



MLP neural network in google colaboratory to predict mechanical properties of manufactured-sand concrete

Munzir Absa¹, Tulus Setiawan², Islami Fatwa³, Amam Taufiq Hidayat⁴

^{1,2}Physics Education Study Program, Malikussaleh University, Indonesia

³Mechanical Engineering Vocational Education Study Program, Malikussaleh University, Indonesia

⁴Mathematics Education Study Program, Malikussaleh University, Indonesia

ARTICLE INFO

ABSTRACT

Article history:

Received Jan 26, 2023

Revised Feb 04, 2023

Accepted Feb 28, 2023

Keywords:

Neural Network
Python
Mechanical properties
Manufactured sand
Concrete

The use of manufactured sand instead of natural sand in concrete can help preserve rivers or beaches where these natural sand is sourced from. The mechanical properties of these manufactured-sand concretes are comparable to those produced using natural sand. To predict and help optimize the mechanical properties of concrete made with manufactured sand, a model of ANN was developed using python 3.8 programming language in Google Colaboratory environment. It was found that ANN model with LBFGS weight optimization algorithm and 50 hidden layer nodes has the best performance, with RMSE = 0.067081. Prediction made with this model is calculated to have an accuracy of up to 95.37%. With this neural network, the relationship of Manufactured Sand Concrete fabrication parameters to its mechanical properties can be approximated.

This is an open access article under the [CC BY-NC](https://creativecommons.org/licenses/by-nc/4.0/) license.



Corresponding Author:

Munzir Absa,
Department of Physics Education,
Malikussaleh University,
Cot Teungku Nie, Reuleut, Muara Batu, North Aceh, Aceh, 24354, Indonesia.
Email: munzir.absa@unimal.ac.id

1. INTRODUCTION

Development of infrastructures such as buildings and roads plays a very significant role for Indonesia as a developing country. The demand for building materials in Indonesia continue to grow yearly to meet the need for development (Drebenstedt, 2021). One of the most dominantly used materials for construction of infrastructures is concrete, which typically made up of water, cement, and aggregate such as sand and gravel (Pradoto et al., 2022). These aggregate materials are usually sourced from rivers or beaches. Continuous extraction of these materials from rivers or beaches can cause environmental harm such as air pollution, damage to the ecosystem, decrease in water quality, shallowing of water level, production of dust, etc. (Firdaus, 2019). Therefore, the use of other resources is needed to minimize and reserve this aggregate sources.

Manufactured sand is a more environmentally friendly type of aggregate that can be used as a substitute for natural sand to produce concrete. Typically, manufactured sand is produced from crushed stones or rocks. Because of the way they are produced, manufactured sand has a rougher texture and more angular shape than natural (river) sand (Shen et al., 2016). This texture and shape meant manufactured sand can improve

interlocking between highly irregular particles in concrete, hence enhancing concrete strength and durability (Ji et al., 2013). Furthermore, the use of rocks with low silica content for making manufactured sand can reduce alkali-silica reactions in concrete, which can lead to expansion, deterioration or failure of concrete structures (Ly et al., 2019). The use of manufactured sand has been found to produce concrete with comparable compressive strength to those produced using natural sand (Hidayawanti et al., 2020).

Mechanical properties of Manufactured Sand Concrete (MSC), such as compressive and tensile strength, are often obtained by doing tests in laboratory. These tests are usually both cost and time expensive. It was often impossible to perform these tests with many variations of concrete ingredients/treatments (i.e. sand ratio, water to cement ratio, water to binder ratio, etc.) (Ly et al., 2019). Because of this, the development of models to predict mechanical properties of MSC can help greatly in producing MSC with optimal mechanical characteristics.

Linear statistical regression methods have been used to predict mechanical properties of concrete (Chopra et al., 2014). However, traditional statistical methods are usually inappropriate to model the highly nonlinear relationship between constituent materials and mechanical properties of MSC (Zhang et al., 2020). To resolve this issue, Machine Learning (ML) tools have been employed in properties prediction to help design high performance concrete. Methods such as Adaptive Neuro Fuzzy Inference System (ANFIS) (Ly et al., 2019), tree-based models (Zhang et al., 2020), and Support Vector Machine (SVM) (Van Dao et al., 2019) have been successfully developed to predict concrete mechanical properties.

