TRIBHUVAN UNIVERSITY

INSTITUTE OF ENGINEERING PULCHOWK CAMPUS

THESIS NO:079/MSPSE/002

OPTIMIZING EV CHARGING STATION WITH RL BASED SOLAR DER INTEGRATION

By

Aryasupurna Timalsina

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING IN

PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN POWER SYSTEM ENGINEERING

DEPARTMENT OF ELECTRICAL ENGINEERING

LALITPUR, NEPAL

April 2025

“OPTIMIZING EV CHARGING STATION WITH RL BASED SOLAR DER INTEGRATION”

By

Aryasupurna Timalsina

Roll No: PUL079MSPSE002

Thesis Supervisor:

Associate Professor Dr Basanta Kumar Gautam

Department of Electrical Engineering Pulchowk, Nepal

A Thesis

Submitted to the Department of Electrical Engineering in partial fulfillment of the requirements for the Degree of

Master of Science in Electrical Engineering in Power System Engineering

Department of Electrical Engineering

Institute of Engineering, Pulchowk Campus Tribhuvan University

Lalitpur, Nepal

April 2025

COPYRIGHT©

Head of Department

Department of Electrical Engineering

Tribhuvan University, Institute of Engineering

Pulchowk Campus, Pulchowk, Lalitpur, Nepal

त्रिभुवन विश्वविद्यालय
TRIBHUVAN UNIVERSITY
इन्जिनियरिङ्ग अध्ययन संस्थान
INSTITUTE OF ENGINEERING
पुल्चोक क्याम्पस
PULCHOWK CAMPUS
DEPARTMENT OF ELECTRICAL ENGINEERING
Pulchowk, Lalitpur

CERTIFICATE OF APPROVAL

The undersigned certify that they have read and recommended to the Institute of Engineering for acceptance, a dissertation entitled **"Optimizing EV Charging station with RL based Solar DER integration"**, submitted by **Aryasupurna Timalsina** in partial fulfillment of the requirement for the award of the degree of **Master of Science in Power System Engineering.**

Assoc. Prof. Dr. Sujan Adhikari
Hillside College of Engineering
(External Examiner)

Asst. Prof. Dr. Bishal Silwal
Program Coordinator
M.Sc. in Power System Engineering Pulchowk Campus, Lalitpur

Assoc. Prof. Dr. Basanta K. Gautam
Head of Department
Department of Electrical Engineering
Pulchowk Campus, Lalitpur
(Supervisor)

April, 2025

Tel: +977-1-5443081, Email: doee@pcampus.edu.np, Website: www.doee.pcampus.edu.np

ABSTRACT

Electric vehicles in Nepal are seen to be increasing, both private and public sector transportation are seen adopting electric vehicles as a means of transport. Also, the number of charging stations has been increasing. Due to this trend, there is a lot of energy demand in EV charging stations. In the present scenario, the energy demand of EVs is fulfilled by the Main Grid Supply.

In the charging station multiple players as MPG, DER, ESS and EVCS can be present. This dissertation presents scheduling model of charging station for three different schemes. The first scheme is a stand-alone model, where the EV receives energy from the main power grid only. Similarly, other models are microgrid-based models and solar-based models. In the microgrid-based model players such as MPG, ESS, and Solar operate in unison to fulfill the demand for charging stations. Besides, the solar-based model has solar as the prime energy source to fulfill the demand for charging stations, if solar energy is insufficient then only MPG and ESS operates. Also, optimization function for each scheme is defined and upon utilizing the concept DDQN algorithm, type of Reinforcement learning the model is operated to obtain the optimal revenue. Besides, data of EV demand, solar plant generation, energy price are used in training the algorithm. While training the model based on real environment data, the microgrid-based model showed maximum revenue, and the standalone-based model showed lowest revenue. The profit for standalone model is found -4000 and for Microgrid based model is 14500. Similarly, the confusion Matrix for microgrid based model showed 78% accuracy. Further for solar based model the profit is 2500 and confusion matrix showed 71% accuracy. Upon visiting the charging station, the present scenario of scheduling is only guaranteed by MPG and the profit is negative. This validates the result that charging station requires such optimal scheduling schemes to improve the revenue.

To predict, the data of EV import information of vehicle count is studied from Department of custom, 12 years past data is studied and 5 years scenario of vehicle count is predicted. During forecasting regression analysis is carried, also necessary policies formulated by government is considered. Most importantly 1 year data of charging volume, Revenue and Electricity bill of Bhaktapur charging station is studied. Then as per the government policy and vehicle number trend obtained from prior forecast is utilized to predict the 5 years scenario of the charging station.

Keywords: **Electric Vehicle, Distributed Energy Resources, Energy Storage system, Deep Q learning, DDQN, Charging station, Microgrid, Main Grid Supply.**

ACKNOWLEDGEMENT

I express my deepest sincerity to the supervisor Associate Prof. Dr. Basanta Kumar Gautam, who has provided continuous support, guidance and perceptive comments throughout the research journey. Besides, I am very much thankful to the Department of Electrical Engineering, IOE, Pulchowk Campus who was there to provide necessary resources and academic environment through the master's tenure.

I am highly grateful to have been a part of student mobility funded by Erasmus International Student Program of the European Union. The tenure of this mobility has augmented my academic journey. Besides, I also got chance to engage myself with diverse culture and academic traditions. The financial support provided under this project made it possible for me to pursue this invaluable learning experience abroad. My sincere thanks go to School of Electrical and Computer Engineering, Electrical Sector, Aristotle University of Thessaloniki, Greece who provided this opportunity.

I would like to extend my appreciation to my husband who was there with his unwavering encouragement and understanding during this challenging yet gratifying endeavour. Also, I would thank my family and friends for their moral support, this gave me strength throughout my academic journey.

