Interdisciplinary School of Management and Technology
Institut Teknologi Sepuluh Nopember

JURNAL TEKNOBISNIS
2023 Vol. 9(2), 96-104

Received 28 Dec, 2022; Revised 8 Feb, 2023; Accepted 1 Mar, 2023 | DOI: 10.12962/j24609463.v9i2.1873

Prediction PCC Cement Compressive Strength Based on Chemical Compounds and Physical Properties with Machine Learning Techniques

Aswamedhika[1,2*], Bagus Jati Santoso[3]

ABSTRACT

Indonesia has a mission to become a developed country by 2045. To achieve this mission, the government is aggressively carrying out equitable development through infrastructure development. Cement is one of the important components in infrastructure development. Compressive strength is one of the quality requirements that must be met by cement products. Compressive strength testing in industry using laboratory equipment takes a long time of up to 28 days for the entire set of test results to be completed. Data collection was taken at PT XYZ with a duration of 5 years. Data was taken from laboratory tests and operational data. In this study, machine learning algorithms used is linear regression, random forest, and neural networks. The modeling of the system obtained is expected to be able to predict the compressive strength of cement aged 3 days, 7 days, and 28 days so that the quality of the cement produced can be estimated quickly and does not take a long time. In addition, it is expected to know chemical compounds and physical properties that can affect the compressive strength of cement. The final result is the decision-making if the parameter changes can be mitigated quickly.

KEYWORDS: Cement, Compressive Strength, Linear Regression, Random Forest, Neural Network

[1]PT Semen Indonesia, Tuban, Indonesia
[2]Business Analytics, Interdisciplinary School of Management and Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[3]Department of Informatics, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia
*Corresponding author: 6032211024@mhs.its.ac.id

1. INTRODUCTION

Indonesia has a mission to become a developed country by 2045 (Ministry of National Development Planning / Bappenas, 2019). To achieve this mission, the government is intensively carrying out equitable development, one of which is through infrastructure development. Cement is one of the important components in infrastructure development. The amount of cement consumption in Indonesia in 2021 reached 62.7 million tons. Compressive strength is one of the quality requirements that must be met by cement products. Testing the compressive strength of cement to completion can take up to 28 days after the cement is produced. 28 days from product finish in production.

The length of time required in compressive strength testing, and corrective action if a product that does not meet specifications appears will be very late (ABB, 2022). Delays in decision-making will result in products being marketed not according to quality targets.

2. LITERATURE REVIEW

Cement is a substance that when applied to the surface of a solid object will make it adhere very firmly (Alsop, 2019). Cement will react and harden if added to water. The compressive strength of cement is formed when cement reacts with water. This reaction is called the hydration reaction (Bye, 2011). The reacting component of the cement composition is the clinker. Assuming the composition of cement is only clinker, then the reaction of clinker with water will form calcium silicate hydrate, calcium hydroxide, and AFm. Calcium silicate hydrate does not have a fixed composition, so it is often called C-S-H and does not have a specific stoichiometry. C-S-H is formed from the reaction of C3S and C2S with water. C2S will react with water into,

$$Ca_2SiO5 + H_2O \rightarrow CaO\text{-}SiO_2\text{-}H_2O$$

$$C_2S + H \rightarrow C\text{-}S\text{-}H$$

C_3S will react with water and become,

$$Ca_3SiO5 + H_2O \rightarrow CaO\text{-}SiO_2\text{-}H_2O + Ca(OH)_2$$

$$C_3S + H \rightarrow C\text{-}S\text{-}H + CH$$

The C3S hydration reaction will produce CaO. CaO will react with water and produce

$$Ca(OH)_2.CaO + H_2O \rightarrow Ca(OH)_2$$

$$C + H \rightarrow CH$$

The next reaction is the reaction of C3A with water which will form,

$$2C_3A + 21H \rightarrow C_4AH_{19} + C_2AH_8$$

The use of machine learning has been proven to predict the compressive strength of cement quickly and without any additional costs (Kumar, 2021). to predict cement compressive strength can be applied to the cement industry in real-time (Naranje, 2020).

