Classification using Deep Learning based on selection optimizing counseling data

A E Budianto1,*, M F I Othman2 and Y A Rahim2

1 Informatics Study Program, Faculty of Science and Technology, Universitas Kanjuruhan Malang, Jl. S. Supriadi No. 48 Malang Jawa-Timur Indonesia
2 Information and Communication Technology, Faculty of Information and Communication Technology Universiti Teknikal Malaysia Melaka Hang Tuah Jaya, 76100 Durian Tunggal Melaka, Malaysia

* endybudianto@unikama.ac.id

Abstract. Counseling is professional assistance given by counselors to clients directly, to solve problems and improve their behavior. In the future. The discussion of the problem in question is in-depth which concerns important matters about the client, broad covering various aspects of the client's problem, and specifically to alleviate the client's pressing problems. The Support Vector Machine and Multi Layer Perceptron algorithm is applied to classify counseling data provided by professionals to clients with an accuracy of 86.39%. Data from Vocational Research in Malang City in 2018. To improve the optimal accuracy of the type of counseling data classification, this study proposes an algorithm optimization feature for the Deep Learning algorithm. The results showed a significant increase in accuracy. The value of classification accuracy in counseling data generated by the Deep Learning algorithm has increased to 97.37% with a better level of accuracy.

1. Introduction
This research aims to help the counseling guidance teachers to carry out their duties and responsibilities optimally counseling guidance both at school and outside the school is required to have insight, knowledge, skills and values and attitudes in achieving the goals of counseling services [1]. Besides that, facilities and infrastructure are also needed, for example the availability of counseling rooms, books, visual media (recording counseling) and instrument application programs that are used to help counseling guidance at school as a model describing counseling services namely e-counseling and utilizing Information Technology in counseling guidance with a good level of accuracy of counseling data in accordance with client problems [2,3].

The legal basis for guidance and counseling is based on Permendikbud No. 111 of 2014, and there are still other higher regulations relating to guidance and counseling, such as: (1) Law no. 14 of 2005 concerning teachers and lecturers, (2) Government regulation No. 74 of 2008 concerning teachers. Then other regulations that are considered in the considerations of this Ministry of Education, are as follows: (1) Regulation of the Minister of National Education no. 27 of 2008 concerning standards of academic qualifications and counselor competencies, (2) Regulation of the Minister of Education and Culture Number 54 of 2013 concerning the standards of competence of primary and secondary education graduates, (3) Regulation of the Minister of Education and Culture Number 64 of 2013 concerning
standards of contents of primary school education and secondary, (4) Regulation of the Minister of Education and Culture No. 57 of 2014 concerning the 2013 elementary school / madrasah ibtidaiyah curriculum, (5) Minister of Education and Culture Regulation No. 58 of 2014 concerning the 2013 junior high school / madrasah tsanawiyah, (6) Regulations Minister of Education and Culture Number 59 of 2014 concerning the 2013 high school / Madrasah Aliyah curriculum, (7) Minister of Education and Culture Regulation Number 60 of 2014 concerning the 2013 curriculum of the Aliyah Vocational high school / Madrasah [4]. Based on these regulations and legal basis, counseling guidance is very important to be explored and reviewed in a study to obtain accurate counseling data that is appropriate for clients.

Related to the issue of accuracy of counseling data, research is focused on accuracy that is only in accordance with the counseling problem. The counseling data examined is the data of counseling of vocational students in Malang with the name counseling Datasets. The results of research conducted in 2017 to 2018 are known as counseling Datasets.

2. Methods

Data mining is a new technology that is very useful for companies to find very important information from their data warehouse. Some data mining applications focus on predictions, they predict what will happen in new situations from data that illustrates what happened in the past [5]. Data mining is part of Knowledge Discovery Data (KDD) which is a process of extracting useful information, unknown previously and hidden from data [6] and also developed a model used to understand phenomena from data analysis and prediction [6-8].

Data mining refers to the extraction of knowledge from the large amount of data stored in a computer, the process of data mining is more based on information technology [10]. Data mining tasks can be classified into two categories namely descriptive and predictive [11].

Descriptive mining tasks characterize the general nature of data in databases. Descriptive mining is used to make predictions [12] Following are the stages of data mining in the process of knowledge discovery as seen in Figure 1 [13]. (a) Data cleaning, done to eliminate noise and inconsistent data, (b) Data integration, where several data sources can be combined, (c) Data selection, where data relevant to the database analysis task will be taken, (d) Data transformation, where data is changed or consolidated into a form suitable for mining by performing summary or aggregation operations, (e) Data mining, the process of identifying really interesting patterns that represent knowledge is based on some interestingness steps, (f) Knowledge presentation, where visualization and representation techniques knowledge used for present knowledge to users.

