Energy Reports 7 (2021) 8465–8475

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Predicting the energy output of hybrid PV–wind renewable energy system using feature selection technique for smart grids

Zakria Qadir [a,*], Sara Imran Khan [b], Erfan Khalaji [c], Hafiz Suliman Munawar [d], Fadi Al-Turjman [e], M.A. Parvez Mahmud [f], Abbas Z. Kouzani [f], Khoa Le [a]

[a] School of Computing Engineering and Mathematics, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia
[b] Faculty of Chemical Energy, University of New South Wales, Kensington, Sydney NSW 2052, Australia
[c] Computer Engineering Department, Middle East Technical University, Northern Cyprus Campus, 99738 Kalkanli, Guzelyurt, Mersin 10, Turkey
[d] Faculty of Built Environment, University of New South Wales, Kensington, Sydney NSW 2052, Australia
[e] Research Center for AI and IoT, Artificial Intelligence Department, Near East University, Nicosia, Mersin 10, Turkey
[f] School of Engineering, Deakin University Locked Bag 20000, Geelong, VIC 3220, Australia

ARTICLE INFO

Article history:
Received 20 September 2020
Received in revised form 22 December 2020
Accepted 11 January 2021
Available online 21 January 2021

Keywords:
Smart grids
Regression models
Feature selection
Prediction accuracy
Renewable energy system
ANN
RFECV

ABSTRACT

In the current technological era, predicting the power and energy output based on the changing weather factors play an important role in the economic growth of the renewable energy sector. Unlike traditional fossil fuel-based resources, renewable energy sources potentially play a pivotal role in sustaining a country's economy and improving the quality of life. As our planet is nowadays facing serious challenges due to climate change and global warming, this research could be effective to achieve good prediction accuracy in smart grids using different weather conditions. In the current study, different machine learning models are compared to estimate power and energy of hybrid photovoltaic (PV)-wind renewable energy systems using seven weather factors that have a significant impact on the output of the PV–wind renewable energy system. This study classified the machine learning model which could be potentially useful and efficient to predict energy and power. The historic hourly data is processed with and without data manipulation. While data manipulations are carried out using recursive feature elimination using cross-validation (RFECV). The data is trained using artificial neural network (ANN) regressors and correlations between different features within the dataset are identified. The main aim is to find meaningful patterns that could help statistical learning models train themselves based on these usage patterns. The results suggest that opting feature selection technique using linear regression model outperforms all the other models in all evaluation metrics having to mean squared error (MSE) of 0.000000104, mean absolute error (MAE) of 0.00083, R^2 of 99.6%, and computation time of 0.02 s The results investigated depict that the sustainable computational scheme introduced has vast potential to enhance smart grids efficiency by predicting the energy produced by renewable energy systems.

1. Introduction

Over recent years the power industry has switched its focus to renewable energy sources to reduce its carbon footprint during energy generation (Sharif et al., 2019; Qadir et al., 2019a). PV and wind have been widely implemented as alternatives in the hybrid power system due to their renewable nature and ease of availability. The hybrid PV–wind power plants have been investigated due to its variability, being technical efficient and cost-effective (Mohammadi and Mehdi, 2018; Fasihi and Breyer, 2020; Groppi et al., 2020; Xu et al., 2020). Due to the variable nature of energy generation from the PV and wind power systems, the managers must control and operate the power plant efficiently (Qadir et al., 2018a). There is a need to forecast the PV–wind energy generation for short- and long-term planning of power transmission. A lot of research has been carried out recently on feasibility analysis and forecasting the power generation by PV–wind systems (Abujubbeh et al., 2019; Zhang et al., 2019). Some studies have used the approach to forecast the input variables (solar irradiance and wind) and calculated the power generated by the hybrid plant using models, while some have focused on direct power forecast of the renewable energy source (Das et al., 2018). Various prediction models such as persistent model, k-nearest

* Corresponding author.
E-mail addresses: Z.Qadir@westernsydney.edu.au (Z. Qadir), s.imrankhan@unsw.edu.au (S.I. Khan), erfan.khalaji@metu.edu.tr (E. Khalaji), h.munawar@unsw.edu.au (H.S. Munawar), fadi.alturjman@neu.edu.tr (F. Al-Turjman), m.a.mahmud@deakin.edu.au (M.A.P. Mahmud), abbas.kouzani@deakin.edu.au (A.Z. Kouzani), K.Le@westernsydney.edu.au (K. Le).

https://doi.org/10.1016/j.egyr.2021.01.018
2352-4847/

Table 1
List of Abbreviations.

Symbol	Definition
PV	Photovoltaic
RFECV	Recursive feature elimination using cross-validation
ANN	Artificial neural network
MSE	Mean squared error
MAE	Mean absolute error
R2	Correlation coefficient
GA	Genetic algorithm
kNNs	K-nearest neighbors
ICT	Information and communications technology
AI	Artificial intelligence
Si	Solar irradiation
Ws	Wind speed
Ta	Ambient temperature
H	Humidity
R	Precipitation
Pa	Atmospheric pressure
Wd	Wind direction
MLP	Multilayer perceptron
IAPS	Invasive alien plant species
METU	Middle east technical university
NCC	North Cyprus campus
GPS	Global positioning system
SVR	Support vector regression
FCME	Forward consecutive mean excision
RH	Relative humidity
ML	Machine learning

neighbors (kNNs), SVM, ANN and genetic algorithm (GA) have been investigated for solar power forecasting (Pedro and Coimbra, 2012; Qadir et al., 2018c; Rashid et al., 2018). The abbreviations used in this study are summarized in Table 1.

The interest of researchers in ANN applications has exploded recently. Many of the new applications are introduced primarily focusing on technological and development issues related to ANN. These applications are not limited to one area but submerge many fields such as agricultural production, environment, energy generation, engineering and science, finance and management, policy, and security (Munawar, 2020a,b; Munawar et al., 2020, 2019a, 2017).

Further examples include stock market, banking, quality prediction of crude oil, money laundering, water treatment, crime detection, etc (Li et al., 2019a). The relationship of ANN applications with classification, pattern-recognition, and prediction and number of publications against different sectors are summarized in Table 2. ANN is a vast field and can solve any problem related to different sectors. The different frameworks, models, algorithms, and schemes are always available to predict, classify, or recognize patterns in any emerging field. ANN applications are applied in different sectors. However, there is an utmost need for robust ANN prediction models related to the energy sector that can be analyzed to utilize energy in a much sustainable and efficient way.

Artificial Neural Network (ANN) is the model inspired by the human nervous system to assimilate nerves in any environment and perform different activities. The architecture of regression in ANN is illustrated in Fig. 1. The ANN regression model comprises of input, weight, error (the difference between an exact value and some approximation to it), transfer function (defines various input–output relationships), activation function (defines the output of node given an input or set of inputs) and output (Ahmad et al., 2014; Saritas and Yasar, 2019). The most common task that an ANN can perform is pattern recognition, prediction, image classification, clustering, signal processing, social networking, machine learning techniques (Koesdwiady et al., 2016; DiCarlo et al., 2012; Park et al., 2009; Kriegeskorte et al., 2008; Kruger et al., 2012). Currently, information and communications technology (ICT) hosts a lot of hot topics related to artificial intelligence (AI), such as machine learning, deep learning, neural networks, cloud computing, big data, wireless communication and information security (Zhang et al., 2018; Luo et al., 2018; Murphey et al., 2012; Brundage et al., 2018; Qadir et al., 2018b; Munawar et al., 2019b, 2020a,b) Under the umbrella of ANNs lie data analysis factors, such as computational time, accuracy, performance, latency, scalability, and fault tolerance. These factors are useful to calculate the prediction accuracy, for example, MSE, MAE, coefficient of determination (R^2), and time (Mozaffari et al., 2018; He and Garcia, 2008). It has high-speed performance capability in massive parallel implementation heightening the need to do comprehensive research in this domain Huang (2017) and Izeboudjen et al. (2014). In the numerical paradigm, ANN is widely used in universal function approximation because of their unique capability of adaptivity, self-learning, fault tolerance, advancement, and non-linearity in input to output mapping (Wang et al., 2017). Moreover, for handling complex and non-complex problems, these data analysis factors provide a clear picture. Therefore, ANNs are preferred to be used (effectiveness, success, and efficiency) in providing high data handling capability.