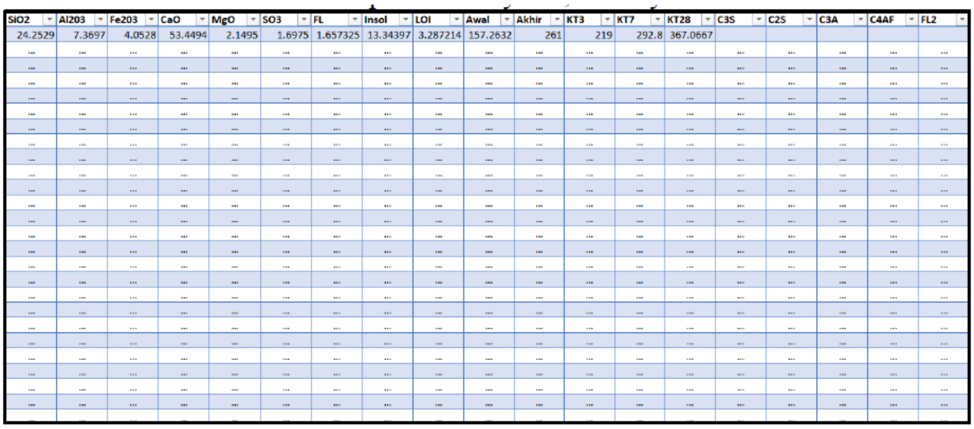
3. METHODS

Data Preprocessing

The initial stage in this study is data preprocessing. Data selection, data cleansing, impute missing value, outlier cleaning and data normalization will be carried out at this stage.

- Data Selection

  The data in this study was taken from production data and laboratory data at PT XYZ factories from four factories. Data were taken over a span of five years, namely 2017 and 2021.



SiO2	Al203	Fe203	CaO	MgO	SO3	FL	Insol	LOI	Awal	Akhir	KT3	KT7	KT28	C3S	C2S	C3A	C4AF	FL2
24.2529	7.3697	4.0528	53.4494	2.1495	1.6975	1.657325	13.34397	3.287214	157.2632	261	219	292.8	367.0667					
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

FIGURE 1. Data Selection

- Data Cleaning

  In the next stage, the data will be treated against empty data. The data that has been taken is raw data, and there are still many empty data. Processing of empty data is done by deleting all raw containing empty data.

- Remove Outliers

  Some data still contain outliers, 0 values, and even negative. Negative values and 0 values in the data, filtered and deleted data through Excel. Data that has been cleared of 0 and negative values are then eliminated outliers. The removal of outliers aims to eliminate values that do not match the conditions they should be, either too small or too large.

Exploratory Data Analysis

TABLE 1. Dataset Description

Column1	SiO2	Al2O3	Fe2O3	CaO	MgO	SO3	FLc	Insol	LOI	Blaine
Count	5433	5433	5433	5433	5433	5433	5433	5433	5433	5433
Mean	19.30	5.12	2.60	58.02	2.36	1.79	1.26	5.48	5.89	382.07
Stdev	2.52	1.03	0.52	2.21	0.58	0.09	0.41	1.73	1.75	22.76
Min	14.84	3.36	1.84	47.03	0.96	1.35	0.21	2.17	2.15	285.18
25%	17.79	4.50	2.29	57.18	1.95	1.76	0.97	3.98	4.38	367.49
50%	18.53	4.82	2.41	58.31	2.27	1.79	1.25	5.44	5.84	381.27

Column1	SiO2	Al2O3	Fe2O3	CaO	MgO	SO3	FLc	Insol	LOI	Blaine
75%	19.65	5.20	2.67	59.38	2.69	1.81	1.48	6.87	7.16	396.04
Max	28.54	8.55	4.48	64.92	6.23	2.31	3.99	11.21	12.81	524.08

Column1	Residu	KT3	KT7	KT28	C3S	C2S	C3A	C4AF	Fl clk
Count	5433	5433	5433	5433	5433	5433	5433	5433	5433
Mean	2.44	192.17	254.55	327.55	59.85	16.49	9.22	10.97	1.35
Stdev	0.74	27.76	32.15	37.82	3.20	3.35	0.59	0.48	0.45
Min	1.00	99.98	157.94	230.80	38.75	2.31	5.41	8.16	0.49
25%	1.83	173.20	233.78	301.00	58.12	14.63	8.85	10.69	1.03
50%	2.36	189.26	250.74	321.80	60.02	16.47	9.21	10.92	1.28
75%	2.97	208.94	271.07	347.00	61.78	18.34	9.57	11.20	1.58
Max	5.86	334.60	416.40	496.60	84.73	35.65	12.63	15.55	4.67