![Figure 1. Data mining stages.](image)

There are four groups in Data mining according to [14]:

- Classification
  The classification process is based on classes and dependent variables. Popular classification models are Decision Tree, Naïve Bayes, Neural Networks, Genetic Algorithms, and Support Vector Machines [15].
• Association
   In principle, the discovery of association rules studies the rules of how to understand identifying rules between different dependencies of group phenomena. Thus the purpose of the association is to look for object relationships based on rules to predict the occurrence of objects based on other events.
• Clustering
   Cluster is finding groups of objects based on similarity, so that each group has differences between one another.
• Prediction
   Prediction is a process related to the ability to predict the best response (output) that is closest to reality based on input data.

Thus, the smaller the difference between what is expected to occur (expected results) and what happens (observed results), the better the prediction, for example the prediction of a weather forecast (for example, for 24 or 48 hours) or the diagnosis for a particular disease that is given to certain patients, which is based on medical data [16]. The concept of data mining, finding valuable patterns in data, is a clear response to the announcement and storage of data volumes [17]. In particular, a collection of methods known as 'data mining' offers methodologies and technical solutions to overcome the analysis of medical data and the construction of predictions [18]. For all data mining applications, prediction accuracy depends on the quality of the prediction attribute [19].

![Figure 2. Research methods.](image)

The data used for training data and testing data in this study is secondary data that has been collected and analyzed by previous researchers and has gone through the stages of data mining. To test the classification model for forest mapping types, Dataset Counseling Data is used [20]. At this stage, it needs exploration of Dataset counseling data. Exploration aims to ensure that all the attributes and classes in the dataset are valid and can be used as research objects. It can achieve so that the objective to find out the best classification results from counseling data [21].

• Population
   The population of the Dataset counseling data is the class in the data itself and is divided into four groups or classes, namely A (good class), B (good enough class), C (poor class), and D (very poor class).
• Sample
   It uses the sample for testing the model produced or can be said to be testing data, while the training data serves as training material for a model. Dataset counseling has a total record of 550,
then it has been divided into two parts, namely 350 records training data and 223 records testing data. Dataset Counseling is secondary data ready for the data mining process. The method to be applied is a method that has been applied by researchers who use this dataset as their previous research object. We have designed dataset counseling through all the stages required as a dataset for the classification model [22].

At this stage, it needs exploration of the datasets used. The first step is to know that the main goal to be achieved is to find out the classification results from mapping the best data types [23].

- **Sampling.** This research uses counseling datasets with 550 records. To test the developed model, we will divide the data into two parts, namely training data and testing data. We use training data for model development, and we use Testing data to test models. We know it that the total amount of data is 550 records and then has been divided into two parts, we use 350 records as training data and 223 records testing data or as sample data.

- **Feature Selection.** To improve accuracy results to a maximum accuracy value, one of them is by applying feature selection optimization. In this research, we will apply optimization of feature selection to each model that has been tested. Optimize selection has the principle of finding and selecting features from the best value of all subsets [24]. The feature selection process in Optimize selection can be seen in the following image.

![Optimize selection method.](image)

**Figure 3.** Optimize selection method.

3. Results and discussion

We intend the application of Deep Learning algorithm with optimization of the Optimize Selection feature on Dataset counseling to improve accuracy in the classification of counseling data from the results of previous studies [25]. Improved accuracy results will be achieved or not, we will see the results in the experimental results. We will carry experiments on the Deep Learning algorithm out in two stages. The first stage of the experiment is an experiment on the Deep Learning algorithm without the application of Optimize Selection optimization features and the second experiment is an experiment on the Deep Learning algorithm with the Optimize Selection Application.

3.1. Results of Deep Learning experiments without Optimize Selection

At this stage the first stage of the experimental process is testing the model using Python Software on the Deep Learning model without Optimize Selection based. Dataset counseling which is a dataset that has been prepared to be implemented in the model testing process, then tested on the Deep Learning algorithm without Optimize Selection. After testing the model the results got are as in table 1.
Table 1. The accuracy of the Deep Learning model based on Optimize Selection [26].

| No | Model (Algorithm) | Accuracy Value |
|----|------------------|----------------|
| 1  | Deep Learning    | 86.39%         |

In table 1 the accuracy results got for the classification of types of counseling data from counseling Dataset have not reached expectations, namely not getting maximum results. These results are still below the accuracy of the algorithm used in previous studies. So from these results in this study will be applied optimization feature selection to improve the accuracy of the results of the Deep Learning model based on Optimize Selection as expressed in the proposed model.

3.2. Results of Optimized Selection based Deep Learning experiments [27]

At this stage we will conduct an experiment with the optimization of the Optimize Selection feature in the Deep Learning model. Optimize Selection is an optimization feature selection that can choose the best attributes so it can increase the level of accuracy in the subsets being tested. After we carry the experiment out, we can see the results in table 2.

