An Innovative Predictor (ZME-DEI) for Generation Electrical Renewable Energy from Solar Energy

Zainab K. Al-Janabi Department of Computer Science Faculty of Science for Women University of Babylon, Babylon, Iraq Samaher Al-Janabi* Department of Computer Science Faculty of Science for Women University of Babylon, Babylon, Iraq samaher@itnet.uobabylon.edu.iq [0000-0003-2811-1493]

I. INTRODUCTION

The natural source of energy can cover 50% of the total need the world of energy if it used in the prefect way. The natural source can split into solar, wind, water, act. Machel mentions about the use of resource energy could reduce premature mortality rate, waste job days, and minimize the total costs for healthcare [Vinoth et. al.,2020]. Thus, the changing to RE, economic, reduce air pollution with harmful gases in addition to that could assist to employ a large number of workers. The energy that is out from RER is changeable (non-dispersible, interrupted, and unreliable) [Omar et. al.,2014] also the lack of sufficient spaces and extensive experience in this field, in addition to the lack of an extensive feasibility study for the implementation of this technology for electric power generation in Iraq, are the obstacles facing this technology. We focus in this work on the two renewable energy resources, (1) solar, (2) wind energy, the reason for using (1) is unlimited, free, and do

extensive feasibility study for the implementation of this technology for electric power generation in Iraq, are the obstacles facing this technology. We focus in this work on the two renewable energy resources, (1) solar, (2) wind energy, the reason for using (1) is unlimited, free, and do not cause air or water pollution the reason of (2) is potentially cheap construction and also do not cause air or water pollution [Omar et. al.,2014]. This study offers, RE filed, and the characteristics of the significant resources, design an optimizer based on Embedded Intelligence System EI-FPGA integrated circuit, to maximize the generation of electric power, and the work applies on smart microgrid model.

Embedded Intelligent System (EI) is a nascent project domain, combining machine learning algorithms such as ("machine learning and neural networks, deep learning, expert systems, fuzzy intelligence, swarm intelligence, self-organizing map and extreme learning") and smart resolution-produce abilities into movable and implanted devices or systems [Seng et. al.,2021]. IoT is the contraction of appliances, software, sensors, operators, and physical objects are embedded in the WSN, vehicles, hous devices, and other outputs that aid these objects to transmissions and data sharing. IoT is rapidly growing with the recent evolutions in wireless technology and embedded devices, with lower energy Microcontrollers that have been

Abstract— Due to the technological development that occurred in different areas of life, the world and the environment became more vulnerable to different types of pollution, such as the increase in the emission of carbon dioxide and gases that contain different toxicity rates, especially in industrial areas and dense residential areas. Therefore, Renewable energy is the best solution to this problem. In general; Renewable Energy (RE) is infinite and natural energy which can be called also environmentally friendly. clean energy, or green energy that meets a clean, pollution-free climate, and supplies people with a novel cost-efficient kind of energy. There are multi types of Renewable Energy Resources (RER), biomass energy (bioenergy) it's the transformation of biomass into the interesting format, like heat, electricity, and liquid fuels. Geothermal energy is an efficient extraction of the earth renewable energy. In hydropower, falling water transformed into energy by utilization turbines. Marine energy (ocean) extract from six roots: waves, tidal range, tidal currents, ocean thermal energy conversion, and salinity gradients, each has the various root and needs various technics for transformation. Solar energy extraction requires, utilization of sun's energy to fit hot water by solar thermal systems or electricity by solar photovoltaic (PV) and concentrating solar power (CSP) system. this paper designs an integration model called "Zero to Max Energy Predictor Model Based on Deep Embedded Intelligence Techniques (ZME_ DEI)" based on the concept of deep embedded intelligent techniques. To determine the main rules for each unit that are effective in generating the maximum electrical energy based on the nature of each dataset. ZME-DEI model contains multi-stages, First stage includes many steps; first step captures data in real-time from multi-sensors. The second step makes merging between solar plant and weather datasets. Then the third step checks missing values. While; second stage conseired of pre-processing contains four steps such as deleting duplicate, adding some features that are useful in prediction, split capture datasets after merging into multi-intervals each interval containing the reading through 15 minutes. In the third stage, the ZME- DEI model creates concerning knowledge constraints and adopts gradient boosting techniques through replace the kernel of GBM represent by DT by new kernel based on multi-objective functions. The dataset divides into two subsets using ten crossvalidation methods, training used to construct the ZME-DEI model, and a testing set for evaluating it. The final stage includes evaluation results based on three measures such as coefficient of determination (R2), Root Mean Square Error (RMSE), and Mean Error (ME), furthermore Max Energy Generation (MEG).

Keywords— Multi objective optimization, Gradian Boosting Machine, Renewable Energy, Deep Embedded Intelligent, Optimization. evaluated that are typical for IoT's remotely located in separated areas to bind and operate for the large period that need not repairing [Yasin et. al.,2021]. Many implementations that could take advantage of EI are in independent systems and edge computing. The challenges of edge computing for independent driving systems to have the ability to carry out great computational power to satisfy the energy-efficient field that is responsible for ensuring the safeness systems on independent driving want to own the ability to handle vast data from multi-sensing and safeness systems in existent time. We mention various

"Many countries in the world suffer from a clear lack of electricity production, which has led most countries to turn towards producing energy from natural sources or environmentally friendly sources that do not cause the emission of carbon dioxide gas while not causing pollution to the environment. The problem of producing electrical energy from environmentally friendly sources with high efficiency and low cost is one of the most important challenges in this field" [Al-Janabi et. Al., 2020]. Therefore; this paper will design an integration model based on the concept of deep embedded intelligent techniques. To determine the Main Rules (Constraints) for each unit that are effective in generating the maximum electrical energy based on the nature of each dataset. Added to that, Zero to Max Energy Predictor Model Based on Deep Embedded Intelligence Techniques (ZME-DEI).

The reminded of this paper is organized as follow: section number one shows the introduction, the second section presents the related work, the third section explains the main tools, the fourth section explain the main stages related to ZME-DEI, fifth section implementation of (ZME-DEI), section sixth show how ZME-DEI suitable for the increase in the production of electrical power compared to other comparable techniques? The final section presents the main conclusion of that work.

