IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 December 2023, accepted 4 February 2024, date of publication 9 February 2024, date of current version 16 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3365080

RESEARCH ARTICLE

ML-Based Energy Consumption and Distribution Framework Analysis for EVs and Charging Stations in Smart Grid Environment

FENIL RAMOLIYA[1], CHINMAY TRIVEDI[1], (Student Member, IEEE), KRISHA DARJI[1], RIYA KAKKAR[1], (Student Member, IEEE), RAJESH GUPTA[1], (Member, IEEE), SUDEEP TANWAR[1], (Senior Member, IEEE), ZDZISLAW POLKOWSKI[2], FAYEZ ALQAHTANI[3], AND AMR TOLBA[4], (Senior Member, IEEE)

[1]Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat 382481, India
[2]Department of Humanities and Social Sciences, The Karkonosze University of Applied Sciences, 58-500 Jelenia Góra, Poland
[3]Software Engineering Department, College of Computer and Information Sciences, King Saud University, Riyadh 12372, Saudi Arabia
[4]Computer Science Department, Community College, King Saud University, Riyadh 11437, Saudi Arabia

Corresponding authors: Sudeep Tanwar (sudeep.tanwar@nirmauni.ac.in) and Rajesh Gupta (rajesh.gupta@nirmauni.ac.in)

This work was supported by King Saud University, Riyadh, Saudi Arabia, under Project RSP2024R509.

ABSTRACT Electric vehicles (EVs) have become a prominent alternative to fossil fuel vehicles in the modern transportation industry due to their competitive benefits of carbon neutrality and environment friendliness. The tremendous adoption of EVs leads to a significant increase in demand for charging infrastructure. But, the scarcity of charging stations (CSs) concerns efficient and reliable EV charging. Existing studies discussed EV energy consumption prediction schemes at the CS without analyzing the affecting parameters such as energy demand, weather, day, etc. In this regard, we have proposed an energy consumption and distribution framework for EVs in a smart grid environment for efficient EV charging after analyzing the affecting parameters such as location, weekday, weekend, and user. Moreover, we have considered EV dataset to perform a detailed and deep analysis of energy consumption patterns based on the aforementioned parameters such as CS (Station ID) within the location (Location ID), weekday, weekend, and user (UserID). The main aim is to understand the smart grid-based electricity distribution to the CS by analyzing energy consumption patterns for reliable EV charging. We have done different analysis on different parameters and present their graphical representations.

INDEX TERMS Smart grid, electric vehicle, charging station, energy consumption, energy distribution, dataset analysis.

I. INTRODUCTION

In emerging nations, electricity as an energy carrier can be utilized to fulfil the people's increasing travelling demand globally. Authentic and dependable electrical power transfer is the foundation of a country's economic growth. According to the Annual Energy Outlook, the United States (US) is going to witness 31% growth in electricity demand by 2035. Towards this goal, people are incorporating electric vehicles (EVs) into their modernized transportation systems across the globe [1]. Moreover, universal CS charges vehicles from many manufacturers with various batteries and charging capacities, boosting demand for EVs and ensuring dependability for charging. This trend towards universal charging stations (CSs) is also being supported by governments and private companies investing in the infrastructure necessary to make electric vehicles a viable option for more people. As a result, we can expect to witness even more drastic growth in the EV market in the coming years [2], [3], [4].

EVs are growing in popularity to reduce carbon emissions, achieve carbon neutrality, reduce operating expenses, have minimal maintenance costs, and financial and tax advantages. Promotion and usage of new energy vehicles, such as electric cars, are becoming more popular due to pressure from carbon

The associate editor coordinating the review of this manuscript and approving it for publication was Ricardo de Castro.

emission reduction and neutrality. An equivalent gasoline or diesel vehicle has substantially higher operating costs than an EV. Moreover, EVs have low maintenance costs since they do not have as many moving parts as internal combustion vehicles [5]. Additionally, EVs produce zero emissions, making them a more environmentally friendly option than traditional vehicles [6]. They also have the potential to reduce dependence on foreign oil, increase energy security, and improve air quality in urban areas [7], [8], [9].

Despite the surge of EVs in the energy market, it becomes crucial for CS to handle the simultaneous arrival of EVs and their energy demand for charging. Therefore, a smart grid needs to be introduced to tackle the colossal energy demand of EVs at the CS, and they can provide energy to the CS in case of the simultaneous arrival of the EVs [10]. The smart grid acts as an intermediary energy supplier to deliver the energy generated at the power substation to the consumers such as homes, buildings, hospitals, etc. Further, it can be considered a sophisticated version of a traditional power infrastructure that offers a more dependable and steady electricity supply [11], [12]. The conventional power grid also referred to as the traditional power grid, is made up of several interconnected electrical power system components, including transformers, alternators, transmission lines, and various electrical loads designed to transmit electricity from a source of production to the consumers, which can be quite complex procedure while providing electricity for the consumption purpose. Thus, with the help of its innovative technologies, the smart grid optimizes energy usage by improving the overall efficiency of the power system through the usage of advanced sensors, meters, and analytic software that provides detailed information on EV energy consumption patterns [13], [14].

Thus, the smart grid can handle EVs arriving with huge energy demand at the CS. But, the energy consumption pattern of EVs varies based on various aspects such as dynamic energy demand, travelling destination, weather conditions, etc. In this regard, EV charging at the CS can be accomplished with the help of two methods, i.e., test-set-based and analysis-based. Trial-and-error, hit-and-miss, prototype, and physical equipment-based methodologies can be used in test-set-based research to predict EV energy consumption patterns using various Machine Learning (ML) and Deep Learning (DL) models. In contrast, an analysis-based method can analyze the Evs' energy consumption based on various aspects (i.e., energy demand, travelling destination, etc.). Many researchers have proposed cognizant solutions for performing the EVs energy prediction for effective and regulated charging at the CS [15], [16], [17], [18], [19]. For instance, Fukushima et al. [20] proposed a machine learning-based energy consumption prediction for EV models. In this work, the authors discussed energy consumption prediction using data-driven models that are highly accurate, and the strategy adopted in this study reduces error from the traditional method by 30%. The authors of [21] presented a new machine learning strategy to improve the range prediction accuracy and lessen the EV range anxiety. This prediction was made using a combination of short-term memory (LSTM) and deep neural networks (DNN), both capable of making long-range predictions while considering various map and traffic data. In order to anticipate the energy of EVs using artificial intelligence (AI), the authors of [22] conducted a research under secure and improved federated learning environment. The authors provided a federated learning system for CSs with security-enhanced mutual authentication. The training outcome demonstrated that their proposed model can produce a more precise energy demand forecast than the differential privacy-based model at the same runtime. Later, Shahriar et al. [23] considered a machine learning algorithm to predict the charging behavior of EVs. They have combined past charging data, weather, traffic, and event data to anticipate the length of an electric car session and energy consumption using well-known machine learning algorithms, producing results that are superior to those of conventional methods.

The above-mentioned prominent EV charging solutions focused on the prototype-based prediction methods to predict the energy consumption of the EVs, but they did not analyze the Evs' energy consumption pattern based on the various aspects such as weekday, weekend, CS location, travelling destination, energy demand, day, etc. This helps the smart grid to fulfil the energy demands of EVs through CS based on varying energy consumptions. So, we proposed a three-layered architecture highlighting the energy distribution to CS through the smart grid with the help of energy consumption analysis performed based on the various parameters. Thus, we have studied the EV energy consumption pattern analysis considering the dataset based on the various parameters such as day, weekday, weekend, and CS location to discover numerous energy consumption patterns of arriving EVs so that the smart grid can analyze and provide energy to the CSs accordingly.

A. RESEARCH CONTRIBUTIONS

Following are the research contributions of the proposed EV energy consumption and distribution analysis framework:

- We proposed an EV energy consumption and distribution framework highlighting that a smart grid can optimize its energy and transfer it to the CS through energy consumption analysis.
- The EV energy consumption analysis helps the smart grid to perform the energy distribution efficiently and reliably to the CS.
- Finally, we have performed a detailed analysis of the EV energy consumption pattern based on the various parameters of the considered dataset, such as weekday, weekend, CSs (Station ID) within the location (Location ID), and user.

B. ORGANIZATION OF THE PAPER

The rest of the paper is organized as follows. Section *II* describes the related work of the proposed analysis

TABLE 1. Comparative analysis of the EV energy consumption prediction schemes with the proposed analysis.

Author	Year	Objective	Methodology	Pros	Cons
[20]	2018	Proposed prediction of energy consumption for new electric vehicle models my machine learning	multiple regression	The rate of prediction error was decreased from the traditional way by 30%	Predict only EV's driving ranges on motorways
[21]	2022	Presented a machine learning method for EV range prediction	LSTM-DNN mixture model	Increase the accuracy of range forecast for electric vehicles (EV) and reduce range apprehension	Suggested strategy on smart grid environments or effective energy use is not covered.
[22]	2020	Presented secure-enhanced federated learning for AI-empowered electric vehicle energy prediction	Federated learning	Incorporates a premium-penalty mechanism for EVs and lightweight authentication and energy demand forecast.	Does not address how to supply effective electricity to charging stations.
[23]	2021	Proposed prediction of EV's charging behavior using machine learning	RF, SVM, Xgboost, ANN	The predictive performance for session length and energy usage, with SMAPE scores of 9.9% and 11.6%, respectively.	Does not discuss the viability of gathering data, nor does it investigate the technological constraints of putting their suggested method into practice.
[24]	2022	Proposed domestic energy Consumption forecasting using Machine Learning	LSTM	Regulate energy usage with the use of electric vehicles and renewable energy sources	Not being compared to other methods and being limited to a certain class of appliances.
[25]	2022	Proposed Machine Learning based Management of hybrid energy storage systems in EV's	Outer-loop adaptive learning and inner-loop reinforcement learning	Vanadium redox flow batteries and solar lead-acid batteries are used that has excellent power control.	Lacks empirical evidence to back up the methods it suggests.
[26]	2019	Smart community grid through blockchain and smart charging infrastructure of EVs	Blockchain and smart agent system for energy grid	Better manage demand while reducing grid load with Blockchain.	Lacks a detailed examination of the effects of EVs on microgrids.
[27]	2023	Proposed electric vehicle disaggregation for residential customer energy efficiency incentives	NILM algortithm	Appropriate for circumstances with little to no sub-metered data.	An actual dataset is sparse and lacks comparison to generally accepted standards, and lacks thorough comparison.
[28]	2023	Presented prediction of EV's annual accessibility to chargers for providing ancillary	Random Forest	The accessibility of EV chargers are predicted	Understanding of ramifications for future research are all excluded.
[29]	2023	Proposed an insight of demand forecasting in smart grids	SD, PLS, ARIMA, ARCH, AR, ARMA, MAM, LR, SS, and ANN-based models	The importance of demand forecasting and other elements in the context of smart grids.	DL algorithms for demand forecasting have certain possible drawbacks or difficulties that are not addressed.
[30]	2021	Presented machine learning based vehicle to grid strategy for improving the energy performance of public building	V2G model, Fuzzy logic classifier method	The study demonstrates energy consumption savings of up to 35% and up to 65% if used constantly.	The necessity for V2G standardization and the difficulties with deployment in terms of regulations and policy.
The proposed framework	2023	Presented an energy consumption and distribution analysis framework for EVs in the smart grid environment	Dataset Analysis considering parameters	Efficient energy distribution to the CS	

framework. Section *III* discusses the proposed framework. Section *IV* elaborates on the dataset description in detail and the association with the attributes. Section *V* presents the detailed dataset analysis based on the CS (Station ID) within the location (Location ID), weekday, weekend, and user (UserID), and finally, section *VI* exhibits the concluding remarks.

II. RELATED WORK

Many researchers have discussed the prominent solutions for EV energy consumption prediction so that EVs can be charged at a CS reliably [13], [31], [32], [33]. The relevance of energy consumption forecasting in demand management for a dependable grid is highlighted by Talwariya et al. [24]. They have estimated demand using a long short-term memory approach and assess predicting errors. The CS energy distribution can be even more effective by integrating the above-mentioned forecasting model with the game theory for energy bidding. Then, the authors of [25] highlighted the management of hybrid energy storage systems for EVs. They have focused on integrating renewable energy sources and improving interactions between vehicles and the grid.

The utilized machine learning model and TD optimal policy algorithm facilitate the best vehicle movement triggering and battery level monitoring. Now, the aforementioned authors did not consider the security aspect while monitoring the state-of-charge (SoC) for EVs charging at the CS.

