A possibility of artificial neural networks to be applied in the predictive test: A systematic literature review and study case

A Asmara[1]

[1] Technological and Vocational Education, College of Humanities and Applied Sciences, National Yunlin University Science and Technology

Email: andikasmara@uny.ac.id[1]

Abstract. Artificial Neural Network (ANN's) was developed continuously and applied to each smart technology. The developing smart tests are continuous with ANN's possibility applied to increase efficiencies and time of decides. The Predictive Test was developed using the Smart Algorithm with the possibility of ANN to apply. Assessments of ANN implementation have been carried out based on a study in past time. The systematic literature review was used to found possibility implementation ANN's in the predictive test and combined with the study case. The result showed as ANN's could implement in the predictive test that used big data inputs. In addition, ANN's can be implantation on the predictive test such as education predictive test, talent and aptitude test or psychological test.

1. Introduction

Developing smart technology is increase toward Artificial Intelligence (AI) implementation, which one part of AI is Artificial Neural Networks (ANN's) which was started in 1943 by Warren McCulloch. He was neurophysiologist, and Walter Pitts wrote a paper explained how neurons might work [1]. Neural Networks was continues developed until now with are optimizing decision times and ability to receive many of input. For predicting output decision, ANN's are implementation is in the computerized system. This system is a branch of AI representative of the human brain behavior that learns jobs or tasks from experience data [10]. The ANN's structure is formed from the processing unit called neurons, which are analogous to biological neurons, and have a connection weight between neurons [14]. The experience data are input connected with neurons systems, being considered a Machine Learning (ML) method.

ANN's are consists of three part, such as Input layer, an output layer, and an activation function layer. The ANN's development model involves the selection of network variable, determining the network structure, and training – testing – validation procedure [14]. Input layer usually is many nodes and form a unique pattern, while the output layer is one of the nodes and two decisions minimum from the activation system. Activation layer where learning system, finding the weight (w) value and output decision. On the learning system, ANN's has the running process used several methods. Asmara A. [2] describe ANN's have several learning methods to founded weight value from each's of input, which one method is called name "Perceptron". Learning systems are critical from ANN's because to take a decision based on this system.

Input layer from ANN's that receive raw information and feedback network. Input nodes are independent variables that collectively affect the output parameter value. The information collected at

the input nodes represents the major condition problem. An output layer represents the response of the network to given input conditions. ANN's system has a hidden layer that connecting the input layer and output layer, and the process depends on the activity of the input layer and condition value or weights. The input, hidden, and output layer nodes are interconnected by tuned connection weight to recognized different problem or pattern of information. In the process, learning until the recognized pattern involves determining network parameters, such as connection weights, threshold values, and an optimum number of hidden nodes [14].

ANN's is a robust system to process multiple inputs and then provide prediction result between input tests and learning weight. David O.S. [5] described that ANN's is a powerful data-driven and has the capability to capture non-linear and complex data as flexible computational tool processing. Neural Network created an internal map that aids with discrimination and can accurately produce predictions [3,12,7]. This system is a soft computation that created to processing multiple data inputs with a unique pattern. The one can predict the result of output based on input learning and decision program. Usefully and robust system in the artificial intelligent to predicting the future decision.

Into components of the education system consist of activity teaching-learning, assessment or test, and synchronize of skill needs with the curriculum. Major discussion based experiences in this study are smart to test development with implement artificial intelligent in the computing system. Focus on the review of possibility artificial intelligent to implement on the predictive test. The predictive test is ones of the test method to know future potentially student to be a success. This test cannot be impacted directly after doing a test is finish, but the result of the impact in the future known. Explanation of type of test this called name is "predictive test".

