JRSSEM 2022, Vol. 02, No. 4, 555 – 561
E-ISSN: 2807 - 6311, P-ISSN: 2807 - 6494

IMPLEMENTING NEURAL NETWORK ON DATA MINING TO PREDICTING KEY PERFORMANCE INDEX FROM EMPLOYEE

Albert Yakobus Chandra[1]
Putri Taqwa Prasetyaningrum[2]
Irfan Pratama[3]
[1,2,3]*Department of Information System, Faculty of Information Technology Universitas Mercu Buana Yogyakarta Indonesia*
[*]*e-mail: albert.ch@mercubuana-yogya.ac.id[1] putri@mercubuanayogya.ac.id[2] irfanp@mercubuana-yogya.ac.id[3]*
**Correspondence: albert.ch@mercubuana-yogya.ac.id[1]*

Submitted: 05 November 2022 *Revised*: 16 November 2022 *Accepted*: 25 November 2022

Abstract: Employee performance could help the company success to achieve its goals. In this era predicting employee performance is a necessity for companies to gain success. In this research, we presented a prediction of the employee performance index using a neural network. In this study, we are using the Cross-industry standard process for data mining (CRISP-DM) as the base framework for the data mining phase. The neural network classification method is employed to create the prediction model. The result shows that using a neural network could get a confusion matrix at 97.777 percent. This model was then applied to get insight from the performance index that showed a couple of factors is has an important role in employee performance.

Keywords: data mining; key performance index; neural network.

INTRODUCTION

Information technology is a new field that combines information science, computer science, telecommunications, and electronics. The evolution of information technology involved several steps leading to different developments (Chitechi & Otanga, 2020). It is well known that information technology affects humans in almost every field, including education, medicine, and especially in the world of work or employment. Information Technology is ubiquitous in the industrialized world (March & Smith, 1995). The use of information technology demands that employees adapt quickly to technology, which will obviously have different impacts on employees. The technology simplifies many functions, which enhances performance and improves professional satisfaction. The employee is an organization most valuable asset. As a result, employee performance has the greatest impact on organizational performance and ability to function. Some employees perceive these as opportunities and are more likely to use them as active job aids to improve job performance and satisfaction (Bala & Venkatesh, 2016). This condition indicates that when an organization decides to use information technology to improve operational performance. At this time, information restores vacant and outdated capabilities and contributes to the development of performance skill of employee. Good performance and employee efficiency are one of the keys to successful organizations (Rahmanidoust & Zheng, 2019). In order to determine the level of employee performance, we need to identify key performance. Key Performance Indicators (KPIs) are performance assessment tools that determine the extent of achievement of desired parameters in industrial production lines, that is of great significance to the success of the manufacturing company. Key performance indicators reflect departmental performance (Singh, 2015). To define this key performance this paper implementing neural network on data mining.

Artificial neural networks or simply "neural networks" as the connection model, the distributed parallel processing model, and the polymorphic neural system. Neural networks reappeared in the mid-1980s after major advances in neuroscience (Mohaghegh et al., 1995). The simply "neuron", and utilizes the massive computer's parallel elements to achieve high-performance speeds. Neural networks are often used for statistical analysis and data modeling, in which their role is perceived as an alternative to standard nonlinear regression or cluster analysis techniques (B. Cheng and D. M. Titterington, 1994). The terminology, they all try to borrow the structure and how the biological nervous system works based on our current understanding of it (Herzog & Almeida, 2018). Neural networks are widely used for effective data mining, transforming raw data into usable information. Data mining has fueled the research and development of methods and algorithms for manipulating vast amounts of data to solve real-world problems (Aggarwal, 2011). Data mining deals with the analysis of large and complex

databases, using machine learning and statistical techniques to discover new, useful and interesting knowledge. Data mining is still in its infancy, but companies in a wide range of industries, including retail, finance, healthcare, manufacturing, transportation, and aerospace, have historically already used data mining tools and techniques to use their data. increase. Data mining uses pattern recognition techniques and statistical and mathematical techniques to sift through stored information. Data mining analysts can identify important facts, relationships, trends, patterns, exceptions, and otherwise overlooked helps identify possible anomalies (Katiyaar & Sharma, 2012).