Artificial Neural Network (ANN) models have also been shown to provide accurate results in predicting concrete properties. It has been shown that an ANN model can predict compressive strength of lightweight concrete under heat treatment more accurately than traditional regression method (Bingöl et al., 2013). ANN and ANFIS has also been used to investigate the compressive strength of brick-mortar masonry (Mishra et al., 2021). A hybrid model involving a relatively simple ANN and multi-objective grey wolves technique has also been able to predict compressive strength of silica fume concrete with accurate result (Behnood & Golafshani, 2018). These results show that AI methods, particularly ANN, are great tools to analyze mechanical properties of concrete based on its constituent materials.

One of the most widely used programming language to develop ANN is Python, because of its many functions and libraries (e.g. numpy, pandas, tensorflow, and matplotlib), because of its convenience and flexibility to build and deploy ANN, and also because of its open sourcedness (Satria et al., 2020). The python language itself can be accessed through many different environments, one of which is Google Colaboratory (or Colab for short) notebook. The use of Google Colab notebook is favorable because it is cloud-based, without the need for setup processes, and interactive, because it can be edited by team members simultaneously (Faizah et al., 2022). The aim of this study is to develop a simple and highly accessible model of Multilayer Perceptron (MLP) ANN using Python programming language from Google Colab notebook to predict the mechanical properties of MSC. This can help optimize the search for material and parameter configurations in fabrication of MSC to achieve desired mechanical properties.

2. RESEARCH METHOD

2.1 Research stages

Stages of this study is shown in Figure 1. The data used for this study are from secondary source (Zhao et al., 2017). A total of 305 data are available with all the properties investigated in this study. This data is then preprocessed with min-max

normalization, to scale the minimum and maximum of each sets of data to be from zero to one, using this equation:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

with:

x' = normalized data

x = data

$\min(x)$ = minimum of each set of data

$\max(x)$ = maximum of each set of data

After the data has been normalized, the data is then split for training and testing of ANN, with 80% of the data used for training and 20% for testing. The ANN is then trained to find the most optimum parameter. The varied parameters in this study are weight optimization algorithm, learning rate, and number of hidden node. The ANN models are evaluated to find the one with lowest error. The accuracy of the model with lowest error is then calculated. Detailed descriptions of this processes are explained in the sections below.

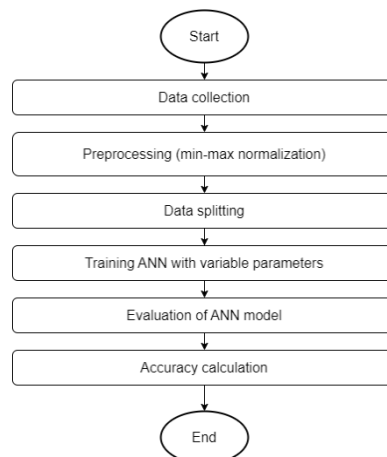


Figure 1. Research stages

2.2 Properties of MSC for input and output layer

Mechanical properties examined in this study are cubic compressive strength (f_{cu}) and splitting tensile strength (f_{st}). Compressive strength of cement is defined as the ability of a material or structure to support a surface load without cracking or deflection. To measure this property, a sample of concrete is cast with certain dimension. This sample is then compressed in a compression testing machine to measure the maximum force that can be exerted to this sample until the sample crack or deflect. Compressive strength can be acquired by dividing the maximum force (load) with the surface dimension of the sample.

Tensile strength is one of the basic and most important property of material. It is defined as the ability of a material to withstand a pulling (tensile) force. One of the method to indirectly measure tensile strength is by splitting tensile strength test. A cylindrical sample with certain dimension is laid down horizontally and compressed using compression testing machine. The load when the sample break (cracked) is measured. From this we can measure splitting tensile strength by dividing the load with the dimension of the sample. It is these two properties that will be predicted using ANN in this study.