II. RELATED WORKS

Many works attempt to handle the same problem present in section number on and the authors of those works used different techniques as follow:

Soydan, O,2020[1] present method to find the perfect location for installing the solar energy plants (SEP), based on analysis eleven layers include ("sunshine duration, solar radiation, slope, aspect, road, water sources, residential areas, earthquake fault line, mine areas, power line, and transformers") through geographic information system (GIS) and analytical hierarchy process (AHP). The result finds the best location for 80% of SEP and fails to find the best location of 20% from SEP. These results evaluated based on consistency ratio (CR) come from divided consistency indicator (CI) on random indicator (RI). This work differs from the evaluation measures (EM) and technologies of our work while using the same type of renewable energy resource (RER) (i.e., solar plant). challenges that should handle to satisfy efficient EI application in hardware like, the requirements for high Computational processing (it is significant since algorithms for EI become so complicated; cost efficiency(to satisfying the need to cost-effective embedded devices against its low batteries power); and scalability to adjust various nets, dimensions and topologies(to solutions detection at EI appliances that make the algorithms and techniques able to be achieved inside architectures whose malleable and able to be scale to conform the computational demands and hardware resources whose ready for applicable)[Seng et. al.,2021].

Guozhou et. al.,2021[2] presented a method to build a dynamic system for generating renewable energy(RE) in real-time with considering the uncertain state, for this hybrid-wind-solar model(HM) was built based on optimal control and multi targets, based on analyzing historical data(HD) of daily wind and solar energy through deep reinforcement algorithm(DNN). The model could be used to reduce costs. This result is evaluated by control policy and cost(C) This work is different from our work by using uncertainty and measures while similar by using solar and wind plants and dynamic systems.

Ahmed et. al.,2020 [3] present method to build math model for maximize supply to satisfy demand, based on optimizer for deterministic and stochastic methods to simulate micro-grid models(MG). The result is evaluated by the minimum cost energy and total cost when the PV solar used, This work different from our work in using DSM and technologies used, while similar in using VP solar plant and predicate economic feasibility(EF).

A.Razmjoo et. al.,2021[4] presented model of hybrid based on gathering data from the meteorological organization(MO) and handled using HOMER software. The result is evaluated by pointers like policies and investments that indicate minimum emissions, maximum energy and good investment. This work is different by using hybrid model and evaluation measures while similar by using the same plant such as PV solar and wind.

Bahareh et. al.,2021[5] presented method to find the best ordering of three RER to be the first one to take on consider and with little constraints, based on massive analysis of five groups ("1. Economic and financial, 2. Social, cultural, & behavioral, 3. Political & regulatory, 4. Technical, and 5. Institutional") using AHP method, The result is solar PV that has first priority with little constrains and then, wind and biomass. This work similar with our work by using the same plants accept one, and different in the measures and techniques.

A. Kaabeche & Y. Bakelli,2019 [6] presented method to find the best performance of four algorithms like("ant lion optimizer algorithm(ALO), grey wolf optimizer algorithm (GWO), krill herd algorithm (KH) and JAYA algorithm") according to less unit electricity cost(UEc) corresponding to simple less power supply probability(LPSp),that applied to solve optimization problems, based on analyzing data set for daily climate, and building hybrid solar/wind system that, and using three battery technologies, for forecasting the good performance many and massive analysis have been used considering the effect of battery life, depth of discharging(DOD) and approximate battery cost on UEc. The result show that JAYA algorithm is the best. This work different from our work by using different alg. and measure while similar in using the solar PV and wind resource.

Shin'ya et. al.,2021 [7] proposed model to minimize cost generation by arrange the RER around places connecting and exchanging power through WAN to handle current changing, analyzing 7 area included 42 city based on GA algorithm, use heat storage tank(HST) as an energy storage instead of battery that is expensive, the result is minimum Cg and maximum system efficiency evaluated through capacity and economic efficiency. This work is different from our work in evaluation measures and techniques while similar in using the solar and wind plant.

Aziah et. al.,2020 [8] presented method to find the best place to build the appropriate type of RER for these places based on analyzing the Sarawak map by using clustering and selection and then image segmentation included ("color thresholding, circular hough transform, and K-means") and regional techniques. The result determine nine from 420 place that is suitable for building hybrid energy (solar PV and hydropower) these result evaluated by HOMER program and the total cost. This work different from our work in the technique used and measures while similar in using the solar PV plant.

Utkucan,2020 [9] suggested a method for predicting total renewable and hydro energy installed capacity (TR-HEIC) and electricity generation (EG) in Turkey from 2019 to 2030, based on a model for analysis data set from 2009 to 2018 through the(" fractional nonlinear grey Bernoulli") model(FANGBM). The result predicts that the TR-HEIC and EG will be minimized from 2019 to 2030 this is evaluated by the mean absolute percentage error (MAPE) and the accuracy(A). This work is similar to our work by using the accuracy measure and different in using hydropower and hydro energy and other measure and techniques.

Ning et. al.,2021 [10] presented method to minimize the TC for setting MG included S&D technique and the energy wasted, a model based on analysis unrealistic dataset through hybrid gravitational search and pattern search (GSA-PS) algorithm. The result finds that using such technologies is high performance than others and the production cost of (GSA-PS) algorithm greatly less. This work is different from our work in using HRE, unrealistic data and different techniques while similar in using microgrid.

III. MAIN TOOLS

A. Internet of Things (IoT)

In SC, IoTs provided excellent facilities for linking different smart appliances with the internet furthermore,

IoTs provided facilities in many disciplines such as academia, industry, healthcare, and business [Xiangdong et. al.,2021]. In an intelligent city climate, a smart grid is attached to the IoT's platform which stand-alone mode and pervasive everywhere, that assist to specify the delivery and distribution of energy [Megahed et. al.,2019].

The development of SC in actual-time is very complex since, the small-scale, energy consumption, and ability of appliance's sensors. Furthermore, the hardness of controlling and detecting battery lifespan of various sensors-base appliances. Therefore, appears of needing to ease the SC program that is achieved by IoT [Sodhro et. al.,2019].

IoT is a network of interconnected objects or appliances that can interact via the Internet to gather, transport, and share data. While the IoT creates a lot of data and requires a lot of capacity to distribute it for various applications like SC, cloud computing services aid by storing and centralization processing data [Ang & Seng,2021].

