International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Novel Machine Learning Technique by Several Meta and Naïve Bayes Method to Predict Full Load Electrical Power Output of Base Load Operated Combined Cycle Power Plant

Dr. Koppisetti Giridhar [1],S. Madhanmohan[2], Dr. A. Karthikayen[3], Dr. P. Jona Innisai Rani[4]
K.V.L.N. Prasad[5] , Dr. A.Thankaraj[6]

Submitted: 12/03/2024 **Revised:** 21/04/2024 **Accepted:** 04/05/2024

Abstract: Avoiding technical concerns, such blackouts, is crucial to the efficient and cost-effective operation of a combined cycle power plant (CCPP). In this paper, we suggest making use of machine learning methods to estimate the hourly electricity output of a CCPP. For this purpose, we take into account the exhaust vacuum, ambient temperature, atmospheric pressure, and relative humidity as basic parameters that affect the generated power. The output power and other parameters are measured and utilized to train and test machine learning models. This paper explores the Logit Boost with Bagging perform well as well it showing an efficient outcome. It has the greatest accuracy result of 85.80%. The Logit Boost with Bagging produces the greatest precision result of 0.87. The Logit Boost with Bagging and Random Committee Bagging produce the maximum recall of 0.87. The Logit Boost with Bagging has the greatest F-Measure result of 0.87. The Logit Boost with Bagging model has the highest MCC value of 0.66. The Logit Boost with Bagging model has the greatest kappa value of 0.67. The Logit Boost with Bagging model has an optimal results compare with other models.

Keywords: *Bagging, Logit Boost, F-Measure, Random Committee, Combined cycle power plant.*

1Assistant Professor, Department of CST, Madanapalle Institute of Technology and Science, Madanapalle, AP, India, kgiridhar562@gmail.com
2Assistant Professor, Department of Computer Applications, Arignar Anna Government Arts College,Villupuram,Tamilnadu, India. nsmadhanmohan@gmail.com
3Professor, Department of ECE, P.T.Lee Chengalvaraya Naicker College of Engineering and Technology,Oovery, Kanchipuram, Tamil Nadu, India. akarthi_mathi@yahoo.co.in
4Associate professor (Computer Science), Department of Agricultural Economics, Anbil Dharmalingam Agricultural College and Research Institute, Trichy, Tamil Nadu, India jir8@tnau.ac.in
5 Assistant Professor, Department of Computer Science, KBN College, Vijayawada, AP, India kvlnprasad@gmail.com
6Assistant Professor, Department of Electrical and Electronics Engineering, Rrase College of Engineering, Padappai, Chennai, Tamilnadu 601301, India, atraj77@gmail.com

I Introduction

One of the most popular fuels nowadays is natural gas, which is used in the production of electricity (to the tune of 23.6% of the entire structure) [1]. CCGT plants are the most effective at transforming the chemical energy of natural gas into electricity, as their working flow consists primarily of the byproducts of combustion of the air and fuel mixture (Brayton cycle) and steam from water (Rankine cycle). Modern CCGT plants have a net electric efficiency of over 63%.At an upstream working flow temperature of around 1550 °C and an exhaust gas temperature of 670 °C, CCGT using Siemens SGT5-9000HL gas turbine plants demonstrates particularly high efficiency values. Although GTE-160 gas turbine plants are popular in Russia, their lower exhaust gas temperature (537 °C) and initial operating flow temperature (about 1100 °C) make them less efficient than their Western counterparts. When used in conjunction with the CCP-220T CCGT power plant, the net efficiency of these GTPs rises to 50.4% from 34.4% in free operating mode.Raising the working flow temperature before the gas turbine is the most direct way to improve CCGT performance. Nevertheless, high-temperature technology for electric power generation [2,3] is required before this approach can be put into practise.Using low-potential heat sources is another option for improving CCGT plant efficiency (by 2%-3%; see [4,5,6]). To better utilise the exhaust gas heat or to lower the cold source temperature [7,8,9], CCGT facilities might incorporate the organic Rankine cycle. Most low-temperature cycle heat carriers (most notably freons) are, however, more hazardous than water and have lower availability and chemical compound stability. However, the capital costs of the power plant rise with the addition of even a single additional cycle. All of these issues combined to make widespread adoption of this technology unlikely [10,11].

The rest of the paper is organized as follow: Section 2 outlines the related work. Section 3 introduces the proposed methodology, and the results and discussion are briefly discussed in Section 4. Finally, we conclude the paper in Section 5.

