abstract **– AES has been widely used in assessing student learning outcomes. However, few studies use Automated Essay Scoring (AES) to simultaneously determine the community academy's competency test scores and levels. This study aims to apply AES to assess essays on the competency certification test. The AES can predict the examinees' scores and classify examinees' competency levels. The method used to build AES uses Back Propagation Neural Networks (BPNN). BPNN was chosen because of its simplicity and ease in building the model. The results showed that the AES for predicting the examinee's competency value showed the MAE value is 0.061621 and the accuracy value is = 97.9665 %. The results of the classification of student competency levels show Accuracy= 0.9063, Precision= 0.9167, Recall= 0.8888, and F1 Score= 0.8857.**

Keywords **– Assessment, competency certification, human rater, neural networks, multimedia IT.**

DOI: 10.18421/TEM114-34
https://doi.org/10.18421/TEM114-34

Corresponding author: I.G.P.A. Buditjahjanto,
Doctoral Program of Vocational Education, Universitas Negeri Surabaya, Indonesia.
Email: asto@unesa.ac.id

Received: 23 August 2022.
Revised: 01 October 2022.
Accepted: 10 October 2022.
Published: 25 November 2022.

1. Introduction

Community Academy is a university that provides vocational education at the level of diploma one or diploma two in one or several branches of particular science and technology [1]. Community Academy in East Java – Indonesia, focuses on multimedia IT expertise. Before graduating from this community academy, the students must have competency in multimedia IT expertise. Cause competency certification is proof to acknowledge the expertise of the workforce. In line with [2], competency certification is an acknowledgment of workforce skills towards mastery of knowledge, skills, and work attitudes by the required work competency standards. Recognition of competence is carried out with competency testing. Competency testing can be carried out through technical and non-technical assessments by collecting relevant evidence to determine whether a person is competent or not in a particular certification scheme [3].

Shavelson [4] stated that competence can be identified as a task, a problem, or a stimulus that is considered to generate these competencies. Engaging in a task can observe a person's behavior or response. Either the presence or absence of a construct can be assessed, or a person's level of performance can be assessed. It is known that several ways can be used to assess a person's abilities, for example, given questions in the form of multiple choice, short answers, and essays [5].In the competency test, to assess a person's ability, the assessment method used is to test in the form of essay questions that competency test participants must answer by writing answers. Essay assessment is a form of competency test used to evaluate the competency value and competency level of the competency test participants. Some researchers consider that essay assessment is considered a potent tool to measure learning outcomes, as well as to observe higher-order thinking skills [6]. However, there are several weaknesses in the assessment using essays. The weaknesses are time-consuming to check the essay answers one by one, too many variations of answers due to the results of each individual's thinking in answering essays, clarity, and neatness of writing. So it makes the human rater understands both the content and quality of the writing and a problem of non-objective assessment [7]. Therefore, human raters must be closely monitored in every administration to ensure the quality and consistency of human raters. The effect of differences between human raters can substantially increase the bias in the final score without careful monitoring [8]. The manual correction makes human rater labor-intensive, time-consuming, and expensive [9]. Based on these problems, a computer assessment is needed to help facilitate the assessment. AES (Automated Essay Scoring) is one way that can help in computer-assisted essay assessment. AES has several advantages, such as being able to assess answers more objectively, more consistently in assessing answers, faster scoring, and lower unit costs [10].

Several studies have used AES based on probability statistics [11] and Artificial Intelligence [12]. The use of AES for essay assessment in Indonesian is still a little using Artificial Intelligence. Most are still based on probability statistics, such as research [13], [14]. Some researchers only focused on the results of the AES assessment, and not many have directly measured the classification of the results of the AES assessment. This study aims to apply AES to predict students' scores and classify students' competency levels simultaneously in assisting the assessment of essays on competency certification tests. The AES used BPNN to assist in predicting and classifying. The advantages BPNN has been widely applied, such as in the financial sector [15], civil engineering [16], wireless sensor networks [17], electricity [18]. There are several reasons for choosing AES based on BPNN, such as BPNN having accuracy and precision in making predictions [19], [20], [21]. BPNN also has advantages in accuracy for determining the classification of a problem [22] [23]. In addition, BPNN can also model a system with only known inputs and outputs [24] [25].