The clean dataset is 5433 rows. The average three-day compressive strength value was 192, the seven-day average compressive strength was 254, and the 28-day compressive strength average was 327. From this average value, the value of 28-day compressive strength is always greater than the compressive strength of three and seven days, and the seven-day compressive strength is greater than the value of three-day compressive strength. This happens because compressive strength growth cannot be negative, it will always be positive. This is also confirmed by the minimum and maximum values, where the minimum values of the compressive strength of three, seven, and 28 days are 99.98, 157.94, and 230.8. While the maximum compressive strength values of 3.7, and 28 days respectively are 334, 416, and 496. Judging from quartile one, quartile two, and quartile three, it can also be seen that the compressive strength of cement always grows positively.

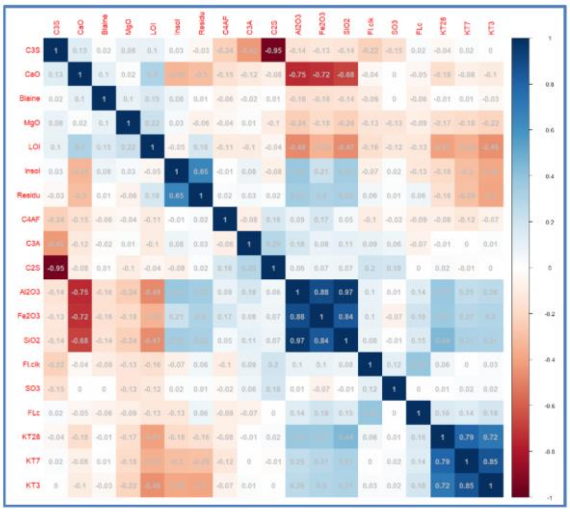


FIGURE 2. Heatmap Pearson Correlation

The parameters that correlate strongly with the compressive strength of cement are not many. In this study, parameters that are more than 0.1 will be drawn. Based on the correlation value on the heatmap above, the physical properties of cement that are correlated are only LOI, Insol, and residue. While the chemical compounds Al2O3, Fe2O3, SiO2, and MgO are seen to both influence compressive strength of three, seven, and 28 days.

Application of Algorithm Models

The algorithms used for model creation are linear regression, random forest, and neural networks. Random Forest algorithm, using the number of trees as many as 100 trees, and with the depth of each tree as many as 25 levels. The neural network algorithm will use a multi-layer perceptron (MLP) architecture with backpropagation. The number of neurons used is 25 neurons, with four hidden layers. The activation function of the hidden layer used is ReLu (the rectified linear unit function). The solver used and available in orange is L-BFGS-B, an optimizer that is still in the quasi-Newton method family. In this algorithm, a maximum iteration of 500x will be carried out.

Algorithm Performance

The sampling distribution of test data and training data was repeated 5 times, 10 times, 20 times, and 50 times. In addition, the accuracy of each model is also compared, namely with all features and seven features with the highest correlation.

4. RESULTS

The sampling distribution of test data and training data was repeated 5 times, 10 times, 20 times, and 50 times. In addition, the accuracy of each model is also compared, namely with all features and seven features with the highest correlation.

TABLE 2. Compressive Strength 3 days Accuracy Data Test. Left (all features) Right (seven features)