Thank you all for believing in me and for being part of this journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

DDQN	**Double Deep Q Network**
DER	**Distributed Energy Resources**
DQN	**Deep Q Network**
ESS	**Energy Storage System**
EV	**Electric Vehicle**
EVCS	**Electric Vehicle Charging Station**
ML	**Machine Learning**
MPG	**Main Power Grid**
RL	**Reinforcement Learning**

CHAPTER ONE: INTRODUCTION

1.1 Introduction

1.1.1 Background

The transportation sector of Nepal is one of the prominent contributors of greenhouse gases, about 44 percent [9] of greenhouse gases and 99 percent of CO_2 emission occurs from the road infrastructure. Due to such a scenario, Electric Vehicles have now gained more popularity as these are gas-free fuels. Also, the Electric Vehicle Association of Nepal has estimated that in 2018, the EVs number reached about 45000[5]. In addition, the EV Act 2019 policy has said that 90 percent private and 60 percent public of transport will be converted to electric vehicles by 2030.[5] Similarly, the number of EV charging stations has also increased. At present there are at least 350 [9] charging stations in Nepal. Moreover, these stations have mostly the CCS2 type of charger while the NEA's charging station has the GB/T type of plugs. From May 2019, it is mentioned that 25 percent of private and 25 percent of public sales will be made electric by 2025. Similarly, from the budget speech 2022/23, it was said that arrangements will be made such that both private as well as public petroleum-based vehicles will be converted to EVs.

Hence, such a growing number of EVs demand more electrical energy. To fulfill the demand, the present scenario of the Grid supply is not enough. Besides, while visiting the Charging station, it is confirmed that the present status of profit for the charging station is very low. If the same topology of energy supply from only the Main Grid exists, then the status of the private charging station in the future will be in danger. Hence, there is a need for a Micro Grid and Distributed Energy sources interaction with the Main Power Grid. The combination of DER, ESS, along with the Main Power Grid will not only boost the environment of the charging station but will also help to manage the energy supply and shift the load from the Main Power Grid.

Now for the interaction of multiple players, i.e. MPG, ESS, EVCS, DER, there needs an appropriate scheduling methodology. Scheduling of charging station is very important because it concerns energy management. The concept of Smart Grid has already flourished, so charging stations cannot be left as an exception. To make the charging station smart, Machine learning based algorithms such as Reinforcement Learning technique could be employed and perform the scheduling of the charging station. This algorithm not only helps in scheduling but also help to obtain the optimal scenario at which the revenue and profit could be maximized. Moreover, instead of using model-based data, the real-world data of energy generation and vehicle count can bring more optimized scheduling of the charging station. In terms of DER, the solar plants have increased in number, so the solar plant could be utilized in the microgrid model to facilitate scheduling of charging stations.

Reinforcement learning is of a variety of types, but among those, the DDQN based algorithm is a popular one. It has lots of advantages over other forms of reinforcement learning. DDQN algorithm is an advanced form of DQN algorithm that has addressed the issue of overestimation bias. Unlike other forms of learning, it uses Q-learning methods that high-dimensional state spaces could be handled easily. It is a sample efficient learning tool compared to policy gradient RL technique. Besides, it is not so complex to understand compared to actor critic methods. Also, this algorithm has a strong baseline for other extended RL based algorithms. The practical application of DDQN is seen in robotics, autonomous driving, and many other real-world applications.

1.1.2 Machine Learning

Artificial Intelligence's one of the most important branches is machine learning. ML mainly concerns the development of algorithms or models that operate without explicit instructions. Moreover, such methodology is data driven. There are several types of ML, these are:

Supervised Learning

It is a form of ML in which labelled data is provided to train algorithms and predict patterns or outputs. In this form of learning, prior knowledge is provided.

Unsupervised Learning

It is a form of ML technique that operates on unlabelled data and helps to discover hidden patterns and features. In this form of ML technique, no prior knowledge is provided.

1.1.3. Reinforcement Learning

Reinforcement learning (RL) is a ML technique where the agent interacts with the environment to make necessary decisions and maximize the reward. In case of supervised learning, the entire model is trained with some defined data, but RL uses features like trial and error, where the agent's actions influence future states and rewards.[8]

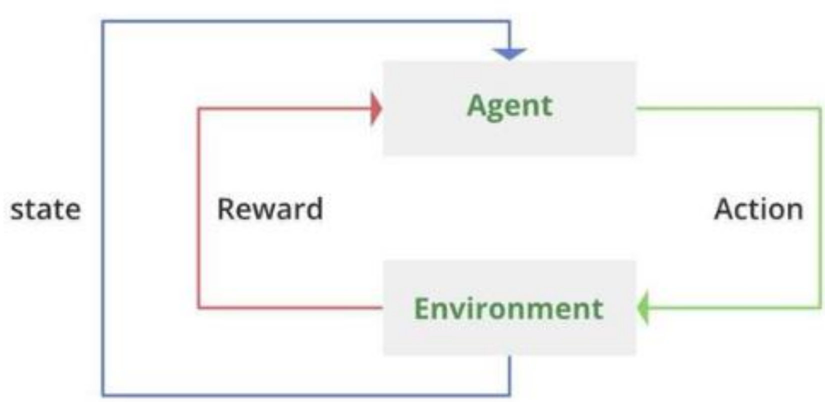

Figure 1: Reinforcement Learning Module

Key elements of RL are:

- Agent: it helps in making necessary decisions
- Environment: it is the system that interacts with the agent.
- State (s): This represents the current or present situation.
- Action (a): it is the decision which is taken by the agent.
- Reward (r): This is the Feedback from the environment evaluating the quality of the action. In simple words it is the profit obtained between the difference of revenue and the cost.
- Value Function(V(s) or Q (s, a)): It helps to estimate the expected cumulative reward.

Some of the Advantages of Reinforcement learning based algorithms for scheduling of charging stations are:

- Decision making is dynamic as Reinforcement learning helps to make Dynamic Decision regarding the energy demand, electricity prices, EV rates etc. Besides, it does not require preset rules as it works under real work-based feedback.
- Optimization of the cost function, RL helps to increase the Revenue by reducing the peak load cost and scheduling the charging station during off peak period.
- RL also helps to carry out the balance of load in Grid with the presence of Distributed energy sources and Energy storage system.
- Uncertainties can be easily handled in comparison to conventional scheduling strategies.
- Multiple objective optimization such as, revenue maximization, cost minimization, ensure customer satisfactions is guaranteed.
- Large, scaled network can be easily adapted using RL algorithms.