2. Method

Data was collected from questions and answers for competency tests with the IT Multimedia scheme at community colleges in East Java - Indonesia, in 2021. The data set consists of 12 questions from 3 competency units, namely Occupational Safety (U1), Software and Hardware (U2), and Create, manipulate and merge 2D and digital images (U3). Table 1 shows the types of questions for each competency unit. Figure 1 shows an example display of questions posed to the examinee.

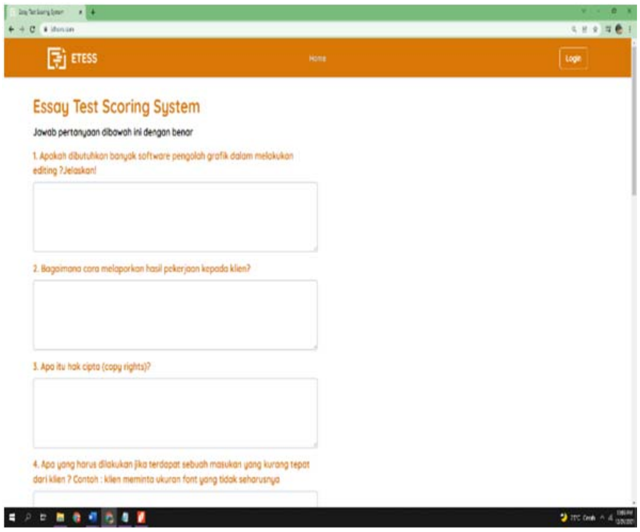

Figure 1. Example of question in AES system

Table 1. Competency unit title and questionnaire list

Item	Competency Unit Title	Questionnaires
1.	Occupational Safety (U1)	1. *Apa tujuan dari Kesehatan dan Keselamatan Kerja dan Lingkungan Hidup?* What is the purpose of Occupational Health and Safety and the Environment? (X1)
		2. *Hal apa saja yang harus diperhatikan dalam menggunakan komputer?* What should things be considered when using a computer? (X2)
		3. *Apa yang harus dilakukan untuk menghindari efek negatif bekerja didepan komputer?* What should be done to avoid the harmful effects of working in front of the computer? (X3)
		4. *Bagaimana posisi tubuh yang tepat untuk menghindari cedera dalam penggunaan komputer?* 5. What is the proper body position to avoid injury in computer use? (X4)
2.	Software and Hardware (U2)	1. *Apakah dibutuhkan banyak software pengolah grafik dalam melakukan editing ?Jelaskan!* Does it take a lot of graphics processing software to do editing? Describe the answer! (X5)
		2. *Apa itu hak cipta?* What is copyright? (X6)
		3. *Bagaimana cara melaporkan hasil pekerjaan kepada klien?* How to report the results of our work to clients? (X7)

		4. *Apa yang harus dilakukan jika terdapat sebuah masukan yang kurang tepat dari klien?* What should we do if there is inappropriate input from the client? (X8)
3.	Create, manipulate and merge 2D and digital images (U3)	1. *Bagaimana caranya menilai kualitas kamera digital?* How to consider the quality of a digital camera? (X9)
		2. *Sebutkan dan jelaskan cara mengambil foto dan upload gambar digital?* Explain how taking photos and uploading digital images works. (X10)
		3. *Bagaimana menggabungkan photografi digital ke dalam multimedia?* How do to merge digital photography into multimedia? (X11)
		4. *Langkah apa sajakah yang diperlukan dalam menyajikan rangkaian video digital?* What steps are required in presenting a digital video series? (X12)

Reference answers come from an assessor, and examinee answers consist of 71. The human rater uses the rubric shown in the table below. Rubric rating scale with a range of 0 to 5.