Table 2. The accuracy of the Deep Learning model based on Optimize Selection.

| No | Model (Algorithm) | Accuracy Value |
|----|------------------|----------------|
| 1  | DL+Optimize      | 97.37%         |

In table 2 the accuracy results got by the application of Optimize Selection for the classification of counseling data types in vocational schools in Malang have increased. So we prove it that the Optimize Selection algorithm can increase the level of accuracy up to 97.37% from the previous 86.39% produced by the Deep Learning model without Optimize Selection. Optimization of feature selection based on Optimize Selection can improve accuracy by 10.98% from the results got by the Deep Learning algorithm without Optimize Selection. At this stage is the stage to test the results of the experiments carried out at stage one and stage two experiments. We carried both experiments out on Python Software that had been implemented the Deep Learning with Algorithm Classifier model [28].

In the first stage experiments the results were not maximal, then in the second stage the results of accuracy could be increased. a. Comparison of Deep Learning Algorithms with Other Algorithms to test whether the accuracy of the Deep Learning algorithm without Optimize selection is the best, when compared with other Classifier models [15]. This evaluation aims only to see if there are other Classifier algorithms able to get accuracy results that are more optimal compared to Deep Learning or whether Deep Learning gets lower results. Evaluation is still being carried out on the same software after testing with the Deep Learning model has been completed. We can see the results of the comparison of Deep Learning algorithms with other algorithms in table 3 [29].

Table 3. Comparison of Deep Learning with other algorithms.

| No | Model (Algorithm) | Accuracy Value |
|----|------------------|----------------|
| 1  | Naïve Bayes      | 77.95%         |
| 2  | Random Forest    | 85.35%         |
| 3  | Deep Learning    | 86.39%         |
| 4  | k-NN             | 88.38%         |
| 5  | Generalize Linier| 89.39%         |

Table 3 shows the accuracy of all the algorithms tested by the Deep Learning algorithm. From this table the Deep Learning algorithm is the algorithm with the lowest acquisition results. So the results is not the best algorithm. The accuracy of the Deep Learning algorithm without Optimize Selection is 86.39%
lower than the accuracy got by the Generalize Linear Model algorithm which is 89.39%. Then Optimize Selection will apply to the Deep Learning model with the aim to improve the accuracy of its results.

![Comparison of Accuracy Results](image1)

**Figure 4.** Graph of evaluation results without Optimize Selection.

![Data Accuracy Level](image2)

**Figure 5.** Graph data accuracy level.

4. Conclusions
In this study, they used the Experiment algorithm based on the classification of counseling datasets. Deep Learning algorithm which is a derivative algorithm of Neural Network, aims to optimize the value of the accuracy of the classification of data types in counseling data at Vocational High Schools in Malang [30]. Based on experimental results in this study, the following conclusions can be drawn:

- Deep Learning Algorithms combined with other algorithms have complete methods in Machine Learning both supervised learning and unsupervised learning. We conducted the first experiment without Optimize Selection, with the resulting accuracy value of 86.39%. These results are not optimal results compared to the results of accuracy in previous studies. To improve the accuracy of the results, we base the second experiment on Optimize Selection, then get an accuracy value of 97.37%. These results not only increase the accuracy value in the first experiment but also increase from the accuracy results obtained in previous studies.

- Optimize Selection is one of the feature selection optimizations that can increase the accuracy value to the maximum value through the best feature selection. Choice Optimization combined with Deep Learning proves to be an excellent combination. The accuracy value of counseling data classification named counseling dataset has increased successfully.
Acknowledgments
This study is supported by Informatics Study Program Universitas Kanjuruhan Malang and Information and Communication Technology at Universiti Teknikal Malaysia Melaka.