II Literature Survey

An accurate prediction of a plant's power output can help avoid power outages, financial losses, and technical obstacles [1, 2]. Specifically due to excessive fuel consumption, erroneous predictions raise the per-unit cost of electric power [3]. Through this research, we hope to reduce the cost per unit of electricity generated [4] by accurately predicting the electric output of a base load CCPP under full load conditions.The efficiency of thermodynamic power plants can be predicted using intricate mathematical models [5]. These models' reliance on a wide range of assumptions and parameters reflect the system's inherent uncertainty. Although time-consuming, these mathematical models are founded on a deterministic approach [5]. Nevertheless, probabilistic approaches rather than mathematical models are used in supervised ML systems for power prediction [6]. Due to the abundance of information, the ML technique of prediction is preferable to other approaches since it eliminates the need to model the complete system. It can also be seen in adjacent domains like groundwater hardness vulnerability forecasting [7], soil erosion susceptibility estimate [8], groundwater level forecasting [9], and groundwater potential forecasting [10]. Our suggested ML algorithms evaluate the power plant's historical data to generate optimal power projections in a shorter amount of time [11]. Unfortunately, ML algorithms are not perfect predictors because of their probabilistic foundation. This is why we suggest trying out a few different algorithms for CCPP power forecasting. In addition, we seek for the parameters of these algorithms that produce the least amount of inaccuracy for the available information.During the course of a year, the amount of electricity produced by a CCPP could fluctuate for a variety of reasons [12]. Some of these reasons include changes in the surrounding environment's temperature, atmospheric pressure, humidity, and exhaust flow. Hence, these parameters have both direct and indirect effects on the CCPP's output power [13]. Controlling these variables effectively allows for increased power generation with less fuel use [14]. This research avoids the common practise of attempting to control the parameters by instead analysing their effect on the expected output power. To do this, we apply various machine learning techniques [4] to these environmental factors in order to forecast electrical output.

There have been previous attempts to anticipate CCPP power using a variety of probabilistic approaches, including bagging and regression ANN [4, 15]. Due to the relatively large prediction error [16], this study presents machine learning algorithms for forecasting electric power of CCPP operated at full load using the aforementioned four criteria. The usage of a gradient-boosted regression tree (GBRT), a linear regression (LR), an artificial neural network (ANN), and a k-nearest neighbour algorithm all lead to more accurate predictions of future power consumption (KNN). We evaluate the effects of each parameter on output power prediction and their interaction using these four machine learning methods. By comparing our algorithms against those of other studies, we can determine which ones do the best. The best performing of these four ML models is chosen based on its RMSE and AE values.

The remainder of the paper's outline is as follows: Section 2 details the associated work. Part 3 provides an introduction of the suggested method, while Section 4 summarizes the findings and discusses them briefly. The results of the paper are summarized in Section 5.

III Materials and Methods

UCI's Combined Cycle Power Plant Data Collection was the repository we were looking for. Throughout a six-year period (2006-2011), 9568 data points were gathered from a Combined Cycle Power Plant while the facility was operating at full load. Predicting the net hourly electrical energy output (EP) of the plant requires features such as the hourly average of ambient variables Temperature (T), Ambient Pressure (AP), Relative Humidity (RH), and Exhaust Vacuum (V).Gas turbines (GT), steam turbines (ST), and heat recovery steam generators (HRSG) make up the components of a combined cycle power plant (CCPP). The CCPP gets its power from a single cycle of gas and steam turbines that exchanges energy with itself. The GT is affected by the Vacuum, which is gathered by the Steam Turbine, and by the other three environmental factors as well. We supply the data shuffled five times to ensure consistency with our baseline investigations and to provide 5x2 fold statistical tests. Ten readings are taken after each randomization cycle thanks to a 2-fold CV procedure.

Features Information:

Features consist of hourly average ambient variables

- Temperature (T) in the range 1.81°C and 37.11°C,

- Ambient Pressure (AP) in the range 992.89-1033.30 milibar,

- Relative Humidity (RH) in the range 25.56% to 100.16%

- Exhaust Vacuum (V) in teh range 25.36-81.56 cm Hg

- Net hourly electrical energy output (EP) 420.26-495.76 MW

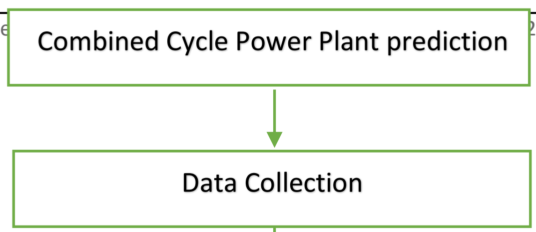
The averages are taken from various sensors located around the plant that record the ambient variables every second. The variables are given without normalization.