Table 2. Human Rating Scale

Scores	Criteria
5 (Very Good)	Examinees answered more than 80% of all questions correctly, according to the reference answers
4 (Good)	Examinees answered less than 79% and more than 60% of all questions correctly, according to the reference answer.
3 (Enough)	Examinees answered less than 59% and more than 40% of all questions correctly, according to the reference answer.
2 (Less)	Examinees answered less than 39% and more than 20% of all questions correctly, according to the reference answer.
1 (Bad)	Examinees answered less than 19% of all questions correctly, according to the reference key
0 (Very Bad)	Examinees are not able to answer at all

Figure 2 shows a research block diagram consisting of several stages. The first stage is the preprocessing stage. This stage consists of several processes to clean up text documents (reference answers and examinee answers) so that important words are obtained and words that are not too meaningful are not included in the following process. The process starts with case folding. This process converts the text document to lowercase.

Furthermore, the tokenizing process converts a text document into a collection of words and removes punctuation marks. The following process is Stopword Removal, which removes words that are not too meaningful. The last process is the stemming process, which changes words with affixes into root words. The AES developed in this study uses Indonesian, so it uses Indonesian rules for stemming. Sastrawi is one of the libraries that can be used for Indonesian stemming [14]. Therefore, for the stemming process, this research uses the Sastrawi library. After passing the preprocessing stage, the reference answer text and examinee answer text were obtained, which were clean and in the form of a collection of words.

The second stage is calculating the weights of the reference answers text and examinee answers text. Each reference answer text and examinee answer text are weighted. One method of calculating text weighting is TF-IDF. Text weighting through TF-IDF is based on comparing the frequency of occurrence of a word in a text document and the number of documents containing that word. Calculation of weights with TF-IDF is done using equation (1).

$$w_{ij} = tf_{ij}\ x\ idf_j$$
$$w_{ij} = tf_{ij}\ x \log\left(D/_{df_i}\right) \qquad (1)$$

W_{ij} is the weight of the term (t_j) for the document (d_i), and tf_{ij} is the number of frequencies of the term (t_j) in the document (d_i). D is the number of all documents, and df_i is the number of documents containing the term (t_j).

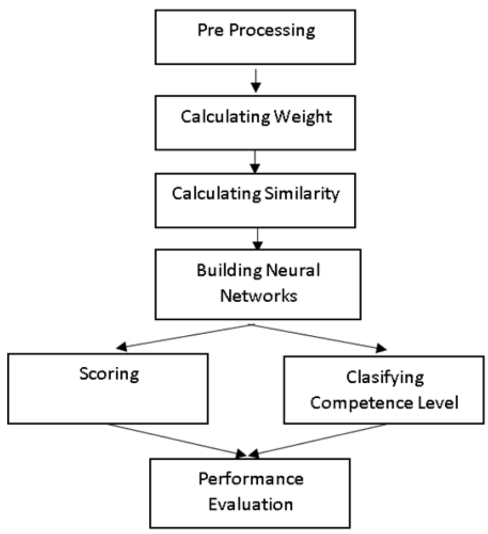

Figure 2. Block diagram of the proposed method

The third step is calculating similarity to measure the similarity between two text documents. This study uses the Cosine Similarity method to measure the similarity between examinee answers and reference answers. The similarity scale of this method is 0-1. The closer to 1, the higher the similarity between the examinee's answer and the reference answer, and vice versa. The closer to 0, the lower the similarity between the examinees and the reference answers. The formula for determining the similarity value using cosine similarity is shown in equation (2).

$$Similarity = \cos(\theta) = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \quad (2)$$

Description :

A_i : The weight of term i on the examinee's answer document
B_i : The weight of term i in the reference answer document
n : Number of terms

The fourth stage is predicting examinee competency scores and classifying examinee competency levels. At the stage of predicting examinee competency scores, this study uses Neural Networks with the type of Backpropagation Neural Networks (BPNN). The built BPNN model consists of 12 inputs, four hidden layers, and one output. The input consists of 12 nodes arranged from each competency unit question, from the first question (X1) to the 12th question (X12). The selected hidden layer consists of one layer with four nodes. The number of outputs is only one node (Y), the examinee's competency value. Figure 3 shows the 12-4-1 architectural model from BPNN to predict examinee competency scores. The prediction of the resulting competency score has a range of values from 0 to 1. It follows the data pattern on the input using TF-IDF weights and the output using the results of the cosine similarity calculation. This BPNN model for prediction has error = 0.0159 with epoch =10000.