Repeat train/test 5x						Repeat train/test 5x					
Model	MSE	RMSE	MAE	MAPE	R2	5x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	370.309	19.243	15.266	0.082	0.493	Linier Regression	267.552	16.357	12.486	0.049	0.732
Random Forest	100.277	10.014	6.049	0.032	0.863	Random Forest	249.717	15.802	11.878	0.047	0.750
Neural Network	92.434	9.614	6.547	0.035	0.874	Neural Network	293.374	17.128	12.825	0.050	0.706
Repeat train/test 10x						Repeat train/test 10x					
Model	MSE	RMSE	MAE	MAPE	R2	10x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	385.024	19.622	15.548	0.082	0.491	Linier Regression	268.776	16.394	12.540	0.049	0.737
Random Forest	102.994	10.149	6.114	0.032	0.864	Random Forest	251.459	15.857	12.007	0.047	0.754
Neural Network	97.428	9.871	6.695	0.035	0.871	Neural Network	292.734	17.109	12.880	0.051	0.714
Repeat train/test 20x						Repeat train/test 20x					
Model	MSE	RMSE	MAE	MAPE	R2	20x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	381.871	19.542	15.442	0.082	0.504	Linier Regression	268.399	16.383	12.573	0.049	0.741
Random Forest	103.71	10.184	6.13	0.032	0.865	Random Forest	254.718	15.960	12.092	0.048	0.754
Neural Network	97.676	9.883	6.721	0.035	0.873	Neural Network	296.651	17.224	13.017	0.051	0.713
Repeat train/test 50x						Repeat train/test 50x					
Model	MSE	RMSE	MAE	MAPE	R2	50x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	382.349	19.554	15.487	0.082	0.505	Linier Regression	264.619	16.267	12.466	0.049	0.744
Random Forest	103.303	10.164	6.133	0.032	0.866	Random Forest	250.998	15.843	12.016	0.047	0.758
Neural Network	97.484	9.873	6.712	0.035	0.874	Neural Network	291.563	17.075	12.914	0.051	0.718

TABLE 3. Compressive Strength 7 days Accuracy Data Test. Left (all features) Right (seven features)

Repeat train/test 5x						Repeat train/test 5x					
5x	MSE	RMSE	MAE	MAPE	R2	5x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	267.552	16.357	12.486	0.049	0.732	Linier Regression	267.552	16.357	12.486	0.049	0.732

Repeat train/test 5x						Repeat train/test 5x					
Random Forest	249.717	15.802	11.878	0.047	0.750	Random Forest	249.717	15.802	11.878	0.047	0.750
Neural Network	293.374	17.128	12.825	0.050	0.706	Neural Network	293.374	17.128	12.825	0.050	0.706
Repeat train/test 10x						Repeat train/test 10x					
10x	MSE	RMSE	MAE	MAPE	R2	10x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	268.776	16.394	12.540	0.049	0.737	Linier Regression	268.776	16.394	12.540	0.049	0.737
Random Forest	251.459	15.857	12.007	0.047	0.754	Random Forest	251.459	15.857	12.007	0.047	0.754
Neural Network	292.734	17.109	12.880	0.051	0.714	Neural Network	292.734	17.109	12.880	0.051	0.714
Repeat train/test 20x						Repeat train/test 20x					
20x	MSE	RMSE	MAE	MAPE	R2	20x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	268.399	16.383	12.573	0.049	0.741	Linier Regression	268.399	16.383	12.573	0.049	0.741
Random Forest	254.718	15.960	12.092	0.048	0.754	Random Forest	254.718	15.960	12.092	0.048	0.754
Neural Network	296.651	17.224	13.017	0.051	0.713	Neural Network	296.651	17.224	13.017	0.051	0.713
Repeat train/test 50x						Repeat train/test 50x					
50x	MSE	RMSE	MAE	MAPE	R2	50x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	264.619	16.267	12.466	0.049	0.744	Linier Regression	264.619	16.267	12.466	0.049	0.744
Random Forest	250.998	15.843	12.016	0.047	0.758	Random Forest	250.998	15.843	12.016	0.047	0.758
Neural Network	291.563	17.075	12.914	0.051	0.718	Neural Network	291.563	17.075	12.914	0.051	0.718

TABLE 4. Compressive Strength 28 days Accuracy Data Test. Left (all features) Right (seven features)