Table 2
Results related to classification, pattern-recognition, and prediction in several ANN Applications.

ANN applications	Classification	Pattern recognition	Prediction	Total
Agriculture	2	3	3	7
Energy	2	15	5	22
Engineering	2	7	22	31
Environmental	2	15	10	27
Finance	2	15	10	27
Management	2	2	40	44
Manufacturing	5	15	12	32
Medical science	2	5	10	17
Mining	2	15	2	19
Policy	2	2	2	6
Science	2	25	25	52
Security	2	18	20	40
Weather/Climate	2	15	2	19
Other fields	10	11	52	71

The choice of training algorithms used to train ANNs in terms of computational time and accuracy depends on many factors, such as size (dimension) of the dataset, weight and biases of the network, number of delays, network complexity and its architecture, splitting of dataset for training, validation and testing purpose and last but not the least is the acceptable errors (error histogram) and autocorrelation between training and test data (Rahmanifard and Plaksina, 2018; Araque et al., 2017; Sheng et al., 2017). ANN energy prediction pipeline consists of five basic steps to select an accurate prediction model as shown in Fig. 2 (Taborda et al., 2015). The input data from the given database is split into training and testing according to the type of problem being addressed. In the second and third steps, feed-forward or feed-back connections are selected along with the parameters to train the ANN model (Xing et al., 2018; Kingston et al., 2005). In the fourth and fifth step, the error values are calculated based on R^2, MSE, MAE, time, and the prediction model is selected based on the least errors and higher accuracy.

There are several problems related to time-consumption convergence, variable quantization and artificial neural system (ANS) using supervised learning that needs to be addressed. This study highlights some of these shortcomings as follows:

• improving the prediction capability of ANNs and making them robust. Additionally, training a generalized range of data to enhance prediction accuracy.

• complete knowledge retrieval from trained ANNs and model transparency to deeply understand the data transfer and processing from input to the output layer.

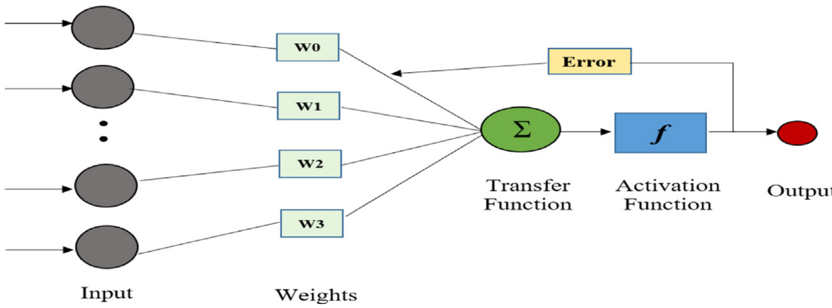


Fig. 1. The architecture of ANN Regression.

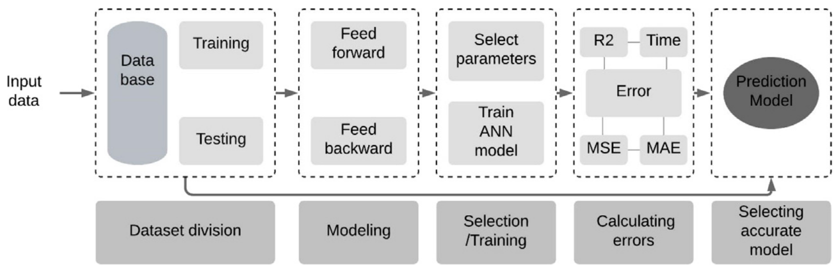


Fig. 2. Generic ANN prediction pipeline.

• enhancing the extrapolation ability of ANNs to design a model that can predict the outward range of data accurately.

• improving the efficacy of ANN prediction algorithms to avoid uncertainty.

Hence the main aim of the study is to enhance the forecast accuracy of hybrid PV–wind system using ANN models by considering different weather factors. The weather factors selected for the study were Solar Irradiation (Si), Wind Speed (Ws), Ambient Temperature (Ta), Humidity (H), Precipitation (R), Atmospheric Pressure (Pa) and Wind Direction (Wd). The intra- and interrelationship of different weather factors and their impacts on PV and wind energy systems are analyzed. This will help in developing a framework for generating a forecast model and compensate for any missing data. This paper is organized as follows: Section 2 discusses the brief literature review on energy forecasting of hybrid PV–wind system and compare different regression models. Section 3 presents a detailed methodology, exploring the collected data using data visualization techniques and a correlation between different features in the dataset. The main goal of the section is to find meaningful patterns that help statistical learning models train themselves based on these usage patterns. Finally, Section 4 concludes the paper along with indications of possible future works.

2. Literature review

The solar irradiance and wind speed play a key role in output power generation from the PV panel or a wind turbine. The weather conditions and the time of the day impact the solar intensity and wind speed reaching the earth surface. This cause rapid change in power output from the PV or wind turbine to the grid stations with time, resulting in instability. This necessitates the need for precise forecasting so that the operator can estimate the difference in the forecasted and generated energy for balancing the grid performance at minimum cost (Al-Turjman et al., 2020). It helps the plant manager and customer to avoid unexpected power shortage, uncertainty, and lower cost of energy (Antonanzas et al., 2016).

Fig. 3 highlights the ANN applications in developing regression models. A keyword analysis was performed using the VosViewer software for the articles retrieved to highlight the focus of the ANN articles published in the last decade. Recent literature studies from the last two decades revolve around keywords such as data mining, feature selection, prediction accuracy, regression, pattern recognition, data processing and others (Kadam, 2020; Li et al., 2019b; Cui et al., 2019; Qadir et al., 2019b). This entirely portrays a central focus on data retrieval in developing regression models, which is in line with recent literature (Jiang et al., 2019; Leng et al., 2018; Jiang and Guoqing, 2017), where it is stated that ANN regression models help in predicting the weather factors based on feature selection.

The most influential input climatic parameters can be directly fed to ANN prediction models. Marquez and Coimbra (Marquez and Coimbra, 2013) selected and analyzed the most relevant input variables from several climatic parameters using the Gamma test-based strategy. Later, a genetic algorithm search is included for speeding the process and finding the relevant combination of input variables. The experimental results depict that the selected inputs are temperature, precipitation, cloud cover, and solar geometry. Moreover, by using these inputs, the values of R^2, MSE, and RMSE comes out to be 94.7%, −0.6, and 0.177, respectively. The use of additional input climatological factors was investigated by Sfetsos and Connick (Sfetsos and Coonick, 2000) for predicting the solar power output. Analyzing the trial-and-error method the two-step technique was considered by the authors. In this technique, minimal errors are achieved initially by training the model. Significantly, the influence of input parameters is abolished by changing it with either zero or mean value.