Suffering from the slow process of gathering and sending data for processing in a centralized location, before that the need for handle and analysis raw data to obtain pure data which useful for a specific implementation, those all are challenges need to solve. But in decentralized processing and enables IoT technologies and mobile computing, edge computing is required for processing data close to IoT appliances or objects sources rather than sending data to cloud server for analysis (performed in realtime), so this will reduce latency and bandwidth usage. Another benefit is that the energy needs for sending and receiving are also reduced between appliances and servers [Ang & Seng,2021].

B. Solar Plants

Solar energy is the most extensively utilized environmental energy source for its ready availability and steady energy scavenging capabilities throughout the day, but it has the disadvantage of being inaccessible at night or in adverse weather conditions. If energy is retrieved properly, harvested energy can even persist through the night. Solar energy is among the most well-established, well-known, and advantageous renewable energy sources [Kaur& Buttar,2019].

Recently the dependence on clean energy generation from solar and wind becoming the main goal to satisfy, because of the obvious shortage in Fossil Fuel Resources (FFR) and environmental degradation [Megahed et. al.,2019].

Fast alteration in the photovoltaic industry, technology, and institutional teams over the last decennium has changed PV's economical feasibility and significant transformation of the energy industrial competitive. Now, solar energy and photovoltaics become universal, many billion-dollar industries giving minimum power cost to millions of individuals across the world in the markets which wide and rising [Vinoth et. al.,2020]. The collection and usage of light and/or heat energy created BY the Sun also uses passive and active methods required to maximize the gain, which is considered the entirety of the solar energy harvesting concept. Mostly, active technology is divided into two classes photovoltaic and thermal technology.

Recently, photovoltaic active technology which uses semiconductor materials to convert photons into ELECTRONS is a popular choice. Solar thermal technology converts solar energy into thermal energy for home and/or commercial uses as drying, heating, pooling, cooking, and so on. On a larger scale, Concentrated Solar Thermal (CST) and Concentrated Solar power meet heating needs, while CSP technologies are used to generate energy. CSP uses big mirrors before converting it to heat energy[Kabir et. al.,2018].

C. Sensor

A Wireless Sensor Network (WSN) is a network in which nodes are distributed randomly or symmetrically, depending on the requirements, to detect data and perform various actions. WSN is one of the world's largest and most widely utilized networks. Solar energy harvesting is highly common in WSN and is also utilized for energy GATHERING [Gao & Lu,2020]. In these WSN applications, the wireless sensor node's life cycle and performance, as well as communication channels, are critical. A sensor node is made up of four major components: a sensing unit, a transceiver unit, a processing unit, and a power unit, as well as supplementary application-specific components including a mobilizer, location detecting system, and power generator[Kaur et. al., 2019].

Sensors are one of the IoTs support technology which contains various kinds of device sensing and actuators, sensors help systems to operate in a real-time environment. Sensors are used to collect a huge amount of data which is used as an input to large data applications, therefore sensors technology is considered as an essential tool for computing and analyzing data, that functions related to IoT. The analysis of huge data resulting from sensing units is performed by using Data Mining and Machine Learning for building models, specifying patterns, relation creation, and deploying the outcomes for support and make-DECISION [Bibri, 2018].

The external effects detection on electronic sensors is divided into active and passive sensing. The active sensing will alter in reaction to external effects and generate a change in current or voltage when an excitation signal is used. In passive sensing, there is no response to external effects. Passive temperature sensing is commonly utilized such as thermocouples. Another type of sensor, which is based on electronic circuit systems, is suitable, inexpensive, and simple but it lacks when used to monitor hostile conditions such as oxidizing fluids, high temperatures, or high pressure, and they need a large amount of space [Lamb et. al., 2020].

D. Embedded Intelligent (EI)

EI is a nascent project domain, combining machine learning algorithms and smart resolution-produce abilities into movable and implanted devices or systems [Seng et. al.,2021]. There exist many challenges to do to achieve efficient EI realization in a device such as a requirement for high computational processing, effective cost, and scalability to accommodate different networks sizes and topologies. The computations challenge is distinguished as the complexity of Embedded Intelligent algorithms and techniques. For example, there is a need for high computations in Deep Neural Networks (DNNs) that use multilayer networks to extract high-level features. The effective cost is required for satisfying the request of costeffective embedded devices with a limited energy supply. The scalability challenges permit the algorithms and techniques to be implemented within architectures that are adaptable and scalable to fit the target computational conditions and hardware [Seng et. al., 2021].

E. Prediction Techniques of Deep Embedded Intelligent from side Data Mining

Prediction is the primary method for energy harvesting systems. Since renewable energy is uncontrollable, so prediction estimation gives approximate efficient outcomes. Solar energy harvesting is highly common in WSN and is also utilized for energy gathering [Gao & Lu,2020].

As a result of the existing and need for analyzing huge data to extract the interest information that is used in various applications, DM has the main role in this domain. DM has more attention in knowledge extraction from the huge amount of data [Tan et. al.,2016].

1) Decision Tree (DT)

DT is a series of logical and mathematical processes based on probabilistic in determining the class label of test records. DT principal working as looking on and comparing the consistency of specific attributes and the threshold node. And it makes classification by assigning the class that occurs more frequently to the test records[Hossny et. al.,2020]. DT collects data in the form of a tree and maybe reformed as a collection of discrete rules. The building block of decision trees includes a root node and a lot of internal branching nodes. The class shown in a leaf node corresponds to an example, whereas features and branching reflect the consistency of features that leads to the classification being performed in internal nodes. The good way of building DT from the training set reflects the effectiveness of DT [Rustam et. al.,2021].

2) Random Forest (RF)

RF is used for classification and regression problems. RF is an ensemble model that utilizes bagging techniques which creates many trees and makes voting. Creating a large number of trees will fit prediction accuracy. RF can handle over-fitting. RF is described as the type of the various prediction trees[Rustam et. al.,2021]. RF is more robust in the election of the training set as compared to the decision tree classifier, RF is hard to interpret, however, its hyper-parameters can be turned with simple[Cotfas et. al.,2021].