There are variety of RL algorithms. Some are listed as:

- Model-Free vs. Model-Based RL
- Value-Based Methods
- Policy-Based Methods
- Actor-Critic Methods
- Deep Reinforcement Learning

In this project Deep Reinforcement learning is utilized.

Deep Q-Learning (DQN)

DQN algorithm is a type of Q learning algorithm, this algorithm utilizes neural network instead of table. The neural network function Q (s, a) is utilized to carry out the approximation. High dimensional state spaces and continuous state spaces evaluation can be easily carried out with this form of algorithm. The major features of DQN algorithm are:

- Experience Replay: It helps to store experiences of the agent.
- Target Network: It is a copy of Q network which is updated so that training is stabilized.
- It uses the same network to carry out function of action selection and action evaluation.

  Loss function: The mathematical equation of loss function is:

$$\boldsymbol{L}(\boldsymbol{\Theta}) = \boldsymbol{E}_{\{s,a,r,s'\}}[\boldsymbol{y} - \boldsymbol{Q}(\boldsymbol{s}, \boldsymbol{a}: \boldsymbol{\Theta}^{-})]^{2}$$

  Where,

- L(ϴ): This is the variable of loss function to predict the wellness of Q values
- Q (s, a: ϴ): This the function of predicted Q values for taking action “a” in state “s” with current parameter “ϴ”
- y: It is the Target Q value; this is calculated as:

$$\boldsymbol{y} = \boldsymbol{r} + \boldsymbol{\gamma}_{\boldsymbol{max}}\, \boldsymbol{Q}(\boldsymbol{s}', \boldsymbol{a}': \boldsymbol{\Theta}^{-})$$

  Where,

- r: It is the reward received after acting “a” in state “s”
- s’: It is the next state.
- y: It is the discount factor that helps to determine the future rewards.
- $\boldsymbol{Q}(\boldsymbol{s}', \boldsymbol{a}': \boldsymbol{\Theta}^{-})$: It is the target Q value using parameter $\boldsymbol{\Theta}^{-}$, from target network.

Double Deep Q learning (DDQN):

This is advanced form of DQN algorithm which addresses the issue of overestimation bias while predicting Q values. In this algorithm Selection and Evaluation process are carried separately on different Q-networks.

- Action Selection: In this process the agent conducts selection of best action as per Q network knowledge and the policy learned by agent so far. The current Q-network $Q(s', a': \Theta)$with parameters " Θ ", is used to select the best action in the next state. This network is updated continuously through training, making it the active Q-network. For a given next state, the agent chooses the action "a" that maximizes the Q-value predicted by the current Q-network.

- Action Evaluation: In this step, DDQN evaluates the value of the action "a" using a separate target Q-network, $Q_{target}(s', a^*: \Theta^-)$ with parameter Θ^-, which is a copy of Q network but a delayed copy. This difference helps to address the over fitting that arises due to same network in traditional DQN learning. Here, the target network is a fixed copy of the Q-network, that is updated less. After selecting action "a*", the agent uses the target network to evaluate the Q-value for taking this action in state "s' ". The target value in DDQN for a transition is:

$$y = r + \gamma Q_{target}(s', a^*: \Theta^-)$$

  Where,

- r: It is the immediate reward obtained after taking the action in the given state
- γ: It is the discount factor
- a*: It is the action selected by the Q network for state s'.

The loss function is evaluated in similar manner as that of DQN learning.

1.1.4 Main Power Grid

It is the network of large electrical platform. Other names given to MPG are power grid, distribution grid, electrical grid or the national grid. This form of network includes the energy supplier or the utility that provides electrical energy to the consumer of every sector.MPG can interconnect with renewable sources, ESS etc.

1.1.5 Distributed Energy Resources

These are the systems which are installed or operated at the point where electrical energy is used. Also, there are charging stations with DER installed. Such DER are operated at very small scale and are generally connected with the MPG. It helps to increase the reliability of energy. There are multiple forms of DER i.e. solar energy, wind energy, fuel cells etc.

Some of the major characteristics of DERs are:

- Decentralization: It concerns on generation of energy at the area of consumption.
- Flexible: DERs are operative both in unison with MPG or can operate in isolated form.
- Scalability: These ranges from small to large scale.
- Integration of Renewable resources: It uses renewable based energies as solar, wind, geothermal etc.

1.1.6 Energy Storage System

This is a form of large batteries called Battery Storage System to store energy and use at the time of need. Moreover, it can operate along with MPG or DER and enhance the network efficiency. The value of ESS ranges from a few kWh to several MWh.

1.1.7 Electric Vehicle Charging Stations

EVCS are the infrastructure that is responsible to power the EV batteries. The charging station has multiple charging point to charge the EV battery. The charging port can be either of AC or DC form. The general features of the charging stations are the following:

- Location and Power Availability
- Accessibility
- DER Integration
- Charging Time
- Billing and Control
- Maintenance and Reliability

1.2 Problem Statement

The problems associated are:

- Increasing demand of EV charging station cannot only be fulfilled by Main Power Grid.
- Necessity of Microgrid along with Distributed energy resources, Energy storage system.
- Multiple players interaction leads to optimal scheduling of charging station
- Scheduling guaranteed with ML-based algorithms, i.e., Reinforcement learning algorithm
- RL needs extensive study and for training data specific resources needed.
- Context to Nepal, Data of solar and EV charging station cannot be achieved in systematic way

1.3 Objectives

The main objective is to develop EV charging station model and perform optimal energy scheduling for microgrids using Reinforcement learning.

- To achieve the objective, models of EV charging stations are framed and DDQN algorithm, a type of RL algorithm, is utilized to visualize the optimal model in response to revenue.
- Also forecast the scenario of charging station by studying vehicle import data from Department of Custom, government policies etc.

1.4 Scopes

The scopes for this project are:

- Use Real world data of Real charging station and real solar plant
- Develop a Scalable Microgrid Model
- Apply advanced DDQN type of reinforcement learning (RL) algorithms and identify optimal energy scheduling decisions
- Comparison of microgrid scenarios: stand-alone, Microgrid based and solar based in terms of Revenue.
- Perform forecast of the charging station scenario depending on present, past scenario, government policies etc.