Towards this goal, Lazaroiu [26] proposed a novel strategy for the smart community grid by combining blockchain technology with intelligent EV charging infrastructure. They have discussed the two-way energy flow and EV owners' involvement in energy trading. The burden on the grid can be decreased using smart charging systems, and the trustworthiness and traceability of the energy transactions are ensured through blockchain during EV charging. Furthermore, the authors in [27] proposed an effective algorithm for NILM (Non-Intrusive Load Monitoring) of EV energy consumption, allowing for precise disaggregation and analysis of power demand. Incorporating NILM approaches in the CS energy distribution framework offer valuable insights into demand patterns and aid in optimizing energy allocation. However, the authors lack the energy consumption-based analysis that can be performed on the real-time dataset. Later, Jahromi [28] focused on implementing the random forest algorithm to predict the EV annual accessibility to the chargers. They have simulated the proposed model so that drivers and EV aggregators, so that charging accessibility can improve the ancillary operations.

Scott et al. [30] investigated a machine learning-based Vehicle-to-Grid (V2G) approach to enhance public buildings' energy efficiency. The V2G strategy implementation within the CS energy distribution framework can promote more environment-friendly energy management by reducing the need for polluting peak power plants. The aforementioned EV energy-efficient solutions for charging at the CS mainly considered the EV energy consumption prediction using machine learning techniques. But, as per the literature, there is no discussion on analyzing the energy consumption of the EVs considering various effecting aspects such as day, energy demand, available energy, destination, etc. So, we have proposed an EV energy consumption framework in the smart grid environment. Further, we have considered the EV dataset to perform the energy consumption analysis based on the various attributes such as CS location, day, weekday, weekend, and users. Table 1 compares the EV energy consumption prediction schemes with the proposed energy consumption analysis framework.

A. STATISTICAL-ML FRAMEWORK

Statistics is the scientific discipline that enables us to gather, examine, interpret, display, and arrange data. Statistical inference stands as a fundamental pillar underlying various technological advancements, especially in the field of ML. Data serves as the basis for the multitude of captivating emerging technologies around us. Utilizing statistical methods allows us to uncover meaningful patterns, relationships, and insights from complex datasets, enhancing the efficiency of any ML tasks [34]. Statistical ML case study enhances the computational capability with statistical inferences and modeling mechanism for any energy consumption and distribution analysis. Our proposed approach focuses on integrating Statistical Learning Theory (SLT) [35], which serves as a fundamental framework for ML and draws inspiration from the fields of statistics and functional analysis. In our proposed energy consumption and distribution framework analysis for EVs and CSs in a Smart Grid Environment, we adopt a statistical approach to transform a conventional smart grid into an intelligent one. Unlike conventional machine learning techniques, our methodology is rooted in SLT [36], [37]. Let X be the vector space of all possible inputs and Y be the vector space of all possible outputs.

SLT [38] views the problem in the context of an unknown probability distribution over the product space $Z = X \times Y$, denoted as $p(z) = p(x_i, y)$. The training set $\mathcal{D}$ consists of n samples from this distribution, where each sample $z_i = (x_i, y_i)$. The goal is to find a function $f : X \rightarrow Y$ such that $f(x) \approx y$. Let $\mathcal{S}$ be the hypothesis space, representing the space of functions the algorithm explores. The loss function $V(f(x), y)$ measures the difference between the predicted value $f(x)$ and the actual value y. The expected risk, denoted as $R(f)$, is defined as the integral over $X \times Y$ of the loss function weighted by the probability distribution:

$$R(f) = \int_{X \times Y} V(f(x_i), y)\, p(x_i, y)\, dx \;\; dy \tag{1}$$

As the true probability distribution $p(x, y)$ is unknown, we rely on the training set for a proxy measure. The empirical risk, denoted as $\hat{R}(f)$, is the average loss over the training samples:

$$\hat{R}(f) = \frac{1}{n} \sum_{i=1}^{n} V(f(x_i), y_i) \tag{2}$$

The empirical risk minimization process, where the learning algorithm selects f_s to minimize $\hat{R}(f)$, underlines our commitment to optimizing energy distribution. This statistical analysis, grounded in SLT, forms the basis for our case study involving Charging Stations, Electric Vehicles, and EV owners. It ensures the smart grid efficiently allocates energy resources to CS based on diverse consumption patterns, facilitating effective energy utilization by EVs.

$$f_s = \arg\min_{f \in \mathcal{S}} \hat{R}(f) \tag{3}$$

III. THE PROPOSED FRAMEWORK

FIGURE 1 depicts the proposed framework encompasses of a 3-layered architecture classified into Smart Grid Layer, Energy Consumption Layer, and Distribution Layer. Thus, smart grid can provide energy to the CS based on the analysis performed considering the EV dataset so that EVs can be charged efficiently.

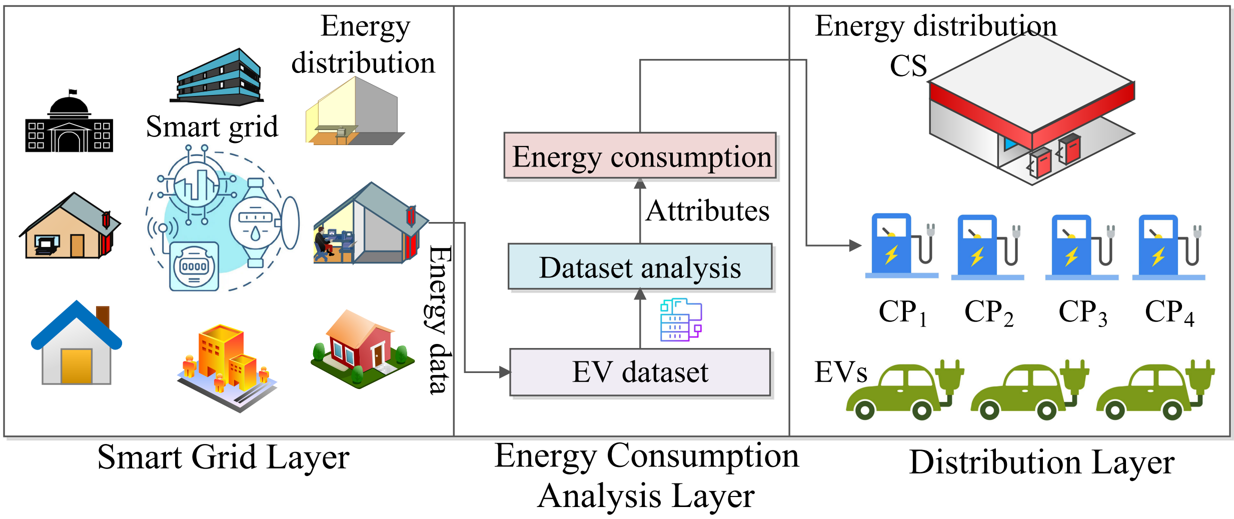


FIGURE 1. The proposed framework.

A. SMART GRID LAYER

The smart grid layer represents energy transmission and distribution, which is the beginning point of the smart grid's energy distribution process. Power-producing facilities create energy transferred through the smart grid as part of the transmission system to various commercial and residential locations such as offices, homes, hospitals, malls, etc. The smart grid as a transmission system is responsible for supplying electricity to numerous areas with the help of various renewable sources such as nuclear power plants, thermal plants, and solar energy, and it can fluctuate based on various circumstances, including environmental conditions, industrial activity, population density, and the number of people using EVs. To prevent power outages and system failures, the transmission system must also guarantee that the energy is supplied at a constant voltage and frequency. Therefore, it is crucial for energy providers to constantly monitor and adjust their systems to meet the changing demands of their customers and have a flexible and adaptable infrastructure that can respond to changing demands in the real-time. Thus, energy distribution through the smart grid should be analysed considering various parameters so that regulated energy can be provided to the CS to avail efficient charging for EVs. The smart grid, enriched with advanced features, employs a statistical-ML methodology to revolutionize energy management. It incorporates predictive energy consumption modeling, enabling precise forecasts for EVs and CSs. Dynamic load balancing optimizes energy distribution, minimizing wastage or shortages. Anomaly detection and adaptive management ensure resilience by responding to unexpected energy consumption patterns. User behavior analysis tailors strategies based on individual preferences. Grid optimization through data integration harnesses diverse data for efficient operation. Real-time decision support, informed by ML, aids grid operators in suggesting optimal energy distribution strategies, considering both current conditions and historical data patterns. This comprehensive approach transforms the smart grid into an adaptive, data-driven ecosystem for enhanced energy efficiency.

B. ENERGY CONSUMPTION ANALYSIS LAYER

The energy consumption analysis layer is considered an intermediary between the smart grid and the distribution layer. The energy distribution and transmission performed through the smart grid is analyzed with the help of the EV dataset, which yields energy consumption patterns considering various dataset attributes. We have closely studied the dataset and its effecting parameters to determine the association between attributes and analyze the Evs' energy consumption pattern information that can be further transferred to the distribution layer. We have considered various dataset features, such as day, weekday, weekend, location, and user. The analysis also considers the energy usage of various CSs with different locations. The main objective is to ensure that energy is distributed effectively and efficiently per the energy requirement of EVs. This data-driven approach helps optimize resource allocation and ensure EVs can be charged efficiently and reliably, even during peak demand periods. Then, the analyzed energy consumption information based on the EVs dataset is transferred to the distribution layer so that the smart grid can provide energy to the CS reliably for EV charging. Moreover, it can prevent energy wastage so that another CS can utilize the energy from the smart grid due to the energy consumption analysis based on the various parameters.

C. DISTRIBUTION LAYER

The distribution of energy to the CS is discussed in the final distribution layer. The EV energy consumption information analyzed in the aforementioned layer is transferred to the CS so that adequate energy distribution can be accomplished without overloading the system and resulting in power shortages. The distribution of energy must also account energy demand of various CSs, which can change based on the day, weekday, weekend, and location. Thus, energy consumption analysis helps to allocate the energy efficiently to the CS based on their requirement, and it is also advantageous for CS at another location. Thus, EVs arriving at the CS can charge their vehicle reliably without overburdening the charging infrastructure, and the smart grid can optimize the energy distribution for another CS. Thus, the 3-layered architecture of the proposed framework ensures the efficient and reliable energy distribution of the smart grid to the CS by analyzing Ev's energy consumption pattern so that EVs can fulfil their energy demand without delay. So in this manner, tradition to smart grid transformation facilitates adaptive measures such as the introduction of new charging stations, ensuring optimal energy utilization. Notably, it addresses potential issues of power waste or shortages at specific charging stations or entire locations by leveraging the power of statistical-ML strategies.

IV. DATASET DESCRIPTION

We have considered and analyzed the EV dataset, which has 24 attributes and 3395 high resolution EV charging sessions, to analyze energy consumption. The data [39] comprises session records from 85 EV drivers who frequented 105 charging stations distributed across 25 location sites participating in a workplace charging program. These workplace locations includes diverse facilities, including research and innovation centers, manufacturing units, testing facilities, and the office headquarters of a firm engaged in the U.S. Department of Energy (DOE) workplace charging challenge. Table 2 shows the name, count, and data type of the attributes required for the dataset analysis. In order to get valuable insights from the dataset, we have mainly focused on determining the association or relation between attributes, which helps in analyzing the dataset for EV energy consumption acquired at the CS. Data about CSs, locations, users, energy consumption, and cost are considered in the data collection.

Before analyzing the dataset considering the energy consumption of the EVs, we performed the data filtration to handle the empty cells in the dataset and modified the datatype to suit our analytical goals. The above-mentioned data filtration is essential for ensuring the precision and thoroughness of the considered dataset analysis. Moreover, we have removed potential biases or inaccuracies from the dataset by tackling the missing numbers and changing the data type accordingly. As a result, we have focused on deriving more insightful conclusions and making data-driven decisions about the dataset considering the association between attributes of the dataset. Thus, we can deeply understand the patterns and trends possible between dataset attributes in the aforementioned data collection step. In this regard, Table 3 shows the updated values of the count and datatype after filtering the dataset.