3) XGBoost (eXtreme Gradient Boosting)

XGBoost is an advanced version of gradient boosting tree growth to concentrate on computational speed and model efficiency. To fit training data XGBoost assume an initial value always 0.5. It builds its regression tree using the individual tree, that fits residuals. A leaf node is used to begin the tree, all residuals are put on a leaf then computing similarity score, and then compute the gain to specify how to split. Based on these scores and the gain, the XGBoost chooses the largest gain assumed threshold for pruning[Hao, J., 2020]. XGBoost has much popularity as considered a tree-based model. XGBoost provides a speed boost since trains the number of poor students (DT) parallelly, different from the gradient that does this sequentially. XGBoost can handle over-fittings, which is unfeasible with Gradient Boosting also Adaboost classifiers. XGBoost can be implemented on a distributed system also can process larger datasets, as a result, it has scalability features. It helps to reduce the loss and enhance accuracy by utilizing a Log Loss function[Rustam et. al.,2021].

4) Extra Tree Classifier(ETC)

ETC is an ensemble learning model same as RF. Increases prediction accuracy by using the meta-estimator, which trains on different samples of the dataset a large number of weak classifiers (i.e. DT). The way of building a tree is different between ETC and RF. ETC creates DT using the original training sample, whereas RF uses samples from the entire dataset. At each iteration, the tree is given a random sample of attributes from the dataset on every test node. Based upon Gini Index criteria the optimal feature must be determined by DT to partition data. Several de-correlated DTs are created because of provided random sample[Rustam et. al.,2021].

5) Tree Net

Tree Net is a specific processing loop of Classification and Regression Tree (CART), including forwarding strategy for providing a series of binary trees. Do not use randomness in selecting variables and bootstrap sampling for transformation to the optimal accuracy. The number of iterations is reduced, to control over-fitting. At each iteration, the current tree is used to fit a residual of the previous one. Tree Net performance is influenced by the restricted wrong size of a tree changes since it does not handle the relationship between variables. Furthermore, with a small number of samples, a weak prediction can occur. Additionally, no need for data conversion or avoidance of outliers and interaction effects among predictors automatically managed [Al-Janabi, S.,2015].

F. Prediction Techniques of Deep Embedded Intelligent from side Neural Network

Deep neural networks (DNN), are the fundamental building blocks of deep learning. Those techniques have allowed significant growth in many fields like sound and image processing, as well as facial recognition, speech recognition, automated language processing, computer vision, text classification [Schmidhuber, J.,2015].

Main prediction techniques from side Neural network and it categorized as deep embedded intelligent prediction techniques.

1) Convolutional Neural Networks (CNN)

CNN uses multi-stage to extract many features that could automatically recognize the representation from data. Show high ability in machine vision techniques and image processing. Can exploit the local or time correlation between data. Consists of a set of convolution layers, nonlinear processing units, and subsampling layers. CNN is a feed-forward algorithm, has a hierarchical learning model, multiple tasks, and sharing weight. It's lacking interpretation and explanation. Not dealing with noise, and may lead to misclassification. Needing a huge of training data to learn. Selection of the hyper-parameter and the little change in its values can affect CNN performance. There are various CNN architectures like LeNet, AlexNet, VGGNet, GoogleNet, ResNet, ZFNet. [Khan et.al.,2020].

2) Long Short Term Memory (LSTM)

LSTM is a spatial kind of Recurrent Neural Network (RNN), results from back-propagation of errors that is infinite inside the cell the ability of LSTM to bridge the long temporal interval. It can deal with noise and distributed forms and continuous values. Can be generalized well. The time complexity for LSTM is O(1) for each weight and step, so it is considered local in both time and place. Suffer from problems that same to that face the feed-forward. Do not rely on random weight estimation but use small weight initialization [Hochreiter & Schmidhuber ,1997] [Al-Janabi et. al.,2021].

3) Multinomial Naive Bayes(MNB)

NBC is a probability classification based on the strong assumption classify data that attributes are autonomous from each other. Although its simplicity, these Bayes theorem is quick, accurate, and reliable in a variety of Natural Language Processing (NLP) classification jobs[Cotfas et. al.,2021]. Handle noise point. Can process missing values during model construction. Robust to unrelated features. Correlated features can damage the performance [Tan ,2016] [Ardianto et. al., 2020].

4) Support Vector classifier(SVC)

SVC is a linear supervised learning algorithms model, which gives data points to each object within n-dimension, the variable n denotes the attribute's number. SVC finds the best hyper-plane that separates between the points, so performs a binary classification that suffers few from overfitting. Besides, it can perform multiple classifications by combining multi-binary classification functions. Furthermore, it can perform other tasks like regression and outlier detection [Rustam et. al.,2021] [Cotfas et. al.,2021].

5) Self-Organizing Map (SOM)

SOM employs unsupervised, competitive learning to output low dimensional, discretized descriptions about given high dimensional data, and maintains similarity relations among the given data items at the same time. That low-dimensional representation is known as a feature map. SOM is an individual-layer NN including units set together with an n-dimensional grid. TABLE I. DEEP EMBEDDED INTELLIGENT Hexagonal grids, some three or higher dimensional spaces are used from Multiple applications. SOMs provide lowdimension projection images to distribution highdimensional data, under preserving similar relations between data items, two-dimensional and rectangular grids, are the most applications used [Miljkovic ,2017] [Mahdi & Al-Janabi, 2019].

ABLE I.	DEEP EMBEDDED INTELLIGENT PR	EDICTION TECHNIQUES FROM SIDE	DATA MINING

TECHNIQUES	ADVANTAGE	DISADVANTAGE
DT [Rustam et. al.,2021]	 Simple and easy to build and interpret. Multi-stage decision-making. Its performance depends on the way of building on the training set. 	 Only one feature is examined at each node. Suitable only for a limited number of features.
RF [COTFAS ET. AL.,2021] [RUSTAM ET. AL.,2021]	 Used to solve both classification and regression problems. Reduces the over-fitting problems. Robust with selecting training samples. Provide more accurate results as compared to another classifier. 	 Hard to interpret, however, its hyper- parameter is an easy turn.
XGBOOST [RUSTAM ET. AL.,2021]	 Used for classification and regression. Popular and fits the number of DT parallelly, unlike GBC which does this sequentially. Control over-fitting Process large datasets. Scalability feature. Use log loss function to enhance accuracy 	 Steps backward require to fix over-fitting. The computational time required for large datasets.
ETC [RUSTAM ET. AL.,2021].	 Higher prediction accuracy by implementing meta-estimator. Generate DT from the original training sample. Same to RF classifier in which both ensembles learning model. Different from RF in the way of building trees. Based on math Gini index criteria, it selects the best feature to split the data. 	 Several de-correlated DTs are created because of provided random sample.
TREE NET [AL-JANABI, S.,2015].	 Using forwarding strategies to create binary trees. Do not use randomness. Control the growing tree. Avoid over-fitting. No need for transformation or outlier prevention. Dealing with the interaction between predictors automatically. 	 Weak prediction can occur if the number of samples is small. Does not consider the relationship between variables.
GBM [AL-JANABI, S.,2015].	 One of the more powerful ML algorithms for prediction. Uses an additive or average model for boosting the error (loss function). Uses many weak learner models for each step, the new model tries to minimize the error of the previous model. The accuracy is the best because its uses a learning rate value (< 1) for reducing the contribution of each BRT. 	 Work in a sequential style therefore it is slow in processing and analyzing data. High computation and time complexity.

