All Sciences Proceedings
http://as-proceeding.com/

2nd International Conference on Contemporary Academic Research

November 4-5, 2023 : Konya, Turkey

ICCAR

https://as-proceeding.com/index.php/iccar

Estimation of the Energy Amount of the Solar Energy System for the House Roof with Adaboost Algorithm

Yasemin Ayaz Atalan*

Department of Mechanical Engineering, Yozgat Bozok University, Türkiye
Department of Energy Management, Çanakkale Onsekiz Mart University, Türkiye

(yasemin.ayaz@yobu.edu.tr)

Abstract – This study focuses on estimating the energy amount of the solar energy system installed on the house roof with the Adaboost Machine Learning (ML) algorithm. Root Mean Square Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Error (MAE) values of the model were calculated as 0.005, 0.068, and 0.015, respectively. The R^2 value of the algorithm was calculated as 0.995, indicating that the independent variables of the model have a very high ability to explain the dependent variable. These results show that this model can be used safely in critical applications such as solar energy production or consumption. This study shows that the Adaboost algorithm is a powerful model for solar energy prediction, and its predictions are incredibly close to actual values. In conclusion, this study is essential for homeowners and energy providers as accurate estimates of electricity consumption will help in more effective energy management and efficient use of resources. This work will lead to critical applications in energy management and effective energy use.

Keywords – Renewable Energy; Solar Energy; Machine Learning; Adaboost Algorithm; House Roof

I. Introduction

Renewable energy is increasingly important in world energy production and consumption and will continue to be an essential carrier of the energy sector in the future [1]. Such energy sources have the potential to reduce carbon emissions by reducing environmental impacts compared to energy production based on fossil fuels [2]. At the same time, renewable energy significantly contributes to energy security because it relies on locally available resources and helps reduce energy imports [3]. In the future, renewable energy technologies are expected to develop rapidly, and their costs to decrease. This will enable broader use of renewable resources such as solar, wind, hydroelectric, geothermal, and marine energy and make energy production more sustainable [4]. Additionally, these technologies will bring new business opportunities that will contribute to local economies [5]. Renewable energy is seen as a critical component to meet energy consumption in a clean and environmentally friendly way and plays a vital role in combating climate change [6]. In the future, renewable energy will be the cornerstone of the energy sector and form the basis of a sustainable energy future [7].

This study discussed solar energy, one of the renewable energy types. Solar energy is a clean and renewable source in which sunlight and heat are used and converted for electrical or heating purposes [8]. Sunlight generates electricity through photovoltaic (PV) panels or to obtain heat energy through solar collectors [9]. These technologies help reduce the environmental impacts of energy production based on fossil fuels by utilizing the free and unlimited energy source of the sun. Additionally, solar energy has great potential to reduce long-term energy costs and provide energy security [10].

As a part of renewable energy, solar energy is essential in combating climate change and sustainable energy production [11]. It reduces atmospheric carbon emissions by reducing greenhouse gas emissions from fossil fuel burning. Additionally, solar energy resources can be obtained locally, thus contributing to reducing energy imports and increasing energy security [12]. Renewable energy makes the energy sector more sustainable and environmentally friendly while offering economic benefits [13]. Solar energy will play an even more significant role in the future as an essential component of this transformation [14]. It will minimize environmental impacts while meeting energy needs [15].

Machine learning (ML) is a branch of artificial intelligence that allows computer systems to gain data analysis and learning capabilities [16]. This discipline is used in many application areas and is essential in processing large data sets, pattern recognition, prediction, decision-making, and automation [17]. ML develops the ability to learn from data using algorithms and statistical methods and thus solve complex problems [18], [19]. Without human intervention, computers know from experience, discover patterns within data, and predict future events [20]. This technology is used in many areas, from online platforms offering customized recommendations to making medical diagnoses and guiding driverless vehicles [21]. It includes subfields such as ML, data mining, deep learning, and tracking learning and is developing day by day, creating a significant impact in different industries [22].