References
[1] Rejeesh E and Anupama M 2017 Social media and data mining enabled pre-counseling session: A system to perk up effectiveness of counseling in distance education 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) pp 153-156
[2] Richards D and Viganò N 2013 Online Counseling: A Narrative and Critical Review of the Literature J. Clin. Psychol. 69(9) 994–1011
[3] Dowling M and Rickwood D 2013 Online counseling and therapy for mental health problems: A systematic review of individual synchronous interventions using chat J. Technol. Hum. Serv. 31(1) 1–21
[4] Ministry of Education and Culture 2014 “Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia tentang Bimbingan dan Konseling pada Pendidikan Dasar dan Pendidikan Menengah,” pp 1–7
[5] Jiawei H and Kamber M 2011 “Data mining: Concepts and techniques, (the morgan kaufmann series in data management systems),”
[6] Zhao X, Zhang L, Ding Z, Xia L, Tang J and Yin D 2018 United Kingdom Reinforcement Learning Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. - KDD ’18 1040–1048
[7] Chen F, Fu Z and Yang Z 2018 Wind power generation fault diagnosis based on deep learning model in internet of things (IoT) with clusters Cluster Comput. 6 1–13
[8] Day O and Khoshgoftaar T M 2017 A survey on heterogeneous transfer learning J. Big Data 4(1)
[9] Sulistianto M P T, Pranata K B, Afandi A N, Sendari S and Sulistiyowati I 2019 Monitoring Electrical Energy in Electronic Energy Audits through Internet of Things Technology Journal of Physics: Conference Series 1402(7)
[10] Li G, Dai J S, Park E M and Park S T 2017 A study on the service and trend of Fintech security based on text-mining: focused on the data of Korean online news J. Comput. Virol. Hacking Tech. 13(4) 249–255
[11] Singhal V 2018 Row-Sparse Discriminative Deep Dictionary Learning for Hyperspectral Image Classification IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11(12) 5019-5028
[12] Hu Z, Liu W, Bian J, Liu X and Liu T-Y 2017 Listening to Chaotic Whispers: A Deep Learning Framework for News-oriented Stock Trend Prediction Proceedings of the eleventh ACM international conference on web search and data mining pp 261–269
[13] Taub M and Azevedo R 2018 How Does Prior Knowledge Influence Eye Fixations and Sequences of Cognitive and Metacognitive SRL Processes during Learning with an Intelligent Tutoring System? Int. J. Artif. Intell. Educ. 29(1) 1–28
[14] Thangaraju P and Nanthini D 2015 An exhaustive study on association rule mining Int. J. Comput. Sci. Mob. Comput. 4(3) 411–417
[15] Kamath C N, Bukhari S S and Dengel A 2018 Comparative Study between Traditional Machine Learning and Deep Learning Approaches for Text Classification Proceedings of the ACM Symposium on Document Engineering 2018 pp 1–11
[16] Mahmud M, Kaiser M S, Hussain A and Vassanelli S 2018 Applications of deep learning and reinforcement learning to biological data IEEE transactions on neural networks and learning systems 29(6) 2063-2079
[17] Han J, Wang C and El-kishky A 2014 Bringing Structure to Text : Mining Phrases , Entities , Topics , and Hierarchies Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD ’14 p 2630804
[18] Zheng S, Wang F and Lu J J 2012 Bridging the unstructured and structured worlds: an adaptive self learning medical form generating system Proceedings of the 2nd international workshop
on Managing interoperability and complexity in health systems pp 59-66

[19] Yi X, Zhang J, Wang Z, Li T and Zheng Y 2018 Deep distributed fusion network for air quality prediction Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining pp 965-973

[20] Balaneshinkordan S, Kotov A and Nikolaev F 2018 Attentive neural architecture for ad-hoc structured document retrieval Proceedings of the 27th ACM International Conference on Information and Knowledge Management pp 1173-1182

[21] Lu Y and Hsiao I-H 2019 Exploring Programming Semantic Analytics with Deep Learning Models Proc. 9th Int. Conf. Learn. Anal. Knowl. - LAK19 pp 155–159

[22] Al-Taei A 2017 Ensemble Classifier for Eye State Classification using EEG Signals arXiv preprint arXiv:1709.08590

[23] Sabanci K, u’lnerseren M F and Polat K 2016 Classification of different forest types with machine learning algorithms Res. Rural Dev. 1 254–260

[24] Wu S, Zhong S and Liu Y 2018 Deep residual learning for image steganalysis Multimedia tools and applications 77(9) 10437-10453

[25] Xiao L, Jiang D, Xu D, Su W, An N and Wang D 2018 Secure Mobile Crowdsensing Based on Deep Learning China Commun. 15 1–11

[26] Bansal T, Belanger D and McCallum A 2016 Ask the gru: Multi-task learning for deep text recommendations Proceedings of the 10th ACM Conference on Recommender Systems pp 107-114

[27] Le T D B and Lo D 2018 Deep specification mining Proceedings of the 27th ACM SIGSOFT International Symposium on Software Testing and Analysis pp 106-117

[28] Xing D, Xue G R, Yang Q and Yu Y 2008 Deep classifier: automatically categorizing search results into large-scale hierarchies Proceedings of the 2008 International Conference on Web Search and Data Mining pp 139-148

[29] Jung K, Zhang B T and Mitra P 2015 Deep learning for the web Proceedings of the 24th International Conference on World Wide Web pp 1525-1526

[30] Mukminin A and Riana D 2017 Komparasi Algoritma C4 . 5 , Naïve Bayes Dan Neural Network Untuk Klasifikasi Tanah J. Inform. 4(1) 21–31