TABLE II.

DEEP EMBEDDED INTELLIGENT PREDICTION TECHNIQUES FROM SIDE NEURAL NETWORK

PREDICTION T.	ADVANTAGE	DISADVANTAGE
CNN [KHAN ET.AL.,2020]	Using multi-stage extraction features. Exploit the local or time correlation between data. Hierarchical learning model. Multiple tasks, and sharing weights.	Hard to interpret and explain. Unhanding noise, and may lead to misclassification. Needing a huge of training data to learn. Hyper-parameter choosing is more sensitive. Ineffective when used to estimate the object's location, orientation, and pose.
LSTM [AL- JANABI ET. AL.,2021]	The ability to bridge the long temporal interval. Can deal with noise, distributed forms, and continuous values. Can be generalized. The time complexity is O(1) for each step. Local in both position and time and space.	Suffer from problems that same to that face the feedforward. do not rely on random weight estimation. using small weight initialization.
MBC [Cotfa set. AL.,202 1]	Simple, quick, accurate in NLP classification tasks. Handle noise. Robust for unrelated features. Process missing value.	Correlated features can damage its performance.
SVM Rusta Met. Al.,2021	binary classification and multi-classification. finding fittest hyper-plane. regression and outlier detection task.	suffer few from over-fitting.
SOM [Мц. коvт с. D., 2017]	low dimensional, discretized descriptions about given high dimensional data, and maintains similarity relations among the given data items at the same time.	Weak prediction can occur if the number of samples is small. Does not consider the relationship between variables.

IV. BUILD MAX ENERGY PREDICTOR MODEL BASED ON DEEP EMBEDDED INTELLIGENCE TECHNIQUES (ZME-DEI)

This section presents a predictor named Zero To Max Energy Predictor Model Based on Deep Embedded Intelligence Techniques (ZME_ DEI) to predict DC-POWER generation.

ZME-DEI model contains multi-stages, stage number one captures data in real-time from multi-sensors. The second step makes merging between solar plant and weather datasets. Then the third step checks missing values.

In stage two pre-processing contains four steps such as deleting duplicate features, adding some features that are useful in prediction, split capture datasets after merging into multi-intervals each interval containing the reading through 15 minutes.

In the third stage, the ZME- DEI model creates concerning knowledge constraints and adopts gradient boosting techniques through adding multi-objective functions to develop it. The dataset divide into two subsets using ten cross-validation methods, training used to Algorithm¹: ZME DEI

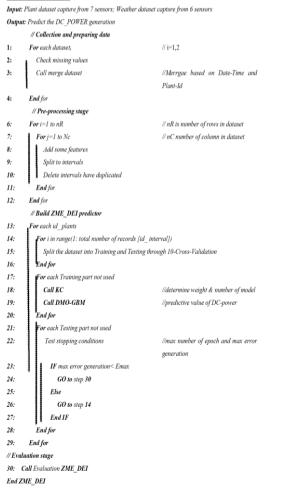
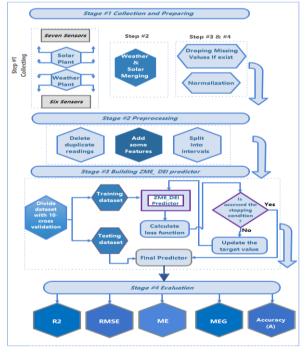


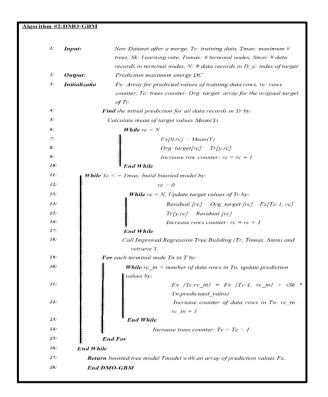
Fig. 1. Block diagram of ZME DEI construction

construct the ZME-DEI model, and a testing set for evaluating it.

The final stage includes evaluation results based on three measures such as coefficient of determination (R2), Root Mean Square Error (RMSE), and Mean Error(ME), furthermore Max Energy Generation(MEG), and Accuracy(A). In figure (2) there is illustrated the sequential stages of the ZME-DEI block diagram, and in (1) the algorithm for building it. The summarization of this study can observe below:

- Collecting the sensor's reading from solar and weather plants.
- Merging between datasets.
- Check missing values that affect prediction
- Delete duplicate features, and adding some features that are important in prediction, and split readings into intervals, all this presents in preprocessing stage.
- ZME- DEI new predictor building that related to gradient boosting techniques with multi-objective function to obtain good accuracy results.
- Evaluation results of ZME-DEI based on five measures.