Methods:

The following method are applied in this research work

- Borrowed dataset
- Data preprocessing
- Apply for Ensemble machine learning algorithms:
- Multi Class Classifier with Bagging (MCC with Bag)
- Logit Boost with Bagging(LB with Bag)
- Random Committee with Bagging(RC with Bag)
- Multi Scheme with Bagging (MS with Bag)
- Input Mapped Classifier with Bagging (IMC with Bag)
- Evaluate models
- Find a best Model

To produce an efficient result, these strategies were applied in python API. This study uses only 10% of the total dataset and uses tenfold cross validation for all categories.



No

Yes

Figure 1: Proposed System

Table 2: Performance of selected classifiers

S.No	Classifiers	Accuracy	Precision	Recall	F-Measure	MCC	Kappa
1	MCC with Bagging	84.15%	0.85	0.85	0.84	0.57	0.56
2	LB with Bagging	85.80%	0.87	0.87	0.87	0.66	0.67
3	RC with Bagging	85.05%	0.86	0.87	0.85	0.59	0.60
4	MS with Bagging	79.07%	0.82	0.81	0.80	0.55	0.55
5	IMC with Bagging	83.90%	0.84	0.84	0.85	0.57	0.59

The above table shows that the various selected ensemble classifiers.

The MCC with Bagging results in an accuracy level of 84.15%, a precision value of 0.85, a recall value of 0.85, an F-Measure value of 0.84, an MCC value of 0.57 and a kappa statistic value of 0.56.

The LB with Bagging has an accuracy level of 85.80%, a precision value of 0.87, a recall value of 0.87, an F-Measure value of 0.87, an MCC value of 0.66 and a kappa statistic value of 0.67.

The RC with Bagging results in an accuracy level of 85.05%, a precision value of 0.86, a recall value of 0.87, an F-Measure value of 0.85,an MCC value of 0.59 and a kappa statistic value of 0.60.

The MS with Bagging produces a yield of 79.07% an accuracy, a precision value of 0.82, a recall of 0.81, an F-Measure of 0.80, an MCC of 0.55 and a kappa statistic of 0.55.

The IMC with Bagging produces accuracy level 83.90%, a precision value 0.84, recall value 0.84, an F-Measure value 0.85,an MCC value 0.57 and a kappa statistic value 0.59.

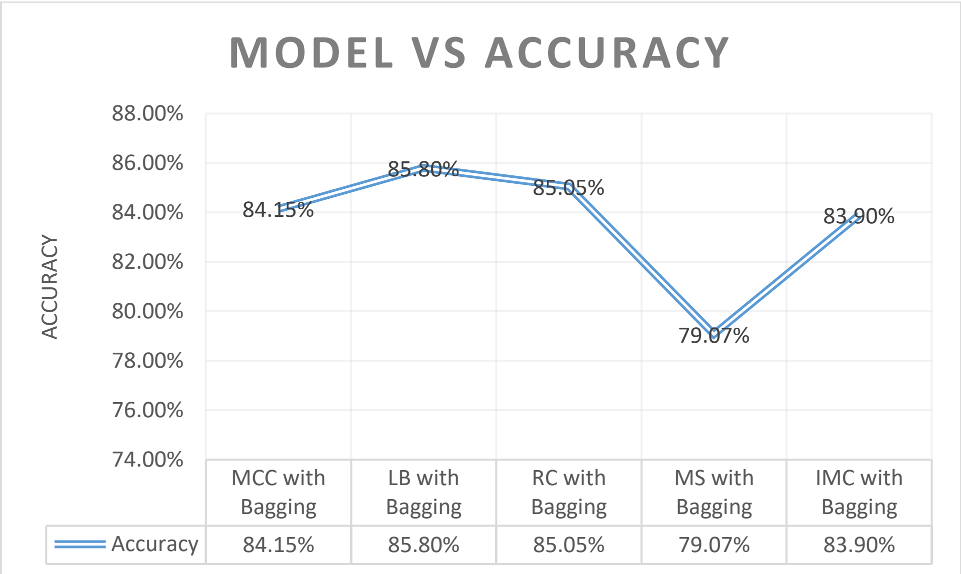


Figure 2: Performance of Ensemble classifiers with their accuracies

The above diagram shows that the accuracy performances of selected models. The LB with Bagging has the greatest accuracy result of 85.80%. The MS with Bagging produces the lowest accuracy result of 79.07%. The accuracy of the IMC with Bagging, MCC with Bagging, and RC with Bagging is 83.90%, 84.15%, and 85.05%, respectively.