TABLE 2. Description of dataset.

index	Column	Non-Null Count	Dtype
0	SessionId	3395 non-null	int64
1	kWhTotal	3395 non-null	float64
2	dollars	3395 non-null	float64
3	created	3395 non-null	object
4	ended	3395 non-null	object
5	startTime	3395 non-null	int64
6	endTime	3395 non-null	int64
7	chargeTimeHrs	3395 non-null	float64
8	weekday	3395 non-null	object
9	platform	3395 non-null	object
10	distance	2330 non-null	float64
11	UserId	3395 non-null	int64
12	StationId	3395 non-null	int64
13	LocationId	3395 non-null	int64
14	managerVehicle	3395 non-null	int64
15	facilityType	3395 non-null	int64
16	Mon	3395 non-null	int64
17	Tues	3395 non-null	int64
18	Wed	3395 non-null	int64
19	Thurs	3395 non-null	int64
20	Fri	3395 non-null	int64
21	Sat	3395 non-null	int64
22	Sun	3395 non-null	int64
23	reportedZip	3395 non-null	int64

TABLE 3. Updated count and datatypes.

index	Column	Non-Null Count	Dtype
8	distance	3395 non-null	float64
21	reportedZip	3395 non-null	int64
22	weekday_n	3395 non-null	int32
23	platform_n	3395 non-null	int32
24	created_n	3395 non-null	int32
25	ended_n	3395 non-null	int32

TABLE 4. List of dropped attributes.

index	Column	Non-Null Count	Dtype
3	created	3395 non-null	object
4	ended	3395 non-null	object
8	weekday	3395 non-null	object
9	platform	3395 non-null	object

We have also dropped some of the attributes based on their irrelevancy or changed datatypes to perform the analysis of the dataset. TABLE 4 shows the dropped attributes (i.e., created, ended, weekday, and platform) considered for the dataset.

Further, we have represented a succinct description of all the attributes of the dataset, which contain numerical values, categorical data, or textual information, as well as any other pertinent attributes that describe the data points and can be used to model the dataset for EVs energy consumption analysis. To fully comprehend the association, pattern, and trend between the data, it is imperative to get insights into the properties of the attributes, which offer a comprehensive perspective of the data in the dataset. EV energy consumption analysis is performed considering all the attributes of the dataset, which improves the overall interpretation and association of the data. Thus, we have explained the attributes of the dataset to understand and analyze it elaborately.

- *SessionID*: It is defined as a unique identification number corresponding to the charging sessions. It can be used to differentiate between several charging sessions that occurred at the same station or by the same user. The dataset contains 105 distinct session IDs.
- *kWhTotal*: The attribute represents the overall EVs energy consumption (in kWh) during a particular charging session. It represents the power the user's vehicle consumes during the charging session. This is a crucial indicator for assessing the user's power costs and the effectiveness of CS energy efficiency.
- *Dollars*: The attribute displays the charging session's overall cost (in dollars). The cost per kilowatt-hour of power and the total energy utilized are considered to compute it. By examining this section, we may learn more about the pricing policies of various charging providers and the overall expense of EV ownership for customers.
- *Created*: This attribute represents the date and time the charging session was created. It can be used to track when the user or EV driver initiated the charging session for charging at the CS.
- *Ended*: This attribute represents the date and time the charging session ended. It can be used to track how long the user's vehicle was charging at a particular CS.
- *StartTime*: This attribute represents the date and time the user's vehicle started charging. It is equivalent to the created attribute.
- *EndTime*: This attribute represents the date and time when the user's vehicle stopped charging. It is equivalent to the ended attribute.
- *ChargeTimeHrs*: Attribute ChargeTimeHrs in the dataset indicates the time it takes for a charging session to complete, measured in hours. It is determined by calculating the difference between the start and end time of the session. The duration of a charging session is influenced by a number of factors, such as the cost of electricity and any additional fees or taxes charged by the CS or platform. This information is useful for analyzing the pricing structures of different charging providers and assessing the overall cost of owning an EV.
- *Weekday*: It represents the day of the week when the charging session occurred (e.g., Monday, Tuesday, etc.). It can be used to identify patterns in charging behaviour across different days of the week.
- *Platform*: It indicates the charging platform or service provider that can be used for the number of charging sessions at the different CS. Further, it can be utilized to track usage across other platforms or to compare the performance of the different providers.
- *Distance*: The distance attribute represents the distance (in miles) that the user's vehicle travelled during the charging session. It can be used to track how far users are driving on a single charge.
- *UserId*: This attribute is a unique identifier for the user who initiated the charging session. It can be used to track usage patterns for individual users.
- *StationId*: This attribute represents a unique identifier for the CS where the charging session occurred. It can be used to track energy usage patterns for the CSs.
- *LocationId*: The attribute provides the location's distinctive identity corresponding to the CSs involved in the various charging sessions. It is further utilized to monitor EV energy consumption trends across several locations or regions. The dataset contains 24 distinct station IDs.
- *ManagerVehicle*: This attribute represents the type of vehicle charging during the session. It can be used to track EVs energy consumption patterns for different types of vehicles.
- *FacilityType*: It defines the type of facility which EVs can avail at the CS (e.g., parking garage, retail store, etc.). It can be used to track energy usage patterns across different types of locations.
- *Day*: The day attribute, i.e., Mon, Tues, Wed, Thurs, Fri, Sat, and Sun, represents binary indicators (0 or 1) for each day of the week. They can be used to identify patterns in charging behaviour across different days of the week.
- *ReportedZip*: This attribute represents the zip code of the location where the charging station is located. It can be used to track usage patterns across different geographic areas.

Researchers can get insights into the EV energy consumption pattern at the CS with the help of detailed dataset analysis, which can also help them to design an infrastructure to fulfil Evs' energy demand. The *dollar* and *kWhTotal* attributes are closely related since the cost of the charging session depends on how much energy the EV consumes for charging. Then, the attributes *created*, *terminated*, *startTime*, and *endTime* offer timestamps for various charging session occurrences. The *startTime* and *endTime* are used to create the *chargeTimeHrs* attribute, which shows how long the charging session lasted. The day on which the user utilizes the CS for energy fulfilment is shown by the *weekday*. The *reportedZip* offers further details on the location of the CS, and attribute, i.e., *Mon-Sun* highlights the binary representation of the day of the week on which the session occurred. Further, other attributes such as *distance*, *user ID*, *Station ID*, and *Location ID* pertain to the CS and the user who started the charging session. In contrast, the *platform* indicates the charging platform that was utilized for the session. The *managerVehicle* and *facilityType* includes details associated with the type of EV arriving for the charging and the type of facility drivers can avail from it.

Now, after getting insights into the attributes and their association with each other by considering the EVs dataset for analyzing the energy consumption based on the various

TABLE 5. Labeling of weekdays along with how many times they occurred.

label	Weekday	Occurrence
4:	Thu	735
6:	Wed	713
5:	Tue	635
1:	Mon	616
0:	Fri	610
2:	Sat	62
3:	Sun	24

TABLE 6. Average consumption for Station ID 369001.

Weekday	Average Consumption
0	5.615789
1	5.613103
2	5.012581
3	5.702000
4	6.016897
5	5.422623
6	5.655156

TABLE 7. Average consumption for Location ID 493904.

Weekday	Average Consumption
0	5.098132
1	5.499186
2	4.999318
3	4.269231
4	5.718817
5	5.325521
6	5.449802

TABLE 8. Total average consumption based on weekday.

Weekday	Total Average Consumption
0	489.073529985215
1	452.992965954921
2	49.9477729528536
3	45.7021785714285
4	521.661990035522
5	534.031772298232
6	530.387259985093

attributes of the dataset. Further, we need to discuss other aspects of the dataset's attributes, such as weekday count and average energy consumption considering Station ID and Location ID. Table 5 shows the weekday occurrence indicated by its label in the dataset, and it describes how frequently each day of the week appears in the dataset for the charging sessions. The term label'' refers to the name of the day of the week for each record in a dataset that includes day information, i.e., Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, or Sunday. The rationale behind this unordered labeling stems from the nature of energy consumption patterns observed in various real-world scenarios. In many contexts, energy usage exhibits distinct patterns during weekdays, with variations on Fridays and Mondays often being more prominent. This choice has enabled us to highlight nuances and trends that might have been obscured with a more conventional labeling approach.

Next, Table 6 displays the average energy consumption per CS for a week, represented in kWh. In the following instances, 0 denotes Friday, 1 signifies Monday, 2 signifies Saturday, 3 signifies Sunday, 4 signifies Thursday, 5 signifies Tuesday, and 6 signifies Wednesday. Further, we have used Station ID 369001 as an example to highlight the average EVs energy consumption in Table 6. On average, the considered CS corresponding to the particular Station ID utilized 5.6157 kWh of energy on Friday, 5.6131 kWh on Monday, etc. Overall, the above-mentioned information indicates that the CS uses energy differently throughout the week, with higher usage during the weekdays and reduced usage during the weekends.

Then, we analyzed the average energy consumption corresponding to the Location ID based on the weekday. Table 7 shows the average EVs energy usage determined for Location ID 493904 considering the weekdays, i.e., 5.098132 kWh on Friday, 5.499186 kWh on Monday, 4.999318 kWh on Saturday, 4.269231 kWh on Sunday, 5.718817 kWh on Thursday, 5.325521 kWh on Tuesday, and 5.449802 kWh on Wednesday. According to the statistics, for this specific location identified using Location ID, power usage is greater on Thursday and Monday and lower on Sunday and Saturday, as shown in Table 8.

Table 9 represents the larger portion of the aforementioned table (which shows average energy consumption for Station ID and Location ID) that details the number of the distinct station IDs present at each location and the requirement of the average energy expressed in kWh by each Station ID based on the day of the week. Moreover, Table 9 also shows the number of stations along with their Station ID within the particular location, which is represented by the Location ID. The information about the number of CSs within a specific location. Now, after discussing the dataset's attributes and their association. Next, we can discuss the analysis of the considered dataset considering the attributes based on the Location ID, Station ID, and user ID.

V. DATASET ANALYSIS

The detailed analysis of the considered EV represents a significant contribution to the field, as it unveils intricate patterns in energy consumption. It covers various attributes, such as CS identified by Station ID, Location ID, weekday and weekend designations, and User ID. This comprehensive approach allows us to discern nuanced energy consumption patterns, providing valuable insights for the smart grid's efficient energy distribution to/from CS. Our findings contribute to the field in three key dimensions: Location ID, Station ID, and User ID. By categorizing the energy consumption analysis into these sections, we present a multifaceted exploration that goes beyond traditional analyses. This categorization enables a more targeted and strategic approach to energy distribution, fostering a deeper understanding of how different factors influence and interact within the broader

TABLE 9. Location - station relation.

index	LocationId	Station No.	StationId	Weekday	Average Energy Consumption in kWh
0	125372	2	445920	0	6.850000
1	125372	2	445920	1	6.622727
2	125372	2	445920	4	6.695000
3	125372	2	445920	5	6.643333
4	125372	2	445920	6	6.822857
5	125372	2	871619	1	5.606667
6	125372	2	871619	5	6.066667
7	125372	2	871619	6	6.518750
8	144857	6	517464	0	5.953333
9	144857	6	517464	1	5.740000
10	144857	6	517464	4	4.166000
11	144857	6	517464	5	5.922500
12	144857	6	517464	6	5.696364

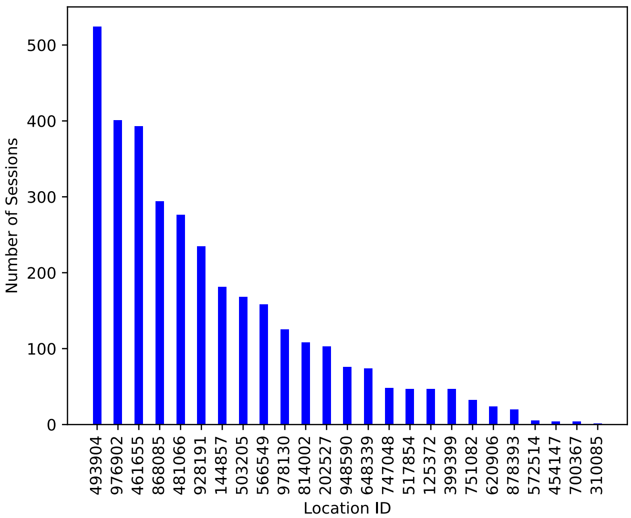


FIGURE 2. Number of sessions by Location ID.

smart grid framework. Thus, our proposed work not only contributes valuable empirical data using statistical ML framework but also introduces a novel analytical framework that can inform future research and advancements in the field of smart grid energy management.