Use predictive test in education to know student potentially on ones of interest or talent. Implementation predictive test can be found on the physiological test, basic capable test, and talent or aptitude tests. All of them to predict of the respondent for collected potential factor that impact in the future. However, developing test now involved computer to changes of the paper test, and also give more advantages such as speed for the decision, data bank and combined with a smart algorithm or artificially intelligent. Lawson A.M [9] said that computer-based testing has increased in popularity over recent years for a various compiling reason. Which one in progress development is a predictive test for several fields? The predictive test has many data input to produces one decision result. Therefore, a simple algorithm is not supported, so need a neural network to solve for this problem.
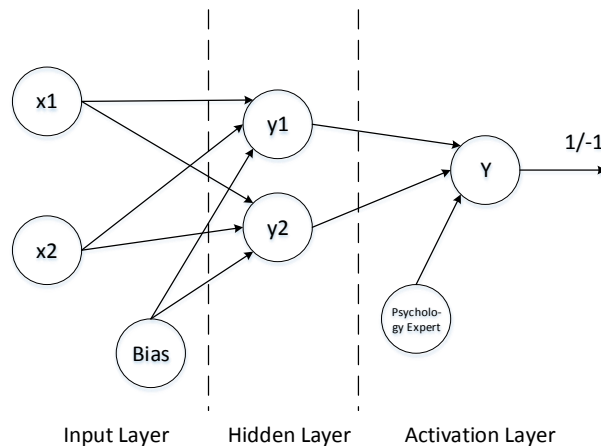
The ANN's are using to solve the problem with complexity formulation or hidden layer that is math algorithm and the problem to solve completely not realized. The system is used available data for collected pattern related to the differentiation of the group. This ability is namely as artificial intelligence because of likely human neuron of the brain. Artificial intelligent is a step in the brain of a human, and furthermore, have the ability to predict human behavior.

Based on the description above, this study was collecting information about the possibility of ANN's implementation of the predictive test. The predictive test is educational pretest before a student decides to enroll in the school, such as aptitude test, and talent test. Collecting supported data for gaining's perception implementation ANN's used literature review method.

1.1 Previous study

Based on experiences in the previous study, the author in this study had experience in buildup predictive test with an implement of ANN's. Model of ANN's "Perceptron" was used to implement in algorithm programing predictive test. Namely of the test software was developed is "Interest and Talent Test" in the robotic field. More detail or describe the construction of the predictive test in the bellow.

ANN's with perceptron type was used in the previous study. This system has network architecture as follows:

Appositions:

x1 = Input node for "Interest" question

x2 = input node for "talent/aptitude" question

Y = Decision/Activation

Input node in the form of interest and talent questions consist of two big group topics such as mechanic- electronic and mathematic-computer. Topics were developed to 45 questions with detail as follows:

Table-1. Scope of question

Type	Indicator	Sum of question
Interest	Mechanical fields	3
	Electronic fields	5
	Programming/computer fields	4
	Mathematics fields	3
Talent	Math capability	5
	Programing language capability	6
	Basic 2D/3D modeling capability	11
	Robotics capability	8
Sum		45

After developing questions, a further step was developed software within which ANN's algorithm. The software was developed using Visual C programming and have three part an important, such as human interface, ANN's algorithm, and machine learning, and data based. The human interface in the application of Interest and Talent Test has built considering user-friendly and humanism. The human interface in these applications as following;

Figure 1. Interest and talent test interface

While for ANN's algorithm and machine learning are programing language which builds to learn and save pattern learner. The output from machine learning is several weights in each way on a neural network system. To be obtained of weight, ANN's was learner used several patterns from experiences of the respondent learner, namely is "Target of the Pattern". Several targets of the pattern showed as following;
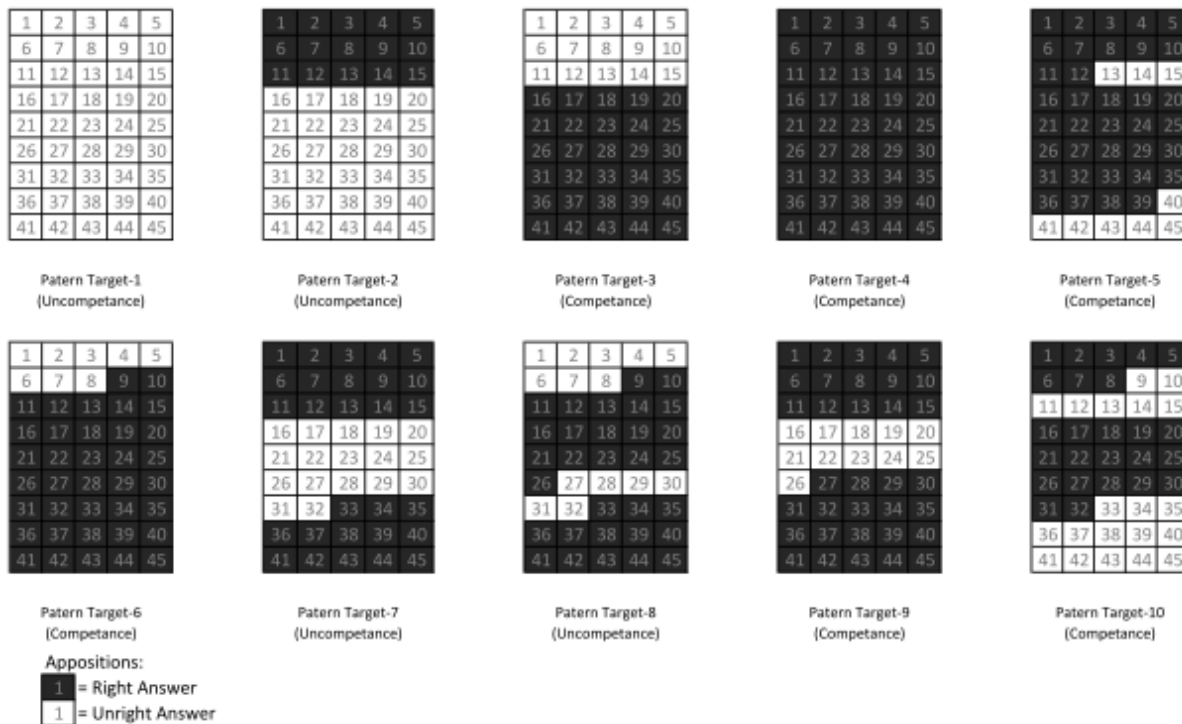