1.5 Limitations

- The Study mainly concerns Solar energy as DER and no other Renewable.
- Solar plant and EV charging station of a particular location is studied. A variety of plants and stations are not studied.
- Customer Behaviour is considered simple, so preference-wise variability is not considered.
- DDQN is used in the research but there are more advanced algorithms, Proximal Policy Optimization, Actor Critic, etc.

1.6 Thesis Organization

The Research is divided into six chapters.

- Chapter 1: It gives a brief introduction to the overall Research. The objectives, Problem statement, limitations, as well as scopes are well briefed.
- Chapter 2: It has explored the necessary literature, theory for the research to proceed. It presents results and findings of variety of research article.
- Chapter 3: It presents the Data Acquisition and Set up for the methodology to be employed
- Chapter 4: It portrays the methodology, flowchart, and algorithm that are employed to develop the model of charging station. Also, the software, tools, and necessary setup for the coding is explained in detail in this section.
- Chapter 5: It discusses the results of the models of charging station in terms of Revenue, State of charge, etc.
- Chapter 6: It presents a 5-year future trend of the charging station in terms of cost, kWh consumption, Revenue, and number of EVs.
- Chapter 7: It concludes the research

Then the thesis ends with the reference and necessary appendix.

CHAPTER TWO: LITERATURS REVIEW

The literature review section helps to give information about the concepts accomplished by the research article that are important for to attain target objective. After going through variety of journal papers and research article concerned with the topic the knowledge horizon is broaden. Multiple research articles are studied to get idea and model the system. Information about present EV status, Vehicle number import status is studied. Furthermore, Reinforcement learning is studied in detail and its types are also studied thoroughly. These are the research article which gave necessary insight.

[1]: This article introduced a novel micro-grid model that incorporated energy generated by DER, possibility of Energy storage system, a load entity in the form of Electric Vehicle Charging station and Main Power Grid as the electricity trader. The system modelled in the article have proven to successfully shift the load from MPG, considering customer satisfactions in a profitable way. Furthermore, instead of using model-based data the system designed in this paper have used real time data. In order to determine the optimal scheduler, Reinforcement Learning have been utilized, so that the profit is maximized. Furthermore, two different versions of DER model are designed in this paper. One model is used to charge the energy storage system using energy obtained from DER i.e. wind energy and sold to the Main Power Grid. Then, in second model the energy from DER is first utilized to charge the EVs present at charging station and only utilized to charge the energy storage system and then sold to MPG. While observing the results, the second model showed more robustness to electricity prices and electricity generation from wind.

[2], the review paper gave insight about different types of Reinforcement Learning algorithm that has been utilized in EV charging management. Among variety of algorithms, concept of deep Q learning algorithm was thoroughly studied. Besides, the algorithms explained in the paper were focused on maximizing the cumulative rewards. Hence, how reward maximization was carried was also explained. Moreover, to give detail information about variety algorithms, comparative analysis between Reinforcement Learning was carried.

[3]: This article focuses on two types of optimal scheduling schemes these are locally optimal scheduling schemes and globally optimal scheduling schemes for charging and discharging of EV. In case of global optimization schemes the main focus was to minimize the cost of all the EVs that were charged and discharged during a day. As a result, this scheme was able to provide a globally minimum total cost. But the scheme was seen to be impractical as it required future load & their arrival time and the charging periods of future arriving EVs. Because of this locally optimal scheduling scenario was proposed in the article, which was a practical scheduling scheme. The prime motive of local optimization scheme was to minimize the total cost of EVs in the current ongoing EV data in the form local group. Upon performing simulation, local optimization is found to achieve close performance than the other scheme.

[4]: This article proposes an RL approach in optimizing the vehicle charging station in terms of scheduling and pricing and maximizing the objective. It has utilized model free data so no reliability on stochastic models. A feature based linear function is proposed as the approximator of the value function. Real world data was utilized and proposed algorithm in the research showed 138.5% higher profit than the target algorithm.

[5]: This website is a pdf document of Nepal's electric mobility. This pdf give insight about socio-economic, transportation and current status of EV adoption in Nepal. Besides, information about vehicle registration, transport activity projections, early initiative on EV mobility, current status of EV fleet, its challenges etc are briefly described. Moreover, government-based policies and target son EV mobility is also briefly stated.

[6]: The government-based policies regarding downpayment status of vehicle is studied from this website. While studying, it is found that both EV as well as Petrol/Diesel based vehicle requires 40 % down payment.

[7]: The website was utilized to study the vehicle import number for the past 12 years. Besides, this website gave information about which EV has considerable rise from beginning and which has reached saturation. While observing the Three-wheeler tempo was seen the pioneer of EV in Nepal. Similarly, this website provided information about the types of EV being import. Also, it gave information about the current status of petrol and diesel-based vehicle in number. Similarly, information about the countries from where different EV are imported is also provided in the website.

[8]: This website is about Deep Learning. It provided a wide range of understanding in context to DQN, DDQN and variety of other RL algorithms. Moreover, the algorithm of deep learning is explained in detail. With the help of Flowchart, the algorithm is explained. Besides, all steps of the algorithm are explained in detail. Hence, through this website detail knowledge about the deep learning algorithm was gained.

These literatures have highlighted how reinforcement learning could be employed in optimizing the Charging station management. Similarly, it provides information about several RL based algorithms that could be utilized in optimization. Similarly, DDQN based algorithm is employed in the paper [7] while studying this article insights regarding the model of charging station is gained. Moreover, information about energy scheduling, pricing strategies is also gained. Besides, how RL based algorithms are more prominent that other traditional based algorithms for efficient management of charging station in known.

CHAPTER THREE: DATA ACQUISITATION, SET UP

3.1 Data Acquisition

- Solar plant of Maiti Nepal is considered, its capacity is 40KW and this capacity is lower to meet the energy demand at charging station. Hence, the capacity of solar plant is scaled to 200KW.
- The data of EV vehicle number, EV revenue, KWh consumption of Bhaktapur charging station is considered. This charging station capacity is 200KW.
- As, the data obtained from solar plant is in cumulative form and the data in this form cannot be analysed. Hence, this data is converted to actual form.
- The data for EV was obtained in the form of bill which showed SOC% initial, SOC% final, Revenue per SOC percent. Upon utilizing these data and performing simple unitary method the value of KWh consumption by each vehicle, Revenue for each vehicle is evaluated.
- Similarly, the data of both solar generation and EV data are not in same time instant so, these data are aggregated at same time instant. During training, data at same time instant is necessary.