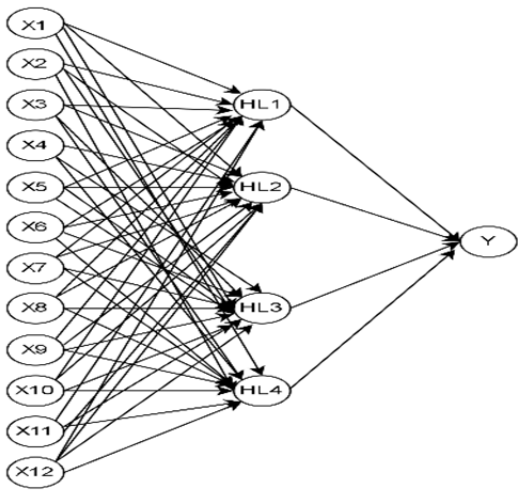

Figure 3. The 12-4-1 architectural model from BPNN to predict examinee competency scores

The competence level is classified into low, medium, and high. The classification stage is carried out by calculating the examinees' average value of each competency unit (U1, U2, and U3). This average value is taken from the Cosine Similarity value from the previous calculation. So U1 is the average value of Cosine Similarity X1, X2, X3, and X4. U2 is the average value of Cosine Similarity X5, X6, X7, and X8. U3 is the average value of Cosine Similarity X9, X10, X11, and X12. The Backpropagation Neural Networks (BPNN) model is shown in Figure 4. The BPNN model consists of 3 inputs, four hidden layers, and one output. The input consists of 3 nodes arranged from each competency unit question U1 to U3. The hidden layer consists of one layer and four nodes, while the output is only one node showing the examinee's competence level. This BPNN model for classification has error = 0.0153 with epoch =10000.

The fifth stage is the performance evaluation stage, which measures the agreement level or proximity value between the score generated by the AES system and the score generated by the assessor to predict the score of the examinee.

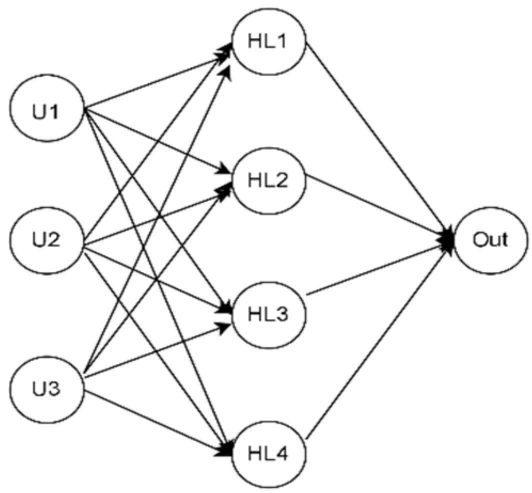

Figure 4. The 3-4-1 architectural model from BPNN to classify the examinee's competency levels

Performance evaluation used is MAE (Mean Absolute Error) and Accuracy. MAE is used to measure the error rate by finding the absolute difference between the score generated by the human rater (X) and the score generated by the AES (Y) and dividing by the number of answers the examinee (n). Equation 3 is the formula for MAE [14]:

$$MAE = \frac{\sum|X-Y|}{n} \quad (3)$$

Then, the following process is to calculate the similarity between the AES score and the human rater score using equation 4 below [26] :

$$Acc = 100\% - \left(\left|\frac{Human\ rater\ score - AES\ score}{100}\right|\right) x100\% \quad (4)$$

