Journal of Social Research

P-ISSN: 2827-9832
E-ISSN: 2828-335x
http://ijsr.internationaljournallabs.com/index.php/ijsr

COAL SPRAY RATE PREDICTION BASED ON FACTOR ANALYSIS AND *NEURAL NETWORK* (NN) ALGORITHM

Hasni Kasim, Muh. Yusuf, Haslinda, Rachmat, Muh. Fahmi Basmar
Universitas Pejuang Republik Indonesia (UPRI)
hasnikasim091@gmail.ccom, yusufsanrang@gmail.com, haslindahs09@gmail.com, rachmat27udinus@gmail.com, fahmi.basmar@gmail.com

ABSTRACT

Given the many factors that affect coal bursts as well as complex nonlinear relationships, this study analyzes the main factors that affect coal bursts in coal seams. Central Sulawesi Province. Using NN algorithms to predict coal bursts. This algorithm uses factor analysis to subtract the attributes of the original high-dimensional sample and obtains three common factors to maintain the correlation characteristics of 82.227% of the original data that had 10 sets. factors affecting coal burst rates can be used as training datasets as well as the last five datasets used as testing data and inputted into the rapid miner application to support NN algorithms to predict results. By comparing the prediction results where the NN algorithm has a non-explosion threshold value of 3.194% and an explosion threshold value of -3.230% and is more suitable for predicting coal burst explosions.

Keywords: *NN algorithm, coal, factor, rapidminer, threshold*

INTRODUCTION

One that is very threatening to work safety in the mining world is where explosions or coal bursts are very destructive (Wang et al., 2014). With the increase in mining intensity and mining depth, work accidents due to coal bursts become more numerous and very dangerous (Li et al., 2021). Therefore, how to predict coal bursts accurately and effectively is a very important problem to be solved in improving work safety in coal mines (Singhal et al., 2020).

In recent years, scientists and researchers both from home and abroad have studied many methods to predict coal bursts, this is very important for advanced researchers to do through several literatures, especially using single-index, multi-index, and comprehensive index methods to predict risks due to coal bursts and determine the risks caused by coal bursts (Wu et al., 2020).

In determining the decision-making model based on the gray theory system of coal bursts whether they are explosions or not according to the visibility between the critical value of the index from the center of the explosion target or coal burst (Perez, 2003). Backpropagation (BP) is used as a model to predict the risk of coal bursts in work surface areas and mine areas that can pose occupational safety risks to employees (Al Shibli et al., 2018).

However, the index analysis method still has many problems, such as difficulty in performing high linear operations and being influenced by human factors (Wiegmann & Shappell, 2017), therefore the gray system theory method cannot work effectively to extract coal and its causative factors. BackPropagation (BP) neural networks have high data corruption where requirements and performance are generally very low. In addition, most of the prediction methods above remain at a static level with a low level of intelligence because they fail in the relationship between variables and need to be increased in value to predict accuracy (Al-Sahaf et al., 2016).

The purpose of the above study uses a reduction model, extracting with a high level of correlation, which is based on a generally high level of good performance and efficiency from a small sample using the Neural Network (NN) algorithm, where the relationship is not linear with causative factors and coal bursts is quite effective.

METHOD

Analysis Factors

Factor Analysis is a method that replaces a number of large variables with several factors as a whole, this method serves to reduce several multi-variant spatial dimensions linearly and general factors obtained and does not depend on each other, has the ability to reflect most of the information from a variable (Yong & Pearce, 2013). Complex variables or factors can be reduced to a few common factors and can simplify the structure of neural networks to increase the efficiency of computational theory and improve its accuracy prediction (Jhony & Firdaus, 2020).

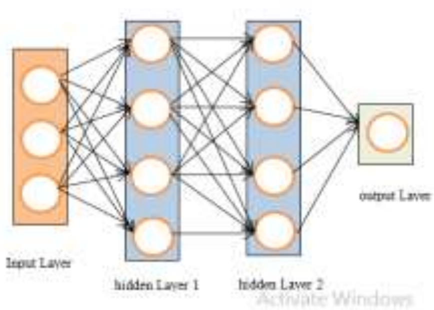
Neural Network (NN) Algorithm

Neural Network (NN) is a development of Multilayer Perceptron (MLP) designed to process two-dimensional data. NN is included in the type of Deep Neural Network because of the high depth of the network and is widely applied to image data. In the case of image classification, MLP is less suitable for use because it does not store spatial information from image data and considers each pixel to be an independent feature that produces poor results (Singhal et al., 2020).