Table 1: Data Description

Aspect	**Description**
Solar Plant Capacity	- Original: 40 kW (Maiti Nepal) - Scaled up to: 200 kW (to meet charging station demand)
Charging Station Capacity	200 kW (Bhaktapur station data used)
EV Data Source	Bhaktapur charging station (vehicle count, revenue, kWh consumption)
Solar Data Processing	-Original data in cumulative form - Converted to actual generation values for analysis
EV Data Processing	- Extracted from billing records (initial SOC%, final SOC%, revenue per SOC%)

Aspect	Description
	- Calculated kWh consumption & revenue per vehicle using unitary method
Data Synchronization	- Solar and EV data were at different time intervals - Aggregated to same time instants for training/analysis

3.2 Set Up

3.2.1 Hyperparameters

Those parameters that are responsible for controlling the performance of DDQN agent of the algorithm are hyperparameters. These are the parameters which are not learned, rather these are the setup for initiating the training procedure. Besides, these parameters are important to ensure the optimality of the model. To obtain optimal hyperparameters, fine-tuning action is carried out. The range at which these hyper parameters are fine-tuned is listed below.

- Gamma (Discount Factor): 0.9 to 0.999 If too high, long-term reward focused. If too low, short-term reward focused
- epsilon_min (Minimum Exploration rate): 0.01 to 0.1 Minimum exploration, if too high, unnecessary exploration, and if too low, early stop of exploration
- epsilon_decay (Exploration rate decay): 0.99 to 0.9999, balance between exploration and exploitation. If slower, more exploration occurs. If faster decay than fast shift to exploitation.
- Batch_size (Sample of experience): 32 to 256, Large size is more stable, but slow training. Small size, fast train, but noisy.
- Lr (Learning Rate):1e-5 to 1e-2(log scale), It affects the speed of convergence. If high, not stable training. If low, slower convergence
- Memory_size (Buffer Size):1000 to 10000, Large buffer, high memory but diversified experience. Small buffer, diversity is limited, which affects learning

To perform the Fine tuning an inbuilt library named Optuna is used. For fine tuning, 20 trials containing 5 episodes each is considered for each model. Then the necessary results of Revenue, SOC percent, Reward is obtained.

The table shows the list of hyper parameters obtained after performing Fine Tuning.

Table 2: Fine Tuned Hyperparameter

S.N.	Hyper parameters	Values
1	Gamma	0.9800490032535142
2	epsilon_min	0.03819949004051562
3	epsilon_decay	0.9912020003120168
4	Batch_size	229
5	Lr (Learning Rate)	0.00030701322052257673
6	Memory_size	4752

3.3 Neural Network

3.3.1. Neural Network Architecture for DQN class

The neural Network structure of DQN class has an input layer, Two hidden layer and a output layer. Here, the hidden layers neuron is activated by Rectified Linear Activation Function i.e. ReLU. ReLU activation function output actual input for positive input and zero for negative input. The architecture of the Neural Network is shown below.

- Input Layer: state as input. Vector size of state is 4
- Layer 1: Fully connected layer with 64 neurons, ReLU activation
- Hidden Layer 2: Fully connected layer with 64 neurons, ReLU activation.
- Output Layer: Fully connected layer with action size (3) neuron

Each neuron represents the Q-value for a specific action. Training Process: Neural network is trained to minimize the Mean Squared Error (MSE) carried between the predicted Q-values and the target Q-values.

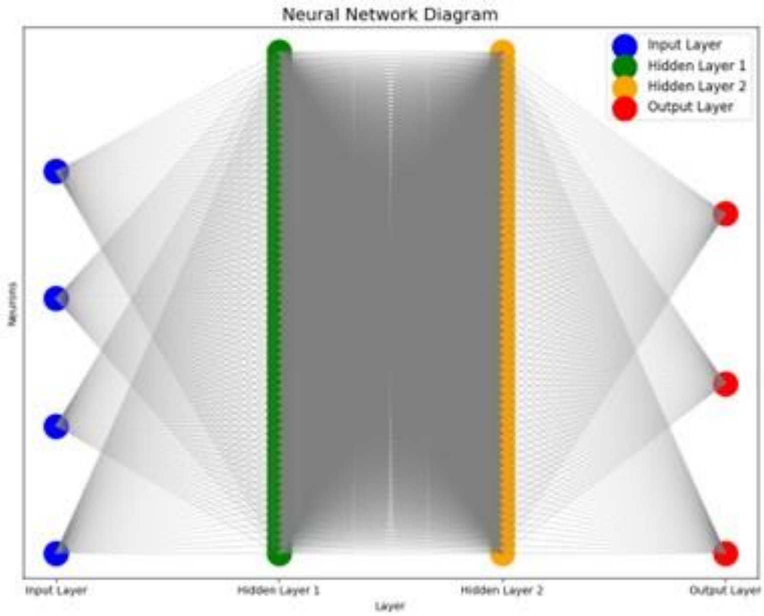

Figure 2:Neural Network Architecture of DDQN class

Target Q-value and loss function is computed using the Bellman equation.

$$Target = r + \gamma\, a'_{max}\ \mathrm{Q}\ (s', a') \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$Loss = 1/N \sum_{i=1}^{N} (\mathrm{Q}(\mathrm{s_i}, \mathrm{a_i}) - Target_i)^2 \ldots\ldots\ldots(2)$$

Where,

- r = reward
- γ = gamma (discount factor)
- a'_{max} = maximum Q value overall in next state
- Q (s', a') = value function
- N = no of batch size

Hyperparameters such as epsilon, learning rate, memory size are not explicitly defined in the mathematical relation. However, Epsilon value is concerned with the type of action taken i.e. exploration, it also influences the data stored in replay memory i.e. the buffer. Similarly, the learning rate is the optimizer concerns with the update in the weight.