Solar power generation forecasting has recently been developed using a variety of intelligent-based forecasting methodologies. However, several parameters like humidity, ambient temperature, Global Horizontal Irradiance (GHI kW/m 2), Direct Normal Irradiation (DNI kW/m 2), cloud variation, seasonal change, and so on impact the accuracy of PV output forecasts. To test the accuracy of the Renewable Energy Forecast (REF) function (F t +1) in predicting renewable energy at time t. The traditional model comprises independent factors for renewable energy forecasting ahead of time (t + 1), whereas the suggested cascaded auxiliary model identifies the model with observation variables and renewable energy forecasting ahead of time (t + 1)[Amir& Khan,2021].

$$\begin{array}{l} h(argmin_h L(y_{t+1}), argmin_g L(REF_{t+1})) & (1) \\ h(argmin_h L(y_{t+1}), R^*(F_{t+1})) & (2) \\ h(F_{t+1}), R^*(F_{t+1}) & (3) \end{array}$$

As a result, a comparison is made between the two models to anticipate overall power generation Y t + 1. The observable value confidence intervals and the predicted renewable energy prediction. The optimum function f for estimating renewable energy forecasts (REF) is:

$$f(F_{t+1}) = f^* = argmin_f L(y_{t+1})$$
(4)
$$R_{adj}^2$$
$$= 1 - \left(\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}\right) X^N - \frac{1}{N - P} - 1$$
(5)

a) Merging Two Datasets

b) Compute the Correlation

To find which parameters in the dataset affect much on the target DC_power the coefficient of correlation method is

$$RMSE = 1/N \sum_{i=1}^{N} (y - y)^2$$
(6)

The R2 value (coefficient of determination is directly reliant on the variance of the dependent variable) and the adjusted of its value are two forms of error measurement [Amir& Khan,2021].

RMSE is usually used to measure the difference between predicted values of the model and the actual values from the system that does being created. The RMSE is determined as the square root about the mean squared error. R2 shows the percentage variety of prediction values. The rate of the R2 is between 0 and 1. The measures are described as following formulas:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_{mod,i} - y_{obs,i})^2}{N}}$$
(7)

$$R^{2} = 1 - \frac{\sum_{i=1}^{N} (y_{obs,i} - y_{mod,i})^{2}}{\sum_{i=1}^{N} (y_{obs,i} - \overline{y_{obs}})^{2}}$$
(8)

where y_{obs} is the original value, and y_{mod} is the value of the model at a time/place i [Basaran et. al.,2019].

V. RESULTS

This section will present the main result extracted from each stage of (ZME_DEI) to predict DC-power which is the maximum energy generation from Renewable Resources that does not cause environmental degradation.

ZME_ DEI model shows many activities which flow sequentially in stepwise style; stage one shows real-time collecting data from two datasets weather and solar plant each having basic features then checking if the dataset contains missing values for dropping.

Stage two pre-processing contains: (a) Merging between two datasets. (b) Using correlation to the final dataset. (c) Splitting readings into intervals every fifteen minutes. (e) Delete duplication intervals. In the third stage, ZME_ DEI model is constructed based on gradient boosting techniques.

The final stage includes evaluation results based on three measures such as coefficient of determination (R2), Root Mean Square Error (RMSE), and Mean Error(ME), furthermore Max Energy Generation(MEG), and Accuracy(A).

A. Collecting datasets

This step shows reading datasets weather and solar plant each having important features, the information of weather plant containing ("Date_Time, Plant_Id, Source_Key, Ambient_Temperature, Module_Temperature, and Irradiation") which has 3182 entries, 0 to 3181. And seven sensors for solar plant such as ("Date_Time, Plant_Id, Source_Key, Dc_Power, Ac_Power, Daily_Yield, and Total_Yield") which has 71358 entries, 0 to 71357, then apply of describing function to each one and the result is:

used to find out the relationship between those parameters. The result shows that the parameters (Ac_Power, Module Temperature, and Irradiation) have high effects on our target.

B. Splitting dataset into Intervals and Deleting Duplications

According to the name of sensors(Source_Key), the data rows are split into intervals every fifteen minutes, then applying compre among intervals with deleting of duplicates, and keeping the difference only

Step1: Create a Base Model /Average Model.

$$\widehat{Y} = \frac{1}{n} \sum_{i=1}^{n} y(i)$$

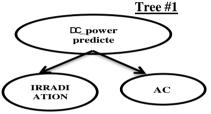
- Set 1.40 as a predicted value for all records.
- Compute the Residuals (Loss function) values

Residuals =
$$y(i) - \hat{Y}$$

Step2:

Create a second decision tree model RM1 to fit on Residuals.

Compute the mean for each node Mean of Residuals (Node1) = 2.26



• RM1 for record split by Node2 = $\hat{Y} + \alpha \times mean(Node2) = 1.9 + (0.1 \times 1.566) = 2.056$

C. Apply the DMO-GBM

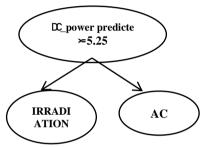
Before push DMO-GBM to predict the Dc-Power need to determined the parmaters such as Goal and features have important affect to take desions that extraction from the privous stages; as explain in the following steps and Table III.

• RM1 for record split by Node3 = $\hat{Y} + \alpha \times mean(Node3) = 1.9 + (0.1 \times 3.3) = 2.23$

Step3:

- Create a third decision tree model RM2.
- Residuals in Tree2= y(i) RM1

Tree#2



Mean= 2.7 Mean= 1.77 Mean of Residuals (Node1) = 2.33 RM2 for record split by Node2 = RM1 + $\alpha \times$ mean(Node2) RM2 for record split by Node3 = RM1 + $\alpha \times$

RM2 for record split by Node3 = RM1 +
$$\alpha \times mean(Node3)$$

 TABLE III.
 SAMPLE OF DATASET BASED ON THE FEATURES MOST IMPORTANT

Ac	MODULE_TEMPERATURE	IRRADIATION	DC-Power
3.585714286	52.12607433	0.848850921	37.14286
5.1625	53.64847207	0.900069661	53.5
5.585714286	54.3455892	0.93859401	58
5.628571429	56.65431327	0.912929421	58.42857
5.25	56.80053313	0.966185457	54.375
0	58.69564164	0.975161304	0
4.228571429	59.36714047	0.967890089	43.85714
5.671428571	58.6883928	0.966922209	58.57143
2.366666667	56.83325393	0.89780282	24.5
5.425	56.27651418	0.902461681	56.125

TABLE IV. RESULTS OF DMO-GBM

Ac	MODULE_TEMPERATURE	IRRADIATION	DC_power predicter	Target (Residuals)
3.585714286	52.12607433	0.848850921	1.40	4.1
5.1625	53.64847207	0.900069661	1.40	1.9
5.585714286	54.3455892	0.93859401	1.40	1.9
5.628571429	56.65431327	0.912929421	1.40	1.9
5.25	56.80053313	0.966185457	1.40	0.9