For the input layer of ANN, we use some data relating to the manufactured sand property. The data for the input layer are: compressive strength and tensile strength of

cement, curing age, maximum diameter of crushed stone used for producing manufactured sand, stone powder content in the sand, fineness modulus of the sand, water-to-binder ratio, amount of water used, and ratio of sand in concrete.

2.3 Tools to develop ANN model

The programming language used to create ANN in this study is python 3.8 for its convenience and many features (libraries) which can run on many operating systems (Adhitya et al., 2022). The interface used is Google Colaboratory or Google Colab in short. With many libraries from python 3.8 such as numpy, pandas, sklearn, and matplotlib, Multilayer Perceptron (MLP) ANN models are created to predict the mechanical properties of MSC. Graphical representation of ANN model developed in this study is shown in Figure 2.

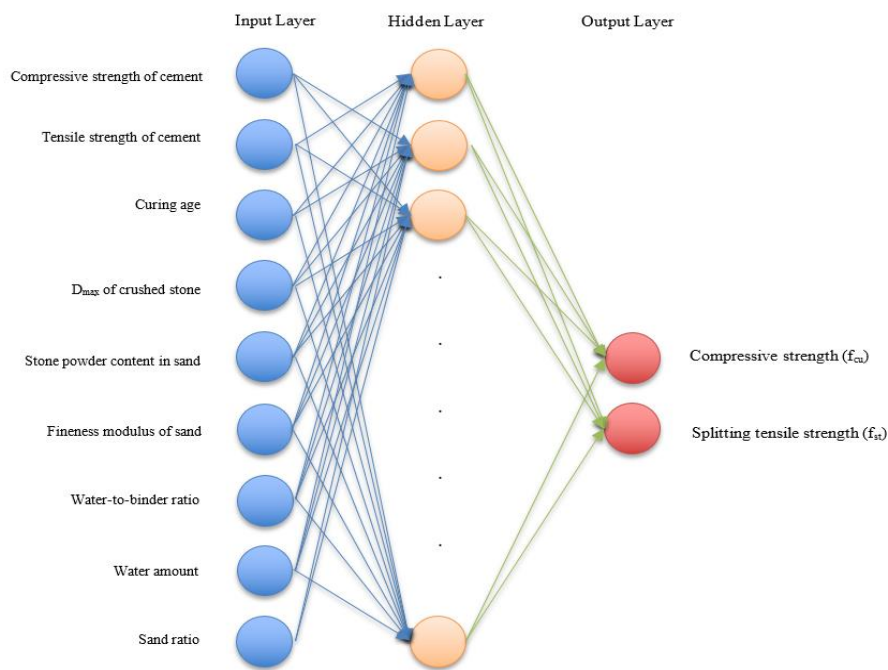


Figure 2. ANN model developed in this study

2.4 Data collection and preprocessing

Data of MSC mechanical properties and its fabrication parameters are collected from secondary source (Zhao et al., 2017). The data were assembled from 41 experimental studies of MSC tensile and compressive strength. The statistics of this data is shown in Table 1 below.

The data acquired are preprocessed and normalized with min-max normalization using Google Colab notebook, specifically by calling minmaxscaler function from sklearn library which used equation 1 as described above. The data is then split for training and testing, with 80% of the data used for training and 20% for testing of ANN.