Implementation issues can be faced while integrating several input variables to an ANN model. The foremost one is the elevation in computation time which is consumed during the training process. Additionally, it can also increase the risk of having redundant parameters which may complicate the training process and can cause a drastic increase in the prediction errors (Xiang et al., 2016). This scenario is most common while using multilayer perceptron (MLP), as each of the hidden neurons is multiplied by the input variables leading to a complex network. To recapitulate, an additional dimension to the output space is caused by the addition of variable in the input data, and to represent the mapping relationship, the training stage will need more data to occupy the

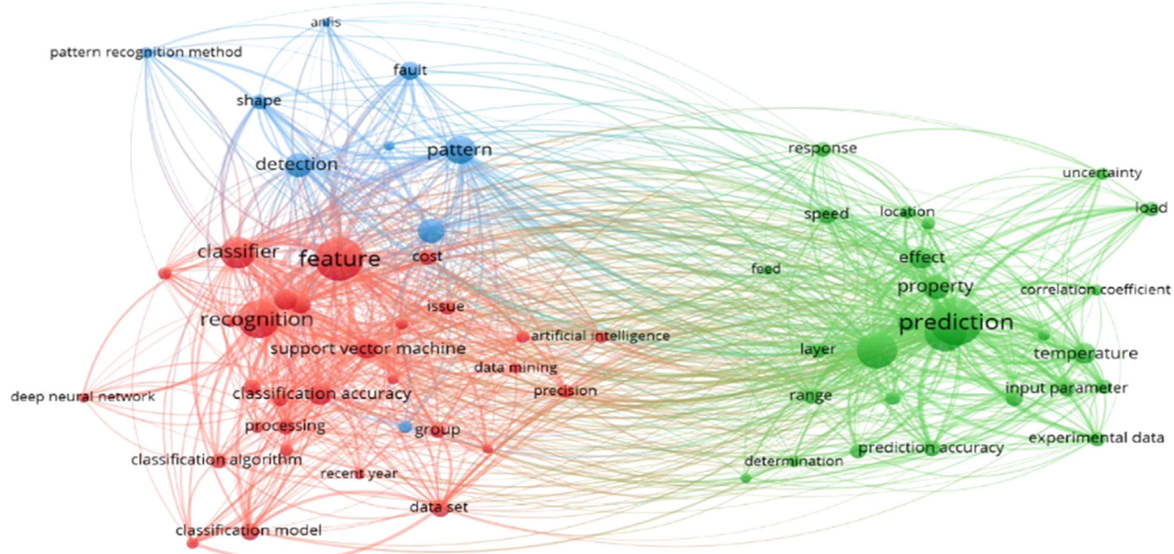


Fig. 3. Most frequent keywords used in ANN-based regression models from 2000 to 2020.

space densely (De Martino and De Martino, 2018). Currently, the development of smart solutions, as well as the implementation of various applications using smart wearables and accessories, is being investigated quite extensively (Jaeger and Haas, 2004; Haykin and Neural Network, 2004; Floreano and Mattiussi, 2008; Guo and Zhu, 2018; Al-Mahasneh et al., 2017; Qadir et al., 2021b). Furthermore, various processes of predicting the energy output of a PV and wind plants as a standalone system has also been considered in various studies (Kulkarni et al., 2004; Yao et al., 2012). In another study (das Merces Machado et al., 2009), the authors mention that if the weak models provide significant changes according to the metrics, we use to evaluate them, the bagging models can provide somewhat better performance. In Hirose and Yoshida (2012), the authors used to boost, and bagging models significantly help improve model accuracy. In Islam et al. (2017), the authors used a linear regression model that provide consistent results, and they use the least square estimation to evaluate the model.

In Marquez and Coimbra (2011), the authors consider linear regression analysis as an important model to be explored. They mention that linear regression enables the identification and characterization of relations between different attributes. We also show in the next sections that the features in our dataset have a sort of distribution that makes a linear regression model capable of learning very well from it. In the results section, we show that this characteristic makes the linear regression outperform all the other regressors. In Priddy and Keller (2005), the authors use linear regression analysis to statistically characterize the relations among the population and gross local products and the request for receiving energy by the residential, transport, and commercial fields. Their main purpose is to determine the energy demand trends over a long period. They mention that the linear regression model for estimating energy demands is an effective method. In Marquez and Coimbra (2011), the author mentions that linear regression is one of the simplest yet popular ways to measure the relationship between continuous predictors. Considering this fact, we show in the results that due to the same reasons that Sfetsos and Connick (2000) mentions, linear regression outperforms all the other predictors, including the most advanced cases such as bagging and boosting approaches. In Priddy and Keller (2005) knowledge engineering, machine learning, and deep learning are used to address sustainable development. It is stated that these tools are highly effective in passing from the obstacles on the path to a sustainable future. In Son et al. (2018) the author claims that as AI moves into environmental health research, near-term opportunities for the technology are arising on several fronts. In Coyle et al. (2007) the authors argue that the combination of massive environmental genomics microbial data with machine learning algorithms can be extremely powerful for biomonitoring programs and pave the way to fill important gaps in our understanding of microbial ecology. In An et al. (2017) the author proposes a machine learning approach to accurately detect the spatial extent of invasive alien plant species (IAPS) to map their spread over time or model their potential invasion area.

As intrigued by the literature review, the importance of using data science in environmental sustainability is undeniable. There are many types of research conducted to address a wide range of question while machine learning and deep learning is their permanent mean of tackling the issues. The reason could be the power of the state-of-the-art machine learning and deep learning models to capture the differences in a huge amount of a collected dataset. This is vital to make sense out of the gathered information by analysis and finding similarities and differences in the pattern which can be found in the data. Finding a potentially useful pattern in a dataset can give us a better insight into the data that we have collected and address relevant issues with a better and a deeper level of knowledge about the gathered data.

3. Methodology

In this study, data visualization techniques were explored. The goal was to find meaningful patterns that help statistical learning models train themselves based on the usage patterns. It selects the relevant feature i.e., the weather factors for the process to make a robust model. Note that here the target feature could be power or energy. The historical hourly metrological dataset was gathered using calibrated sensors deployed at Middle East Technical University (METU), NCC, from 1st Jan till 26th Dec 2015. The weather factors selected for the study were Solar Irradiation (Si), Wind Speed (Ws), Ambient Temperature (Ta), Humidity (H), Precipitation (R), Atmospheric Pressure (Pa) and Wind Direction (Wd). The global positioning system (GPS) module of Raspberry Pi was used for data collection and storage in clouds. The collected data were processed with and without data manipulation. While data manipulations were carried out using RFECV. Data were explored using ANN prediction algorithms and correlations between different features within the dataset were identified. The main aim was to find meaningful patterns that could help statistical learning models train themselves based on these usage patterns.

Fig. 4 illustrates the development model for predicting the power output of PV–wind system based on different regressor models. The 1st stage determines the weather data extracted from the calibrated sensors. The extracted raw data is then pre-processed, and feature engineering is applied in stage 2 to get the real values, eliminating any garbage value which can malfunction our model. Feature selection is opted in stage 3 to select the most relevant feature to save computational time and minimize error. The features are also adjusted, and zero paddings are applied where necessary to make the model robust and efficient. In stage 4, the data is split and trained using RFECV to eliminate the repeating feature which can alter the model results. Moreover, the processed data is analyzed based on the seven most common regressor model. In the final stage 5, the power and energy output of the PV–wind system is evaluated based on MSE, MAE, R^2 and computational time. The regressors used for the study are Extra Trees regressor (meta estimator that fits several randomized decisions parameters over subsamples of the population), AdaBoost (Boost the performance of decision trees), SVR (evaluates real-value function), K-Neighbors regressors (solves classification and regression issues), Gaussian Process Regressor (provide uncertainty measurements on the predictions), MLP regressor (prevent overfitting of the model and update parameters) and Linear Regression (explain the relationship between scalar response and explanatory variables).