Repeat train/test 5x						Repeat train/test 5x					
5x	MSE	RMSE	MAE	MAPE	R2	5x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	448.299	21.173	15.903	0.048	0.680	Linier Regression	457.229	21.383	16.105	0.049	0.674
Random Forest	386.744	19.666	14.886	0.045	0.724	Random Forest	403.999	20.100	15.149	0.046	0.712
Neural Network	440.075	20.978	15.918	0.048	0.686	Neural Network	426.530	20.653	15.419	0.047	0.696
Repeat train/test 10x						Repeat train/test 10x					
10x	MSE	RMSE	MAE	MAPE	R2	10x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	450.344	21.221	15.908	0.048	0.684	Linier Regression	458.202	21.406	16.078	0.049	0.678
Random Forest	395.441	19.886	15.008	0.046	0.722	Random Forest	410.204	20.254	15.202	0.046	0.712
Neural Network	447.470	21.153	16.034	0.049	0.686	Neural Network	435.297	20.864	15.557	0.047	0.694
Repeat train/test 20x						Repeat train/test 20x					
20x	MSE	RMSE	MAE	MAPE	R2	20x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	457.212	21.383	15.911	0.048	0.685	Linier Regression	463.978	21.540	16.073	0.049	0.680
Random Forest	404.253	20.106	15.063	0.046	0.721	Random Forest	422.456	20.554	15.317	0.046	0.709
Neural Network	453.225	21.289	16.072	0.049	0.688	Neural Network	438.919	20.950	15.578	0.047	0.698
Repeat train/test 50x						Repeat train/test 50x					
50x	MSE	RMSE	MAE	MAPE	R2	50x	MSE	RMSE	MAE	MAPE	R2
Linier Regression	455.357	21.339	15.905	0.048	0.684	Linier Regression	461.979	21.494	16.052	0.049	0.680
Random Forest	402.180	20.054	15.058	0.046	0.721	Random Forest	419.473	20.481	15.309	0.046	0.709
Neural Network	454.552	21.320	16.064	0.049	0.685	Neural Network	436.374	20.890	15.583	0.047	0.697

Compressive strength predictions, both three-day, seven-day, and 28-day represented by the top seven features with the highest correlation of feature availability are not much different when compared to predictions with the overall available features. From these results, it can be said that the addition of features in compressive strength prediction does not have too much impact on modeling accuracy.

The best three-day compressive strength prediction performance is obtained from neural network algorithms, while the best prediction performance for predicting seven-day and 28-day compressive strength is best achieved by random forest algorithms.

TABLE 5. Accuracy at Data Train

Model	MSE	RMSE	MAE	R2
Neural Network (KT3)	75.094	8.666	6.052	0.903
Random Forest (KT7)	35.260	5.938	4.461	0.966
Random Forest (KT28)	57.495	7.583	5.650	0.960

The summary accuracy results of the test data for three-day compressive strength are MSE of 92.434, RMSE of 9.614, MAE of 6.547, MAPE of 0.035, and R2 of 0.874. Seven-day compressive strength accuracy, MSE of 250.730, RMSE of 15.834, MAE of 11.957 MAPE of 0.047, and R2 of 0.749. Strong accuracy of 28 days, MSE of 403.999, RMSE of 20.1, MAE of 15.149 MAPE of 0.046, and R2 of 0.712.

Comparison of accuracy in train data and test data states that the accuracy of the best model produced meets the criteria of goodness of fit, where the difference in accuracy between the data train and test data is not too far.

5. CONCLUSIONS

Machine learning modeling using linear regression, random forest, and neural networks can predict the compressive strength of cement aged three days, seven days, and 28 days.

Neural networks are the most accurate algorithm in predicting three-day compressive strength. The best seven and 28-day compressive strength predictions were obtained from the random forest algorithm. Cement compressive strength prediction algorithm can be applied in the field of quality control and quality assurance in the cement industry. The best algorithm that has been obtained can be deployed into the factory's internal system. Data sources can be directly obtained from OPC servers, and from laboratory servers. Machine learning used in predicting compressive strength can be combined with factory process expert systems, and become input in determining the parameters of cement production operations. On the other hand, quality assurance and quality control can quickly see the compressive strength of cement without any delay in testing time.

REFERENCES

ABB, (2022), “How cement producers are using AI to transform their operations — and future trajectory - Bestpractices”. Retrieved from https://library.e.abb.com/public/eba69de0f26147fa8ea84ded7f708d91/25206%20ABB%20AI%20ebo ok_cement_20220223.pdf

Abbott, Dean, (2014), Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst, John Wiley & Sons, Inc, Indiana.

Alsop, Philip, (2019), The Cement Plant Operations Handbook, 7th Edition, Tradeship Publications Ltd. Aman, Kumar., et al., (2022), “Compressive Strength Prediction of Lightweight Concrete: Machine Learning

Models”. Sustainability 2022, Vol. 14, No. 2404.

Apostolopoulou, Maria., et al., (2018), “Prediction of Compressive Strength of Mortars using Artificial Neural Networks”, Proceedings of 1st International Conference (TMM_CH) Transdisciplinary Multispectral Modelling and Cooperation for the Preservation of Cultural Heritage, Athens, Greece.