A. ENERGY CONSUMPTION ANALYSIS BASED ON LOCATION ID

In this section, we have performed the EVs energy consumption analysis considering the Location ID attribute. There are a number of CSs represented by the Station ID within a CS location denoted by Location ID. FIGURE 2 visualizes the summary of the number of charging sessions initiated by the user and utilized by the CS at that specific Location ID. Further, the graph displays the charging sessions broken down by Location ID. It can be observed from the figure that 524 charging sessions have been recorded for Location ID 493904, which seems to be the highest number of sessions from which EVs are utilizing energy based on the CS location. With 401 sessions, Location ID 976902 has the second-highest number of sessions, while Location ID 10085 has the lowest number of sessions at 1. The recorded and analyzed charging sessions for Location ID indicates the EVs energy demand for each CS location.

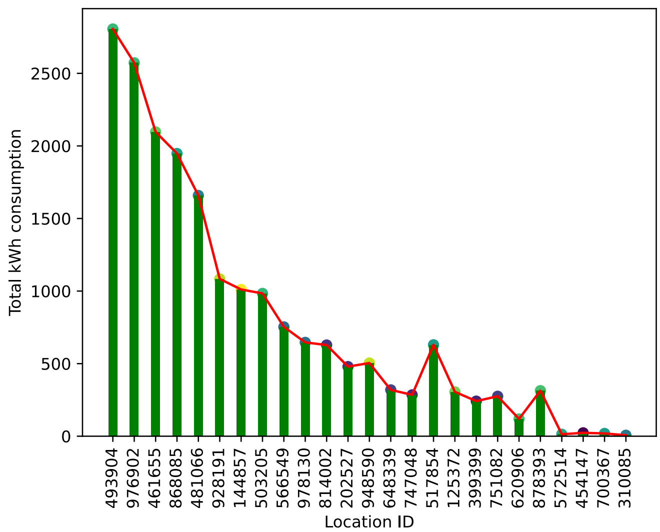
1) STATISTICS OF EACH LOCATION ID

In this section, we have provided details about the total and average EVs energy consumption for every CS location (denoted by Location ID) as mentioned in the dataset. FIGURE 3a illustrates the analysis of the total energy consumption (in kWh) pattern for each of the Location ID. It can be perceived from the graph that the CS location with the highest energy usage is Location ID 493904 at which EV's energy fulfillment results into energy utilization of 2805.85999 kWh. On other hand, Location ID 310085 reflects lowest energy usage by the EVs based on the analysis performed on the dataset.

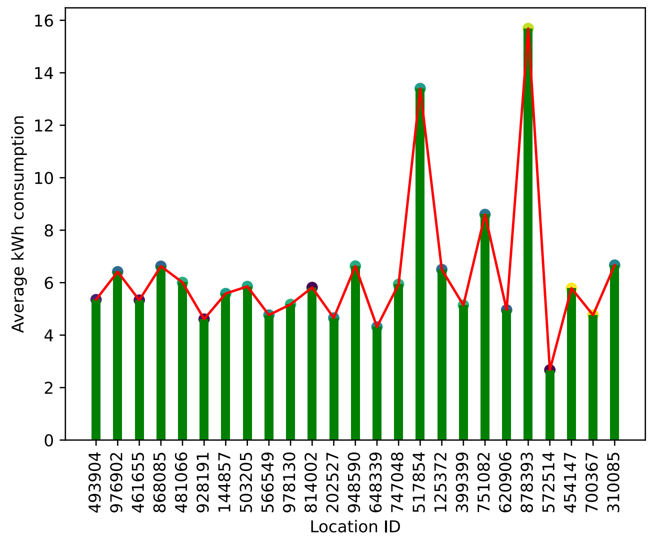
FIGURE 3b shows the location's average kWh analysis of EVs energy consumption corresponding to the various Location IDs at which various CSs are located within that particular region. Location ID 878393 reflects the highest energy utilization by the EVs, i.e., approximately 15.6905 kWh, while Location ID 572514 shows the lowest energy consumption of 2.6719 kWh, which represents that EVs are not opting for the particular location more often to fulfil their energy demand. Thus, energy usage corresponding to the specific Location ID helps to get insights into the requirement of the energy distribution based on the user's energy demand and considering cost parameters to make the charging energy efficient.

2) TOTAL AND AVERAGE ENERGY CONSUMPTION BY LOCATION ID BASED ON WEEKDAY

This section analyzes the overall weekly energy consumption (in kWh) for a specific Location ID. In this context, FIGURE 4a and FIGURE 4b show the visualization of total energy consumption for all the Location IDs based on the weekday. For that, we have considered two types of visualizations, i.e., heatmap and stacked chart. FIGURE 4a depicts the energy

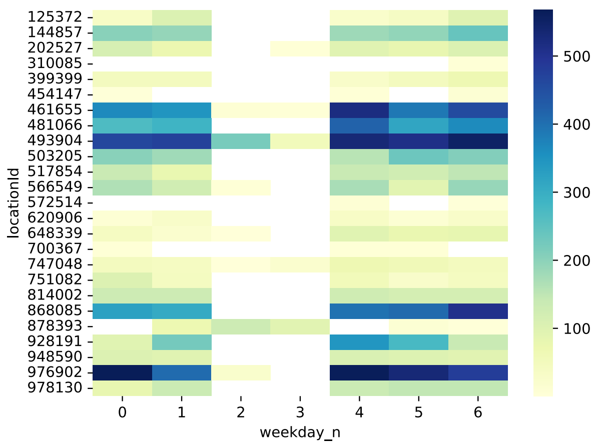


(a) Total kWh Consumption for each Location ID.

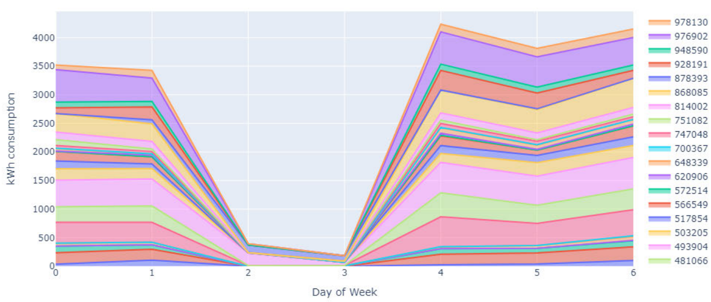


(b) Average kWh Consumption for each Location ID.

FIGURE 3. Total and average energy consumption analysis based on Location ID.



(a) Daily total kWh consumption for each Location ID based on weekday

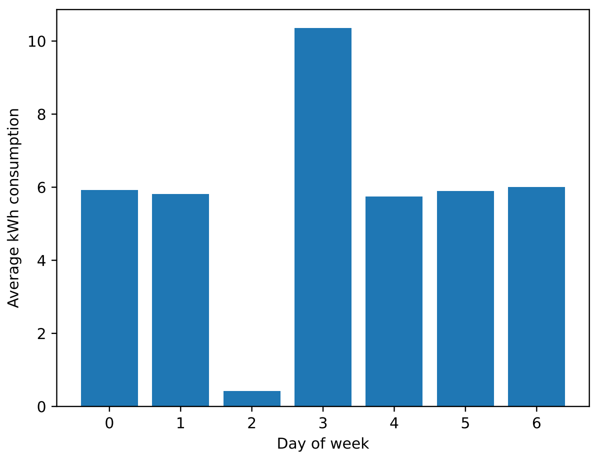


(b) Breakdown of kWh consumption based on weekday

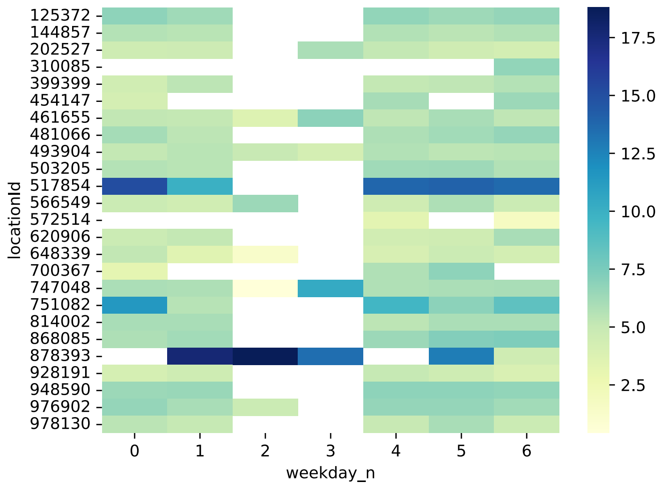
FIGURE 4. Total energy consumption for Location ID based on the weekday.

consumption analysis in kWh in the form of a heatmap based on the days of the week. With the help of a heatmap, it seems easy to determine which location users may find busiest due to the high energy demand on the working days. Further, another visualization chart, i.e., FIGURE 4b, shows the stacked chart that displays the total energy usage of the particular location (with multiple CSs) denoted by Location ID for the weekdays. After examining the graph, we have identified that Thursday(denoted by label 4) has the highest energy consumption, and Sunday(denoted by label 3) has the lowest energy consumption.

We have conducted an analysis of the average daily kWh consumption for each location and present our findings with the two visualization charts, i.e., FIGURE 5a and FIGURE 5b. Foremost, we can focus on FIGURE 5b, which depicts the detailed breakdown of the typical daily average kWh usage for each location by day of the week. The x-axis shows the weekdays, from Monday through Sunday, and the y-axis shows the mean daily kWh consumption. This number enables us to spot any patterns or outliers in each location's daily energy consumption patterns. For instance, we might see that some places use more energy throughout the week than they do on the weekends, while other places display a more regular energy usage pattern. We pay particular attention to location 747048 in FIGURE 5a. The graph shows the typical daily kWh usage for the aforementioned location based on the weekday. The big departure from the mean line indicates a significant difference in energy

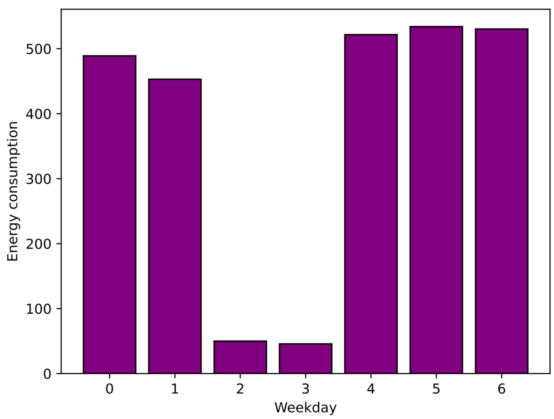


(a) Average kWh energy consumption based on the weekday for Location 747048

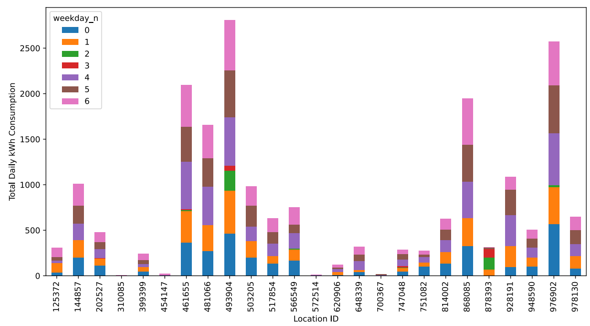


(b) Daily average kWh consumption for all locations

FIGURE 5. **Average energy consumption analysis for Location ID based on the weekday.**