3.3.2. Neural Network Architecture for DDQN class

As, DDQN, is the improvised version of DQN. So, its neural network architecture is same as that of DQN. DDQN helps to address the issue of overfitting as seen in DQN with the help of two separate network.

- Online Network: Updated during training.
- Target Network: Copy of the online network, updated less frequently (few episodes).

In order to compute the value of Target Q value and loss following equation is utilized:

$$Target = r + \gamma\, Q_{target}(s', arga'_{max}\ \mathrm{Q_{online}}\ (s', a'))\ldots\ldots\ldots\ldots\ldots(3)$$

$$Loss = 1/N \sum_{i=1}^{N} (\mathrm{Q_{online}}(s_i, a_i) - Target_i)^2\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

Where,

Q_{target} = is the target Q network, $\mathrm{Q_{online}}$ = is the online Q network

3.2.2 Code Set Up

The necessary parameters required during coding are enlisted in the table below. This table consist of the list of hyperparameters, other parameters with its specifications. It includes the capacity of ESS, maximum charging and discharging capacity of ESS and energy price data obtained from NEA. Similarly, it consists of the range of hyper parameters which is considered during the coding. It also consists of the excel file for the data of EV and solar to be utilized during coding.

Table 3: Parameters defined in code

Parameter	Value	Description
ESS capacity	150KW	Total capacity of ESS
Max rate	100KW	Charging/discharging rate
Energy Price	{17<h<23:7,23<h or h<5:3.7, else: 5.5}	Energy price based on time of day
State Size	4	No of state variables, State variables as: battery soc, solar energy, energy price, EV KWh demand
Action size	2 or 3	Depending on the model
Episodes	1500 or 2500	Number of episodes for training
Discount factor gamma	0.9 to 0.99	Discount for rewards
Epsilon min	0.01 to 0.1	Min exploration rate
Epsilon decay	0.99 to 0.9999	Decay rate of exploration
Batch-size	32 to 256	Sample of experience used during training
Memory size	1000 to 10000	Maximum size of replay memory
EV KW usage	Data from excel	Energy demand of EVs
EV revenue	Data from excel	Revenue from EVs
Solar energy	Data from excel	Solar energy at each time step
timesteps	30min	time at which data are taken

CHAPTER FOUR: METHODOLOGY AND FLOWCHART

4.1 System Block Diagram

The entire system for the model is explained by the figure, which consist of a block representation of the model. In this diagram, DER, ESS, EVCS and MPG connection is shown, and their interaction is also portrayed. As, the entire model operates in three different ways, but general operation for the charging station is shown by this diagram.

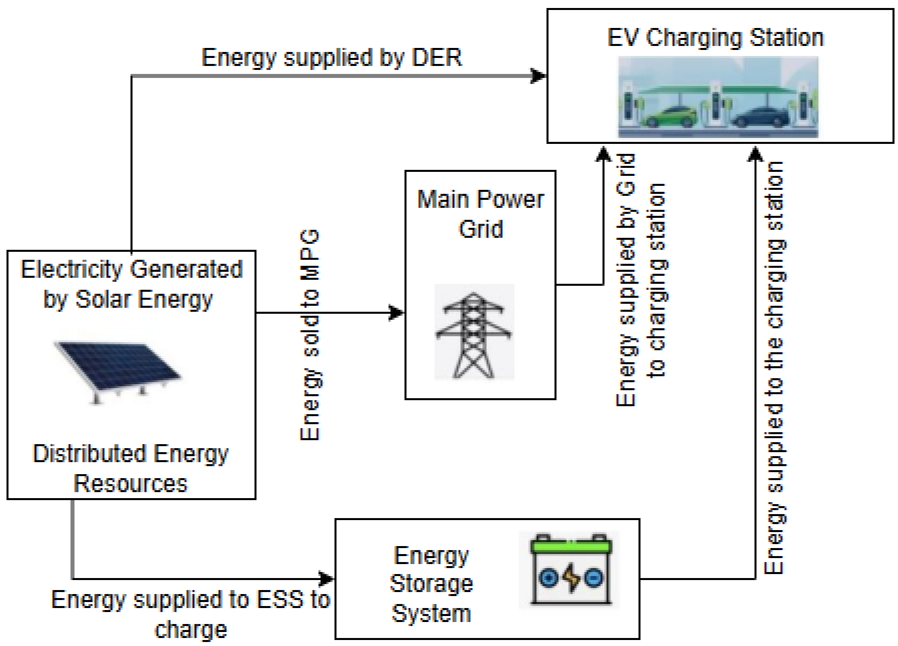

Figure 3: System block diagram

4.1.1 Standalone charging station:

This is the charging station scheduling model where, entire demand of the charging station is only met by Main Power Grid. There are no other sources except MPG to provide energy to the EV. Hence, this scheduling model is termed as standalone charging station model.

4.1.2 Micro Grid based Modelling

Scheme 1

This is the charging station scheduling model where, the demand of the charging station is met by collective effort of DER, MPG and the ESS [1]. But, its mode of operation follows two types of action explained below:

Action 1: In this action first, the ESS provides energy to the charging station and then respectively the DER sells energy to the Main Power Grid.

- ❖ ESS Powers the charging station: The ESS supplies energy to the Charging Station.
- ❖ DER Sells Energy to the MPG: The electricity generated from Solar plant is sold to the Main Grid.
- ❖ Grid Supplies Additional Energy: Whenever, the energy is deficit in ESS then MPG provides the necessary energy to the charging station.

Action 2: In this action first, the solar plant energy i.e. DER is utilized to charge the ESS then charging station receives necessary energy from the MPG.

- ❖ DER Powers the charging station: Electricity obtained from solar plant is used to charge the Energy storage system.
- ❖ Energy from DER is excess then traded to Grid: If the electricity generated from DER is available after charging ESS, then is traded to MPG.
- ❖ Charging station receives energy from Main Grid: All the energy requirement of the charging station is fulfilled by MPG.