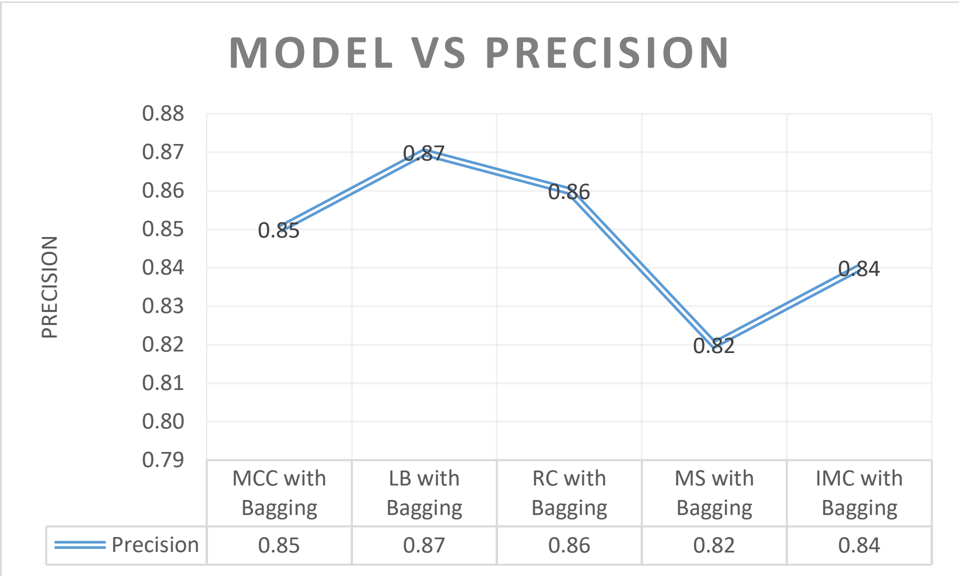


Figure 3: Performance of Ensemble Classifiers with their Precision values

The precision performances of selected models are depicted in the diagram above. The LB with Bagging produces the greatest precision result of 0.87. The MS with Bagging produces the lowest accuracy result of 0.82. The accuracy levels of the IMC with Bagging, MCC with Bagging, and RC with Bagging are 0.84, 0.85, and 0.86, respectively.

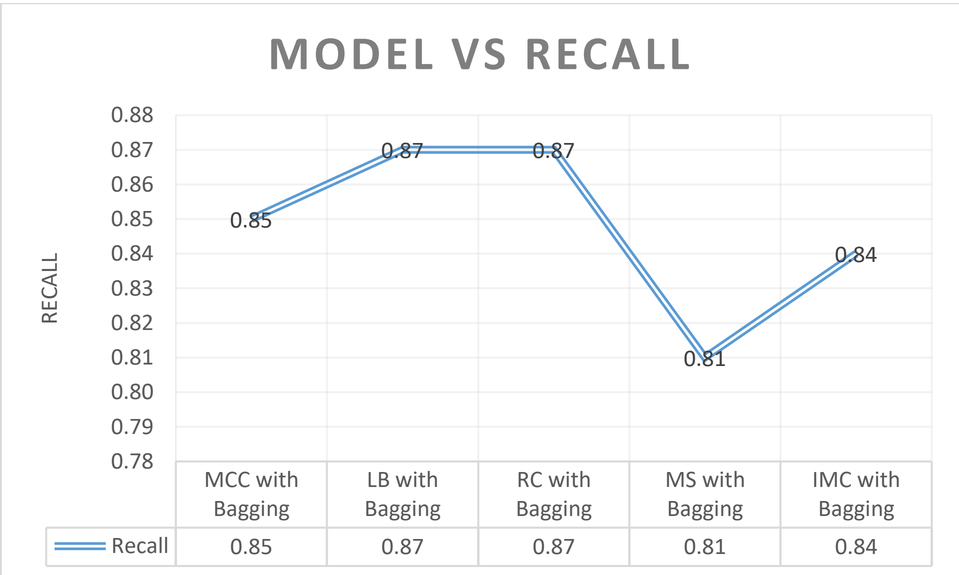


Figure 4: Performance of Ensemble Classifiers with their Recall values

The graph above depicts the recall performances of selected models. The LB with Bagging and RC with Bagging produce the maximum recall of 0.87. The MS with Bagging produces the lowest recall result of 0.81. The recall levels for the IMC with Bagging and the Extreme MS with Bagging are 0.84 and 0.85, respectively.