Performance evaluation for the classification of competency levels of examinees is to use a confusion matrix to measure accuracy, precision, recall, and F1. Accuracy is the division between the sum of True Positive (TP) and True Negative (TN) with all samples (N). It can be calculated as:

$$Accuracy = \frac{TP+TN}{N} \quad (5)$$

Precision is the division between True Positive with the sum of True Positive and False Positive (FP). It is calculated as:

$$Precision\ (p) = \frac{TP}{TP+FP} \quad (6)$$

Recall is the division between True Positive with the sum of True Positive and False Negatif (FN). It can be calculated as:

$$Recall\ (r) = \frac{TP}{TP+FN} \quad (7)$$

A good model needs to strike the right balance between Precision and Recall. For this reason, the F-score (F-measure or F1) is used by combining Precision and Recall to obtain a balanced classification model. The F-score is calculated by the average of the Precision and Recall harmonics as in the following equation.

$$F1\ score = 2 \text{ x } \frac{p \text{ x } r}{p+r} \quad (8)$$

3. Result and Discussion

The research data consisted of 71 test results data from the examinees. This data is divided into two parts, namely 38 data for BPNN training and 33 data for BPNN testing. In the prediction stage, the weight of the BPNN training results with the 12-4-1 architectural model is used to predict the examinee's value. Table 3 shows the results of testing using BPNN, which also calculated the magnitude of the error compared to the calculation of the human rater whose assessment used a rubric, as shown in Table 2.

Table 3. Test results for predicting competency scores

No	Human rater (X)	BPNN Prediction (Y)	Error (X-Y)
1	0.2	0.1828	0.0172
2	0.3	0.2378	0.0622
3	0.2	0.2098	0.0098
4	0.2	0.3042	0.1042
5	0.3	0.2192	0.0808
6	0.3	0.3370	0.037
7	0.3	0.2074	0.0926
8	0.3	0.2449	0.0551
9	0.3	0.2435	0.0565
10	0.2	0.1941	0.0059
11	0.3	0.2899	0.0101
12	0.4	0.3285	0.0715
13	0.5	0.4412	0.0588
14	0.5	0.5409	0.0409
15	0.6	0.6377	0.0377
16	0.3	0.3171	0.0171
17	0.4	0.4251	0.0251
18	0.6	0.6685	0.0685
19	0.8	0.7394	0.0606
20	0.6	0.7202	0.1202
21	0.7	0.7834	0.0834
22	0.4	0.3265	0.0735
23	0.5	0.4425	0.0575
24	0.4	0.4833	0.0833
25	0.5	0.5917	0.0917
26	0.6	0.7296	0.1296
27	0.7	0.7881	0.0881
28	0.6	0.5444	0.0556
29	0.7	0.6392	0.0608
30	0.6	0.7207	0.1207
31	0.7	0.7797	0.0797
32	0.7	0.7657	0.0657
33	0.8	0.8121	0.0121
		Total	2.0335

The result of the MAE calculation is the comparison between the total error and the number of answers the examinee is as follows: MAE is =2.0335/33 = 0.061621. This MAE value is relatively small, so the predicted value from BPNN is close to the value of the human rater. The accuracy calculation is 97.9665%. The results of the accuracy calculation show good results that the predicted value of BPNN also shows results that are not much different from the results of the human rater calculation.

Table 4. Test results for classifying competency scores

No	U1	U2	U3	BPNN Classifi -cation	Classifi- cation
1	0.1483	0.2570	0.1469	0.0934	0
2	0.2026	0.3893	0.1579	0.1090	0
3	0.1390	0.3467	0.1724	0.1008	0
4	0.2975	0.4285	0.2171	0.1343	0
5	0.0934	0.3538	0.2508	0.1059	0
6	0.3338	0.4198	0.2859	0.1494	0
7	0.1549	0.2747	0.2345	0.1046	0
8	0.1705	0.4451	0.1617	0.1089	0
9	0.3017	0.2316	0.2280	0.1221	0
10	0.2212	0.2438	0.2674	0.1028	0
11	0.3347	0.3981	0.3112	0.1373	0
12	0.2323	0.2513	0.655	0.1604	0
13	0.3214	0.3723	0.7883	0.2028	1
14	0.2432	0.6222	0.8278	0.2263	1
15	0.3765	0.7553	0.9234	0.2682	1
16	0.2925	0.6211	0.2543	0.1290	0
17	0.3457	0.7951	0.3256	0.1669	0
18	0.6663	0.8567	0.6547	0.2739	1
19	0.7538	0.9522	0.7734	0.3088	2
20	0.6882	0.8223	0.8175	0.3051	2