Figure 2. The target of pattern

After the system was finished and ready to implement, the next step is software testing. Software testing was used white box testing method and then does the black box testing method. The expert judgment was involved in the testing of this system. Consist of four expert people in the fields related, such as an expert on the psychology, expert on the education measurement, expert on the ANN's, and expert on the Robotics. Furthermore, testing used respondent divide two groups that are limited testing and final testing.

The results were showed perceptron model of ANN's which developed have two activations (bipolar) model that is competence and incompetence in the robotic field. Grade of the measurement resulted in application proper to the beginning project of implementation ANN's in predictive test. However, machine learning was limited learner that is used 10 targets of the pattern. Next to future study need's to learning more than 10 targets of pattern to have a better result.

2. Method

The systematic Literature review was the major approach for this study. This method to increased result being adopts in the social field and science [4]. In addition, reduced subjective bias and risk of overlooking literature in a reliable and rigorous process [6]. The data were collected from several journals taken from web-science, with years of range from 2015 until 2019. The keyword is predictive, Artificial Neural Networks, and ANN's. Use the data for supporting possibility ANN's implementation in predictive test on the education field. The conceptual framework in this study as following;
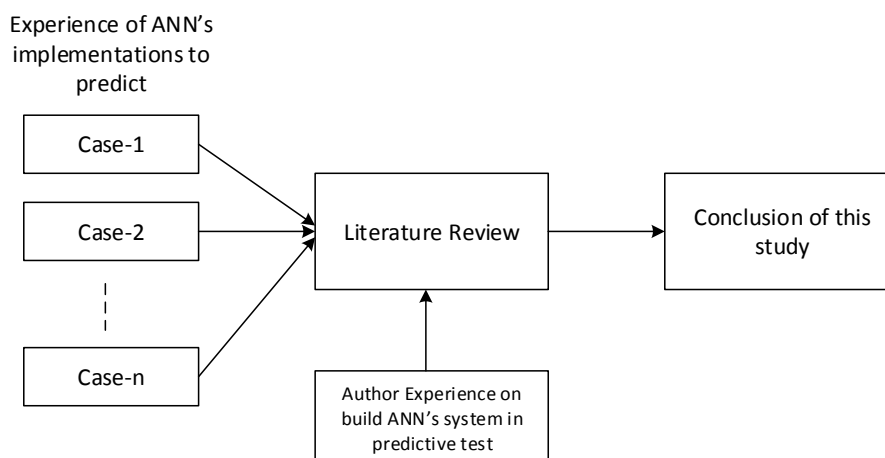
Figure 3. Conceptual framework study

Every case (case-1, case-2 until case-n) based on experience in deferent implementation fields. However literature review focus on finding success experience to implementation ANN's. Besides, the finding excludes method or type are used.