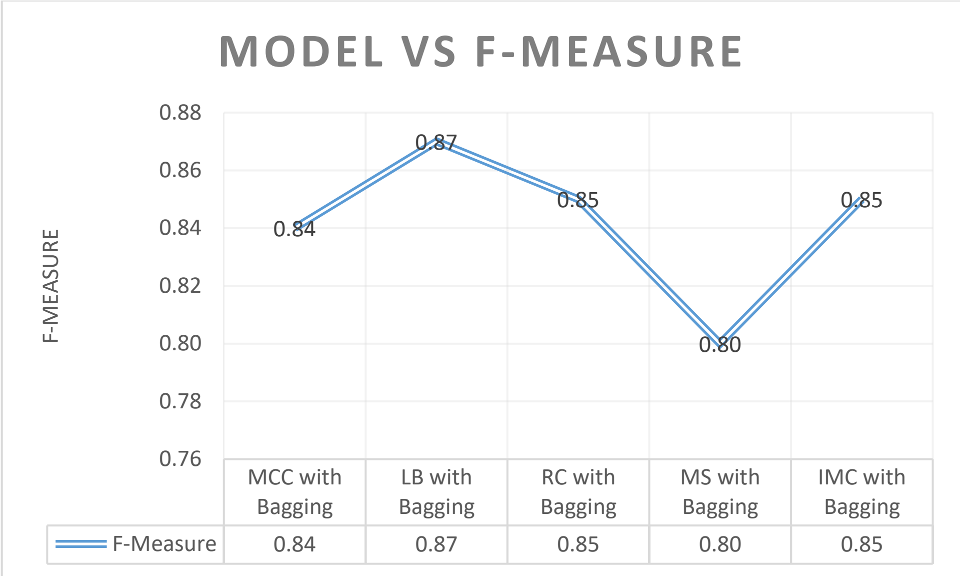


Figure 5: Performance of Ensemble Classifiers with their F-Measure values

The graph above depicts the F-Measure performances of selected models. The LB with Bagging has the greatest F-Measure result of 0.87. The MS with Bagging produces the lowest F-Measure result of 0.80. The MCC with Bagging has an F-Measure of 0.84, whereas the IMC with Bagging and RC with Bagging have the same value of 0.85.