0			1.40	
-	58.69564164	0.975161304		1.9
4.228571429	59.36714047	0.967890089	1.40	0.1
5.671428571	58.6883928	0.966922209	1.40	1.9
2.366666667	56.83325393	0.89780282	1.40	0.9
5.425	56.27651418	0.902461681	1.40	7.1

		-	Target	
Ac	MODULE_TEMPERATURE	IRRADIATION	Residuals	RM2
3.585714286	52.12607433	0.848850921	3.77	2.4
5.1625	53.64847207	0.900069661	2.056	2.32
5.585714286	54.3455892	0.93859401	2.056	2.32
5.628571429	56.65431327	0.912929421	2.056	2.23
5.25	56.80053313	0.966185457	1.056	2.23
0	58.69564164	0.975161304	2.23	2.5
4.228571429	59.36714047	0.967890089	0.23	2.4
5.671428571	58.6883928	0.966922209	2.056	2.32
2.366666667	56.83325393	0.89780282	1.056	2.32
5.425	56.27651418	0.902461681	6.77	2.5

TABLE V.	RESULTS AFTER APPLYING RM2

D. Evaluation

This stage is used for evaluation results based on three grid measures such as coefficient of determination (R2), Root Mean Square Error (RMSE), and Mean Error (ME), furthermore Max Energy Generation (MEG) and Accuracy(A). results of all masseurs present in table 9.

VI. CONCLUSION

This paper produces a predictor named Zero to Max Energy Predictor Model based on Deep Embedded Intelligence Techniques (ZME- DEI) to predict DC-power which is the maximum energy generation from Renewable Resources that does not cause environmental degradation. ZME- DEI model shows many stages; can summarization the main benfit of this research as follow; today the friendly environment energy such as solar plant energy become the main source of energy and very important in multi countries for the following reasons: Economic and political independence: Renewable energy is good for maintaining the local economy, as reliance on imported fossil fuels leads the country to be subject to the economic and political goals of the supplying country. As for renewable energy represented in wind, sun, water, and organic materials, it exists all over the world. On the other hand, renewable energy needs more labor compared to other energy sources that rely mostly on technology, where there will be workers to install solar panels, and technicians to maintain Wind farms, and other jobs that increase employment. Low prices: Renewable energy is

witnessing a continuous decrease in costs, despite the progress made in its development, as the equipment used in it has become more efficient, and technology and engineering work has also become more advanced in this field, unlike gas, fossil fuels, and other energy sources that despite Its advantages are that prices fluctuate periodically. Improving public health: Coal and natural gas plants lead to air and water pollution, which leads to many health problems; Such as breathing disorders, nervous problems, heart attacks, cancer, premature death, and other serious problems, and it is noteworthy that the majority of these negative health effects, resulting from water and air pollution, are not caused by the use of renewable energy technologies, as wind, solar, and energy systems Hydroelectricity all generate electricity without any emissions causing air pollution, although some types of renewable energy can cause pollution; Such as geothermal energy systems and biomass, but the total polluting emissions in them are generally much less than the total emissions from power plants that use coal and natural gas. Inexhaustibility: Renewable energy is inexhaustible, compared to other energy sources; Such as coal, gas, and oil, and this means that they are always available, such as: the sun, which produces energy, and falls within the natural cycles, and this makes renewable energy an essential element in a sustainable energy system that is capable of development and development without risking, or harming future generations.

RATE OF TRAINING	RATE OF TESTING	RESULTS OF THE TRAINING DATASET		RESULTS OF THE TESTING DATASET		DATASET	
DATASET	DATASET	R2	RMSE	ACCURACY	R2	RMSE	ACCURACY
80% =19 intervals	20% =5 intervals	0.423	0.497	0.766	0.167	0.328	0. 620
60% =14 intervals	40% =10 intervals	0.562	0.526	0.853	0.433	0.334	0.741
50% =12 intervals	50% =12 intervals	0.301	0.792	0.870	0.010	0.421	0.800
40% =5 intervals	60% =19 intervals	0.296	0.787	0.891	0.020	0.203	0.825
20% =10 intervals	80% =14 intervals	0.097	0.0167	0.945	0.136	0.190	0.883

TABLE VI. RESULTS OF EVALUATION MEASUERS FOR TRAINING AND TESTING DATASETS

REFERENCES

 Al-Janabi, S., Alkaim, A.F. & Adel, Z. An Innovative synthesis of deep learning techniques (DCapsNet & DCOM) for generation electrical renewable energy from wind energy. Soft Comput 24, 10943–10962 (2020). https://doi.org/10.1007/s00500-020-04905-9

- [2] A. Razmjoo, L. Gakenia Kaigutha, M.A. Vaziri Rad, M. Marzband, A. Davarpanah, M. Denai, (2021), A Technical analysis investigating energy sustainability utilizing reliable renewable energy sources to reduce CO2 emissions in a high potential area, Renewable Energy, Volume 164,Pages 46-57, https://doi.org/10.1016/j.renene.2020.09.042.
- [3] Ahmed M.A. Haidar, Adila Fakhar, Andreas Helwig,(2020),Sustainable energy planning for cost minimization of autonomous hybrid microgrid using combined multi-objective optimization algorithm,Sustainable Cities and Society,Volume 62,102391,https://doi.org/10.1016/j.scs.2020.102391.
- [4] Al_Janabi, S. (2015). A Novel Agent-DKGBM Predictor for Business Intelligence and Analytics toward Enterprise Data Discovery. Journal of Babylon University/Pure and Applied Sciences, 23(2), 482-507.
- [5] Al-Janabi, S., & Mahdi, M. A. (2019). Evaluation prediction techniques to achievement an optimal biomedical analysis. International Journal of Grid and Utility Computing, 10(5), 512-527.
- [6] Al-Janabi, S., Alkaim, A., Al-Janabi, E., Aljeboree, A., & Mustafa, M. (2021). Intelligent forecaster of concentrations (PM2. 5, PM10, NO2, CO, O3, SO2) caused air pollution (IFCsAP). Neural Computing and Applications, 1-31.
- [7] Amir, M., & Khan, S. Z. (2021). Assessment of renewable energy: status, challenges, COVID-19 impacts, opportunities, and sustainable energy solutions in Africa. Energy and Built Environment. https://doi.org/10.1016/j.enbenv.2021.03.002
- [8] Ang, K. L.-M., & Seng, J. K. P. (2021). Embedded Intelligence: Platform Technologies, Device Analytics, and Smart City Applications. IEEE Internet of Things Journal, 8(17), 13165–13182. doi:10.1109/jiot.2021.3088217
- [9] Ardianto, R., Rivanie, T., Alkhalifi, Y., Nugraha, F. S., & Gata, W. (2020). Sentiment Analysis on E-Sports For Education Curriculum Using Naive Bayes and Support Vector Machine. Jurnal Ilmu Komputer dan Informasi, 13(2), 109-122.
- [10] Aziah Khamis, Tamer Khatib, Nur Amira Haziqah Mohd Yosliza, Aimie Nazmin Azmi, (2020), Optimal selection of renewable energy installation site in remote areas using segmentation and regional technique: A case study of Sarawak, Malaysia, Sustainable Energy Technologies and Assessments, Volume