NN was first developed under the name NeoCognitron by Kunihiko Fukushima, a researcher from NHK Broadcasting Science Research Laboratories, Kinuta, Setagaya, Tokyo, Japan. The concept was then matured by Yann LeChun, a researcher from AT&T Bell Laboratories in Holmdel, New Jersey, USA. The NN model under the name LeNet was successfully applied by LeChun to his research on number recognition and handwriting. In 2012, Alex Krizhevsky with his NN application won the ImageNet Large Scale Visual Recognition Challenge 2012 competition. This achievement is a moment of proof that Deep Learning methods, especially NN (Zhang & Kabuka, 2021).

NN Model Concept

The way NN works has similarities to MLP, but in NN each neuron is presented in two dimensions, unlike MLP where each neuron is only one dimensional.



Picture. 1. Simple MLP Architecture

An MLP as in Figure. 1. It has an i layer (red and blue squares) with each layer containing ji neurons (white circle). MLP accepts one-dimensional input data and propagates that data on the network to produce output. Each connection between neurons on two adjacent layers has a one-dimensional weight parameter that determines the quality of the mode. In each input data on the layer, a linear operation with an existing weight value is performed, then the computational results will be transformed using a non-linear operation called the activation function (Yu et al., 2015).

In NN, the data propagated on the network is two-dimensional data, so the linear operation and weight parameters on NN are different. In NN linear operations use convolution operations, while weights are no longer one-dimensional, but are four-dimensional which is a collection of convolution kernels as in Figure 2.

The weight dimensions on NN are:

X neuron input neuron output x
Height x *width*

Due to the nature of the convolution process, NN can only be used on data that has a two-dimensional structure such as image and sound.

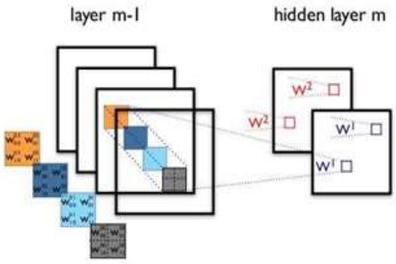


Figure.2. Convolution Process on NN

NN Network Architecture

JST consists of various layers and several neurons on each layer. These two cannot be determined using definite rules and apply differently to different data.

In the case of MLP, a network without hidden layers can map any linear equation, while a network with one or two hidden layers can map most equations to simple data.

But on more complex data, MLP has limitations. In the problem of the number of hidden layers below three layers, there is an approach to determine the number of neurons in each layer to approach optimal results. The use of layers above two is generally not recommended because it will cause overfitting and significantly reduced backpropagation strength (Chen et al., 2019).

With the development of deep learning, it was found that to overcome MLP's shortcomings in handling complex data, functions are needed to transform input data into a form that is easier for MLP to understand. This triggers the development of deep learning where in one model is given several layers to transform data before the data is processed using the classification method. This triggered the development of neural network models with the number of layers above three. However, due to the function of the initial layer as a feature extraction method, the number of layers in a DNN does not have universal rules and applies differently depending on the dataset used (Zeng et al., 2021).

Because of this, the number of layers in the network and the number of neurons in each layer are considered hyperparameters and optimized using a searching approach. An NN consists of several layers. Based on the LeNet5 architecture, there are four main layers in an NN but only three layers are applied to this TRAN:

The purpose of convolution of image data is to extract features from the input image. Convolution will produce linear transformations of the input data according to the spatial information in the data. The weight of the layer specifies the convolution kernel used, so the convolution kernel can be trained based on input to NN (Bernadó-Mansilla & Garrell-Guiu, 2003).

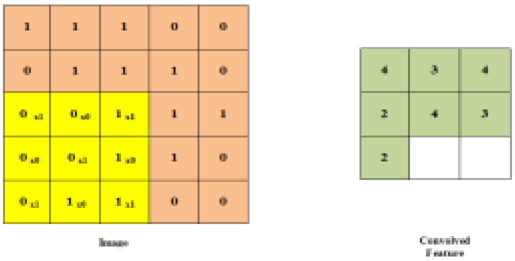
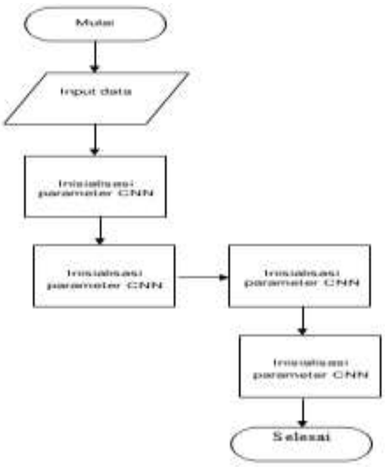


Figure.3. Convolution Operations

SOFTWARE DESIGN

Training Process

The training process is the stage where NN is trained to obtain high accuracy from the classification carried out. This stage consists of a feed forward process and a backpropagation process. To start the feedforward process requires the number and size of layers to be formed, the size of the subsampling, the vector image obtained. Where the vector image will go through a process of convolution and Max pooling to reduce the size of the image and multiply its neurons. So many networks are formed which add variant data to be studied.