ML is a powerful tool vital in solar energy forecasting [23]. Solar energy forecasting involves predicting electricity production based on sunlight so that photovoltaic energy systems can operate efficiently [24]. These predictions are critical to providing the right amount of energy to energy producers, distribution networks, and consumers. By analyzing the data used in solar energy forecasting, ML can predict future energy production by considering sunlight intensity, weather conditions, seasonal variables, and other factors [25]. This helps energy companies and facility owners balance energy production and consumption, energy plans, and manage resources more effectively. ML can also improve the performance of solar systems by creating continuously improving models and predictions. As a result, ML contributes to sustainable energy production by increasing accuracy and efficiency in solar energy forecasting.

II. Materials and Method

This study aims to obtain prediction data using the Adaboost model, one of the ML algorithms, from the data obtained from solar energy panels installed on the roof of a house. Data from a PV system located on the roof of a house in an area of 18 square meters was considered. In this system, the energy obtained between 6:00 AM and 1:00 PM during the day is accepted as the sum of the data. Two independent and one dependent variables were used for this study. Daily temperature and sunshine hours were treated as independent variables. The amount of energy produced daily was considered the dependent variable. One-year data for the dependent variable is given in Figure 1.

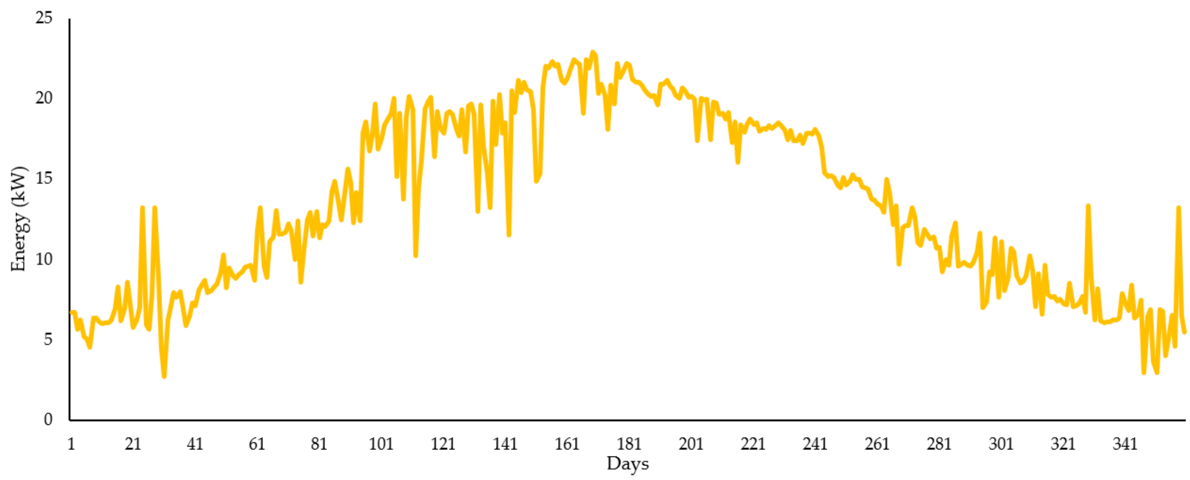

Figure 1. One-year data for the dependent variable

ML is a subfield of artificial intelligence that enables computer systems to perform specific tasks by analyzing data and automatically learning patterns and relationships [26]. ML uses large amounts of data to improve a model's ability to perform a particular task. The first step involves data collection and cleaning processes. Then, this data is fed into ML algorithms. These algorithms use statistical methods and mathematical calculations to identify patterns and relationships between data. Once the model analyzes and learns from the data, it can make predictions or decisions by processing new inputs. ML is significant in solving complex problems using image recognition, natural language processing, automation, recommendation systems, and many other areas [27]. It offers new opportunities with constantly developed techniques.

This study ran the Adaboost model, one of the ML algorithms. AdaBoost (Adaptive Boosting) is an ensemble learning algorithm that aims to create a robust classifier by combining weak learners (usually weak learning algorithms or classifiers) [28]. AdaBoost is a frequently used method in data mining and ML. The algorithm first builds a base weak classifier on the data and then strengthens this classifier through iterations, focusing on poorly classified examples [29]. Each iteration increases the weights of misclassified samples, and the subsequent weak classifier is trained based on these weights. As a result, AdaBoost allows combining a set of weak learners to obtain a higher accuracy and robust classifier. AdaBoost has been used successfully in many application areas, such as face recognition, object detection, and spam filter, and is an essential example of ensemble learning methods. The ML model developed for this study is given in Figure 2.