The first case from Machado C.F [10] with the title is "A catalyst selection method for hydrogen production through Water-Gas Shift Reaction using artificial neural networks". The Water-Gas Shifts (WGS) reaction mainly routes for the production of hydrogen and depending on the operating process condition in the used different catalysts. In order to predict optimal catalyst composition for the WGS reaction, ANN's were used to build a model from literature catalysts data. The model ANN's has used three-layer feedforward neural network with active phase composition and the input variable was supported type as some, and Carbon Monoxide (CO) conversion as the output variable. Three-layer neural network architecture drawn as following;
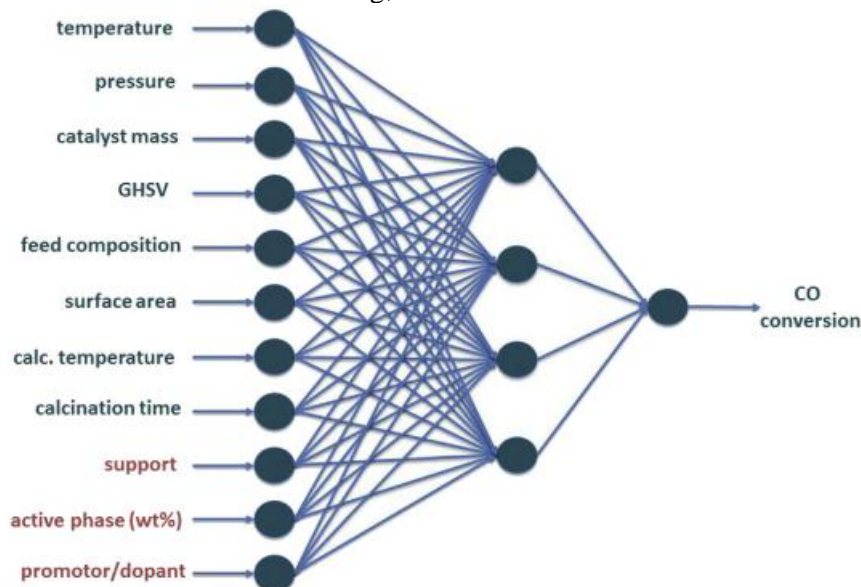

Figure 4. The architecture of Three-layer Feedforward Neural network [10]

The number of neuron in the hidden layer is determined to use partitioned 10 times the data in different training and testing sets for each NH value considered. The ANN's developed contributes to analyzing the past of Water-Gas Shift reaction catalysis to design new catalysts for a more efficient process. The results from this study were shown the power of ANN's for prediction better catalysts and

conditions for this vital process in the environmental field. Furthermore, ANN's was possible to predict some of the most relevant variables that influence CO conversion that is the temperature and surface area. The novelty in relation to previous similar ANN's studies of the insertion of this latter variable into the model was essential for its good predictive performance.

The Second case from Jifei Deng [8] with the title is "Application of neural networks for predicting hot-rolled strip crown". The problem of strip crown defect is always serious, and strip with an undesirable crown control has always been a daunting task. Factor influencing the performance of strip can be divided into rolling conditions and material parameters. Production data of 10,133 coils collected from a hot rolling plant are used to establish models, considering the mean square error (MSE) and correlation coefficient (R). A non-dominated sorting genetic algorithm II (NSGA II) optimized ANN, and deep neural network (DNN) is applied to evaluate the prediction performance. The model of ANN's was used in this study is Backpropagation using a gradient descent algorithm. For ANN's developed in this paper, the neurons of the hidden layer, the learning rate, the activation function of the hidden layer and output layer, and the training algorithm are seriously important.
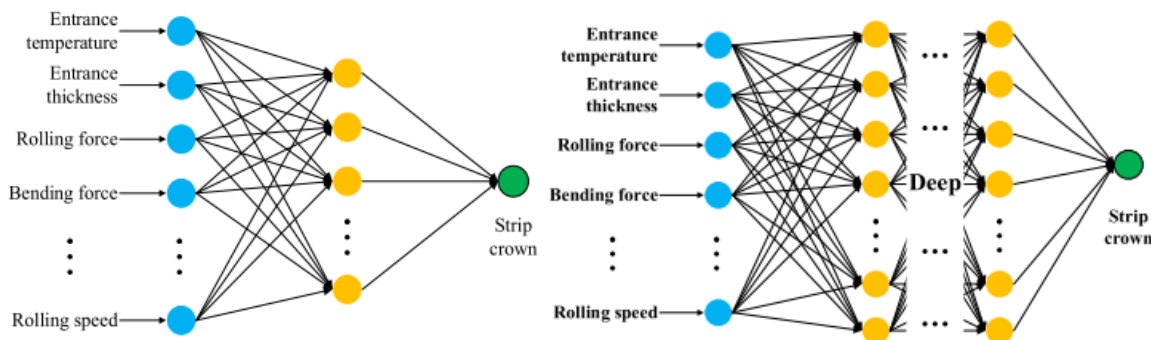