In RFECV method, an estimator that is a linear regression model was selected for the current study that estimates one or more target features containing continuous values. The discrimination function was given by the equation as follows:

$$y = wx + b \tag{1}$$

where w is the coefficient that the model identifies, and b signifies the bias of the model. In each iteration the model runs, the estimator uses the data with all the features to provide a set of scores at the end of each iteration. Each element in this set of scores will be associated with a feature. For example, the top 5 features of the data which could provide the best contribution to the model accuracy need to be identified and let us assume that we have 20 features in general. In this case, the algorithm starts removing the features recursively at each iteration, only if their gained score is less than expected by the algorithm. After training an algorithm the accuracy of the model is evaluated through the metrics enabling to judge the precision of the learning model. Cross-validation which is explained by Arlot and Celisse (2010) can be used. In this study k-fold cross-validation, where k is the number of folds that divide the dataset into different partitions. In each fold, the training set and the test set will be separated from each other. In each fold, the learning and evaluation process occurs, and at the end, the average accuracy, precision, recall and F1 score is reported.

The comparison of different regressors based on MSE, MAE, R2 and computational time was carried out for the yearly datasets using ANN. Based on the weather parameters and hybrid power system output a total of 77,000 samples were collected and there pattern is observed in Qadir et al. (2021a). The regressors were trained, tested and validated based on the collected data. The datasets were adjusted by removing outliers, errors and missing values. The network simplification was carried out using the validation process until no further improvement can be observed. Finally, the performance is tested of the developed model. The datasets were divided into three sets for training, validating and testing. Around 75% of the data was used for training while 15% of the dataset was used for the validating and testing each. Through network training, the values of MSE and R was evaluated for the output values based on input datasets. The average mean square difference between the outputs and target values is given by MSE, and correlation is indicated by the R. MSE with zero indicates no error and R-value close to one indicates a strong relationship and zero is a weak relationship.

An interesting relationship between temperature and humidity was observed for the analyzed dataset in Fig. 5(a). With the increase in temperature fluctuation in humidity was observed, however after 0.4 °C a linear increase in humidity was evident with consecutive rise in temperature. The relationship between temperature and relative humidity (RH) has been investigated for estimating the amount of moisture in the air. It has been widely accepted by the meteorological community that for a 1 °C decrease in temperature there is a 5% decrease in relative humidity. For RH > 50% the relationship between temperature and RH become linear (Spoel et al., 2005). The linear relationship between temperature and RH is given as below:

$$RH = b - a(t - t_d) \tag{2}$$

where b is intercept, a is a slope and t = initial temperature and td= new temperature

Fig. 5(b) shows that, although the humidity is fluctuation all along with the change in solar irradiation, it was observed that when the solar irradiation increases, the humidity gradually decreases. At the first stage of the experiment, the missing values from the dataset were removed, and then the values were scaled between 0 to 1. After completing the data cleansing phase, we define ten classifiers and evaluated them based on the Coefficient of determination (R^2), Mean Squared Error (MSE), and Mean Absolute Error (MAE). These findings are in line with the findings of Thornton et al. (2000) who estimated the relationship between solar irradiation and humidity of daily observations recorded from 24 stations located in Austria. The algorithms were used for estimating radiation and humidity that was based on temperature and precipitation variables. The MAE for combined radiation and humidity estimates were 2.52 MJ m-2 per day and 85.6 Pa, respectively

4. Results and discussion

The weather factor dataset (Solar Irradiation (Si), Wind Speed (Ws), Ambient Temperature (Ta), Humidity (H), Precipitation (R), Atmospheric Pressure (Pa) and Wind Direction (Wd)) are formulated based on varying climatic conditions and the output power is derived as shown in Eq. (3). Once the power is derived, we simply find the energy required by multiplying it by a varying value α, here the value of α is 0.9 based on the load as shown in Eq. (4). The weather factor measured reading from the calibrated sensors provides the amount of power and energy produced by the PV–wind system in hours as shown in Table 3. As the weather factors vary the values of Power and Energy varies. In this study, the calculated values of Power and Energy are considered as target values and is compared with the values simulated by the regressor models.

$$\begin{aligned} ActivePower\,[0-1] = (&+2.68Si + 0.027Ws + 0.1003Ta \\ &- 0.1766H + 0.0842R - 0.0813Pa + 0.0022Wd \\ &+ 0.5228) \end{aligned} \tag{3}$$

$$Energy = ActivePower * \alpha \tag{4}$$

After scaling the data, a further step was taken for developing the ML models. Machine learning (ML) experiments that were cover in this study can be categorized as the followings: (1) Feeding learning models without manipulating the dataset. (2) Performing RFECV before applying any learning models. Recursive

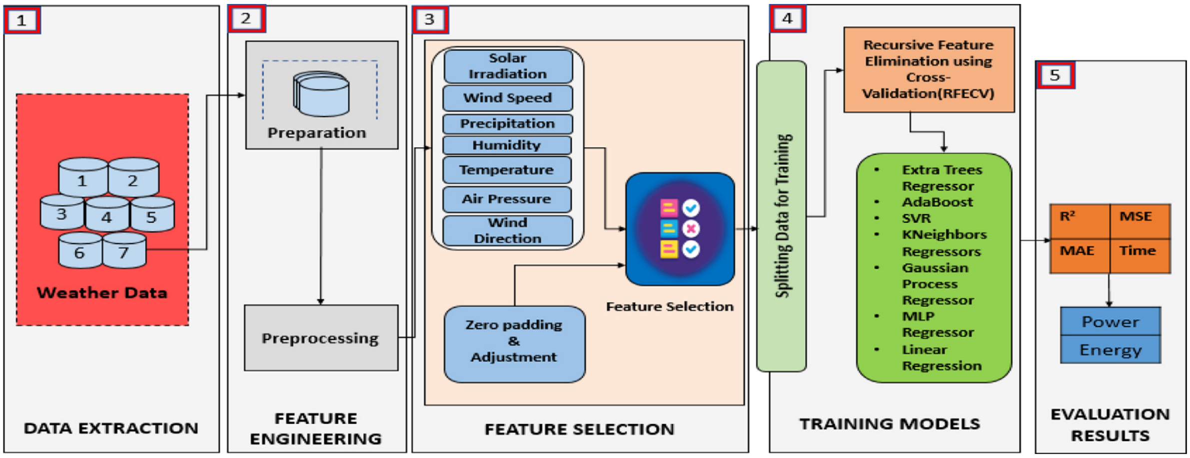
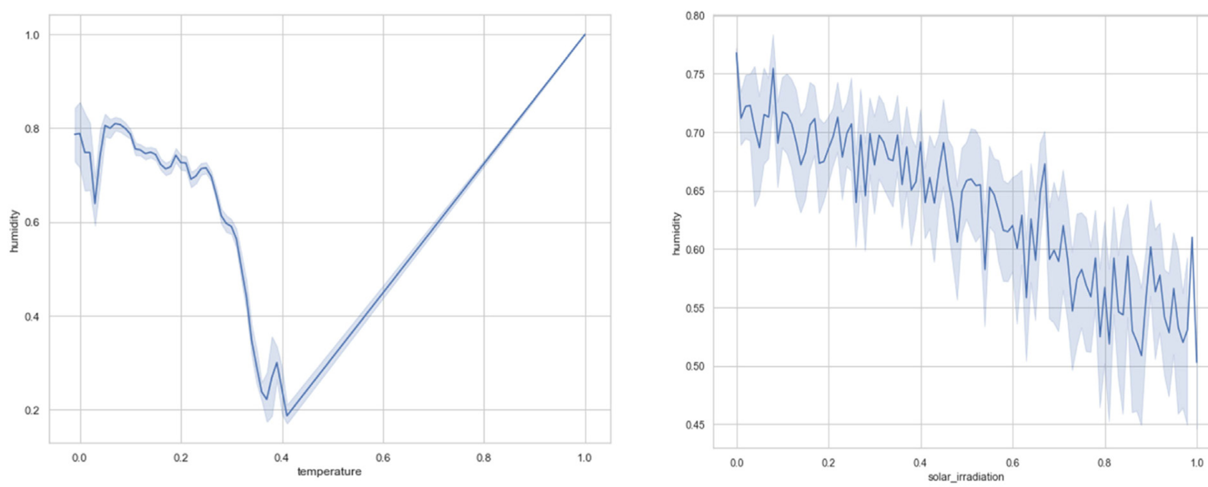


Fig. 4. Development model stages for predicting power output of PV–wind system.