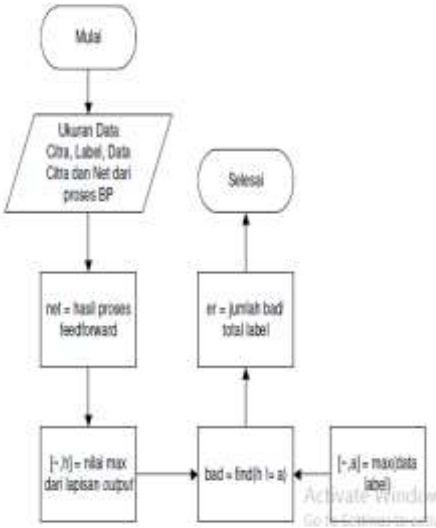


Training Process Flow

Testing Process

The testing process is a classification process using weights and biases from the results of the training process. This process is not much different from the training process which distinguishes that there is no backpropagation process after the feedforward process. So that the end result of this process produces the accuracy of the classification performed, the data that failed to be classified, the image number that failed to be classified, and the form of the network formed from the feedforward process.

With new weights and biases, a feedforward process is applied which then produces an output layer. The output layer is fully connected with the labels provided. The fully connected results obtained data that failed and were successfully classified.



Testing Process Flow

Measurable sample data of factors affecting coal bursts obtained from datasets

Table 2.1 Sample Data

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	Whether the outburst
1	5,2	1,2	357,8	21,64	9,2	1,6	19,2	1,58	2,58	2	outburst
2	9,2	1,9	361,9	26,24	9,8	1,77	22,6	1,71	2,39	3	outburst
3	6,3	1,32	208,32	16,65	7,9	1,53	17,6	1,78	2,21	2	outburst
4	5,4	1,41	211,44	16,5	6,8	1,23	9,78	1,98	2,13	2	not outburst
5	6,5	1,42	305,2	11,51	8,2	1,52	10,56	1,99	2,35	3	not outburst
6	7,2	1,36	265,45	16,5	7,9	1,48	11,07	2,06	2,42	3	not outburst
7	10,2	2,56	378.92	31,44	11,2	1,96	22.03	1,46	1,56	3	outburst
8	8,7	1,98	365,75	28,76	10,3	1,82	20,3	1,68	1,63	3	outburst
9	6,3	1,32	208,32	16,65	7,9	1,53	17,6	1,78	2,21	2	outburst
10	5,4	1,41	211,44	16,5	6,8	1,23	9,78	1,98	2,13	2	not outburst
11	6,5	1,42	305,2	11,51	8,2	1,52	10,56	1,99	2,35	3	not outburst
12	10,2	2,56	378.92	31,44	11,2	1,96	22.03	1,46	1,56	3	outburst
13	8,7	1,98	365,75	28,76	10,3	1,82	20,3	1,68	1,63	3	outburst

RESULTS AND DISCUSSION

Preprocessing

The initial process before testing the proposed method, namely the Neural Network preprocessing stage. The preprocessing stage used in this study is data normalization. Using the formula equation (3.1), namely:

$$\hat{x}_{ik} = \frac{x_{ik} - \min(x_k)}{max(x_k) - \min(x_k)} \qquad \hat{x}_1 = \frac{57056-57056}{225084-57056} = 0$$

Neural Network (NN) Algorithm Design

By using rapidminer application by using Neural Network (NN) design algorithm

Figure 4.1 Neural Network

Test Results with Neural Network (NN)

Figure 4.2 NN Algorithm Design results

Table 4.1 Outburst (Sigmoid) Class

No	Node	Value	Thershold
1	Node 1	1.602	
2	Node 2	0,379861	
3	Node 3	1.285	
4	Node 4	-0.661	**-3.230**
5	Node 5	0,6125	
6	Node 6	1.277	
7	Node 7	2.006	
8	Node 8	1.917	

Where in Table 4.1 Class Outburst obtained a Threshold value of -3.230% in the coal explosion class.