This paper is expected give solution for the company leader to identify the strengths and weaknesses in their performance. Once the company leader obtains Implementing Neural Network on Data Mining to define Key Performance Index, it will be easier to make decisions that will have a great impact on the organization or the company. Currently, the neural network is very suitable to solve data mining problems because of its good features, adaptive self-organization, parallel treatment, distributed storage, and high fault tolerance. Key Performance Indicators (KPIs) are generally used to measure the performance of the management process to recommend appropriate future directions (de Andrade & Sadaoui, 2017). The result using a neural network, on data mining it will be more efficient and the Key Performance Index has a maximum level of accuracy. Neural networks have high acceptance capacity for noisy data and high precision and are preferred in data mining (Krieger, 1996).

MATERIALS AND METHODS

This research following Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology. CRISP-DM act as framework that provide standardized steps on how to develop a data mining project. CRISP-DM has six phases which is: business understanding, data understanding, data preparation, modelling, evaluation and deployment [21]. In this research we are using CRISP-DM as framework and customize it to match the process and result we wanted on this research. The research phases diagram based on customize CRISP-DM as follow:

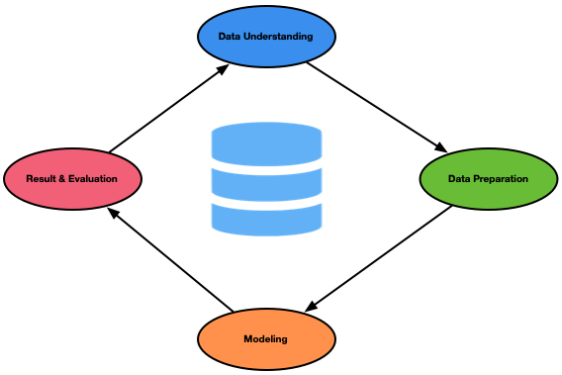

Figure 1. Research Phase Diagram

Data Understanding

The data from (*BRI Data Hackathon - People Analytics | Kaggle*, n.d.) is used as dataset in this research. There are total 2205 rows data in this dataset. The attributes in this dataset shown as below:

```
['job_level',
 'job_duration_in_current_job_level',
 'person_level',
 'job_duration_in_current_person_level',
 'job_duration_in_current_branch',
 'Employee_type',
 'Employee_status',
 'gender',
 'age',
 'marital_status_maried(Y/N)',
 'number_of_dependences',
 'number_of_dependences (male)',
 'number_of_dependences (female)',
 'Education_level',
 'GPA',
 'year_graduated',
 'job_duration_as_permanent_worker',
 'job_duration_from_training',
 'branch_rotation',
 'job_rotation',
 'assign_of_otherposition',
 'annual leave',
 'sick_leaves',
 'Best Performance',
 'Avg_achievement_%',
 'Last_achievement_%',
 'Achievement_above_100%_during3quartal',
 'achievement_target_1',
 'achievement_target_2',
 'achievement_target_3']
```

Figure 2. Attribute from dataset

Data Preparation

In this research, the data is prepared to be used on the process, start from handling the missing values, remove the duplicate rows and load the data to be processing on the modeling using neural network.

1. Modeling

In this phase the neural network algorithm is employed. The data contained numerical data and continuous data. The neural network multilayer perceptron was used to train the dataset and making the neural network classifier. Using python to process the dataset with neural network classifier we get accuracy for neural network is 97.777 confusion matrix.

```
Accuracy for Neural Network is: 97.77777777777777
Confusion Matrix
[[37  0  0  0  0  0  0  1  0  0]
 [ 0 32  0  0  0  0  0  0  0  0]
 [ 0  0 41  0  0  0  0  0  0  0]
 [ 0  0  0 32  0  1  0  0  0  0]
 [ 0  1  0  0 36  0  0  0  0  0]
 [ 0  0  0  0  0 38  0  0  0  0]
 [ 0  0  0  0  0  0 34  0  1  0]
```

Figure 3. Neural network confusion matrix

RESULTS AND DISCUSSION

With implementing neural network method on dataset, it shown classification data based on indicators that affect the performance of employees. The neural network models need to be refined so as to increase their performance. In choosing the neural network, considerations should be given to selecting a model based upon the prediction risk. Figure 4 displays performance the employee base on gender, it shown best female at work 0.127 and best male at work 0.12 It's seems that male are more productive than female. If we see on the chart that has shown male has got more productive than female.