(a) (b)

Fig. 5. Relationship between temperature, relative humidity and solar irradiation (a) abrupt Change in temperature based on humidity, (b) changes in the solar irradiation based on humidity.

Fig. 6. Correlation between different features in the dataset.

Table 3
Weather factor dataset and output power and energy produced by the PV–wind system.

Hourly reading	Wind speed	Temperature	Humidity	Air pressure	Wind direction	Precipitation	Solar irradiation	Power	Energy
0	0.61	0.13	0.73	1.0	0.17	0.16	0.0	0.6897	0.6208
1	0.64	0.13	0.74	1.0	0.21	0.16	0.0	0.6909	0.6218
2	0.56	0.13	0.75	1.0	0.20	0.16	0.0	0.6791	0.6112
3	0.58	0.14	0.75	1.0	0.27	0.16	0.0	0.6731	0.6058
4	0.52	0.12	0.75	1.0	0.22	0.16	0.0	0.6703	0.6033

Table 4
Statistical measure of the weather parameters and output variables.

Measurements	Wind speed	Temperature	Humidity	Air pressure	Wind direction	Precipitation	Solar irradiation	Power	Energy
count	8632	8632	8632	8632	8632	8632	8632	8632	8632
mean	0.407853	0.186580	0.709743	0.996968	0.182660	0.254109	0.207660	0.679220	0.611298
std	0.229877	0.074352	0.152935	0.005119	0.077407	0.101936	0.284296	0.067437	0.060694
min	0.040000	0.010000	0.130000	0.830000	0.000000	0.090000	0.000000	0.514000	0.462600
25%	0.220000	0.130000	0.640000	0.990000	0.120000	0.170000	0.000000	0.628200	0.565400
50%	0.360000	0.180000	0.730000	1.000000	0.190000	0.230000	0.000000	0.665700	0.599150
75%	0.560000	0.240000	0.830000	1.000000	0.240000	0.330000	0.390000	0.721225	0.649125
max	1.000000	1.000000	1.000000	1.000000	0.360000	0.560000	1.000000	0.910200	0.819200

feature elimination is a feature selection model that fits a model and eliminates the weakest feature until a specified number of features is reached. However, it is often not clear in advance how many features are valid. To find a favorable number of features the method of cross-validation is used to score numerous features and choose the best scoring collection of features. It normally plots the number of features in the model along with their cross-validated test score and diversity and envisages the selected number of features. It leads to fewer features and less MSE loss. Furthermore, the time needs to train the learning model significantly decreases. The data was visualized using a different statistical measure of mean, median, standard deviation, minimum and maximum values for weather parameters and output variables as shown in Table 4. The mean variation in the values of wind speed was found to be 0.4 m/s, the temperature was 0.18 ^{o}C, air pressure 0.99, for wind direction it was 0.18, precipitation around 0.25 and solar irradiation was at 0.2 (KW/m2). While for power the mean value was 0.67 and 0.611 for energy.

The linear regression analysis presented in Fig. 6 depicts the correlation between the input and the output variables obtained from the ANN method. The agreement between the measured and predicted values is represented by R^2 with the highest value of 1. During the training and testing of the ANN model, the R^2 value was high indicating adequate performance capacity of the ANN model. The properties of a good fit model are that the training and the testing errors should be low, while training errors are slightly lower than the testing errors (Duncan and Fiske, 2015).

Power and energy being the target values could be predicted separately. Therefore, once energy is removed and power as the target variable can be considered, and vice versa. It was observed from the linear regression analysis that weak correlations were observed for temperature, humidity, air pressure, wind direction and precipitation for PV output. While power and energy were highly correlated with the solar irradiation and wind with an R2 value of 0.79 and 0.74 as shown in Fig. 6. It indicates that these factors will play a significant role in forecasting the output variables. Any change in solar irradiation and wind speed will directly impact the PV out power. The power system controller should consider these factors when running the prediction models and manage it for maintaining grid stability, maximum unit commitment, and regulations. It was also observed that solar irradiation, humidity, temperature, and wind speed are correlated with each other. Therefore, these can be combined linearly to create a new feature. A pair-plot depicting the relationship of each feature with other feature was plotted in Fig. 7. The diagonal plot showed the histograms depicting the probability distribution of each weather factor. While the upper and lower triangles showed the scatter plots indicating the relationships between the features. From Fig. 7 it was observed that each feature can follow a particular distribution. The diameter of the plot illustrates the distribution of each feature, and each feature demonstrates the distribution/relationship with other features. This pair plot aimed for showing the changes in one feature based on all other features.

Due to the variable nature of solar energy the output power of a hybrid PV power plant is subjected to ramping (Chiteka and Enweremadu, 2016). The difference in energy generation within a time interval is represented by ramping and is expressed in a relative unit to the actual demand. The gap between the measured and forecasted value significantly affects the end-user especially the network controllers as it helps to determine an abrupt change in the power output. Therefore, for short term and long-term solar forecasting ramp event is significant for solar power management. For large ramps, it is important that the forecasting is done accurately for time and rate so that the power grids are operated safely (Zhang et al., 2017).

The datasets were subjected to different regressor models such as ExtraTrees Regressor, AdaBoost Regressor, SVR, K-Neighbours Regressor, Gaussian process Regressor, MLP Regressor and Linear Regressor. The datasets were analyzed without any feature selection and with feature selection. The computational time reduced with feature selection for all the tested regressors. Tables 5 and 6 illustrate the results of the MSE, MAE, R^2 and computational time for the tested regressors. According to the data visualization results, the attributes are linearly dependent on each other. Therefore, the linear regression could be a suitable one to feed the data to it. Results also suggest that the linear regression outperforms the rest of the models. Furthermore, concerning the obtained results, it takes less time for a linear regression model to train itself, compared to the other models.

The results show that reducing the number of features is quite helpful. Finding the most favorable features to use for Machine learning model training is often a difficult task to achieve. The feature selection model has been applied by different researchers. In Iwata et al. (2020), the author investigated high-efficiency energy detection-based spectrum measurements. The noise floor estimation was found to be elemental for energy detection-based measurement. A forward consecutive mean excision (FCME) algorithm was selected which was found to be computationally inefficient and unable to gauge the slowly varying time-dependent