Asteris, Panagiotis G., et al., (2020), "Prediction of cement-based mortars compressive strength using machine learning techniques", Neural Computing and Applications, Vol. 33, Hal. 13089–13121.

Ayaz, Ahmad., et al., (2022), "Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques", Case Studies in Construction Materials, Vol. 16.

Bye, Gerry, (2011), Portland Cement, 3rd Edition, Thomas Telford Limited, Eastbourne.

Europan Standart, (2011), Cement Part 1: Composition, specifications and conformity criteria for common cement BS EN 197-1:2011, BSI.

Ghosh, SN, (1983), Advances in Cement Technology Critical Reviews and Case Studies on Manufacturing, Quality Control, Optimization and Use, 1st Edition, Pergamon Press Ltd., New York.

Hamid, Eskandari-Naddaf., Ramin, Kazemi., (2017), "ANN prediction of cement mortar compressive strength influence of cement strength class", Construction and Building Materials, Vol. 138, Hal. 1- 11.

Huang, Jiandong., et al., (2022), "Prediction of the Compressive Strength for Cement-Based Materials with Metakaolin Based on the Hybrid Machine Learning Method", Materials, Vol. 15.

Kemenperin, (2007), Pemberlakuan Standar Nasional Indonesia (SNI) Semen Secara Wajib, No. 35/M- IND/PER/4/2007, Jakarta.

Kemensekneg, (2020), Peraturan Presiden Republik Indonesia nomor 109 Tahun 2020 Tentang Perubahan Ketiga Atas Peraturan Presiden Nomor 3 Tahun 2016 Tentang Percepatan Pelaksanaan Proyek Strategis Nasional, No. 109 tahun 2020, Jakarta.

Kementerian PPN, (2019), Indonesia 2045: Berdaulat, Maju, Adil, dan Makmur, Bappenas, Jakarta. Mauliauwan, HN., et al., (2020), "Prediction of Concrete Compressive Strength Using Artificial Intelligence

Methods", Journal of Physics: Conference Series, Vol. 1625.

Meade, MS., Richard, K., (1926), Portland Cement Its Composition, Raw Materials, Manufacture, Testing and Analysis, 3rd Edition, The Chemical Public Company, London.

Mohri, Mehryar., Rostamizadeh, Afshin., Talwalkar, Ameet., (2018), Foundations of Machine Learning, 2nd Edition, The MIT Press, London.

Mukhiya, S Kumar., Usman, Ahmed., (2020), Hands-On Exploratory Data Analysis with Python, Packt Publishing Ltd., Birmingham.

Muller, C Andreas., Guido, Sarah, (2017), Introduction to Machine Learning with Python, O'Reilly Media Inc., California.

Narenje, Vishal., Salunke, Sachin., (2020), "Cement strength prediction using cloud-based machine learning techniques", Journal of Structural Integrity and Maintenance, Vol. 5, No. 4, Hal. 244-251.

PT Semen Indonesia (Persero) Tbk, (2021), Better Contribution To Deliver Sustainability, Jakarta.

PT Semen Indonesia (Persero) Tbk, (2021), Presentasi Kinerja Perusahaan Public Expose Live 2021, Jakarta. PT Semen Indonesia (Persero) Tbk, (2021), Strengthening Capabilities To Sustain Growth, Jakarta.

Reddy, K Sumanth., et al., (2021), "Modelling and Prediction of Concrete Compressive Strength Using Machine Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Vol. 7, No. 3, Hal. 526-532.

SNI 7064:2014, (2014), Semen Portland Komposit, Badan Standarisasi Nasional, Jakarta. SNI-2049-2015, (2015), Semen Portland, Badan Standarisasi Nasional, Jakarta.

Tsamatsoulis, D., (2021), "Prediction of Cement Compressive Strengthby Combining Dynamic Models of Neural Networks", Chemical and Biochemical Engineering Quarterly, Vol. 35 No. 3, Hal. 295-318.

How to cite this article:

Aswamedhika., & Santoso, B. J. (2023). Prediction PCC Cement Compressive Strength Based on Chemical Compounds and Physical Properties with Machine Learning Techniques. *Jurnal Teknobisnis, 9*(2): 96-104. DOI: 10.12962/j24609463.v9i2.1873