Table 1. Statistics of data

Input/Target	Data size	Min	Max	Standard Deviation
Compressive strength of cement, f_{ce} (MPa)	305	38.2	55.2	3.74
Tensile strength of cement, f_{ct} (MPa)	305	6.9	9.4	0.54
Curing age (day)	305	3	388	100.39
Dmax of crushed stone (mm)	305	20	80	11.70

Stone powder content in sand (%)	305	0	20	4.69
Fineness modulus of sand	305	2.3	3.34	0.26
Water to binder ratio (%)	305	0.25	0.56	0.08
Water (kg/m ³)	305	104	291	20.28
Sand ratio (%)	305	28	44	4.25
Compressive strength $f_{cu,t}$ (MPa)	305	18.4	92.2	16.80
Splitting tensile strength $f_{st,t}$ (MPa)	305	1.13	5.48	1.00

2.5 ANN model parameters variation

The parameters modified in this study are:

- a. Weight optimization algorithms (solver), of which two are available in Google Colab:
 - 1) L-BFGS (Limited-memory Broyden-Fletcher-Goldfarb-Shanno), an optimization method in nonlinear programming to find global minima value with relatively minimum memory (Wibawa, 2017). In the Google Colab environment, this algorithm is named as “lbfgs” solver.
 - 2) Stochastic gradient descent, an iterative method to optimize a function with differentiable properties (Tanggih & Islami, 2022). This algorithm is denoted as “sgd” in Google Colab.
 - 3) Adaptive moment estimation, a method that computes individual adaptive learning rates for different parameters from estimates of first and second moments of the gradients (Kingma & Ba, 2015). This is called “adam” solver in Google Colab.
- b. Learning rate for sgd and adam solver from 0.01 to 0.001
- c. Number of nodes for hidden layer from 10 to 100

2.6 Evaluation of ANN model performance

To evaluate the ANN model with the best performance, mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE) are used. The equations for each of these evaluation methods are given below:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i| \quad (3)$$

$$RMSE = \sqrt{MSE} \quad (4)$$

with:

n = number of data points

Y_i = observed values

\hat{Y}_i = predicted values

The accuracy of MSC mechanical properties prediction using the ANN model can be calculated by measuring R-squared value:

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (5)$$

$$Accuracy = R^2 * 100\% \quad (6)$$

with:

n = number of data points

Y_i = observed values

\hat{Y}_i = predicted values

Another way to calculate the accuracy of regression method is by finding the percentage of mean absolute error (MAE), then subtracting it from 100% (Harsiti et al., 2022):

$$\text{Accuracy} = 100\% - \text{Error}(\%) \quad (7)$$

$$\text{Accuracy} = 100\% - (\text{MAE} * 100\%) \quad (8)$$

3. RESULTS AND DISCUSSIONS

Table 2 shows the results of MSE, MAE, RMSE, and R-squared for each possible configurations of parameters in the ANN. It is shown in the table that generally, ANN with LBFGS algorithm results in better prediction compared to SGD and Adam algorithm. This result agrees with previous research which shows that LBFGS are generally a better algorithm used for relatively small dataset (Najafabadi et al., 2017). Amongst ANN model using LBFGS algorithm, the one with 50 nodes in the hidden layer is the best, with the lowest MSE, MAE, and RMSE, as well as the highest R-squared.