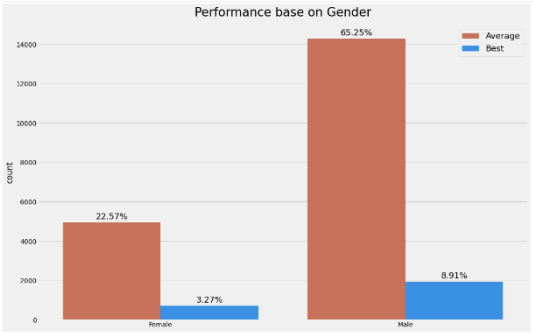

Figure 4. performance based on Gender

Figure 4 shown significant difference performance base on marital status, already married at work 0,139 and have not married at work 0.078. That significant difference, married people are more productive at workplace. Married people have a family that depending on them, this makes them more serious & productive at workplace, this make sense.

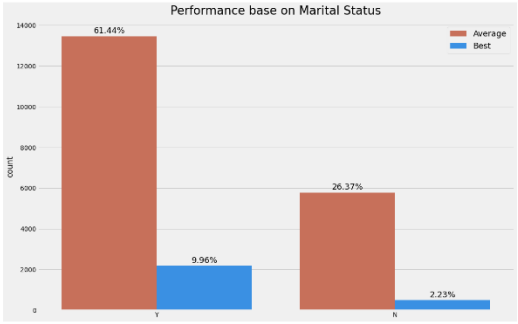

Figure 5. Performance base on marital status

Figure 3 shows performance base on division, this dataset we have three division. The employee on Mantri Kupedes have percentage 50.4 % , Mantri KUR 46.1% and the smallest percentage is 3.4% from Mantri Briguna. Dataset the employee on Mantri Kupedes have best performance mor than Mantri KUR and Mantri Briguna.

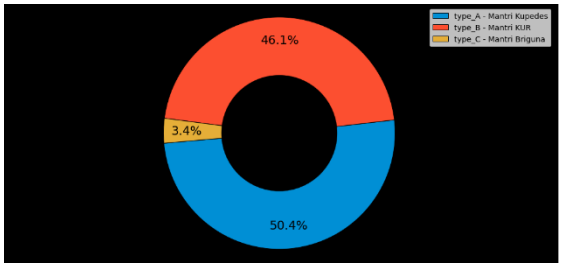

Figure 6. Performance base on type of division

Figure 6 shows surprisingly on performance base on employee status, that contract employee only have standard performance compared to best permanent employee 12.18%.

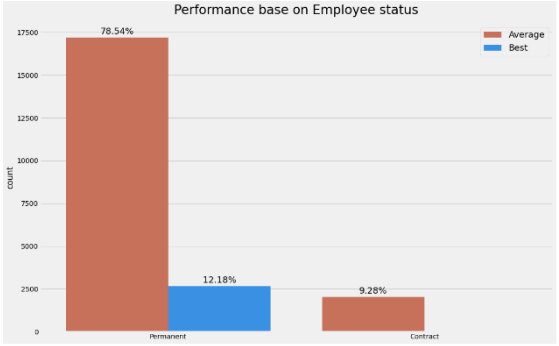

Figure 7. Performance base on Employee status

Figure 7 shows performances base on age with best age density above 0.12. Age attribute showed a positive effect on performance. This could be due to newly working employees have energy and high enthusiasm for work. On the other hand, older employees may have had much experience that would influence their performance. It was observed that employees between the age of 24 years and 45 years showed better performance.

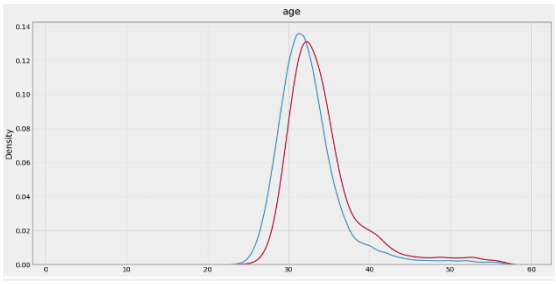

Figure 8. Performance base on age

Figure 8 shows best performance with density above 1.6. Employees without dependent prior to having one dependent get the best performance from the dataset. Because more and more dependents will make employees much more serious in pursuing their work targets to stay afloat.