42,100858,https://doi.org/10.1016/j.seta.2020.100858.

- [11] Bahareh Oryani, Yoonmo Koo, Shahabaldin Rezania, Afsaneh Shafiee,(2021), Barriers to renewable energy technologies penetration: Perspective in Iran, Renewable Energy,Volume 174,Pages 971-983, https://doi.org/10.1016/j.renene.2021.04.052.
- [12] Samaher Al-Janabi, Ayad Alkaim, A novel optimization algorithm (Lion-AYAD) to find optimal DNA protein synthesis, Egyptian Informatics Journal, 2022, https://doi.org/10.1016/j.eij.2022.01.004.
- [13] Bibri, S. E. (2018). The IoT for smart sustainable cities of the future: An analytical framework for sensor-based big data applications for environmental sustainability. Sustainable Cities and Society, 38, 230–253. doi:10.1016/j.scs.2017.12.034
- [14] C. Liu, C. Wang and J. Luo, (2020), "Large-Scale Deep Learning Framework on FPGA for Fingerprint-Based Indoor Localization," in IEEE Access, vol. 8, pp. 65609-65617, doi: 10.1109/ACCESS.2020.2985162.
- [15] Cotfas, L. A., Delcea, C., Roxin, I., Ioanăş, C., Gherai, D. S., & Tajariol, F. (2021). The Longest Month: Analyzing COVID-19 Vaccination Opinions Dynamics from Tweets in the Month following the First Vaccine Announcement. IEEE Access, 9, 33203-33223. DOI: 10.1109/ACCESS.2021.3059821.
- [16] Das, H. S., & Roy, P. (2019). A deep dive into deep learning techniques for solving spoken language identification problems. In Intelligent Speech Signal Processing (pp. 81-100). Academic Press.
- [17] Dhanraj, Joshuva A., Ali Mostafaeipour, Karthikeyan Velmurugan, Kuaanan Techato, Prem K. Chaurasiya, Jenoris M. Solomon, Anitha Gopalan, and Khamphe Phoungthong. (2021). "An Effective Evaluation on Fault Detection in Solar Panels" Energies 14, no. 22: 7770. https://doi.org/10.3390/en14227770
- [18] Dogan, A., & Birant, D. (2020). Machine Learning and Data Mining in Manufacturing. Expert Systems with Applications, 114060. doi:10.1016/j.eswa.2020.114060
- [19] Gao, T., & Lu, W. (2020). Machine learning toward advanced energy storage devices and systems. Iscience, 101936<u>https://doi.org/10.1016/j.isci.2020.101936</u>
- [20] Guozhou Zhang, Weihao Hu, Di Cao, Wen Liu, Rui Huang, Qi Huang, Zhe Chen, Frede Blaabjerg, (2021),Data-driven optimal energy management for a wind-solar-diesel-battery-reverse osmosis hybrid energy system using a deep reinforcement learning approach,Energy Conversion and Management,Volume

227,113608,https://doi.org/10.1016/j.enconman.2020.113608.

- [21] Guozhou Zhang, Weihao Hu, Di Cao, Wen Liu, Rui Huang, Qi Huang, Zhe Chen, Frede Blaabjerg, (2021), Data-driven optimal energy management for a wind-solar-diesel-battery-reverse osmosis hybrid energy system using a deep reinforcement learning approach, Energy Conversion and Management, Volume 227,113608, https://doi.org/10.1016/j.enconman.2020.113608.
- [22] Hao, J. (2020). Deep Reinforcement Learning for the Optimization of Building Energy Control and Management (Doctoral dissertation, University of Denver).
- [23] Al-Janabi, S. & Alkaim, A.F. A nifty collaborative analysis to predicting a novel tool (DRFLLS) for missing values estimation, Springer, Soft Comput (2020), Volume 24, Issue 1, pp 555–569. DOI 10.1007/s00500-019-03972-x
- [24] Hossny, K., Magdi, S., Soliman, A. Y., & Hossny, A. H. (2020). Detecting explosives by PGNAA using KNN Regressors and decision tree classifier: A proof of concept. Progress in Nuclear Energy, 124, 103332. doi:10.1016/j.pnucene.2020.103332
- [24] Kabir, E., Kumar, P., Kumar, S., Adelodun, A. A., & Kim, K. H. (2018). Solar energy: Potential and future prospects. Renewable and Sustainable Energy Reviews, 82, 894-900. https://doi.org/10.1016/j.rser.2017.09.094
- [25] Kaur, H., & Buttar, A. S. (2019). A review on solar energy harvesting wireless sensor network. Int. J. Comput. Sci. Eng, 7(2), 398-404. DOI: https://doi.org/10.26438/ijcse/v7i2.398404
- [26] Khan, A., Sohail, A., Zahoora, U. et al. (2020) A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53, 5455–5516. <u>https://doi.org/10.1007/s10462-020-09825-6</u>
- [27] Lamb, J. J., & Pollet, B. G. (Eds.). (2020). Micro-Optics and Energy. doi:10.1007/978-3-030-43676-6
- [28] Lei Yin, Yu Zhang, (2020) ,Village precision poverty alleviation and smart agriculture based on FPGA and machine learning, Microprocessors and Microsystems, 103469,
- [29] Mahdi, M. A., & Al-Janabi, S. (2019). A novel software to improve healthcare base on predictive analytics and mobile services for cloud data centers. In International conference on big data and networks technologies (pp. 320-339). Springer, Ch