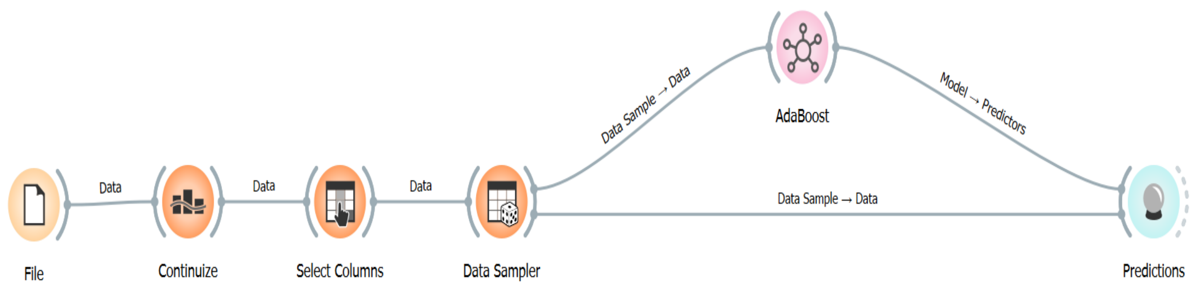

Figure 2. The flow chart of the ML model

Root Mean Square Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2) are performance measures and error metrics frequently used in the field of ML. MSE calculates how far predictions are from actual values and measures errors by squaring these differences. RMSE is the square root of MSE and provides a more understandable value on the scale of the errors in the original data. MAE measures error magnitude by averaging the absolute values of forecast errors and is considered a more robust metric than RMSE. R-squared (R^2) measures how well the independent variables explain the dependent variable and shows what percentage of the variation they can present. RMSE, MSE, MAE, and R^2 are essential tools to evaluate model performance and compare different models. Each contributes to the model selection and improvement processes by examining various aspects of the model, such as accuracy, error size, and explanatory power. In this study, RMSE, MSE, MAE, and R^2 values were calculated to measure the operating performance of the Adaboost algorithm.

III. RESULTS

The performance values obtained by the Adaboost algorithm for solar energy estimation seem impressive. Calculating the RMSE value as 0.005 indicates that the model's predictions are, on average, only 0.005 units away from the actual values. The MSE value is 0.068, representing the average sum of squares of the model's prediction errors. Since this value is relatively low, the model's predictions are generally consistent with the actual values with low error. The MAE value is 0.015, which represents the average of the absolute values of the model's prediction errors. A low

MAE value indicates that the model's predictions are generally close to actual values and avoid significant errors. Table 1 contains the performance values of the Adaboost algorithm.

Table 1. The performance values obtained by the Adaboost algorithm

Model	**RMSE**	**MSE**	**MAE**	**R2**
Adaboost	0.005	0.068	0.015	0.995

It shows that the ability of the model's independent variables to explain the dependent variable is relatively high when the R^2 value is calculated as 0.995. The closer the R^2 value is to 1, the better the model describes the data. These comments show that the Adaboost algorithm is a powerful model for solar energy prediction, and its predictions are incredibly close to the actual values. These results indicate that this model can be used reliably in critical applications such as solar energy production or consumption. For this study, the Adaboost algorithm was used to estimate the electricity consumption of a house. The prediction data and actual data based on the Adaboost algorithm are shown in Figure 3.