Figure 5. The architecture of ANN's (left side) and DNN's (right side) Model [8]

The result was 83 data points have been removed and 10.050 coils area selected for examination, and totally are 10.133 pairs of coil data. In the modeling process, the initial thresholds (activation) and weight of the ANN are randomly selected bay software, and the training process is undertaken to update them, to achieve the optimal model. The end of this study were several important results, such as first, the prediction accuracy of the models can be significant influences by different hidden neuron, learning rate, parameter setting, training function, and hidden structure of ANN's or DNN's. In addition, in the most recommended algorithm in this study was DNN's. Second, using the parameter settings obtained from the proposed model, and then the prediction of strip crown is close to the target value. Furthermore, the result is the DNN's are proved to predict strip crown effectively. The DNN's method has strong generalization and reliability, and it is worth applying in hot rolling lines for predicting strip crown.

The Third case from Sinshaw T.A. [13] with the title is "Artificial Neural Network for Prediction of Total Nitrogen and Phosphorus in US Lakes". The important aspect of water quality management is modeling because it saves material and labor costs. The nonlinearity of water quality variables due to the complex chemical and physical processes in the body of water. The ANN's was a powerful computational tool for nonlinear relationships, to develop a model that estimates the summer concentration, with detail as total nitrogen (TN) and total phosphorus (TP) in US lake using interrelated and easily measurable water quality parameter. Three inputs (pH, Conductivity, and turbidity) were trained, tested, and validating in ANN's model, and that is statistically correlated with the output. Dataset downloaded from USEPA is 1.217 record data sets from approximately 1.000 US lake. In addition, the data sets are used as input from ANN's system. The network was trained using the TR-SEQ1 program (regional and national data sets), which was built on the back-propagation algorithm (Najjar, 1999).
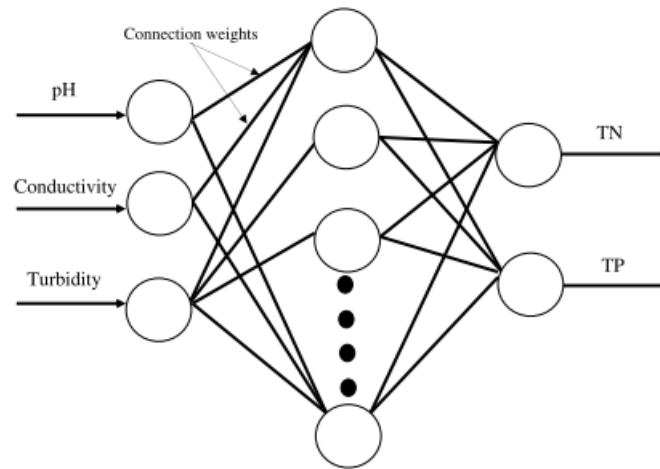
Figure 6. The architecture of a feed-forward back-propagation neural network [13]

The result of this study was the ability of ANN's model to predict TN and TP concentrations using easily measurable water quality variables. The name of the variable are pH, conductivity and turbidity were explore and justified in the case of US lakes during summer months. The ANN's models were found to be suitable to estimate TN and TP concentration in both the national and the regional data sets compared with counterpart linear regression models. The purposed ANN's model is a useful alternative tool on the laboratory analysis of TP and TN for lake in the US in the summer season.