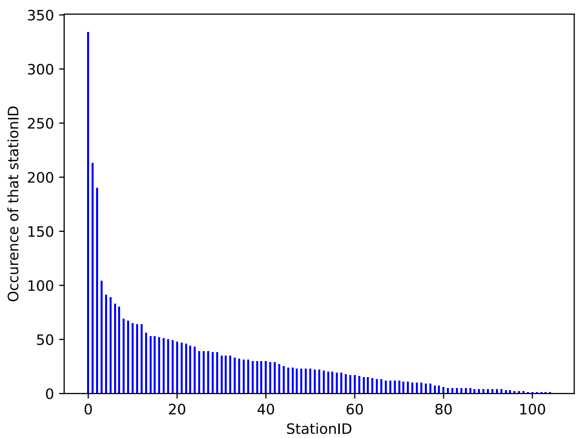


(a) The total daily kWh consumption for all locations

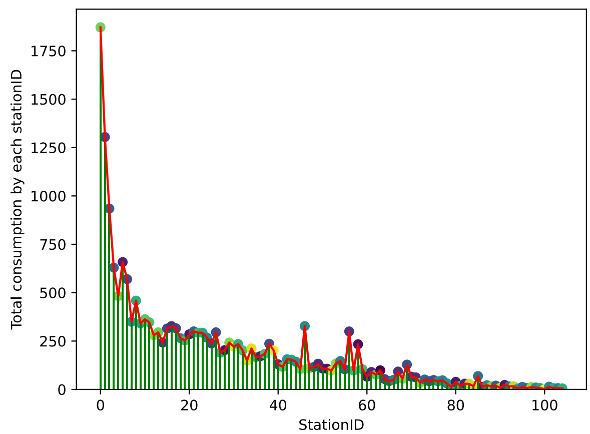


(b) Total kWh consumption

FIGURE 6. **Total kWh consumption by time of day for each location.**



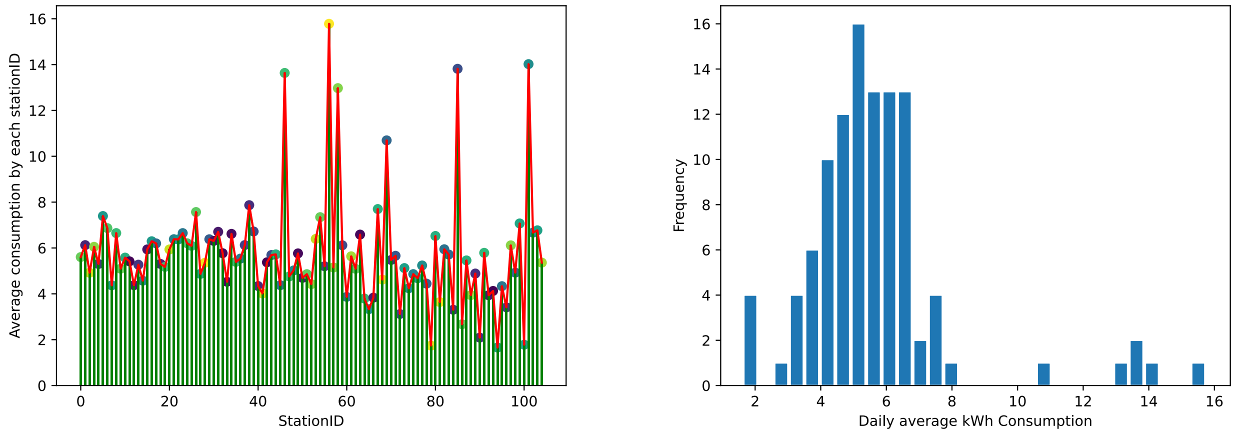
(a) Station ID and its occurrence



(b) Total consumption for each Station ID

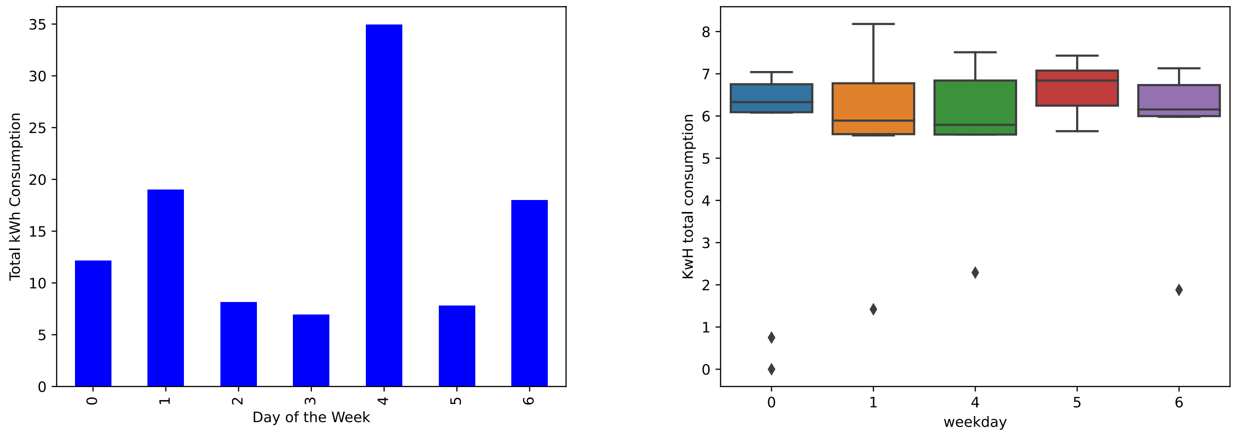
FIGURE 7. **Station ID occurrence and total energy consumption analysis based on the Station ID.**

use between Saturday (2) and Sunday (3) than the other weekdays. Most weekdays have the same average energy consumption, ranging from 5.5 to 6 kWh. This finding hints that some operational processes use more energy on weekends, or it might be a chance to optimize energy usage over the weekend.



(a) Average consumption for each Station ID (b) Histogram of Daily average KwH consumption for all stations

FIGURE 8. Total and average energy consumption analysis based on the Station ID.



(a) Total kWh Consumption on weekday for Station ID 582873 (b) Daily Total kWh Consumption for Station 129465 in Box plot

FIGURE 9. Energy utilized at each CS on weekday for Station ID.

Overall, FIGURE 5a and 5b provide useful insights into the daily average energy usage patterns of each location, which can be used to design customized and successful energy management plans. Industries may optimize their energy usage and reduce costs while reducing their environmental footprint by analyzing consumer trends and abnormalities at each site.

3) AVERAGE OF TOTAL ENERGY CONSUMPTION ANALYSIS BY LOCATION ID BASED ON WEEKDAY

FIGURE 6a depicts the total daily kWh consumption for all locations considering the weekdays. Analyzing the total energy consumption data at each location provides the average kWh consumption statistic for each day of the week. The graph's bar chart displays the number of days of the week on the x-axis and the total number of kWh consumed on the y-axis. The height of each bar represents the total daily kWh use for that weekday across all locations.

Figure 6 shows the proportion of the total kWh used for each location on every day of the week, from Monday to Sunday. As we analyze the stack bar graph, we have determined that the specific location, 493904, has the highest total consumption over each day of the week, and locations 310085, 454147, 572514, and 700367 have significantly lower total energy consumption. This graph can give us insight into how we have to distribute the energy to each location based on their previous weekday-wise analysis. By identifying the location with the highest energy consumption, we can allocate more resources to meet their needs and ensure they do not experience any power outages. Additionally, we can investigate why the other locations have lower energy consumption and see if there are any opportunities to optimize their energy usage.

In the next section, we can observe the Station IDs, which are quite important in our data collection and analysis performed for the energy consumption. Understanding the relevance of Station IDs can be incredibly beneficial when working with large amounts of data.

B. ENERGY CONSUMPTION ANALYSIS BASED ON STATION ID

In this section, we have observed the frequency of occurrence of each of the 105 Station IDs as shown in Figure 7a, which gives us useful information about the most commonly utilized and power-consuming CSs according to the considered dataset. This knowledge can be used to predict and analyze data sets. For example, identifying stations with fewer occurrences allows us to remove them from the dataset, making it more manageable and focused. To make the data more understandable, the Station IDs are sorted by frequency of occurrence, and the resulting graph is displayed. The x-axis shows the Station IDs, while the y-axis indicates the frequency of occurrence. This sorting enables us to discover the most commonly occurring Station IDs rapidly. By analyzing the frequency of station IDs, we can identify the most power-consuming stations and investigate the reasons behind their high energy consumption.

1) STATISTICS OF EACH STATION ID

This section highlights how each CS behaves for their total and average energy consumption for charging the EVs. FIGURE 7b shows the total energy consumption by each station which is represented in kWh. We have focused on determining which CS consumes more energy than others, and it is clear that a station's frequency affects how much energy is utilized overall. The station's overall energy usage rises when its frequency increases. It gave us an estimate of how much energy we had overall supplied to a particular station over the analysis period. FIGURE 8a displays an assortment of the EV CS's average kWh (kilowatt hours) usage for the Station IDs. The x-axis shows the unique identity of each charging station, while the y-axis shows the usual kWh utilized by each station. On the scatter plot, a dot represents each station, and the dot colour is a random value used to distinguish across stations. A green bar in the graph shows each station's typical kWh. The red line, which also shows the normal pattern of kWh consumption across all stations, indicates the trend of the data points. Average energy consumption is calculated by dividing the total consumption by the total number of instances in the dataset.

FIGURE 8b visualizes the daily average kWh consumption frequency distribution for all stations. The daily average kWh consumption is shown on the x-axis, and the frequency of that consumption level is shown on the y-axis. The width of each of the 30 bins in the histogram, which has 30 of them, is 0.4.Â The graph shows how the daily average kWh consumption for each station in the data set is distributed overall. Here, the minimum energy consumption is approximately 2 and is provided by four stations, while the maximum consumption is around 15 and is provided by just one station. Almost all stations need an average usage of between 3 and 8 kWh.

2) STATISTICS REGARDING PARTICULAR STATION IDS

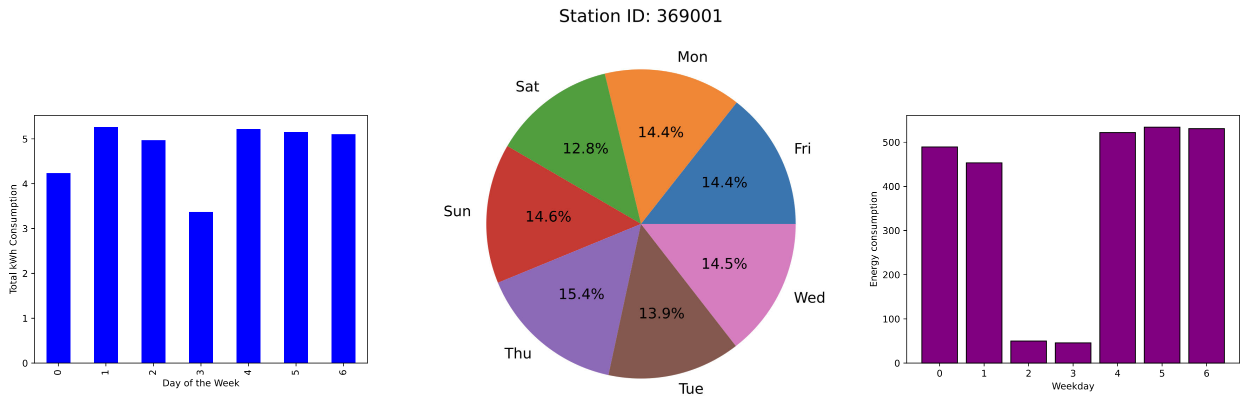
In this subsection, we have focused on the expanded version of section I, which discusses the overall station IDs' average and total energy consumption with the help of dataset analysis. In contrast, we have covered particular Station IDs characteristics in this section. FIGURE 9a depicts the total energy consumption (in kWh) analysis performed for Station ID 582873 based on the weekdays. CS corresponding to Station ID 582873 reflects energy usage of 4.745 kWh on Monday, 7.780 kWh on Tuesday, and 3.598 kWh on Wednesday, among other days. FIGURE 9b shows the summary of the number of kWh utilized at each charging station on each day of the week. The box plots make it easy to compare the mean and range of the consumption distribution for each charging station by day of the week. FIGURE 9b demonstrates how the distribution of kWh usage at each charging station changes dramatically depending on the day of the week. The box plots show that although certain stations' weekly consumption patterns are quite stable, those of other stations are more variable. In order to assure optimal energy utilization, outlier identification and control measures are required. Outliers are obvious in the box plot.

Each day of the week's average consumption is shown by the bar at station 955429 in FIGURE 10a. The average daily consumption of Station ID 955429 is 5.2632 kWh on Monday, 5.1562 kWh on Tuesday, 5.0945 kWh on Wednesday, 5.2248 kWh on Thursday, 4.2303 kWh on Friday, 4.9677 kWh on Saturday, and 3.3737 kWh on Sunday. The energy consumption on Monday(1) is higher than on any other day of the week; on Sunday(3), it is significantly lower. 10b shows the visualization of average kWh usage for each day of the week in the form of pie charts for each CS with Station ID 369001. The data shows that certain CSs use the most energy on weekdays, while others use the most on weekends.