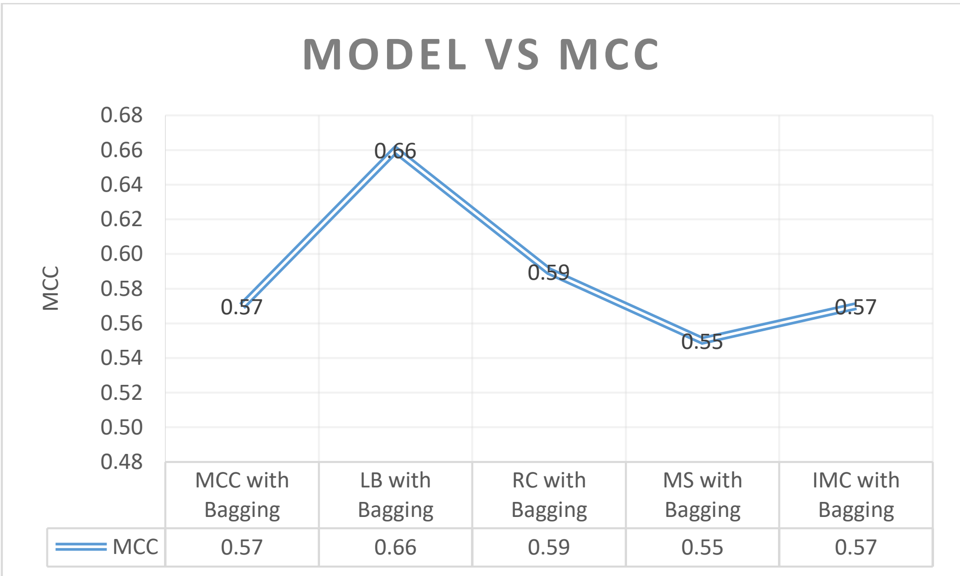
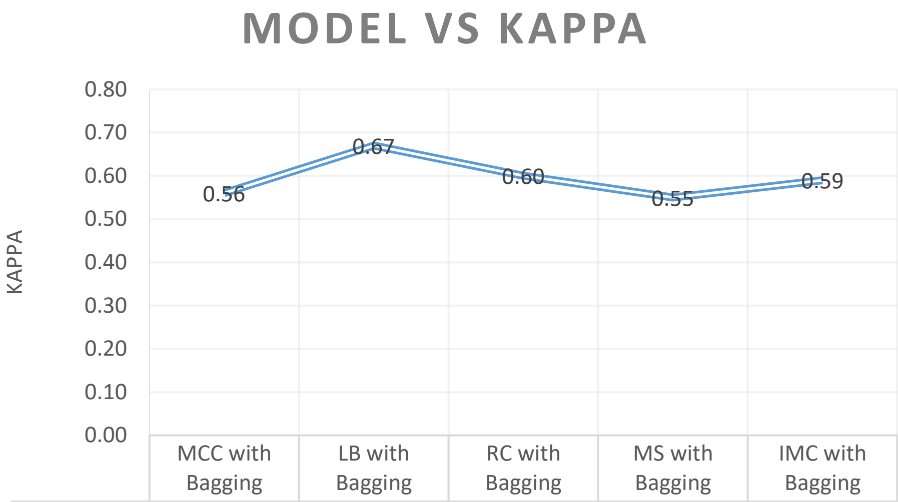


Figure 6: Performance of Ensemble Classifiers with their MCC values

The graphic above depicts the MCC performance of selected models. The LB with Bagging model has the highest MCC value of 0.66. The MS with Bagging produces the lowest MCC result (0.55). The remainder of the models, such as the MCC with Bagging model and the Light Gradient Boosting Machine with NB Decision Trees model, have the same MCC value of 0.57. The MCC value for RC with Bagging 0.59.



	MCC with Bagging	LB with Bagging	RC with Bagging	MS with Bagging	IMC with Bagging
Kappa	0.56	0.67	0.60	0.55	0.59

Figure 7: Performance of Ensemble classifiers with their Kappa statistic values

The graph above depicts the kappa value performances of selected models. The LB with Bagging model has the greatest kappa value of 0.67. The MS with Bagging produces the lowest kappa result of 0.55. Other models with kappa values between 0.56 and 0.60 are MCC with Bagging, IMC with Bagging, and RC with Bagging.

V Conclusions

Based on this study's findings, the MCC with Bagging yields an 84.15% accuracy rate, 0.85 precision, 0.85 recall, 0.84 F-Measure, 0.57 MCC, and 0.56 kappa statistic. The LB with Bagging achieves an impressive 85.80% accuracy, 0.87 precision, 0.87 recall, 0.87 F-Measure, 0.66 MCC, and 0.67 kappa statistic. The resulting RC with Bagging has an accuracy of 85.05%, precision of 0.86, recall of 0.87, precision of 0.85, recall of 0.87, F-Measure of 0.85, MCC of 0.59, and kappa statistic of 0.60. With Bagging applied to the MS, we get an accuracy of 79.07%, precision of 0.82, recall of 0.81, sensitivity of 0.80, specificity of 0.80, mean correlation coefficient of 0.55, and kappa of 0.55. Accuracy of 83.90 percent, precision of 0.84, recall of 0.84, F-measure of 0.85, mean correlation coefficient of 0.57, and kappa statistic of 0.59 are all generated by the IMC with Bagging. The LB with Bagging has the greatest accuracy result of 85.80%, a precision result of 0.87, a recall of 0.87, an F-Measure result of 0.87, an MCC value of 0.66 and a kappa value of 0.67.This model recommends the LB with Bagging compare with other models.

References

[1] Siddiqui, Raheel,Anwar, Hafeez,Ullah, Farman,Ullah, Rehmat,Rehman, Muhammad Abdul,Jan, Naveed,Zaman, Fawad,TI - Power Prediction of Combined Cycle Power Plant (CCPP) Using Machine Learning Algorithm-Based Paradigm, 9966395,2021,ISSN:1530-8669,Wireless Communications and Mobile Computing, Hindawi, https://doi.org/10.1155/2021/9966395, DOI - 10.1155/2021/9966395.