21	0.7336	0.9123	0.9324	0.3358	2
22	0.6862	0.2768	0.2658	0.1603	0
23	0.7156	0.3886	0.3664	0.2028	1
24	0.6987	0.2699	0.6677	0.2288	1
25	0.7125	0.3342	0.7436	0.2698	1
26	0.8738	0.6562	0.8325	0.3222	2
27	0.9882	0.7222	0.9324	0.3557	2
28	0.8661	0.6456	0.2786	0.2256	1
29	0.9133	0.7122	0.3859	0.2674	1
30	0.8225	0.8453	0.6685	0.3041	2
31	0.9324	0.9232	0.7237	0.3374	2
32	0.8567	0.8657	0.8527	0.3328	2
33	0.9111	0.9322	0.9159	0.3624	2

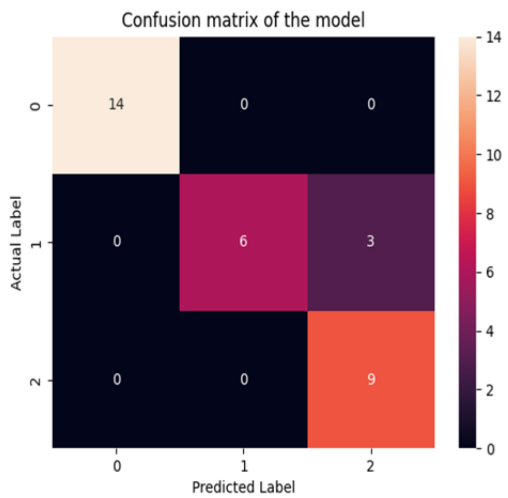

Figure 5. Confusion matrix of the classification model

The results of the AES performance evaluation for the examinee's competency level classification as a whole show the following results: Accuracy level: 0.9062, showing good results to determine whether the examinee's competency level is low, medium, and high. Precision Performance: 0.9167 and Recall Performance: 0.8888 showed good results in correctly classifying students' competency levels. The performance of F1 Score: 0.8857 indicates that the model formed is good because the closer to 1, the better the model. Figure 5 shows the confusion matrix of the classification model. Table 4 shows test results for classifying competency scores.

The use of AES for student assessment has been widely developed and has provided much help for users in conducting student assessments. Many methods are used in building AES, including the BPNN method. BPNN for education has been widely used to assess student learning outcomes and determine student skills [27] [28]. This study also uses BPNN to predict the examinee's value and competency level. Research conducted [29] using BPNN to evaluate the effect of STEAM-graded teaching shows the results of assessing the accuracy of student answers with significant results with small errors. This study also resulted in the examinee's assessment results being well and significantly. In addition, the study result shows the examinee's value prediction with small errors and high accuracy values. As for the classification, it produces a high precision value so that it is precise in classifying, and the F1 score value shows that the classification system model that has been built is good.

4. Conclusion

The development of Automated Essay Scoring (AES) based on BPNN to predict competency scores and classify examinee competency levels in the IT Multimedia field shows good and precise results. The performance results of competency score prediction show the MAE value = 0.061621, which is quite low, and the accuracy value is high, namely Acc = 97.9665 %. The results of the classification of student competency levels show Accuracy: 0.9063, Precision: 0.9167, Recall: 0.8888, F1 Score: 0.8857. The study results indicate that the development of Automated Essay Scoring (AES) based on BPNN, when applied to real conditions, can help reduce the need for much energy and reduce the time to check the examinee's answers more efficient than doing it manually.