The Fourth case from Wood D.A [15] with the title is "Transparent open-box learning network provides auditable predictions". The algorithm was applied here to predict the pool boiling heat transfer coefficient (PBHTC). The material to detection are alumina-water-based Nano-fluids from a large dataset, that is 870 data records with PBHTC varying from 0.33 to 65.68 kW m-2 K-1. The dataset, compiled from a published source and listed in full in a supplementary file, involves four relatively easy-to-measure input variables, such as Nano size, operating pressure, nanoparticle concentration in water, and excess temperature. Called TOB, that means transparent open-box, this is learning network algorithm adds the useful dimensions to machine learning of audibility and interrogation of each prediction made. The methods are making available for instant inspection the exact calculations and relationship it applies to its prediction for each data records in data sets. The ANN's model was construct based on MATLAB model and then tuner to generate PBHTC predictions for the compiled dataset. Toward minimization of the cost, function exploits the gradient descent multilayer perceptron method. This involves a backpropagation algorithm applied to train the network.
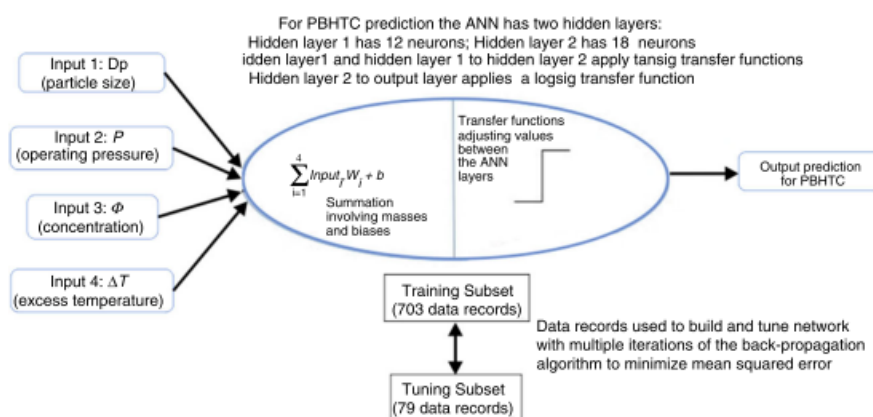

Figure 7. The architecture of ANN's applied to the prediction of Pool-boiling Heat Transfer Coefficient (PBHTC) [15]

From this study, to predict the pool boiling heat transfer coefficient (PBHTC) from four poorly correlated and unevenly distributed input variable was used Transparent Open-Box (TOB) learning network algorithm. The method used in this study has provided a credible and accurate prediction of PBHTC. It has two stage process approach describes as stage-1 machine test data records to the training subset, and stage-2 optimizing the contribution of the high ranking matching records to fine the prediction. In addition, the process is shown to be effective and transparent.

3. Result and discussion

The ability to predict the phenomenon in the future with many data input (dataset) has ANN's owned. This similar to the literature review has explained result in their studies that ANN's prediction has been carried out successfully using data set [5,16,17], and addition used three key steps such as optimum value of neuron, appropriate training algorithm, and trial-validation of the model [5]. Leaning and training process needed to recognize the event as usual or repeat the process. Furthermore, the algorithm used the weight of training to processing dataset and then produced specific value which next step to decide in the activation layer. The vital step in ANN's process is training with many datasets to provide the weight value of the neuron network. Value of weight depends on the pattern or dataset characteristics [18].

Based on several of the literature reviews can be drawn finding that as following;

a. ANN's to process many input or dataset

   The datasets are formed information's about characteristic to influence the next data or future data. Datasets in literature reviews is a set of data with each variable characteristics, which each group have a pattern or characteristic. Training used this pattern (characteristic) to produces weight in each neurons network.

b. Used computer programing to running the algorithm

   The algorithm of ANN's is mathematic formulas which working repeatedly. As a concept, ANN's algorithm has a similar principle on the build the program. Computer programing have many kinds of languages that can be used, such as C, C++, Visual C, Visual Basic, and MATLAB more easily. Although each ANN's programming developed used different programming languages, the purpose of development is similar, that is to be user-friendly and easy to use for prediction.

c. The whole process ANN's are a cross-training or learning process

   The kinds of ANN's such as perceptron, backpropagation, and other deep learning have one process which a similar that is training systems. Other journal described used term of training, and other side used the term of the learning process. The meaning of the both is similar that is pattern recognizing to produce weight each neuron network. The training system using many dataset or input, come from phenomena or scientific data was collected.

d. The whole conclusions have shown success in predict

   Based on the literature review, in conclusion, each study describes successful in predict. The prediction is producing empirical data or decision to triggering the next action.