Scheme 2

This is the charging station scheduling model where, the demand of the charging station is met solely by DER, but there occur two actions explained below:

Action 1: When there is excess energy (Solar Generated energy > EV demand):

If there is extra energy even after charging EV is first utilized to charge the ESS, then remaining energy is utilized to trade with MPG.

Action 2: When there is a deficit energy (Solar Generated energy < EV demand):

First the ESS supplies necessary energy to the EV charging station and still if demand is not met in that case MPG provides the remaining energy to the charging station.

The tabular representation of above methodology is shown below, this table explains the methodology of the entire system.

Table 4: Tabular representation of the system methodology

Model	Energy Sources	Action	Power Flow
Standalone	MPG only	-	MPG → EV Station
Microgrid (Scheme 1)	DER + ESS + MPG	**Action 1**	1. ESS → EV Station 2. DER → MPG (sell) 3. If ESS deficit → MPG → EV Station
		Action 2	1. DER → ESS (charge) 2. Excess DER → MPG (sell) 3. MPG → EV Station
Microgrid (Scheme 2)	DER + ESS + MPG (MPG and ESS as backup)	**Action1 (Excess Solar)**	1. DER → EV Station 2. Excess DER → ESS (charge) 3. Remaining DER → MPG (sell)
		Action2 (Deficit Solar)	1. ESS → EV Station 2. If still deficit → MPG → EV Station

4.2 Mathematical Model

The mathematical model for each scheduling scenario is provided below. This model presents the optimization function for each model with their respective constraints considered as per the methodology described earlier.

4.2.1 Standalone model

The Optimization function for the system is:

$$\boldsymbol{Min\ Z} = \sum_{t=1}^{T} P_{buy} * e_b^t \text{.............(5)}$$

Where,

P_{buy} is the price at which energy is bought from Main Grid.

e_b^t is the energy supplied by the grid to the charging station at t instant of time.

Similarly, the constraints for above objective function are,

- Energy Demanded by charging station: Energy provided by MPG should be sufficient to meet demand of charging station.

$$e_{EVCS}^t = e_b^t$$

e_{EVCS}^t is the energy demanded by charging station for t instant of time.

4.2.2 Microgrid Model

Scheme 1

For this scenario, there are two potential actions, as described in methodology. As this model consist of microgrid ESS, DER, and the grid (MPG), hence, there energy flow within these players.

Action 0: (First ESS powers the charging station then DER sells to the MPG)

Objective Function: The goal is to maximize the net revenue by selling DER energy to the grid and minimizing the energy bought from the grid.

$$\boldsymbol{Max\ Z} = \sum_{t=1}^{T} P_{sell} * e_s^t - P_{buy} * e_b^t \text{............(6)}$$

P_{sell}, is the price at which energy is sold to the grid.

P_{buy} , is the price at which energy is bought from the grid.

e_s^t, is the energy sold to the grid by the DER.

e_b^t, is the energy supplied by the grid to the EVCS.

The constraints are:

- Energy Demanded by charging station: The energy storage system and Grid supplies necessary energy demanded by charging station.

$$e_c^t + e_b^t = e_{EVCS}^t$$

e_c^t, is the energy supplied by ESS and e_b^t is the energy supplied by the grid.

- Energy Balance for energy storage system: ESS provides energy based on its charging limits.

$$0 \leq e_c^t \leq \min(s_t, r_{max})$$

s_t is the state of charge of ESS and r_{max} is the maximum charging/discharging limit.

- Energy trade from Solar plant: The generated energy from solar plant is traded to MPG.

$$e_s^t = e_g^t$$

e_g^t, is the energy generated by the solar plant.

- If energy supply not enough: When ESS is not able to meet the demand, then MPG supplies the remaining.

$$e_b^t = e_{EVCS}^t - e_c^t$$

Action 1: (First ESS is charged from DER energy then charging station receives power from MPG)

Objective Function: The goal is to optimize the operation of the solar plant for charging the energy storage system and minimizing the purchase cost from grid.

$$\boldsymbol{Max\ Z} = \sum_{t=1}^{T} P_{sell} * e_s^t - P_{buy} * e_b^t \text{............(7)}$$

e_s^t ,energy sold by the DER to the MPG.

e_b^t, is the energy supplied by the grid to charging station

The constraints are:

- Energy Balance for energy storage system: The solar energy charges the ESS to its maximum limit.

$$e_r^t \leq r_{max}$$

- Energy trading by DER: The excess energy generated from solar plant is used to charge ESS then sold to Grid.

$$e_s^t = e_g^t - e_r^t$$

- Grid Supplies to charging station: Overall energy demand of charging station provided by MPG.

$$e_b^t = e_{EVCS}^t$$

Scheme 2

In the DER-Based Modelling scenario, the EVs are directly charged from the solar energy generated by the DER. There are two cases to consider: excess energy and energy deficit.

Objective Function: The goal is to optimize the energy management based on the generation from solar plant and minimize the reliance towards MPG and further maximizing the energy stored in ESS.

$$\boldsymbol{Max\ Z} = \sum_{t=1}^{T} P_{sell} * e_s^t - P_{buy} * e_b^t \text{......................}(8)$$

e_s^t, is the energy sold to the grid by the DER.

e_b^t, is the energy supplied by the grid to the EVCS.

The Constraints are:

Action 0: When the energy generated by the Solar plant exceeds the EV demand at charging station, the extra energy is used to charge the ESS then further traded to MPG.

- If excess energy is used to charge the ESS:

$$a_e^t = 1$$

,$e_r^t = \min\,(e_g^t, r_{max})$, e_r^t is the energy used to charge the ESS

- If excess energy is sold directly to the grid.

$$a_e^t = 0$$

, $e_s^t = \mathrm{e}_{\mathrm{g}}^{\mathrm{t}}$

Action 1: When the energy generated by the Solar plant is less than the EV demand at charging station, First ESS supplies the energy then if still not met MPG supplies the energy.

- If ESS provides necessary energy to charging station:

$$a_d^t = 1$$

$e_d^t = \min\,(e_g^t, r_{max})$, e_d^t is the energy discharged from the ESS.