Table 2. MSE, MAE, RMSE, and R² for each ANN model

Solver	Learning Rate	Hidden Layer Nodes	MSE	MAE	RMSE	R ²
Lbfgs	N/A	10	0.006469	0.053797	0.080428	0.864343
		20	0.004753	0.048738	0.068942	0.900623
		30	0.004905	0.048289	0.070039	0.897262
		40	0.005248	0.050500	0.072440	0.890044
		50	0.004500	0.046874	0.067081	0.905837
		60	0.005128	0.051153	0.071610	0.892297
		70	0.004984	0.049977	0.070600	0.895857
		80	0.005613	0.052990	0.074923	0.882060
		90	0.005850	0.053310	0.076485	0.877300
		100	0.005739	0.054572	0.075759	0.879563
sgd	0.01	10	0.021595	0.117540	0.146953	0.543948
		20	0.020796	0.112204	0.144209	0.561350
		30	0.017383	0.103613	0.131846	0.634473
		40	0.018868	0.111252	0.137360	0.602783
		50	0.016830	0.103804	0.129731	0.644762
		60	0.016884	0.103414	0.129940	0.642654
		70	0.015701	0.099604	0.125304	0.667700
		80	0.016408	0.104111	0.128094	0.654210
		90	0.128094	0.098740	0.125652	0.667993
		100	0.015091	0.099193	0.122846	0.681219
	0.005	10	0.026959	0.131173	0.164191	0.429424
		20	0.023405	0.118638	0.152986	0.505630
		30	0.021162	0.114170	0.145473	0.555143
		40	0.021102	0.115106	0.145264	0.555063
		50	0.019657	0.109099	0.140203	0.584087
		60	0.019527	0.110070	0.139737	0.586260
		70	0.017992	0.106201	0.134135	0.618112
		80	0.018928	0.111182	0.137579	0.600306
		90	0.018873	0.106632	0.137378	0.603353
		100	0.017201	0.103952	0.131152	0.636343
0.001	10	0.059862	0.198178	0.244667	-0.272254	
	20	0.043570	0.169364	0.208735	0.073924	
	30	0.051142	0.183969	0.226145	-0.078022	
	40	0.038773	0.160484	0.196909	0.178180	
	50	0.035035	0.142864	0.187177	0.258853	
	60	0.044526	0.174035	0.211012	0.051932	
	70	0.037640	0.157674	0.194011	0.193251	

Solver	Learning Rate	Hidden Layer Nodes	MSE	MAE	RMSE	R ²
		80	0.032603	0.149435	0.180562	0.306145
		90	0.035190	0.143080	0.187590	0.259156
		100	0.032040	0.144342	0.178998	0.321060
		10	0.019298	0.109821	0.138918	0.595408
		20	0.015935	0.098021	0.126233	0.663741
		30	0.013358	0.092677	0.115579	0.718555
		40	0.014557	0.094751	0.120651	0.694011
	0.01	50	0.014365	0.095682	0.119854	0.698248
		60	0.013480	0.094499	0.116102	0.716060
		70	0.013955	0.093885	0.118132	0.705798
		80	0.014009	0.095432	0.118362	0.705052
		90	0.014411	0.097549	0.120047	0.696104
		100	0.013315	0.089972	0.115390	0.719435
		10	0.023416	0.118739	0.153023	0.510223
		20	0.018109	0.105233	0.134569	0.617152
		30	0.015075	0.098164	0.122780	0.682272
		40	0.015446	0.099464	0.124282	0.675136
adam	0.005	50	0.014210	0.094992	0.119207	0.701399
		60	0.013885	0.096854	0.117834	0.707743
		70	0.013980	0.095650	0.118237	0.705161
		80	0.014655	0.099003	0.121060	0.691468
		90	0.014190	0.096690	0.119122	0.701022
		100	0.013724	0.094956	0.117147	0.711042
		10	0.051748	0.168465	0.227482	-0.083007
		20	0.021884	0.115723	0.147932	0.536924
		30	0.019339	0.110197	0.139064	0.591539
		40	0.020403	0.116089	0.142838	0.569092
	0.001	50	0.016876	0.102290	0.129909	0.643420
		60	0.017250	0.106212	0.131339	0.636524
		70	0.019144	0.110494	0.138361	0.592285
		80	0.018427	0.109601	0.135745	0.608997
		90	0.017752	0.104297	0.133238	0.626310
		100	0.015317	0.101343	0.123761	0.677132

Using the ANN model with the best performance, the difference between predicted values of MSC compressive strength with its measured values can be plotted. This plot is shown in Figure 3. It is shown that from 20 data points predicted, the ANN can predict this property with quite a high accuracy. After calculating the R-squared score in Google Colab using equation 6, it is found that the accuracy of this ANN model is:

$$R^2 = 0.9058$$

$$Accuracy = 90.58\%$$

The mean absolute error of this ANN model is 0.046874, so the accuracy can also be acquired by calculating equation 8:

$$Accuracy = 100\% - (MAE * 100\%)$$

$$Accuracy = 100\% - 4.69\%$$

$$Accuracy = 95.37\%$$

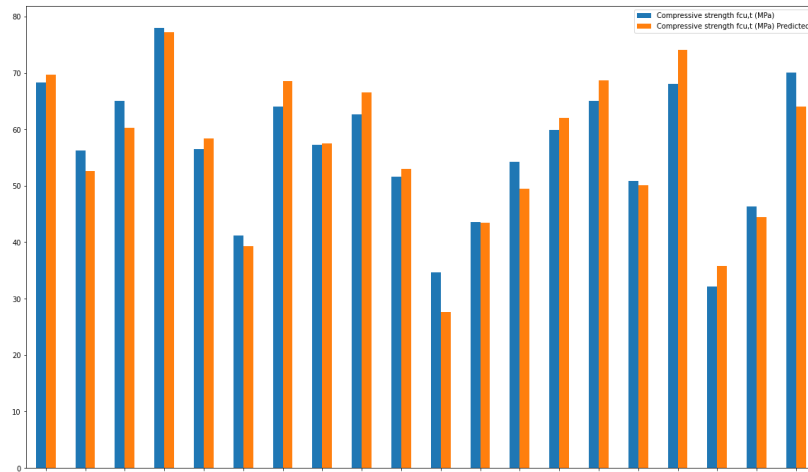


Figure 3. Compressive strength measured (blue) vs predicted (orange)

4. CONCLUSION

A simple and accessible model of MLP neural network has been developed to predict mechanical properties of Manufactured Sand Concrete from its fabrication parameters. It was found that LBFGS are the optimum weight optimization algorithm to use. The optimum number of hidden nodes is 50. The accuracy of the best ANN model developed was calculated by means of R-squared value and MAE value, which results in 90.58% and 95.37% accuracy. Using the MLP neural network developed in this study, the relationship between fabrication parameters of MSC to its resulting mechanical properties can be approximated. To further improve the accuracy of this prediction, the dataset used in this study can be expanded with other secondary sources or by doing tests independently. Further variation to the neural network parameters can also be applied, such as the variation of hidden layer and activation function.

ACKNOWLEDGEMENTS

The author would like to express gratitude for the contributions of the second, third, and fourth author to this study.

REFERENCES

- Adhitya, R., Devi, I., & Sari, I. N. (2022). Prediction of Tidal Data Using Cloud Computing-Based Support Vector Regression in Alor , East Nusa Tenggara. *Jurnal Mandiri IT*, 11(1), 1–9.
- Behnood, A., & Golafshani, E. M. (2018). Predicting the compressive strength of silica fume concrete using hybrid artificial neural network with multi-objective grey wolves. *Journal of Cleaner Production*, 202, 54–64. <https://doi.org/10.1016/j.jclepro.2018.08.065>
- Bingöl, A. F., Tortum, A., & Gül, R. (2013). Neural networks analysis of compressive strength of lightweight concrete after high temperatures. *Materials and Design*, 52, 258–264. <https://doi.org/10.1016/j.matdes.2013.05.022>
- Chopra, P., Sharma Professor, R. K., & Kumar Professor, M. (2014). Regression Models for the Prediction of Compressive Strength of Concrete with & without Fly ash. *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, 3(March 2014), 400.
- Drebenstedt, C. B. X. L. C. (2021). Proceedings of the International Conference on Innovations for Sustainable and Responsible Mining. In *International Conference of Innovations for Sustainable and Responsible Mining* (Vol. 109). <https://link.springer.com/10.1007/978-3-030-60839-2>
- Faizah, N. M., Fabrianto, L., Nurcahyo, W., & Trisnawati, H. (2022). Natural Language Processing