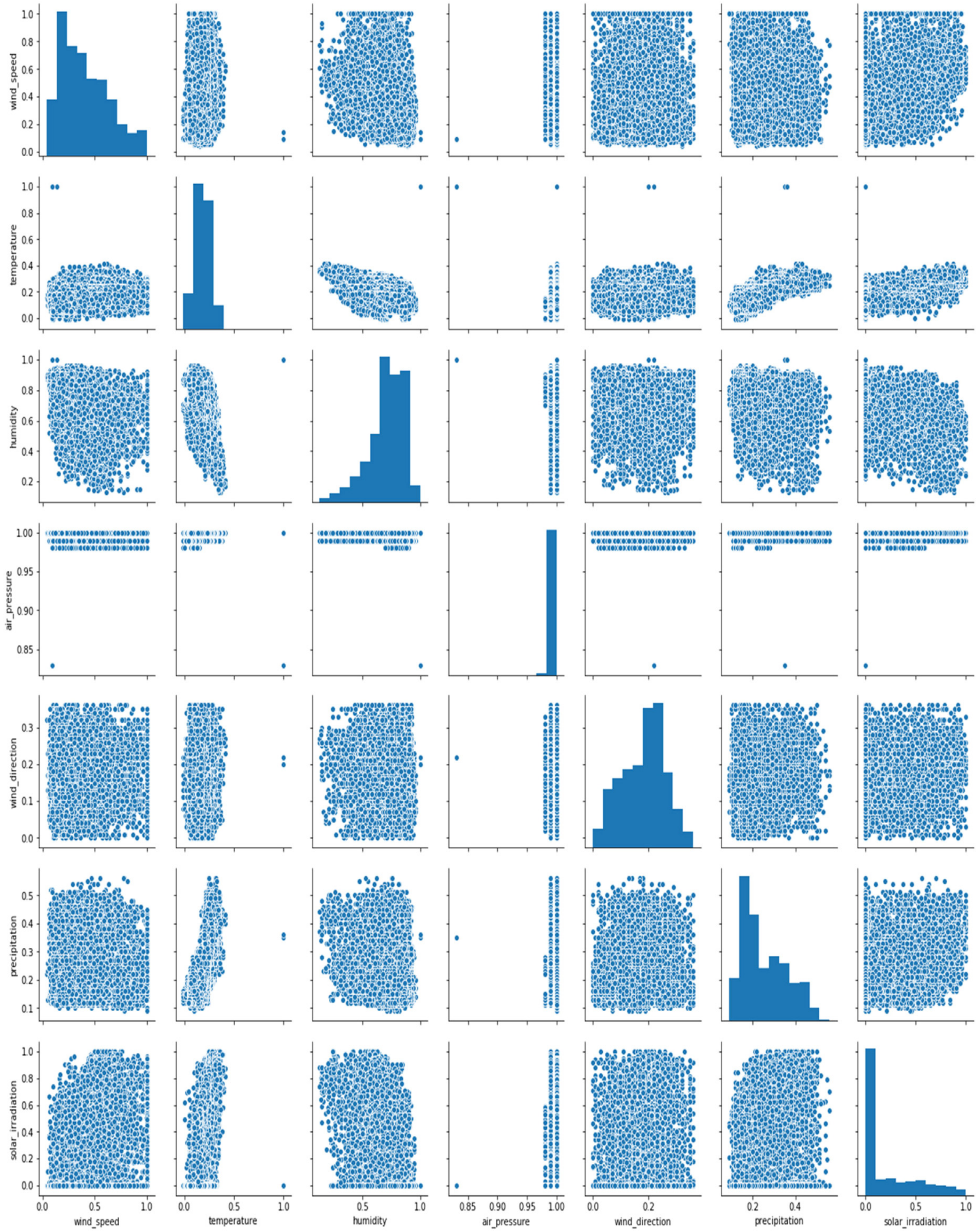


Fig. 7. Pair-plot, showing the distribution of each feature based on the other feature.

property of noise floor. Therefore, a computational complexity reduction algorithm was selected which considered NF change detection and skipped NF estimation when there was no estimation. This proposed computational complexity reduction algorithm was found to be more efficient than FCME. Similarly, studies have been carried out in improving energy detection in cognitive radio systems using machine learning techniques and has been widely used in spectrum sensing techniques. The factors such as multipath fading and shadowing are limiting factors resulting in errors. The computational efficiency was enhanced using weighted KNN which improved the overall accuracy to about 92% with feature elimination at various SNR conditions. Therefore, it can be stated that a feature selection method can improve the overall computational efficiency of the model.

Hence, apart from the feature selection techniques, due to the linearity attribute of the dataset, linear regression outperforms all the other classifiers in all evaluation metrics. It is worth noting that estimators such as Extra Trees model are much more

Table 5
Results without any feature selection technique.

Regressor	MSE	MAE	R^2	Time (s)
ExtraTrees	5.39e−05	0.0055	0.98	1.56
AdaBoost	0.0004	0.016	0.911	0.44
SVR	0.001	0.026	0.765	0.016
K-Neighbors	5.09e−05	0.0053	0.98	0.045
Gaussian Process	1.71e−06	0.00098	0.9996	11.19
MLP	2.90e−05	0.0039	0.9935	0.69
Linear	1.043e−06	0.00087	0.99	0.04

Table 6
Results with feature selection technique.

Regressor	MSE	MAE	R^2	Time (s)
ExtraTrees	4.93e−05	0.005	0.98	1.4
AdaBoost	0.0003	0.015	0.91	0.42
SVR	0.001	0.028	0.75	0.01
K-Neighbors	5.45e−05	0.005	0.98	0.04
Gaussian Process	1.3e−06	0.001	0.99	10.6
MLP	2.86e−05	0.004	0.99	0.69
Linear	1.041e−06	0.00083	0.996	0.02

advanced, compared to linear regression. However, in this case, linear regression is outperforming all of them. This shows that linear regression can potentially be an excellent estimator based on the dataset that we use.

5. Conclusion

Advancement in renewable energy technologies is coming up with several challenges and introducing new machine learning approaches is the best way to predict the accurate output generated by these technologies. The main contribution and results in this paper are stated as follows:

- The solar irradiance, wind speed, temperature and humidity have the most significant impact on the power output of the PV–wind system relative to atmospheric pressure, precipitation and wind direction for the data recorded at METU, NCC.
- According to the data visualization results, the abrupt elevation in humidity with the consecutive rise of temperature and the continuous descend of humidity with increased solar irradiance depicts the linear behavior.
- The feature selection technique applied results in increased prediction accuracy and shows a significant reduction in computational time and errors. The state-of-the-art regression models are compared and the linear regressor model outperforms all others with an MSE of 0.0000001, MAE of 0.00083, R^2 and computational time of 99.6% and 0.02 s
- The results investigated depict that the sustainable computational scheme introduced has the potential to enhance smart grids by efficiently predicting the energy produced by renewable energy systems.

The main challenge in this study was the collection of precise datasets, as inaccurate datasets may affect the prediction application by the neural networks. To overcome this issue, linear regression and zero paddings were applied to the collected datasets.

For future research, further investigation can be carried out to deploy new strategies and implement the developed model based on the predicted results. Particularly evaluating the performance parameters using deep learning techniques to predict the power generated by the PV–wind farms on a larger scale, that can be used for commercial and industrial purposes. Tapping into renewable energy sources in the way forward to reduce carbon footprint and overcome the energy crisis. Improved prediction and forecasting tools can make these PV-hybrid systems more efficient and reliable.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

All authors each made a significant contribution to the research reported and have read and approved the submitted manuscript.