Table 4.2 Class no outburst (Sigmoid)

No	Node	Value	Thershold
1	Node 1	-1.637	
2	Node 2	-0.484	
3	Node 3	-1.332	
4	Node 4	0,5	**3,194**
5	Node 5	-0.831	
6	Node 6	-1.322	
7	Node 7	-1.980	
8	Node 8	-1.904	

Where in Table 4.2 Class no outburst obtained a Threshold value of 3.194% in the class where there is no coal explosion.

CONCLUSION

Research on Prediction of Coal Spray Rate Based on Factor Analysis and Neural Network Algorithm is expected to determine the level of new bursts to maintain the safety of coal mine workers in the Morowali area, Central Sulawesi, where the level of coal bursts can be known by using the rapidminer application based on past datasets with the level of bursts reaching a therhold value of -2.230% compared to the rate of no bursts reaching The thershold value is 3.194%, it is also expected for researchers to use other methods so that the level of coal bursts can be known.

REFERENCES

Al Shibli, M., Marques, P., & Spiridon, E. (2018). Artificial intelligent drone-based encrypted machine learning of image extraction using pretrained Convolutional Neural Network (CNN). *ACM International Conference Proceeding Series*, 72–82. https://doi.org/10.1145/3293663.3297155

Al-Sahaf, H., Zhang, M., & Johnston, M. (2016). Binary image classification: A genetic programming approach to the problem of limited training instances. *Evolutionary Computation*, *24*(1), 143–182. https://doi.org/10.1162/EVCO_a_00146

Bernadó-Mansilla, E., & Garrell-Guiu, J. M. (2003). Accuracy-Based Learning Classifier Systems: Models, Analysis and Applications to Classification Tasks. *Evolutionary Computation*, *11*(3), 209–238. https://doi.org/10.1162/106365603322365289

Chen, Z., Ai, S., & Jia, C. (2019). Structure-aware deep learning for product image classification. *ACM Transactions on Multimedia Computing, Communications and Applications*, *15*(1s), 1–20. https://doi.org/10.1145/3231742

Jhony, J., & Firdaus, W. (2020). Pemanfaatan Drone Untuk Monitoring Akurasi Perencanaan Tambang Batubara Terbuka. *Prosiding Temu Profesi Tahunan PERHAPI*, *1*(1), 181–190. https://doi.org/10.36986/ptptp.v1i1.61

Li, X., Chen, S., Wang, E., & Li, Z. (2021). Rockburst mechanism in coal rock with structural surface and the microseismic (MS) and electromagnetic radiation (EMR) response. *Engineering Failure Analysis*, *124*, 105396.

Perez, C. (2003). *Technological revolutions and financial capital*. Edward Elgar Publishing.

Singhal, D., Gupta, A., Tripathi, A., & Kothari, R. (2020). CNN-based Multiple Manipulation Detector Using Frequency Domain Features of Image Residuals. *ACM Transactions on Intelligent Systems and Technology*, *11*(4). https://doi.org/10.1145/3388634

Wang, L., Cheng, Y.-P., & Liu, H.-Y. (2014). An analysis of fatal gas accidents in Chinese coal mines. *Safety Science*, *62*, 107–113.

Wiegmann, D. A., & Shappell, S. A. (2017). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Routledge.

Wu, J., Sheng, V. S., Zhang, J., Li, H., Dadakova, T., Swisher, C. L., Cui, Z., & Zhao, P. (2020). Multi-Label Active Learning Algorithms for Image Classification: Overview and Future Promise. *ACM Computing Surveys*, *53*(2). https://doi.org/10.1145/3379504

Yong, A. G., & Pearce, S. (2013). A beginner’s guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, *9*(2), 79–94.

Yu, K., Ding, W., Simovici, D. A., Wang, H., Pei, J., & Wu, X. (2015). Classification with streaming features: An emerging-pattern mining approach. *ACM Transactions on Knowledge Discovery from Data*, *9*(4), 1–31. https://doi.org/10.1145/2700409

Zeng, S., Zhang, B., Gou, J., Xu, Y., & Huang, W. (2021). Fast and Robust Dictionary-based Classification for Image Data. *ACM Transactions on Knowledge Discovery from Data*, *15*(6). https://doi.org/10.1145/3449360

Zhang, D., & Kabuka, M. R. (2021). Protein Family Classification from Scratch: A CNN Based Deep Learning Approach. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *18*(5), 1996–2007. https://doi.org/10.1109/TCBB.2020.2966633