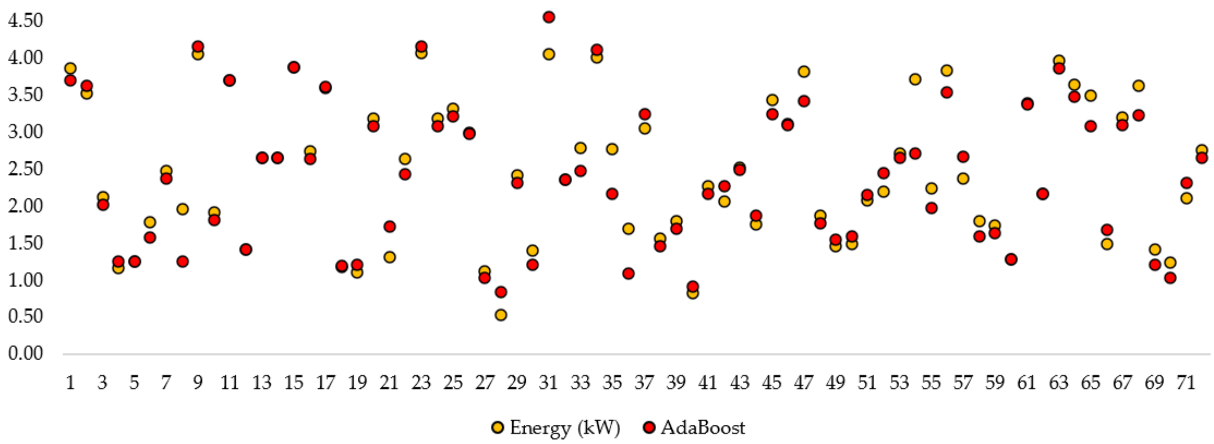

Figure 3. The prediction data and actual data based on the Adaboost algorithm

Table 2 shows the descriptive statistics results of the actual and prediction data. With these data, similarities between forecast and actual data emerge.

Table 2: The descriptive statistics results of the actual and prediction data.

Variable	Actual Data	Prediction Data
Total Count	72.00	72.00
Mean	2.478	2.411
SE Mean	0.114	0.112
StDev	0.963	0.951
Variance	0.928	0.904
CoefVar	38.87	39.43
Minimum	0.535	0.835
Q1	1,706	1,594
Median	2,396	2,372
Q3	3,380	3,192
Maximum	4,065	4,559
Range	3,530	3,724
IQR	1,674	1,598
Skewness	0.08	0.24
Kurtosis	-1.15	-0.93

There are 72 data points in both the actual and prediction data, indicating that the two data sets are the same size. While the average value of the actual data is 2.478, the average value of the prediction data is slightly lower at 2.411. This indicates that the predictions could be underestimated compared to the actual values. The average error for both data sets is quite similar and ranges from approximately 0.112 to 0.114. This shows that the intermediate error level of the predictions is identical. While the standard deviation of the actual data is 0.963, the standard deviation of the prediction data is slightly lower at 0.951. The variance for both data sets is similar and was calculated as 0.928 for the actual data and 0.904 for the prediction data.

The coefficient variability of the two data sets is also at close values. It was calculated as 38.87% for real data and 39.43% for forecast data. While the minimum value in real data is 0.535 and the maximum value is 4.065, in the prediction data, the minimum value is 0.835, and the maximum value is 4.559. Quartile values for the two data sets are similar. While Q1 = 1.706, Median = 2.396 and Q3 = 3.380 in real data, it is calculated as Q1 = 1.594, Median = 2.372 and Q3 = 3.192 in prediction data. The range and IQR (Quarterly Range) are at similar values for both data sets. The range is calculated as 3.530 in actual data and 3.724 in forecast data, while the IQR is calculated as 1.674 in actual data and 1.598 in forecast data.

Skewness and kurtosis values for the two data sets are also close. Skewness was calculated as 0.08 in actual data and 0.24 in forecast data. Kurtosis was calculated as -1.15 in actual data and -0.93 in forecast data. This shows that the distributions of both data sets have similar shapes. As a result, a generally identical distribution and central tendency is observed between the actual data and the forecast data. Still, the average value of the forecasts is slightly below the actual values. These results suggest that the model does an excellent job on a particular prediction task, but there are opportunities for improvement.

In this study, using the Adaboost algorithm to estimate the electricity consumption of a house is a very successful application. RMSE, MSE, MAE, and R^2 performance values calculated based on the data show that the model's predictive ability is relatively high, and its predictions are incredibly close to the actual data. These results are essential for homeowners and energy providers because an accurate estimate of electricity consumption will help with more effective energy management and efficient use of resources.

The use of Adaboost shows that it is resistant to the complexity and variability of data content, especially since it is among the ensemble learning methods. This means the ability to make robust predictions even in scenarios that consider factors such as different weather conditions, seasonal changes, or changes in the characteristics of the house. As a result, it is emphasized in this study that using the Adaboost algorithm to estimate the electricity consumption of a house is a valuable tool to increase energy efficiency, reduce energy costs, and contribute to sustainable energy use. This study will lead to critical applications in energy management and the effective use of energy resources.