References

Abujubbeh, M., Marazanye, V.T., Qadir, Z., Fahrioglu, M., Batunlu, C., 2019. Techno-economic feasibility analysis of grid-tied PV-wind hybrid system to meet a typical household demand: Case study - amman, Jordan. In: 2019 1st Global Power, Energy and Communication Conference. GPECOM, Nevsehir, Turkey,, pp. 418–423. http://dx.doi.org/10.1109/GPECOM.2019.8778539.
Ahmad, A.S., et al., 2014. A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renew. Sustain. Energy Rev. 33, 102–109.
Al-Mahasneh, A.J., Anavatti, S.G., Garratt, M.A., 2017. Altitude identification and intelligent control of a flapping wing micro aerial vehicle using modified generalized regression neural networks. In: 2017 IEEE Symposium Series on Computational Intelligence. SSCI, IEEE, pp. 1–6.
Al-Turjman, F., Qadir, Z., Abujubbeh, M., Batunlu, C., 2020. Feasibility analysis of solar photovoltaic-wind hybrid energy system for household applications. Comput. Electr. Eng. 86, 106743.
An, Byeong Wan, Shin, Jung Hwal, Kim, So-Yun, Kim, Joohee, Ji, Sangyoon, Park, Jihun, Lee, Youngjin, et al., 2017. Smart sensor systems for wearable electronic devices. Polymers 9 (8), 303.
Antonanzas, Javier, et al., 2016. Review of photovoltaic power forecasting. Sol. Energy 136, 78–111.
Araque, O., Corcuera-Platas, I., Sanchez-Rada, J.F., Iglesias, C.A., 2017. Enhancing deep learning sentiment analysis with ensemble techniques in social applications. Expert Syst. Appl. 77, 236–246.
Arlot, Sylvain, Celisse, Alain, 2010. A survey of cross-validation procedures for model selection. Stat. Surv. 4, 40–79.
Brundage, Miles, et al., 2018. The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. arXiv preprint arXiv:1802.07228.
Chiteka, K., Enweremadu, C.C., 2016. Prediction of global horizontal solar irradiance in Zimbabwe using artificial neural networks. J. Cleaner Prod. 135, 701–711.
Coyle, Shirley, Wu, Yanzhe, Lau, King-Tong, Brady, Sarah, Wallace, Gordon, Diamond, Dermot, 2007. Bio-sensing textiles-wearable chemical biosensors for health monitoring. In: 4th International Workshop on Wearable and Implantable Body Sensor Networks. BSN 2007, Springer, Berlin, Heidelberg, pp. 35–39.
Cui, H., Liu, Q., Zhang, J., Kang, B., 2019. An improved deng entropy and its application in pattern recognition. IEEE Access 7, 18284–18292.
Das, Utpal Kumar, et al., 2018. Forecasting of photovoltaic power generation and model optimization: A review. Renew. Sustain. Energy Rev. 81, 912–928.
De Martino, A., De Martino, D., 2018. An introduction to the maximum entropy approach and its application to inference problems in biology. Heliyon 4 (4), e00596.
DiCarlo, J.J., Zoccolan, D., Rust, N.C., 2012. How does the brain solve visual object recognition? Neuron 73 (3), 415–434.
Duncan, Starkey, Fiske, Donald W., 2015. Face-To-Face Interaction: Research, Methods, and Theory. Routledge.
Fasihi, M., Breyer, C., 2020. Baseload electricity and hydrogen supply based on hybrid PV-wind power plants. J. Cleaner Prod. 243, 118466.
Floreano, D., Mattiussi, C., 2008. Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies. MIT Press.
Groppi, D., Pfeifer, A., Garcia, D.A., Krajačić, G., Duić, N., 2020. A review of energy storage and demand-side management solutions in smart energy islands. Renew. Sustain. Energy Rev. 135, 110183.
Guo, A.J., Zhu, F., 2018. Spectral-spatial feature extraction and classification by ann supervised with centre loss in hyperspectral imagery. IEEE Trans. Geosci. Remote Sens. 57 (3), 1755–1767.
Haykin, Simon, Neural Network, 2004. A comprehensive foundation. Neural Netw. 2 (2004), 41.