The author experiences in the previous study have a similar process and propose with the whole of the literature review. Purpose of these studies is developed predict system for big data input (dataset). The application author developed is a predictive test, which called 'interest and talent test" have a lot of data set, consisting of a group of 45 questions, and each respondent has 10 patterns. In addition, in the future study developing question to predict interest and talent more than 45 patterns. The dataset has similar with dataset in the whole literature review, that is fist literature have 11 variable, second have 10.133 coils dataset, third have 1.217 dataset and last, have 870 datasets. It has shown a group question on the test with a certain pattern can be used as learning (training) in ANN's algorithm.

Ability to predict based on training has shown on predictive test interest and talent, also forth literature review. Although different purpose on the subject of prediction, as the principle from the beginning until deciding on the result is similar. Start from a collected dataset, programming algorithm, training process, and then getting a result from the predicted process.

4. Conclusion

Possibility to predicted of result in the future has been own by ANN's algorithm. It is can be helping the user (human) to facilitate the process of many datasets. The ability to learning from the training of dataset makes ANN's are powerful in the prediction. This ability not only used in the engineering predicting data but can implementation on education, especially on the smart test. In the finding and discussions showed to have a similar process between ANNs used to prediction in the engineering field and used to prediction in education, called name is a predictive test.

The similar process in predictive test starts from developing a question to collect learning dataset, developing algorithm programming, collecting pattern of learning dataset, processing learning data set and then in the final is trial on the subject.

5. References

[1] _____ (access on 2019), https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history1.html

[2] Asmara, A., & Haryanto, H. (2015). Pengembangan tes minat dan bakat dengan metode jaringan syaraf tiruan (JST) untuk memprediksi potensi siswa bidang robotika. *Jurnal Pendidikan Vokasi, 5*(3), 273-286. doi: https://doi.org/10.21831/jpv.v5i3.6483

[3] Balderas-Silva D., 2018. Convolutional Long Short Term Memory Deep Neural Networks for Image Sequence Prediction. doi: https://doi.org/10.1016/j.eswa.2018.12.055

[4] Burrows, R. (2000). Methodologies for Socially Useful Systematic Reviews in Social Policy. End of Award Report, ESRC.

[5] David O.S. (2018) Comparison of Response Surface Methodology (RSM) and Artificial Neural Network (ANN) in modelling of waste coconut oil ethyl esters production

[6] Elaine V., Thais, & Robertson, Maxine (2017). A systematic literature review of UK university-industry collaboration for knowledge transfer: A future research agenda. Science and Public Policy, 2017, 1–12. doi: 10.1093/scipol/scx086

[7] Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

[8] Jifei Deng (2019) with title is "Application of neural networks for predicting hot-rolled strip crown"

[9] Lawson A.M (2018). Reliability and validity of the revised Gibson Test of Cognitive Skills, a computer-based test battery for assessing cognition across the lifespan. Psychology Research and Behavior Management 2018:11 25–35

[10] Machado C.F. 2019. A catalyst selection method for hydrogen production through Water-Gas Shift Reaction using artificial neural networks. Journal of Environmental Management 237 (2019) 585–594

[11] Najjar, Y. M. 1999. Quick manual for TR-SEQ1. Manhattan, KS: Dept. of Civil Engineering, Kansas State Univ.

[12] Schmidhuber, J., 2015. Deep learning in neural networks: An overview. Neural Networks 61, 85 – 117. URL http://www.sciencedirect.com/science/article/pii/S0893608

[13] Sinshaw T.A. (2019) with title is "Artificial Neural Network for Prediction of Total Nitrogen and Phosphorus in US Lakes

[14] Tadesse A.S. 2019. Artificial Neural Network for Prediction of Total Nitrogen and Phosphorus in US Lakes. J. Environ. Eng., 2019, 145(6): 04019032

[15] Wood D.A (2018). Transparent open-box learning network provides auditable predictions. Journal of Thermal Analysis and Calorimetry. https://doi.org/10.1007/s10973-018-7722-9

[16] Xie, Yunyun. (2018). Predicting lightning outages of transmission lines using generalized regression neural network. Applied Soft Computing Journal. https://doi.org/10.1016/j.asoc.2018.09.042

[17] Yang, Yuanchao. (2018) Study on tire-ice traction using a combined neural network and secondary development finite element modeling method. Concurrency Computat Pract Exper. 2018;e5045. https://doi.org/10.1002/cpe.5045

[18] Y.D. Koo (2018). Nuclear reactor vessel water level prediction during severe accidents pusing deep neural networks, Nuclear Engineering and Technology, https://doi.org/10.1016/j.net.2018.12.019