[2] Rogalev, A.; Rogalev, N.; Kindra, V.; Komarov, I.; Zlyvko, O. Research and Development of the Combined Cycle Power Plants Working on Supercritical Carbon Dioxide. Inventions 2022, 7, 76. https://doi.org/10.3390/inventions7030076

[3] World Energy Balances—International Energy Agency. Available online: https://www.iea.org/data-and-statistics/data-product/world-energy-balances.

[4] Uvarov, A.; Antonova, A.; Vorobjev, A. The Analysis of Initial Parameters of Steam in the Combined-Cycle Plant with High Temperature Gas Turbine. MATEC Web Conf. 2015, 37, 01062.

[5] Rogalev, A.; Komarov, I.; Kindra, V.; Osipov, S. Methods for Competitiveness

Improvement of High-Temperature Steam Turbine Power Plants. Inventions 2022, 7, 44.
[6] Vélez, F.; Segovia, J.J.; Martín, M.C.; Antolín, G.; Chejne, F.; Quijano, A. A technical, economical and market review of organic Rankine cycles for the conversion of low-grade heat for power generation. Renew. Sustain. Energy Rev. 2012, 16, 4175–4189.
[7] Rogalev, N.; Kindra, V.; Komarov, I.; Osipov, S.; Zlyvko, O.; Lvov, D. Comparative Analysis of Low-Grade Heat Utilization Methods for Thermal Power Plants with Back-Pressure Steam Turbines. Energies 2021, 14, 8519.
[8] Rogalev, A.; Rogalev, N.; Kindra, V.; Zlyvko, O.; Vegera, A. A Study of Low-Potential Heat Utilization Methods for Oxy-Fuel Combustion Power Cycles. Energies 2021, 14, 3364.
[9] converter-wec-array-simulator), MATLAB Central File Exchange. Retrieved April 4, 2023.
[10] Raj, S.., Shankar, G., Murugesan, S., Raju, M. N. ., Mohan, E., & Rani, P. J. I. (2023). Exploratory Data Analysis on Blueberry yield through Bayes and Function Models. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(11s), 634–641. https://doi.org/10.17762/ijritcc.v11i11s.8299
[11] Thivakaran, T. K. ., Priyanka, N. ., Antony, J. C. ., Surendran, S. ., Mohan, E. ., & Innisai Rani, P. J. . (2023). Exploratory Data Analysis for Textile Defect Detection. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(9s), 121–128. https://doi.org/10.17762/ijritcc.v11i9s.7403
[12] Rani, P. J. I. ., Venkatachalam, K. ., Sasikumar, D. ., Madhankumar, M. ., A., T. ., Senthilkumar, P. ., & Mohan, E. . (2024). An Optimal Approach on Electric Vehicle by using Functional Learning . *International Journal of Intelligent Systems and Applications in Engineering*, *12*(13s), 197–206
[13] Cao, Y.; Mihardjo, L.W.; Dahari, M.; Tlili, I. Waste heat from a biomass fueled gas turbine for power generation via an ORC or compressor inlet cooling via an absorption refrigeration cycle: A thermoeconomic comparison. Appl. Therm. Eng. 2021, 182, 116117.
[14] Kindra, V.; Rogalev, N.; Osipov, S.; Zlyvko, O.; Naumov, V. Research and Development of Trinary Power Cycles. Inventions 2022, 7, 56.
[15] Braimakis, K.; Karellas, S. Energetic Optimization of Regenerative Organic Rankine Cycle (ORC) Configurations. Energy Convers. Manag. 2018, 159, 353–370.
[16] Sun, W.; Yue, X.; Wang, Y. Exergy Efficiency Analysis of ORC (Organic Rankine Cycle) and ORC-Based Combined Cycles Driven by Low-Temperature Waste Heat. Energy Convers. Manag. 2017, 135, 63–73.
[17] Shi, L.; Shu, G.; Tian, H.; Deng, S. A Review of Modified Organic Rankine Cycles (ORCs) for Internal Combustion Engine Waste Heat Recovery (ICE-WHR). Renew. Sustain. Energy Rev. 2018, 92, 95–110.
[18] J. L. Lobo, I. Ballesteros, I. Oregi, J. del Ser, and S. Salcedo-Sanz, "Stream learning in energy IoT systems: a case study in combined cycle power plants," Energies, vol. 13, no. 3, 2020.
[19] Castillo.A, "Risk analysis and management in power outage and restoration: a literature survey," Electric Power Systems Research, vol. 107, pp. 9–15, 2014.
[20] Çetin.A, S. H. S. Erdem, S. H. Sevilgen, and A. V. Akkaya, "Electricity production cost analysis of a combined cycle power plant," Energy Sources, Part B: Economics, Planning, and Policy, vol. 3, no. 3, pp. 224–232, 2008.
[21] P. Tüfekci, "Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods," International Journal of Electrical Power & Energy Systems, vol. 60, pp. 126–140, 2014.