- Analysis Of Frequently Used Words On Indonesia Website Names. *Jurnal Mantik*, 6(3).
- Firdaus. (2019). Dampak Lingkungan Dan Sosial Penggalian Pasir Sepanjang Aliran Sungai Di Kota Bima (Studi Di Kelurahan Rabadompu Timur Kota Bima). *Jurnal Komunikasi dan Kebudayaan*, 6(1), 9–26.
- Harsiti, Muttaqin, Z., & Srihartini, E. (2022). Penerapan Metode Regresi Linier Sederhana Untuk Prediksi Persediaan Obat Jenis Tablet. *JSiI (Jurnal Sistem Informasi)*, 9(1), 12–16. <https://doi.org/10.30656/jsii.v9i1.4426>
- Hidayawanti, R., Yuhanah, Mayasari, D., & Wicaksono, B. (2020). The effect of m-sand and waste marble for strength of concrete. *IOP Conference Series: Materials Science and Engineering*, 930(1). <https://doi.org/10.1088/1757-899X/930/1/012024>
- Ji, T., Chen, C. Y., Zhuang, Y. Z., & Chen, J. F. (2013). A mix proportion design method of manufactured sand concrete based on minimum paste theory. *Construction and Building Materials*, 44, 422–426. <https://doi.org/10.1016/j.conbuildmat.2013.02.074>
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.
- Ly, H. B., Pham, B. T., Van Dao, D., Le, V. M., Le, L. M., & Le, T. T. (2019). Improvement of ANFIS model for prediction of compressive strength of manufactured sand concrete. *Applied Sciences (Switzerland)*, 9(18). <https://doi.org/10.3390/app9183841>
- Mishra, M., Bhatia, A. S., & Maity, D. (2021). A comparative study of regression, neural network and neuro-fuzzy inference system for determining the compressive strength of brick–mortar masonry by fusing nondestructive testing data. *Engineering with Computers*, 37(1), 77–91. <https://doi.org/10.1007/s00366-019-00810-4>
- Najafabadi, M. M., Khoshgoftaar, T. M., Villanustre, F., & Holt, J. (2017). Large-scale distributed L-BFGS. *Journal of Big Data*, 4(1). <https://doi.org/10.1186/s40537-017-0084-5>
- Pradoto, R. G. K., Soemardi, B. W., Gazali, A., Putri, A. T., Purba, R. P., & Mahardika, I. (2022). The Technology Landscape of Construction Material in the Indonesian Construction Industry. *IOP Conference Series: Earth and Environmental Science*, 1022(1). <https://doi.org/10.1088/1755-1315/1022/1/012017>
- Satria, W. A., Fitri, I., & Ningsih, S. (2020). Prediction of Customer Churn in the Banking Industry Using Artificial Neural Networks. *Jurnal Mantik*, 4(1), 10–19.
- Shen, W., Yang, Z., Cao, L., Cao, L., Liu, Y., Yang, H., Lu, Z., & Bai, J. (2016). Characterization of manufactured sand: Particle shape, surface texture and behavior in concrete. *Construction and Building Materials*, 114, 595–601. <https://doi.org/10.1016/j.conbuildmat.2016.03.201>
- Tangguh, F., & Islami, Y. (2022). Analisis performa algoritma Stochastic Gradient Descent (SGD) dalam mengklasifikasi tahu berformalin. *Indonesian Journal of Data and Science*, 3(1), 1–8.
- Van Dao, D., Ly, H. B., Trinh, S. H., Le, T. T., & Pham, B. T. (2019). Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete. *Materials*, 12(6). <https://doi.org/10.3390/ma12060983>
- Wibawa, M. S. (2017). Pengaruh Fungsi Aktivasi, Optimisasi dan Jumlah Epoch Terhadap Performa Jaringan Saraf Tiruan. *Jurnal Sistem dan Informatika (JSI)*, 11(1), 167–174. <http://archive.ics.uci.edu/ml/datasets/Wine>
- Zhang, J., Li, D., & Wang, Y. (2020). Toward intelligent construction: Prediction of mechanical properties of manufactured-sand concrete using tree-based models. *Journal of Cleaner Production*, 258, 120665. <https://doi.org/10.1016/j.jclepro.2020.120665>
- Zhao, S., Hu, F., Ding, X., Zhao, M., Li, C., & Pei, S. (2017). Dataset of tensile strength development of concrete with manufactured sand. *Data in Brief*, 11(136), 469–472. <https://doi.org/10.1016/j.dib.2017.02.043>