3) AVERAGE AND TOTAL ENERGY CONSUMPTION BASED ON WEEKDAY FOR STATION ID

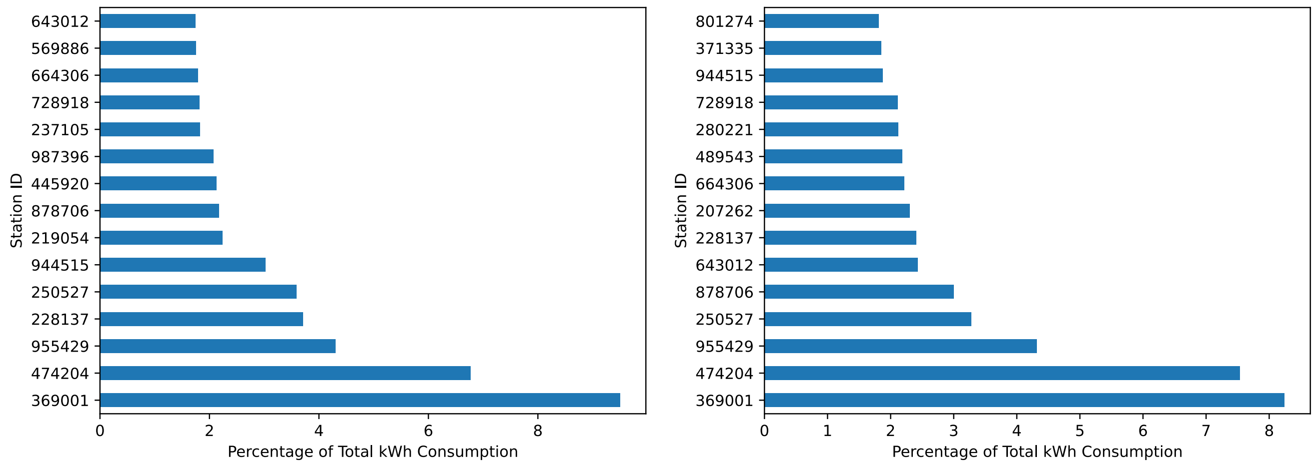
This section is an extended version of subsection II, which discusses the total and average consumption for a particular station. We calculated the total energy consumed by stations on certain weekdays by adding the average energy consumption of each station.

FIGURE 10c depicts the overall kWh usage considering the weekday. This statistic is the result of processing data from all stations, and it represents the average kWh usage for each day of the week. The graph is a bar chart with the y-axis indicating total kWh usage and the x-axis displaying the days of the week. The bars are black-bordered and purple in tone. Each bar's height shows the total kWh usage for that weekday across all sites. This graph depicts the overall energy consumption trends for each day of the week across all stations. We crosschecked the graph using Location ID information from FIGURE 6b and Table 9 to corroborate further the insights provided by FIGURE 10c. We discovered that all stations utilized a total of 452.99 kWh on Monday, 534.032 kWh on Tuesday, 530.387 kWh on Wednesday, 521.662 kWh on Thursday, 489.073 kWh on Friday, 49.94 kWh on Saturday, and 45.702 kWh on Sunday, according to the data shown in Figure 6b and TABLE 9. These results are consistent with the overall trends shown in 10c.



(a) Average kWh Consumption on weekday for station 955429 (b) Average kWh Consumption on weekday for station 389001 (c) Total energy consumption on weekday for all stations

FIGURE 10. **Average and Total energy consumption analysis based on the weekday considering Station ID and all stations.**



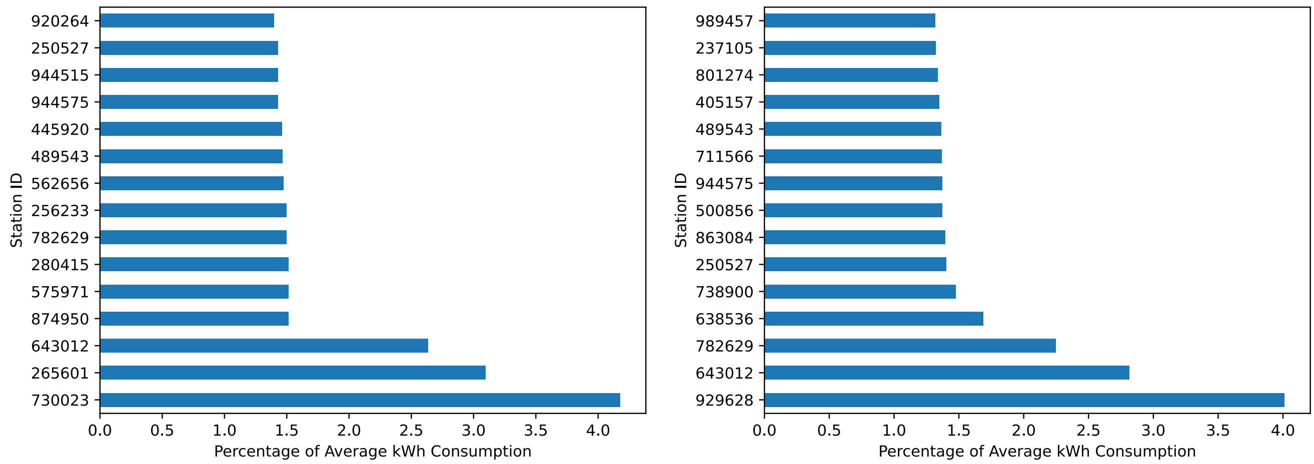
(a) Percentage of Total kWh Consumption on Day 1 [Monday] for Top 15 Station (b) Percentage of Total kWh Consumption on Day 4 [Friday] for Top 15 Station

FIGURE 11. **Percentage of total energy consumption analysis of top 15 Station IDs for a particular day.**

4) PERCENTAGE-BASED ENERGY CONSUMPTION FOR STATION IDS

Here, the top 15 out of 105 station IDs are discussed according to the percentage of total and average energy consumption while analyzing the dataset. FIGURE 11a indicates the top 15 stations' percentage of overall energy usage on Monday as a horizontal bar. The dataset's largest proportion of total consumption is represented by Station ID 369001 and is around 8.9 %. The percentage of overall consumption decreases somewhat as we move up the horizontal bar. Because Monday is regarded as a weekday, there is significantly more consumption than Sunday and Saturday. The top 15 stations' percentage of overall energy usage on Friday is shown in Figure 11b as a horizontal bar. The dataset's largest proportion of total consumption is represented by Station ID 369001 and is around 8.1 %. As we can see, aside from the top three stations, every station shifts its place as the day changes.

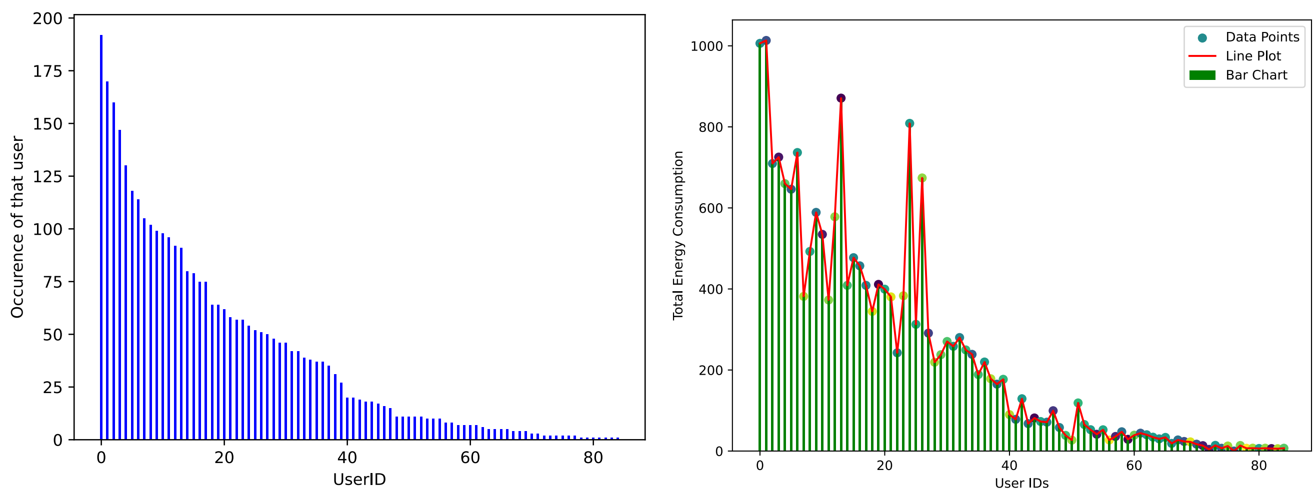
FIGURE 12a and FIGURE 12b displays the average weekly percentage of kWh consumed for the top 12 stations along with their Station IDs for a particular day. The data has been normalized to percentages after being sorted by weekday and Station ID. The horizontal bar graphs show the percentage of the average kWh usage for each station. The stations are listed in descending order, with the top station at the top of the chart. The graph can reveal which stations use the most energy each day of the week, which may assist in pinpointing regions in need of energy-saving measures. Because average consumption is equal to total consumption divided by the frequency of that particular Station ID, we can see a significant variation in Station IDs. It is also practical that consumption at a specific Station ID is higher in just



(a) Percentage of Average kWh Consumption on Day 1 [Monday] for Top 15 Station

(b) Percentage of Average kWh Consumption on Day 4 [Friday] for Top 15 Station

FIGURE 12. **Top 15 stations by the percentage of average kWh consumption for a particular day.**



(a) User ID and its Occurrence

(b) Total Energy Consumption with respect to User IDS

FIGURE 13. **Users ID frequency and total energy consumption for each User ID.**

one day and that Station ID does not occur elsewhere, which would give us a high average consumption by that Station ID.

C. ENERGY CONSUMPTION ANALYSIS BASED ON USER ID

This section discusses the energy consumption pattern by considering the aspect of users. The energy consumption analysis based on User ID information helps identify which stations and locations consume the most energy by considering their users' behaviour.

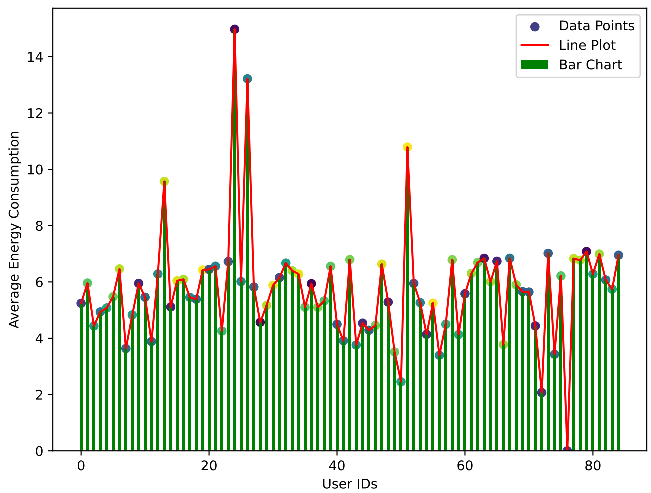
1) STATISTICS OF EACH USER ID

FIGURE 13a shows how frequently each of the 84 UserID sessions occurred, giving us a visual representation of which User IDs utilize the most power and are used the most often. FIGURE 13a on the x-axis represents 84 distinct UserIDs, and the y-axis shows the frequency with which that session occurred. This will help us determine which UserID is using the energy from the system most. Users with an incidence of less than 10 can be removed for the sake of data set analysis.

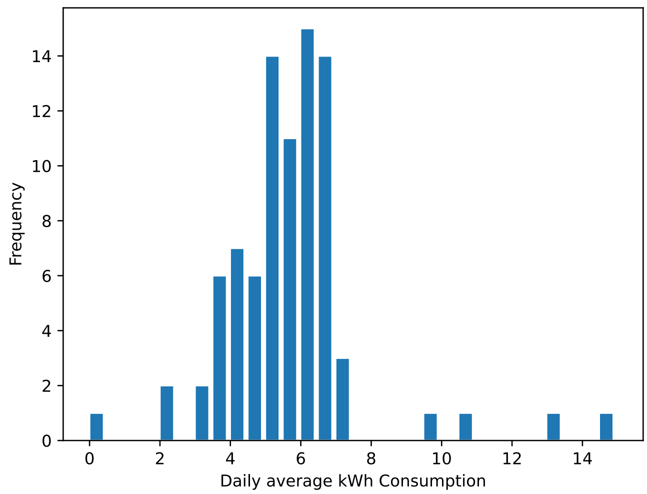
2) TOTAL AND AVERAGE ENERGY CONSUMPTION BY USER ID BASED ON WEEKDAY

The average and overall energy use by users is briefly summarized in this section.