IV. CONCLUSION

In this study, the Adaboost model, one of the ML algorithms, was used to obtain prediction data by considering the solar energy system to meet the electricity consumption of a house. Using the Adaboost algorithm to estimate the electricity consumption of a house has produced very successful results. The calculated RMSE, MSE, MAE, and R^2 performance values show that the predictive ability of the model is quite strong, and its predictions are incredibly close to the actual data. These findings are essential for homeowners and energy providers because precise estimates of electricity consumption help with more effective energy management and efficient allocation of resources.

Additionally, the choice to use the Adaboost algorithm emphasizes its robustness against data complexity and variability, especially since it is among the ensemble learning methods. This robustness means the model can make robust predictions even in scenarios involving different weather conditions, seasonal changes, or home characteristics. In conclusion, this study highlights that the Adaboost algorithm is a valuable tool for predicting a home's electricity consumption. It reveals its potential to increase energy efficiency, reduce energy costs, and contribute to sustainable energy use. This research will lead to critical applications in energy management and effective use of energy resources, promising hope for a more efficient and environmentally friendly future.

REFERENCES

[1] Y. A. Atalan and A. Atalan, “Integration of the Machine Learning Algorithms and I-MR Statistical Process Control for Solar Energy,” *Sustainability*, vol. 15, no. 18, p. 13782, Sep. 2023, doi: 10.3390/su151813782.

[2] N. Hagumimana *et al.*, “Concentrated Solar Power and Photovoltaic Systems: A New Approach to Boost Sustainable Energy for All (Se4all) in Rwanda,” *Int. J. Photoenergy*, vol. 2021, pp. 1–32, Jun. 2021, doi: 10.1155/2021/5515513.

[3] B. Dong, C. Cao, and S. E. Lee, “Applying support vector machines to predict building energy consumption in tropical region,” *Energy Build.*, vol.

37, no. 5, pp. 545–553, May 2005, doi: 10.1016/J.ENBUILD.2004.09.009.

[4] Y. Fang, "Economic welfare impacts from renewable energy consumption: The China experience," *Renew. Sustain. Energy Rev.*, vol. 15, no. 9, pp. 5120–5128, Dec. 2011, doi: 10.1016/J.RSER.2011.07.044.

[5] Y. A. Atalan, A. Keskin, and A. Atalan, "Factors Affecting Clean Energy Investments for Zero Emissions: OECD Countries," in *Clean Energy Investments for Zero Emission Projects: An Analysis on How to Reduce the Carbon Footprint*, Springer, 2022, pp. 203–219.

[6] Y. Ayaz Atalan, M. Tayanç, K. Erkan, and A. Atalan, "Development of Nonlinear Optimization Models for Wind Power Plants Using Box-Behnken Design of Experiment: A Case Study for Turkey," *Sustainability*, vol. 12, no. 15, p. 6017, Jul. 2020, doi: 10.3390/su12156017.

[7] A. Atalan and Y. A. Atalan, "Nonlinear Optimization Models of Box-Behnken Experimental Design: Turbine Simulation for Wind Power Plant".

[8] M. Rajendran, M. Chinnasamy, S. Muthusamy, and M. K. Nair, "Introduction to Solar Energy Conversion," *Mater. Sol. Energy Convers. Mater. Methods Appl.*, pp. 1–31, 2021.

[9] W. G. J. van Helden, R. J. C. van Zolingen, and H. A. Zondag, "PV thermal systems: PV panels supplying renewable electricity and heat," *Prog. photovoltaics Res. Appl.*, vol. 12, no. 6, pp. 415–426, 2004.

[10] E. Hache, "Do renewable energies improve energy security in the long run?," *Int. Econ.*, vol. 156, pp. 127–135, 2018.

[11] P. A. Owusu and S. Asumadu-Sarkodie, "A review of renewable energy sources, sustainability issues and climate change mitigation," *Cogent Eng.*, vol. 3, no. 1, p. 1167990, 2016.

[12] A. Shahsavari and M. Akbari, "Potential of solar energy in developing countries for reducing energy-related emissions," *Renew. Sustain. Energy Rev.*, vol. 90, pp. 275–291, 2018.

[13] S. R. Bull, "Renewable energy today and tomorrow," *Proc. IEEE*, vol. 89, no. 8, pp. 1216–1226, 2001.