[22] U. Kesgin and H. Heperkan, “Simulation of thermodynamic systems using soft computing techniques,” International Journal of Energy Research, vol. 29, no. 7, pp. 581–611, 2005.

[23] E. Rich and K. S. Knight, “Chapter 3,” in Artificial Intelligence, McGraw Hill, New York, USA, 2nd Ed. edition, 1991.

[24] Mosavi A, F. S. Hosseini, B. Choubin et al., “Susceptibility prediction of groundwater hardness using ensemble machine learning models,” Water, vol. 12, no. 10, 2020.

[25] Mosavi A, F. Sajedi-Hosseini, B. Choubin, F. Taromideh, G. Rahi, and A. Dineva, “Susceptibility mapping of soil water erosion using machine learning models,” Water, vol. 12, no. 7, p. 1995, 2020.

[26] Choubin B, F. S. Hosseini, Z. Fried, and A. Mosavi, “Application of bayesian regularized neural networks for groundwater level modeling,” in 2020 IEEE 3rd International Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE), pp. 000209–000212, Budapest, Hungary, 2020.

[27] Mosavi A, F. Sajedi Hosseini, B. Choubin, M. Goodarzi, A. A. Dineva, and E. Rafiei Sardooi, “Ensemble boosting and bagging based machine learning models for groundwater potential prediction,” Water Resources Management, vol. 35, no. 1, pp. 23–37, 2021.

[28] F. Jiménez-Espadafor Aguilar, M. T. García, E. C. Trujillo, J. A. Becerra Villanueva, and F. J. Florencio Ojeda, “Prediction of performance, energy savings and increase in profitability of two gas turbine steam generator cogeneration plant, based on experimental data,” Energy, vol. 36, no. 2, pp. 742–754, 2011.

[29] E. A. Elfaki and A. H. Ahmed, “Prediction of electrical output power of combined cycle power plant using regression ANN model,” Engineering, vol. 6, no. 12, pp. 17–38, 2018.

[30] G. Şen, M. Nil, H. Mamur et al., “The effect of ambient temperature on electric power generation in natural gas combined cycle power plant-a case study,” Energy Reports, vol. 4, pp. 682–690, 2018.

[31] González-Díaz A, A. M. Alcaráz-Calderón, M. O. González-Díaz, Á. Méndez-Aranda, M. Lucquiaud, and J. M. González-Santaló, “Effect of the ambient conditions on gas turbine combined cycle power plants with post-combustion CO2 capture,” Energy, vol. 134, pp. 221–233, 2017.

[32] F. Chu, J. Wang, L. Nannan, T. Tan, and F. Wang, Eds.“Prediction of ccpp output based on improved fuzzy analytical hierarchy process,” in 2017 29th Chinese Control And Decision Conference (CCDC), F. Chu, J. Wang, L. Nannan, T. Tan, and F. Wang, Eds., pp. 3636–3641, Chongqing, China, 2017.

[33] Lorencin I, N. Anđelić, V. Mrzljak, and Z. Car, “Genetic algorithm approach to design of multi-layer perceptron for combined cycle power plant electrical power output estimation,” Energies, vol. 12, no. 22, 2019.

[34] Pınar Tüfekci, Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods, International Journal of Electrical Power & Energy Systems, Volume 60, September 2014, Pages 126-140, ISSN 0142-0615,

[35] Heysem Kaya, Pınar Tüfekci , Sadık Fikret Gürgen: Local and Global Learning Methods for Predicting Power of a Combined Gas & Steam Turbine, Proceedings of the International Conference on Emerging Trends in Computer and Electronics Engineering ICETCEE 2012, pp. 13-18.