FIGURE 13b is a scatter plot with each point denoting the total consumption data for each UserID. A bar graph and a line graph are also layered on top of the scatter plot in addition to it. Each vertical bar in the bar graph is centered on the associated UserID and displays the identical total consumption figures for each UserID. The bars are green

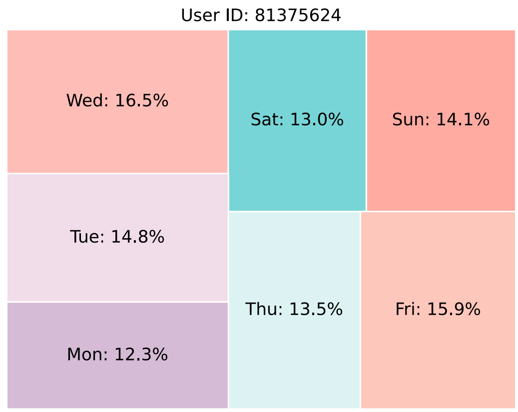


(a) Average Energy Consumption for User IDS

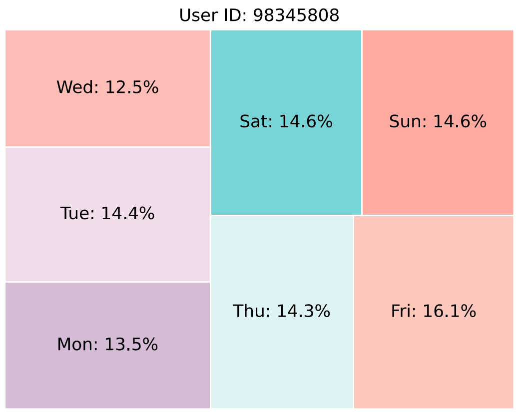


(b) Histogram of Daily average kWh consumption for all Users

FIGURE 14. **Average energy consumption analysis for User IDs.**



(a) Average kWh consumption of User ID 81375624 on weekday



(b) Average kWh consumption of User ID 98345808 on weekday

FIGURE 15. **Average Energy consumption of a specific UserID on weekday.**

in color. The scatter plot's continuous line joining the data points displays the same total consumption figures as the line graph. Overall, utilizing a combination of scatter, line, and bar, the graph seems to adequately convey the overall total consumption (Kwh) data for each UserID.

Based on the user IDs, FIGURE 14a displays the average energy usage of various users. User IDs are shown on the x-axis, while average energy use is shown on the y-axis. There are three different plot kinds in the graph: scatter plots, bar charts, and line graphs. The average energy consumption of each user is shown by a bar, with the height of the bar corresponding to the average energy consumption amount. Overall, this graph gives a visual depiction of how much energy is used on average by various users. The daily average kWh consumption frequency distribution for all users are displayed in figure 14b, there in total 84 userIDs. The daily average kWh consumption is shown on the x-axis, and the frequency of that consumption level is shown on the y-axis. The histogram is with width 0.4 and bin size of 30 for visual presentation. The graph shows how the daily average kWh consumption for each user in the data set is distributed overall. Here, as can be seen, the minimum average energy consumption is approximately between 0 to 1 and is provided by 1 user, while the maximum average consumption is approximately 15 and is provided by just 1 user. Almost all stations need an average usage of between 4 and 7 kWh. These findings suggest that there is a significant variation in energy consumption among users and that most stations require a moderate amount of energy to operate efficiently.

3) STATISTICS REGARDING PARTICULAR USER IDS

This section represents the average energy usage as a percentage for a certain User ID. Here, two user IDs have been considered for visual presentation out of all users. In Figure 15a, we have focused on User ID 81375624 into account and given a graphical analysis of that specific user based on which day of the week they consume a percentage of how much electricity over a week. By finding patterns in the user's energy consumption behaviour, this analysis may help the user optimize their energy consumption. Comparing this user's consumption to others in the same region or demographic might be done through further analysis. The average daily kWh usage for user 98345808 is displayed on treemap charts in Figure 15b. The information shows that this specific user consumes the most energy on weekdays. On Friday, the highest day in the chart, the user used 16.1% of the total average energy.

VI. CONCLUSION AND FUTURE SCOPE

In this paper, we presented an extensive energy consumption and distribution analysis framework for EVs with the help of a smart grid environment. Furthermore, we focused on smart grid-based distribution framework which transfers the energy to the CS reliably and efficiently based on the analyzed EVs energy consumption pattern. For that, we have considered an EV dataset to analyze the energy consumption pattern based on the various attributes such as weekday, weekend, and CS within the location (Station ID, Location ID) and user (User ID). Moreover, we have focused on getting insights into the association between attributes and deeply performing the energy consumption analysis based on the various parameters. The deep analysis of the energy consumption helps the smart grid to optimize its energy to the CS efficiently and reliably. It contain various real-world implementation in industrial and electric automotive domain. It offers practical insights into optimizing energy distribution within smart grids. Findings from CS and location based insights has practical implications for EV charging station planning. Stakeholders involved in the establishment and expansion of charging infrastructure can use our insights to strategically position CS based on anticipated energy demand patterns. Our research contributes to the formulation of informed government policies related to smart grid development and sustainable energy usage. The identification of energy consumption patterns at both the location and station levels provides a basis for predictive maintenance strategies. This has practical implications for enhancing the reliability of EV charging infrastructure, reducing downtime, and ensuring a seamless experience for users.

In the future, we will explore deep learning and reinforcement learning techniques to optimize the energy consumption for EVs arriving at the CS. Also, explore the development of dynamic pricing models that take into account real-time demand, energy availability, and grid conditions. Extend the optimization framework to cater to fleets of EVs, such as those used in ride-sharing services or commercial operations. Investigate cybersecurity measures to ensure the resilience and security of the smart grid infrastructure. Investigate the feasibility and benefits of decentralized energy management systems. It will encourage collaboration between researchers, industry stakeholders, and policymakers to address interdisciplinary challenges in the adoption of smart grid technologies.

REFERENCES

[1] C. B. Saner, A. Trivedi, and D. Srinivasan, "A cooperative hierarchical multi-agent system for EV charging scheduling in presence of multiple charging stations," *IEEE Trans. Smart Grid*, vol. 13, no. 3, pp. 2218–2233, May 2022.
[2] B. Sohet, Y. Hayel, O. Beaude, and A. Jeandin, "Hierarchical coupled driving-and-charging model of electric vehicles, stations and grid operators," *IEEE Trans. Smart Grid*, vol. 12, no. 6, pp. 5146–5157, Nov. 2021.
[3] E. Mocanu, D. C. Mocanu, P. H. Nguyen, A. Liotta, M. E. Webber, M. Gibescu, and J. G. Slootweg, "On-line building energy optimization using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3698–3708, Jul. 2019.
[4] L. Hou, C. Wang, and J. Yan, "Bidding for preferred timing: An auction design for electric vehicle charging station scheduling," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 8, pp. 3332–3343, Aug. 2020.
[5] P. Dixit, P. Bhattacharya, S. Tanwar, and R. Gupta, "Anomaly detection in autonomous electric vehicles using AI techniques: A comprehensive survey," *Expert Syst.*, vol. 39, no. 5, Jun. 2022, Art. no. e12754.
[6] S. Tanwar, R. Kakkar, R. Gupta, M. S. Raboaca, R. Sharma, F. Alqahtani, and A. Tolba, "Blockchain-based electric vehicle charging reservation scheme for optimum pricing," *Int. J. Energy Res.*, vol. 46, no. 11, pp. 14994–15007, 1002.
[7] K. M. Muttaqi, O. Rahman, D. Sutanto, M. S. H. Lipu, M. G. M. Abdolrasol, and M. A. Hannan, "High-frequency ripple injection signals for the effective utilization of residential EV storage in future power grids with rooftop PV system," *IEEE Trans. Ind. Appl.*, vol. 58, no. 5, pp. 6655–6665, Sep. 2022.
[8] Z. Yang, X. Huang, T. Gao, Y. Liu, and S. Gao, "Real-time energy management strategy for parking lot considering maximum penetration of electric vehicles," *IEEE Access*, vol. 10, pp. 5281–5291, 2022.
[9] A.-M. Koufakis, E. S. Rigas, N. Bassiliades, and S. D. Ramchurn, "Offline and online electric vehicle charging scheduling with V2V energy transfer," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 5, pp. 2128–2138, May 2020.
[10] R. Kakkar, R. Gupta, S. Agrawal, S. Tanwar, A. Altameem, T. Altameem, R. Sharma, F.-E. Turcanu, and M. S. Raboaca, "Blockchain and IoT-driven optimized consensus mechanism for electric vehicle scheduling at charging stations," *Sustainability*, vol. 14, no. 19, p. 12800, Oct. 2022.
[11] Y. Zhang, X. Yang, B. Li, B. Cao, T. Li, and X. Zhao, "Two-level optimal scheduling strategy of electric vehicle charging aggregator based on charging urgency," in *Proc. 4th Int. Conf. Smart Power Internet Energy Syst. (SPIES)*, Dec. 2022, pp. 1755–1760.
[12] C. L. Shek, A.-K. Manoharan, S. Gampa, T. Chandrappa, and V. Aravinthan, "A diversity-based clustering technique for implementing decentralized node level charge scheduling of electric vehicles," in *Proc. North Amer. Power Symp. (NAPS)*, Oct. 2019, pp. 1–6.
[13] S. Das, P. Acharjee, and A. Bhattacharya, "Charging scheduling of electric vehicle incorporating grid-to-vehicle and vehicle-to-grid technology considering in smart grid," *IEEE Trans. Ind. Appl.*, vol. 57, no. 2, pp. 1688–1702, Mar. 2021.
[14] A. Mohammad, R. Zamora, and T. T. Lie, "Integration of electric vehicles in the distribution network: A review of PV based electric vehicle modelling," *Energies*, vol. 13, no. 17, p. 4541, Sep. 2020.
[15] Z. Wang and A. Ben Abdallah, "A robust multi-stage power consumption prediction method in a semi-decentralized network of electric vehicles," *IEEE Access*, vol. 10, pp. 37082–37096, 2022.
[16] Z. Wang, M. Ogbodo, H. Huang, C. Qiu, M. Hisada, and A. B. Abdallah, "AEBIS: AI-enabled blockchain-based electric vehicle integration system for power management in smart grid platform," *IEEE Access*, vol. 8, pp. 226409–226421, 2020.
[17] X. Xu, Y. Jia, Y. Xu, Z. Xu, S. Chai, and C. S. Lai, "A multi-agent reinforcement learning-based data-driven method for home energy management," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3201–3211, Jul. 2020.