He, H., Garcia, E.A., 2008. Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. (9), 1263–1284.
Hirose, A., Yoshida, S., 2012. Generalization characteristics of complex-valued feedforward neural networks about signal coherence. IEEE Trans. Neural Netw. Learn. Syst. 23 (4), 541–551.
Huang, T.-J., 2017. Imitating the brain with neurocomputer a new way towards artificial general intelligence. Int. J. Autom. Comput. 14 (5), 520–531.
Islam, M.M., Lee, G., Hettiwatte, S.N., Williams, K., 2017. Calculating a health index for power transformers using a subsystem-based grnn approach. IEEE Trans. Power Deliv. 33 (4), 1903–1912.
Iwata, Hiroki, et al., 2020. A study on high-efficiency energy detection-based spectrum measurements. In: 2020 IEEE Wireless Communications and Networking Conference Workshops. WCNCW, IEEE.
Izeboudjen, N., Larbes, C., Farah, A., 2014. A new classification approach for neural networks hardware: from standards chips to embedded systems-on-chip. Artif. Intell. Rev. 41 (4), 491–534.
Jaeger, H., Haas, H., 2004. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. Science 304 (5667), 78–80.
Jiang, Yan, Guoqing, Huang, 2017. Short-term wind speed prediction: Hybrid of ensemble empirical mode decomposition, feature selection and error correction. Energy Convers. Manage. 144, 340–350.
Jiang, Yan, et al., 2019. A new hybrid framework for probabilistic wind speed prediction using deep feature selection and multi-error modification. Energy Convers. Manage. 199, 111981.
Kadam, V., 2020. Thematic issue intelligent data mining for data analysis and knowledge discovery. In: Recent Advances in Computer Science and Communications(formerly: Recent Patents on Computer Science), Vol. 13. (3), pp. 433–434.
Kingston, G.B., Maier, H.R., Lambert, M.F., 2005. Calibration and validation of neural networks to ensure physically plausible hydrological modelling. J. Hydrol. 314 (1–4), 158–176.
Koesdwiady, Arief, Soua, Ridha, Fakhreddine, Karray, 2016. Improving traffic flow prediction with weather information in connected cars: A deep learning approach. IEEE Trans. Veh. Technol. 65 (12), 9508–9517.
Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., Bandettini, P.A., 2008. Matching categorical object representations in the inferior temporal cortex of man and monkey. Neuron 60 (6), 1126–1141.
Kruger, N., Janssen, P., Kalkan, S., Lappe, M., Leonardis, A., Piater, J., Rodriguez-Sanchez, A.J., Wiskott, L., 2012. Deep hierarchies in the primate visual cortex: What can we learn for computer vision? IEEE Trans. Pattern Anal. Mach. Intell. 35 (8), 1847–1871.
Kulkarni, S.G., Chaudhary, A.K., Nandi, S., Tambe, S.S., Kulkarni, B.D., 2004. Modeling and monitoring of batch processes using principal component analysis (PCA) assisted generalized regression neural networks (grnn). Biochem. Eng. J. 18 (3), 193–210.
Leng, Hua, et al., 2018. A new wind power prediction method based on ridgelet transforms, hybrid feature selection and closed-loop forecasting. Adv. Eng. Inf. 36, 20–30.
Li, Y., Hu, J., Zhang, C., Yu, D.J., Zhang, Y., 2019b. ResPRE: high-accuracy protein contact prediction by coupling precision matrix with deep residual neural networks. Bioinformatics 35 (22), 4647–4655.
Li, Xuerong, Shang, Wei, Shouyang., Wang, 2019a. Text-based crude oil price forecasting: A deep learning approach. Int. J. Forecast. 35 (4), 1548–1560.
Luo, Changqing, et al., 2018. Channel state information prediction for 5G wireless communications: A deep learning approach. IEEE Trans. Netw. Sci. Eng..
Marquez, R., Coimbra, C.F., 2011. Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments, and the new database. Sol. Energy 85 (5), 746–756.
Marquez, Ricardo, Coimbra, Carlos F.M., 2013. Intra-hour DNI forecasting based on cloud tracking image analysis. Sol. Energy 91, 327–336.
das Merces Machado, R.N., Bezerra, U.H., Pelaez, E.G., de Oliveira, R.C.L., de Lima Tostes, M.E., 2009. Use of wavelet transform and generalized regression neural network (grnn) to the characterization of short-duration voltage variation in electric power system. IEEE Latin Am. Trans. 7 (2), 217–222.
Mohammadi, Amin, Mehdi, Mehrpooya, 2018. A comprehensive review of coupling different types of electrolyzer to renewable energy sources. Energy 158, 632–655.
Mozaffari, A., Emami, M., Fathi, A., 2018. A comprehensive investigation into the performance, robustness, scalability and convergence of chaos-enhance evolutionary algorithms with boundary constraints. Artif. Intell. Rev. 1–62.
Munawar, Hafiz Suliman, 2020a. Flood disaster management: Risks, technologies, and future directions. In: Machine Vision Inspection Systems: Image Processing, Concepts, Methodologies and Applications, Vol. 1. pp. 115–146.
Munawar, Hafiz Suliman, 2020b. Image and video processing for defect detection in key infrastructure. In: Machine Vision Inspection Systems: Image Processing, Concepts, Methodologies and Applications, Vol. 1. pp. 159–177.
Munawar, Hafiz Suliman, et al., 2017. Version Management by Time Based Approach in Modern Era. Int. J. Educ. Manag. Eng..
Munawar, Hafiz Suliman, et al., 2019a. Mining multispectral aerial images for automatic detection of strategic bridge locations for disaster relief missions. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, Cham.
Munawar, Hafiz Suliman, et al., After the flood: A novel application of image processing and machine learning for post-flood disaster management. In: Proceedings of the 2nd international conference on sustainable development in civil engineering. 2019b.
Munawar, Hafiz Suliman, et al., 2020. Big data and its applications in smart real estate and the disaster management life cycle: A systematic analysis. Big Data Cogn Comput 4 (2), 4.
Munawar, Hafiz Suliman, et al., 2020a. Flood Disaster Management: Risks, Technologies, and Future Directions. Concepts Methodol. Appl..
Munawar, Hafiz Suliman, et al., 2020b. Image and Video Processing for Defect Detection in Key Infrastructure. Concepts Methodol. Appl..
Murphey, Yi Lu, et al., 2012. Intelligent hybrid vehicle power control—Part I: Machine learning of optimal vehicle power. IEEE Trans. Veh. Technol. 61 (8), 3519–3530.
Park, Jungme, et al., 2009. Intelligent vehicle power control based on machine learning of optimal control parameters and prediction of road type and traffic congestion. IEEE Trans. Veh. Technol. 58 (9), 4741–4756.
Pedro, Hugo T.C., Coimbra, Carlos F.M., 2012. Assessment of forecasting techniques for solar power production with no exogenous inputs. Sol. Energy 86 (7), 2017–2028.
Priddy, K.L., Keller, P.E., 2005. Artificial Neural Networks: an Introduction, Vol. 68. SPIE Press.
Qadir, Z., Abujubbeh, M., Mariam, A., Fahrioglu, M., Batunlu, C., 2019a. Hydropower capacity of different power sectors in Pakistan. In: 2019 1st Global Power, Energy and Communication Conference. GPECOM, Nevsehir, Turkey, pp. 408–412. http://dx.doi.org/10.1109/GPECOM.2019.8778596.
Qadir, Z., Al-Turjman, F., Khan, M.A., Nesimoglu, T., 2018b. ZIGBEE based time and energy efficient smart parking system using IOT. In: 2018 18th Mediterranean Microwave Symposium. MMS, Istanbul, pp. 295–298. http://dx.doi.org/10.1109/MMS.2018.8611810.
Qadir, Zakria, Ever, Enver, Batunlu, Canras, 2021a Use of neural network based prediction algorithm for powering up smart portable accessories. Neural Process. Lett. http://dx.doi.org/10.1007/s11063-020-10397-3, https://link.springer.com/article/10.1007/s11063-020-10397-3#article-info.
Qadir, Z., Tafadzwa, V., Rashid, H., Batunlu, C., 2018a. Smart solar micro-grid using zigbee and related security challenges. In: 2018 18th Mediterranean Microwave Symposium. MMS, Istanbul, pp. 299–302. http://dx.doi.org/10.1109/MMS.2018.8611886.
Qadir, Zakria, et al., 2018c. Design and Development of a Low Cost Device for Bone Fracture Detection Using FFT Technique on MATLAB. In: 18th Mediterranean Microwave Symposium (MMS). Istanbul, pp. 321–324. http://dx.doi.org/10.1109/MMS.2018.8611970.
Qadir, Zakria, et al., 2019b. Time and Cost Efficient Vein Pattern Recognition and Injection Point Suggestion using Machine Vision Technology. Eur. J. Eng. Appl. Sci..
Qadir, Zakria, et al., 2021b. Addressing disasters in smart cities through UAVs path planning and 5G communications: A systematic review. Comput. Commun. In press.
Rahmanifard, H., Plaksina, T., 2018. Application of artificial intelligence techniques in the petroleum industry: a review. Artif. Intell. Rev. 1–24.
Rashid, Haroon, et al., 2018. Wireless Monitoring of ECG Signal in Infants Using SWM and DWT Techniques. In: 18th Mediterranean Microwave Symposium (MMS). Istanbul, pp. 317–320. http://dx.doi.org/10.1109/MMS.2018.8611926.
Saritas, Mucahid Mustafa, Yasar, Ali, 2019. Performance analysis of ANN and Naive Bayes classification algorithm for data classification. Int. J. Intell. Syst. Appl. Eng. 7 (2), 88–91.
Sfetsos, A., Connick, A., 2000. Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques. Sol. Energy 68 (2), 169–178.
Sfetsos, A., Coonick, A.H., 2000. Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques. Sol. Energy 68 (2), 169–178.
Sharif, Arshian, et al., 2019. The dynamic relationship of renewable and non-renewable energy consumption with carbon emission: a global study with the application of heterogeneous panel estimations. Renew. Energy 133, 685–691.
Sheng, W., Shan, P., Mao, J., Zheng, Y., Chen, S., Wang, Z., 2017. An adaptive memetic algorithm with rank-based mutation for artificial neural network architecture optimization. IEEE Access 5, 18895–18908.
Son, J., Park, Y., Lee, J., Kim, H., 2018. Sensorless PV power forecasting in grid-connected buildings through deep learning. Sensors 18 (8), 2529.
Spoel, Van Der, David, et al., 2005. GROMACS: fast, flexible, and free. J. Comput. Chem. 26 (16), 1701–1718.
Taborda, M., Almeida, J., Olivier-Lima, J.A., Martins, J.F., 2015. Towards a web-based energy consumption forecasting platform. In: 2015 9th International Conference on Compatibility and Power Electronics. CPE, IEEE, pp. 577–580.

Thornton, Peter E., Hasenauer, Hubert, White, Michael A., 2000. Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: an application over complex terrain in Austria. Agric. Forest Meteorol. 104 (4), 255–271.
Wang, D., He, H., Liu, D., 2017. Intelligent optimal control with critic learning for a nonlinear overhead crane system. IEEE Trans. Ind. Inf. 14 (7), 2932–2940.
Xiang, K., Li, B.N., Zhang, L., Pang, M., Wang, M., Li, X., 2016. Regularized Taylor echo state networks for predictive control of partially observed systems. IEEE Access 4, 3300–3309.
Xing, F.Z., Cambria, E., Welsch, R.E., 2018. Natural language-based financial forecasting: a survey. Artif. Intell. Rev. 50 (1), 49–73.
Xu, X., Hu, W., Cao, D., Huang, Q., Chen, C., Chen, Z., 2020. Optimized sizing of a standalone PV-wind-hydropower station with pumped-storage installation hybrid energy system. Renew. Energy 147, 1418–1431.
Yao, S., Guo, D., Yang, G., 2012. Three-dimensional aerodynamic optimization design of high-speed train nose based on ga-grnn. Sci. China Technol. Sci. 55 (11), 3118–3130.
Zhang, Hongyi, et al., 2017. mixup: Beyond empirical risk minimization. arXiv preprint arXiv:1710.09412.
Zhang, Yongzhi, et al., 2018. Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. IEEE Trans. Veh. Technol. 67 (7), 5695–5705.
Zhang, Weiping, et al., 2019. Sizing a stand-alone solar-wind-hydrogen energy system using weather forecasting and a hybrid search optimization algorithm. Energy Convers. Manage. 180, 609–621.