Acknowledgements

The authors would like to thank the DRTPM Kemdikbudristek-Indonesia, who has funded this research with contract number B/29577/UN38.9/LK.04.00/202

References

[1]. Moore, S. D., Brennan, S., Garrity, A. R., & Godecker, S. W. (2000). Winburn Community Academy: A University-Assisted community school and professional development school. *Peabody Journal of Education*, *75*(3), 33-50. https://doi.org/10.1207/S15327930PJE7503_3

[2]. Watkins, S. C. (2020). Simulation-based training for assessment of competency, certification, and maintenance of certification. In *Comprehensive healthcare simulation: InterProfessional team training and simulation* (pp. 225-245). Springer, Cham. https://doi.org/10.1007/978-3-030-28845-7_15.

[3]. Chi, T. P., Tu, T. N., & Minh, T. P. (2020). Assessment of information technology use competence for teachers: Identifying and applying the information technology competence framework in online teaching. *Journal of Technical Education and Training*, *12*(1), 149-162. https://doi.org/10.30880/jtet.2020.12.01.016

[4]. Shavelson, R. J. (2010). On the measurement of competency. Empirical Research in Vocational Education and Training, 2 (1), 41–63. https://doi.org/10.1007/BF03546488

[5]. Burrows, S., Gurevych, I., & Stein, B. (2015). The eras and trends of automatic short answer grading. *International Journal of Artificial Intelligence in Education*, *25*(1), 60-117.

[6]. Stephen, T. C., Gierl, M. C., & King, S. (2021). Automated essay scoring (AES) of constructed responses in nursing examinations: An evaluation. *Nurse Education in Practice*, *54*, 103085. https://doi.org/10.1016/j.nepr.2021.103085.

[7]. Wang, Z., Zechner, K., & Sun, Y. (2018). Monitoring the performance of human and automated scores for spoken responses. *Language Testing*, *35*(1), 101-120. https://doi.org/10.1177/0265532216679451

[8]. Williamson, D. M., Xi, X., & Breyer, F. J. (2012). A framework for evaluation and use of automated scoring. *Educational measurement: issues and practice*, *31*(1), 2-13.

[9]. Zhang, M., Breyer, F. J., & Lorenz, F. (2013). Investigating The Suitability Of Implementing The E-Rater® Scoring Engine In A Large-Scale English Language Testing Program. *ETS Research Report Series*, *2013*(2), i-60.

[10]. Ramineni, C. (2013). Validating automated essay scoring for online writing placement. *Assessing Writing*, *18*(1), 40-61.

[11]. Liang, C., Zhang, B., Gong, X., Li, M., Guo, H., & Li, R. (2021). A Survey of Automated Essay Scoring System Based on Naive Bayes Classifier. In *Artificial Intelligence in Education and Teaching Assessment* (pp. 247-250). Springer, Singapore. https://doi.org/10.1007/978-981-16-6502-8_21.

[12]. Bin, L., Jun, L., Jian-Min, Y., & Qiao-Ming, Z. (2008, December). Automated essay scoring using the KNN algorithm. In *2008 International Conference on Computer Science and Software Engineering* (Vol. 1, pp. 735-738). IEEE. https://doi.org/10.1109/CSSE.2008.623

[13]. Amalia, A., Gunawan, D., Fithri, Y., & Aulia, I. (2019, June). Automated Bahasa Indonesia essay evaluation with latent semantic analysis. In *Journal of Physics: Conference Series* (Vol. 1235, No. 1, p. 012100). IOP Publishing. https://doi.org/10.1088/1742-6596/1235/1/012100

[14]. Hasanah, U., Permanasari, A. E., Kusumawardani, S. S., & Pribadi, F. S. (2019). A scoring rubric for automatic short answer grading system. *Telkomnika (Telecommunication Computing Electronics and Control)*, *17*(2), 763-770. https://doi.org/10.12928/telkomnika.v17i2.11785

[15]. Peng, H. (2021). Research on credit evaluation of financial enterprises based on the genetic backpropagation neural network. *Scientific Programming*, *2021*, 1-8. https://doi.org/10.1155/2021/7745920.