[14] S. Jacobsson, B. Sandén, and L. Bångens, "Transforming the energy system—the evolution of the German technological system for solar cells," *Technol. Anal. Strateg. Manag.*, vol. 16, no. 1, pp. 3–30, 2004.

[15] S. Kanwal, M. T. Mehran, M. Hassan, M. Anwar, S. R. Naqvi, and A. H. Khoja, "An integrated future approach for the energy security of Pakistan: Replacement of fossil fuels with syngas for better environment and socio-economic development," *Renew. Sustain. Energy Rev.*, vol. 156, p. 111978, 2022.

[16] A. Atalan, "Forecasting drinking milk price based on economic, social, and environmental factors using machine learning algorithms," *Agribusiness*, vol. 39, no. 1, pp. 214–241, Jan. 2023, doi: 10.1002/agr.21773.

[17] A. Atalan, C. Ç. Dönmez, and Y. Ayaz Atalan, "Yüksek-Eğitimli Uzman Hemşire İstihdamı ile Acil Servis Kalitesinin Yükseltilmesi için Simülasyon Uygulaması: Türkiye Sağlık Sistemi," *Marmara Fen Bilim. Derg.*, vol. 30, no. 4, pp. 318–338, Dec. 2018, doi: 10.7240/marufbd.395255.

[18] A. Atalan, H. Şahin, and Y. A. Atalan, "Integration of Machine Learning Algorithms and Discrete-Event Simulation for the Cost of Healthcare Resources," *Healthcare*, vol. 10, no. 10, p. 1920, Sep. 2022, doi: 10.3390/healthcare10101920.

[19] Y. Ayaz Atalan and A. Atalan, "A Statistical Analysis of the Relationship Between Meteorological Parameters and the Spread of COVID-19 Cases: Comparison Between Turkey and Italy," *J. Stat. Appl. Sci.*, vol. 1, no. 2, pp. 76–84, 2020.

[20] A. Atalan, "Effect of Healthcare Expendıture on the Correlatıon Between the Number of Nurses and Doctors Employed," *Int. J. Heal. Manag. Tour.*, vol. 6, no. 2, pp. 515–525, Jul. 2021, doi: 10.31201/ijhmt.949500.

[21] A. Atalan, "Logistic Performance Index of OECD Members," *Akad. Araştırmalar ve Çalışmalar Derg.*, vol. 12, no. 23, pp. 608–619, Nov. 2020, doi: 10.20990/kilisiibfakademik.720604.

[22] A. Atalan, "Neural Network and Random Forest Algorithms for Estimation of the Waiting Times Based on the DES in ED," in *International Conference on Contemporary Academic Research*, 2023, pp. 14–20.

[23] C. Voyant *et al.*, "Machine learning methods for solar radiation forecasting: A review," *Renew. energy*, vol. 105, pp. 569–582, 2017.

[24] C. Wan, J. Zhao, Y. Song, Z. Xu, J. Lin, and Z. Hu, "Photovoltaic and solar power forecasting for smart grid energy management," *CSEE J. Power Energy Syst.*, vol. 1, no. 4, pp. 38–46, 2015.

[25] J. Devaraj, R. Madurai Elavarasan, G. M. Shafiullah, T. Jamal, and I. Khan, "A holistic review on energy forecasting using big data and deep learning models," *Int. J. energy Res.*, vol. 45, no. 9, pp. 13489–13530, 2021.

[26] A. Atalan, "Statistical optimization of forecast data from Adaptive Boosting and Support Vector Machine Algorithms," in *International Conference on Engineering, Natural and Social Sciences*, 2023, pp. 571–579.

[27] J. Chen *et al.*, “A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide,” *Environ. Int.*, vol. 130, p. 104934, 2019, doi: https://doi.org/10.1016/j.envint.2019.104934.

[28] X. Li, L. Wang, and E. Sung, “A study of AdaBoost with SVM based weak learners,” in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, IEEE, 2005, pp. 196–201.

[29] P. Pandey and R. Prabhakar, “An analysis of machine learning techniques (J48 & AdaBoost)-for classification,” in *2016 1st India International Conference on Information Processing (IICIP)*, IEEE, Aug. 2016, pp. 1–6. doi: 10.1109/IICIP.2016.7975394.