[18] M. H. K. Tushar, A. W. Zeineddine, and C. Assi, "Demand-side management by regulating charging and discharging of the EV, ESS, and utilizing renewable energy," *IEEE Trans. Ind. Informat.*, vol. 14, no. 1, pp. 117–126, Jan. 2018.
[19] R. Kakkar, S. Agrawal, R. Gupta, and S. Tanwar, "Blockchain and zero-sum game-based dynamic pricing scheme for electric vehicle charging," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, May 2022, pp. 1–6.
[20] A. Fukushima, T. Yano, S. Imahara, H. Aisu, Y. Shimokawa, and Y. Shibata, "Prediction of energy consumption for new electric vehicle models by machine learning," *IET Intell. Transp. Syst.*, vol. 12, no. 9, pp. 1174–1180, Nov. 2018.
[21] D. Kim, H. G. Shim, and J. S. Eo, "A machine learning method for EV range prediction with updates on route information and traffic conditions," in *Proc. AAAI Conf. Artif. Intell.*, vol. 36, Jun. 2022, pp. 12545–12551.
[22] S. Shahriar, A. R. Al-Ali, A. H. Osman, S. Dhou, and M. Nijim, "Prediction of EV charging behavior using machine learning," *IEEE Access*, vol. 9, pp. 111576–111586, 2021.
[23] W. Wang, F. H. Memon, Z. Lian, Z. Yin, T. R. Gadekallu, Q.-V. Pham, K. Dev, and C. Su, "Secure-enhanced federated learning for AI-empowered electric vehicle energy prediction," *IEEE Consum. Electron. Mag.*, vol. 12, no. 2, pp. 27–34, Mar. 2023.
[24] A. Talwariya, P. Singh, M. D. Brinda, J. Jobanputra, and M. L. Kolhe, "Domestic energy consumption forecasting using machine learning," in *Proc. 7th Int. Conf. Smart Sustain. Technol. (SpliTech)*, Jul. 2022, pp. 1–4.
[25] S. Chelladurai, D. Krishna, A. Maheswari, and V. R. R. Sivakumar, "Machine learning-based management of hybrid energy storage systems in e-vehicles," *J. Nanomaterials*, vol. 2022, Aug. 2022, Art. no. 8270718.
[26] C. Lazaroiu and M. Roscia, "New approach for smart community grid through blockchain and smart charging infrastructure of EVs," in *Proc. 8th Int. Conf. Renew. Energy Res. Appl. (ICRERA)*, Nov. 2019, pp. 337–341.
[27] C. Mariño, G. Cossio, P. Massaferro, M. Di Martino, A. Gómez, and A. Fernandez, "NILMEV: Electric vehicle disaggregation for residential customer energy efficiency incentives," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Jan. 2023, pp. 1–5.
[28] S. N. Jahromi, "Prediction of electric vehicle's annual accessibility to chargers for providing ancillary services using an efficient random forest method," in *Proc. 8th Int. Conf. Technol. Energy Manage. (ICTEM)*, Feb. 2023, pp. 1–6.
[29] J. M. Aguiar-Pérez and M. Á. Pérez-Juárez, "An insight of deep learning based demand forecasting in smart grids," *Sensors*, vol. 23, no. 3, p. 1467, Jan. 2023.
[30] C. Scott, M. Ahsan, and A. Albarbar, "Machine learning based vehicle to grid strategy for improving the energy performance of public buildings," *Sustainability*, vol. 13, no. 7, p. 4003, Apr. 2021.
[31] S. Ahmad, A. Ahmad, A. Shoukat, B. Behram, S. S. Khan, and R. Yaqub, "Consumer energy management in residential distribution power system considering consumer's privacy," in *Proc. 16th Int. Conf. Emerg. Technol. (ICET)*, Dec. 2021, pp. 1–7.
[32] X. Xu, H. M. A. Aziz, H. Liu, M. O. Rodgers, and R. Guensler, "A scalable energy modeling framework for electric vehicles in regional transportation networks," *Appl. Energy*, vol. 269, Jul. 2020, Art. no. 115095.
[33] R. O. Kene and T. O. Olwal, "Energy management and optimization of large-scale electric vehicle charging on the grid," *World Electr. Vehicle J.*, vol. 14, no. 4, p. 95, Apr. 2023.
[34] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
[35] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning* (Springer Series in Statistics), 2nd ed. New York, NY, USA: Springer, 2009.
[36] G. Sidhu and B. Caffo, "MONEYBaRL: Exploiting pitcher decision-making using reinforcement learning," *Ann. Appl. Statist.*, vol. 8, no. 2, pp. 1–21, Jun. 2014.
[37] T. Poggio and L. Rosasco, "Statistical learning theory and applications, 2012," Massachusetts Inst. Technol., USA, Tech. Rep. 9.520, 2012.
[38] M. Hardy et al., "Statistical learning theory," USA, Tech. Rep., 2023. [Online]. Available: https://pdfcoffee.com/statistical-learning-theory-pdf-free.html
[39] O. I. Asensio, C. Z. Apablaza, M. C. Lawson, and S. E. Walsh, "Replication data for: A field experiment on workplace norms and electric vehicle charging etiquette," Georgia Inst. Technol., USA, Tech. Rep., 2020, doi: 10.7910/DVN/NFPQLW.

FENIL RAMOLIYA is currently pursuing the Bachelor of Technology degree in computer engineering from the Institute of Technology, Nirma University. His academic pursuits and research interests include cutting-edge domains, including deep learning, data science, explainable AI, blockchain, healthcare 4.0, and network security.

CHINMAY TRIVEDI (Student Member, IEEE) received the Master of Technology degree in computer engineering from the Institute of Technology, Nirma University. His research interests include deep learning, data science, and blockchain. He is working on presenting solutions to integrate generative adversarial networks in adversarial learning techniques in wide range of domains, such as healthcare, vehicular networks, and emerging communication networks.

KRISHA DARJI is currently pursuing the Bachelor of Technology degree in computer engineering with the Institute of Technology, Nirma University. Her research interests include deep learning, data science, explainable AI, and blockchain.

RIYA KAKKAR (Student Member, IEEE) is currently pursuing the Ph.D. degree with the Computer Science and Engineering Department, Institute of Technology, Nirma University. Her research interests include electric vehicles, smart grid, and machine learning. She has several journals and conference papers to her credit.

RAJESH GUPTA (Member, IEEE) received the Bachelor of Engineering degree from the University of Jammu, India, in 2008, the master's degree in technology from Shri Mata Vaishno Devi University, Jammu, India, in 2013, and the Ph.D. degree in computer science and engineering from Nirma University, Ahmedabad, Gujarat, India, in 2023, under the supervision of Dr. Sudeep Tanwar. He is currently an Assistant Professor with Nirma University. He has authored/coauthored more than 150 publications (including articles in SCI indexed journals and IEEE ComSoc sponsored international conferences). Some of his research findings are published in top-cited journals and conferences, such as IEEE Transactions on Industrial Informatics, IEEE Transactions on Network and Service Management, IEEE Transactions on Network Science and Engineering, IEEE Transactions on Green Communications and Networking, IEEE Transactions on Computational Social Systems, *IEEE Network* magazine, IEEE Internet of Things Journal, *IEEE Internet of Things Magazine*, *Computer Communications*, *Computer and Electrical Engineering*, *International Journal of Communication Systems* (Wiley), *Transactions on Emerging Telecommunications Technologies* (Wiley), *Physical Communication* (Elsevier), IEEE ICC, IEEE INFOCOM, IEEE

GLOBECOM, and IEEE CITS. His research interests include device-to-device communication, network security, blockchain technology, 5G communication networks, and machine learning. His Google Scholar citations are more than 4200, H-index 32, and i10-index 68. He is also a recipient of Doctoral Scholarship from the Ministry of Electronics and Information Technology, Government of India, under the Visvesvaraya Ph.D. Scheme. His name has been included in the list of Top 2% scientists worldwide published jointly by Elsevier and Stanford University, USA, consecutively, in 2021, 2022, and 2023. He is a recipient of Travel Grant from IEEE ComSoc to attend the IEEE ICC 2021 held in Canada and from SERB DST to attend ICC 2023 held in Rome, Italy. He has been selected and participated the fully-funded the most prestigious ACM's Heidelberg Laureate Forum 2023, which was held in Heidelberg University, Heidelberg, Germany. He has been awarded the Best Research Paper Awards from IEEE ECAI 2021, IEEE ICCCA 2021, IEEE IWCMC 2021, and IEEE SCIoT 2022. He has been felicitated by Nirma University for their research achievements bagged, in 2019, 2020, 2021, and 2022. He is also an active member of ST Research Laboratory (www.sudeeptanwar.in).

SUDEEP TANWAR (Senior Member, IEEE) is currently a Professor with the Computer Science and Engineering Department, Institute of Technology, Nirma University, India. He is also a Visiting Professor with Jan Wyzykowski University, Polkowice, Poland; and The University of Piteşti, Pitesti, Romania. He has authored two books, edited 13 books, and more than 270 technical papers, including top journals and top conferences, such as IEEE Transactions on Network Science and Engineering, IEEE Transactions on Vehicular Technology, IEEE Transactions on Industrial Informatics, IEEE Wireless Communications, *IEEE Network* magazine, ICC, GLOBECOM, and INFOCOM. He initiated the research field of blockchain technology adoption in various verticals, in 2017. His H-index is 68. He actively serves his research communities in various roles. His research interests include blockchain technology, wireless sensor networks, fog computing, smart grid, and the IoT. He is a member of the Technical Committee on Tactile Internet of the IEEE Communication Society. He is a Senior Member of CSI, IAENG, ISTE, and CSTA. He has been awarded the Best Research Paper Awards from IEEE GLOBECOM 2018, IEEE ICC 2019, and Springer ICRIC-2019. He has served many international conferences as a member of the organizing committee, such as the Publication Chair for FTNCT-2020, ICCIC 2020, and WiMob2019; a member of the Advisory Board for ICACCT-2021 and ICACI 2020; the Workshop Co-Chair for CIS 2021; and the General Chair for IC4S 2019 and 2020 and ICCSDF 2020. He is serving on the editorial boards for *Frontiers of Blockchain*, *Cyber Security and Applications*, *Computer Communications*, the *International Journal of Communication Systems*, and *Security and Privacy*.

ZDZISLAW POLKOWSKI received the Engineering degree in computer systems in industry from the Technical University of Zielona Góra, the master's degree in microcomputer systems in management from the Wroclaw University of Economics, Wrocław, the master's degree in IT education from the University of Economics in Katowice, and the Ph.D. degree in computer science and management from the Wroclaw University of Technology. He is currently a Professor with the Faculty of Technical Sciences, Jan Wyzykowski University, Polkowice, where he is also a Rector's Representative for International Cooperation and Erasmus+ Program. Since 2019, he has been an Adjunct Professor with the Department of Business Intelligence in Management, Wroclaw University of Economics and Business, Poland. Since 2021, he collaborates with The Karkonosze State University of Applied Sciences, Jelenia Góra, Poland. Moreover, he is a Visiting Professor with The University of Piteşti, Romania; and WSG University Bydgoszcz, Poland; and an Adjunct Professor with Marwadi University, India. He is the former Dean of the Technical Sciences Faculty, UZZM, Lubin, from 2009 to 2012. He has published more than 120 articles in journals, 25 conference proceedings, including more than 50 articles in journals indexed in the Web of Science, Scopus, IEEE. He is the co-editor of five books which have been published in Springer and CRC Press Taylor & Francis. He served as a member of technical program committee in many international conferences, journals in Poland, India, China, Iran, Syria, U.K., Romania, Turkey, and Bulgaria. Till date, he has delivered over 60 invited talks at different international conferences across various countries. He is also a member of the Board of Studies and an expert member of the doctoral research committee in many universities in India. He is also a member of the editorial board of several journals and he served as a reviewer in a wide range of international journals. His research interests include IT in business, the IoT in business, and education technology. He has successfully completed a research project on developing the innovative methodology of teaching business informatics funded by the European Commission. He also owns an IT SME consultancy company in Polkowice and Lubin, Poland.

FAYEZ ALQAHTANI is currently a Full Professor with the Software Engineering Department, College of Computer and Information Siences, King Saud University (KSU). He was appointed as the Director of the Computer Division, Deanship of Student Affairs. He is also a member of a number of academic and professional associations, such as the Association for Computing Machinery (ACM), the Australian Computer Society, and the Association for Information Systems. He has conducted research projects in several areas of information and communication technology, such as Web 2.0, information security, enterprise architecture, software process improvement, the Internet of Things, and fog computing. He has participated in several academic events.

AMR TOLBA (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees from the Mathematics and Computer Science Department, Faculty of Science, Menoufia University, Egypt, in 2002 and 2006, respectively. He is currently a Full Professor in computer science with King Saud University (KSU), Saudi Arabia. He has authored/coauthored over 180 scientific articles in top-ranked (ISI) international journals, such as IEEE Internet of Things Journal, IEEE Access, IEEE Systems Journal, *ACM TOIT*, *IEEE Consumer Electronics Magazine*, *FGCS*, *JNCA*, *NC&A*, *JAIHC*, *COMNET*, *COMCOM*, *P2PNET*, *VCOM*, and *WWWJ*. He served as a Technical Program Committee (TPC) Member at several conferences, such as DSIT 2022, CICA2022, EAI MobiHealth 2021, DSS 2021, AEMCSE 2021, ICBDM 2021, ICISE 2021, DSS 2020, NCO 2020, ICISE2019, ICCSEA 2019, DSS 2019, FCES 19, ICISE 2018, ESG'18, Smart Data'17, NECO 2017, NC'17, WEMNET'17, NET'17, and Smart Data'16. He has been included in the list of the top 2% of influential researchers globally (prepared by scientists from Stanford University, USA) during the calendar years 2020, 2021, 2022, and 2023, respectively. He served as an associate editor/a guest editor for several ISI journals. He has translated four books into the Arabic language. His main research interests include artificial intelligence (AI), the Internet of Things (IoT), data science, and cloud computing.