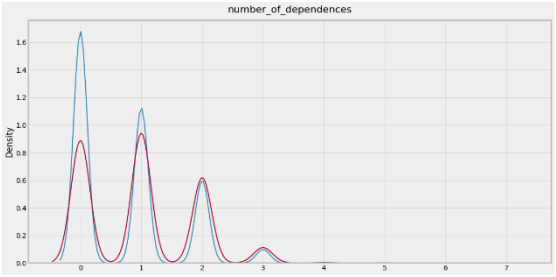

Figure 9. Performance base on number of dependences

Figure 9 shows Performance base on GPA best employee have 0.175 – 0.200, seems a reasonable predictor of effective job performance – a high GPA signals the individual has a considerable degree of competence.

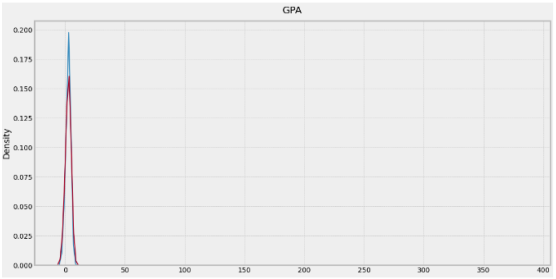

Figure 10. Performances base on GPA

Figure 10 shows The Training attribute had a slight effect on performance. The percentage below 12 months 11.9%, 12-24 months 14.2%, and above 24 22.4%. The employee who above 24 months professional training showed better performance compared to those below 12 months and 12 - 24 months. Generally, the number of employees who had outstanding performance after the training was higher than those below 12 months and 12 - 24 months.

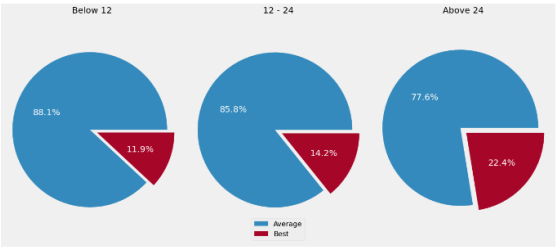

Figure 11. Performance base on training

CONCLUSIONS

In this research shown that from the dataset with using neural network classifier we cloud get insight from the data and doing exploratory data analysis (EDA) resulted from the neural network classifier. Couple insight has important role on employee performance index such as gender, marital status, GPA, division and training. Using neural network to process the data shown that this technique is very good at predicting key performance index from employee.

REFERENCES

Aggarwal, C. C. (2011). An introduction to social network data analytics. In *Social network data analytics* (pp. 1–15). Springer.

B. Cheng and D. M. Titterington. (1994). "RNA2. pdf,." *Statistical Science*, *9*(1), 2–54.

Bala, H., & Venkatesh, V. (2016). Adaptation to information technology: A holistic nomological network from implementation to job outcomes. *Management Science*, *62*(1), 156–179.

BRI Data Hackathon - People Analytics | Kaggle. (n.d.). Retrieved October 30, 2022, from https://www.kaggle.com/competitions/bri-data-hackathon-pa

Chitechi, K. V., & Otanga, D. A. (2020). *Current Trends In Information Technology: Which Way For Modern It Experts*.

de Andrade, P. R. M., & Sadaoui, S. (2017). Improving business decision making based on KPI management system. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1280–1285.

Herzog, H., & Almeida, P. (2018). Introduction to PET. In *Hybrid MR-PET Imaging* (pp. 129–146).

Katiyaar, V., & Sharma, I. K. (2012). Use of Data Mining & Neural Network in Commercial Application. *Vinodhini Katiyaar et al,/(IJCSIT) International Journal of Computer Science and Information Technologies*, *3*(3), 4041–4049.

Krieger, C. (1996). *Neural Networks in Data Mining*.

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266.

Mohaghegh, S., Arefi, R., Bilgesu, I., Ameri, S., & Rose, D. (1995). Design and development of an artificial neural network for estimation of formation permeability. *SPE Computer Applications*, *7*(06), 151–154.

Rahmanidoust, M., & Zheng, J. (2019). Evaluation of Factors Affecting Employees' Performance Using Artificial Neural Networks Algorithm: The Case Study of Fajr Jam. *International Business Research*, *12*(10), 86–97.

Singh, K. (2015). Leadership Style and Employee Productivity: A Case Study of Indian Banking Organizations. *Journal of Knowledge Globalization*, *8*(2).