[16]. Bu, J., Zhang, F., Zhu, M., He, Z., & Hu, Q. (2019). Prediction of punching capacity of slab-column connections without transverse reinforcement based on a backpropagation neural network. *Advances in Civil Engineering*, *2019*, 1-9. https://doi.org/10.1155/2019/7904685

[17]. Sun, Y., Zhang, X., Wang, X., & Zhang, X. (2018). Device-free wireless localization using artificial neural networks in wireless sensor networks. *Wireless Communications and Mobile Computing*, *2018*, 1-8. https://doi.org/10.1155/2018/4201367

[18]. Shi, J., & Chang, D. (2022). Assessment of Power Plant Based on Unsafe Behavior of Workers through Backpropagation Neural Network Model. *Mobile Information Systems*, *2022*, 1-12. https://doi.org/10.1155/2022/3169285

[19]. Zhang, Q., Yan, L., Hu, R., Li, Y., & Hou, L. (2022). Regional Economic Prediction Model Using Backpropagation Integrated with Bayesian Vector Neural Network in Big Data Analytics. *Computational Intelligence and Neuroscience*, *2022*, 1-10. https://doi.org/10.1155/2022/1438648

[20]. Anam, S., Maulana, M. H. A. A., Hidayat, N., Yanti, I., Fitriah, Z., & Mahanani, D. M. (2021). Predicting the Number of COVID-19 Sufferers in Malang City Using the Backpropagation Neural Network with the Fletcher–Reeves Method. *Applied Computational Intelligence and Soft Computing*, *2021*, 1-9. https://doi.org/10.1155/2021/6658552

[21]. Nguyen, T. A., Ly, H. B., & Pham, B. T. (2020). Backpropagation neural network-based machine learning model for prediction of soil friction angle. *Mathematical Problems in Engineering*, *2020*, 1-11. https://doi.org/10.1155/2020/8845768

[22]. Elgin Christo, V. R., Khanna Nehemiah, H., Minu, B., & Kannan, A. (2019). Correlation-based ensemble feature selection using bioinspired algorithms and classification using backpropagation neural network. *Computational and mathematical methods in medicine*, *2019*, 1-17. https://doi.org/10.1155/2019/7398307

[23]. Xin, C. W., Jiang, F. X., & Jin, G. D. (2021). Microseismic Signal Classification Based on Artificial Neural Networks. *Shock and Vibration*, *2021*, 1-14. https://doi.org/10.1155/2021/6697948

[24]. Ma, Z., & Wang, Y. (2022). Analysis and Prediction of Body Test Results Based on Improved Backpropagation Neural Network Algorithm. *Advances in Multimedia*, *2022*, 1-9. https://doi.org/10.1155/2022/1701687

[25]. Liu, Y., Jing, W., & Xu, L. (2016). Parallelizing backpropagation neural network using MapReduce and cascading model. *Computational intelligence and neuroscience*, *2016*, 1-11. https://doi.org/10.1155/2016/2842780

[26]. Ratna, A. A. P., Arbani, A. A., Ibrahim, I., Ekadiyanto, F. A., Bangun, K. J., & Purnamasari, P. D. (2018, November). Automatic essay grading system based on latent semantic analysis with learning vector quantization and word similarity enhancement. In *Proceedings of the 2018 International Conference on Artificial Intelligence and Virtual Reality* (pp. 120-126). https://doi.org/10.1145/3293663.3293684

[27]. Hussein, M. A., Hassan, H., & Nassef, M. (2019). Automated language essay scoring systems: A literature review. *PeerJ Computer Science*, *5*, e208. https://doi.org/10.7717/peerj-cs.208

[28]. Filho, A. H., Concatto, F., do Prado, H. A., & Ferneda, E. (2021). Comparing Feature Engineering and Deep Learning Methods for Automated Essay Scoring of Brazilian National High School Examination. *Proceedings of the 23rd International Conference on Enterprise Information Systems,* 575-583. https://doi.org/10.5220/0010377505750583

[29]. Shi, Y., & Rao, L. (2022). Construction of STEAM Graded Teaching System Using Backpropagation Neural Network Model under Ability Orientation. *Scientific Programming*, *2022*, 1-9. https://doi.org/10.1155/2022/7792943