- If the ESS cannot supply all the demand, the grid supplies the remainder: $e_b^t = e_{EVCS}^t - e_d^t$
- If no energy is discharged from the ESS and the grid fully supplies the EV demand.

$$a_d^t = 0$$
$$e_b^t = e_{EVCS}^t$$

4.3 Flowchart

4.3.1 Flow chart of the system

This flow diagram explains the entire system's flow.

- The process begins with start
- Energy price, energy generation from solar and EV data are provided.
- The system constraints are initialized
- Based on the options the three models are operated individually

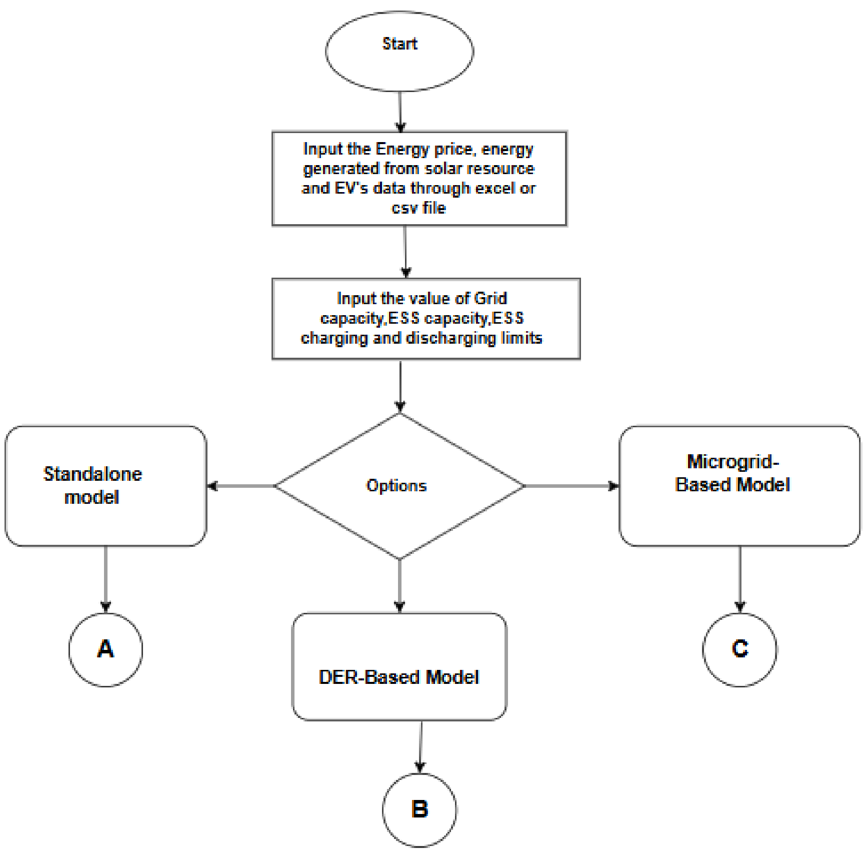

Figure 4: Flowchart of the entire system

4.3.2 Flow chart of each methodology

For A methodology: First checks the EV demand and MPG capacity, if demand is less energy supplied from grid. if not proceed to other option

- Then energy is supplied to EVs
- Otherwise, next option is chosen

For B methodology: if Action 1: proceed with DER and ESS for ''NO" DER charge the ESS and if any excess energy MPG trades it. For" YES" ESS charges EV and DER sells energy to the MPG

- Check demand is completely met
- Proceed to other option

For C methodology: All energy demand met by solar generation

- if generation excess, trade to MPG and charge ESS and then supply the charging station through ESS.
- if generation deficit supplies all demand by MPG

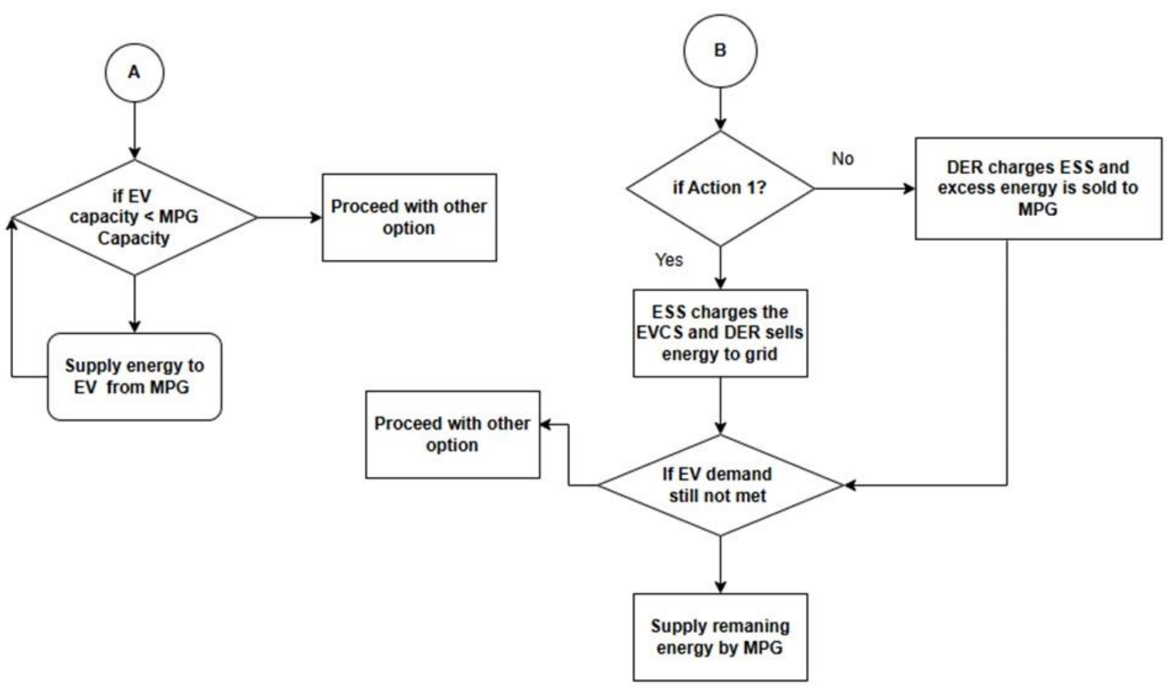

Figure 5: Flowchart of methodology A and B

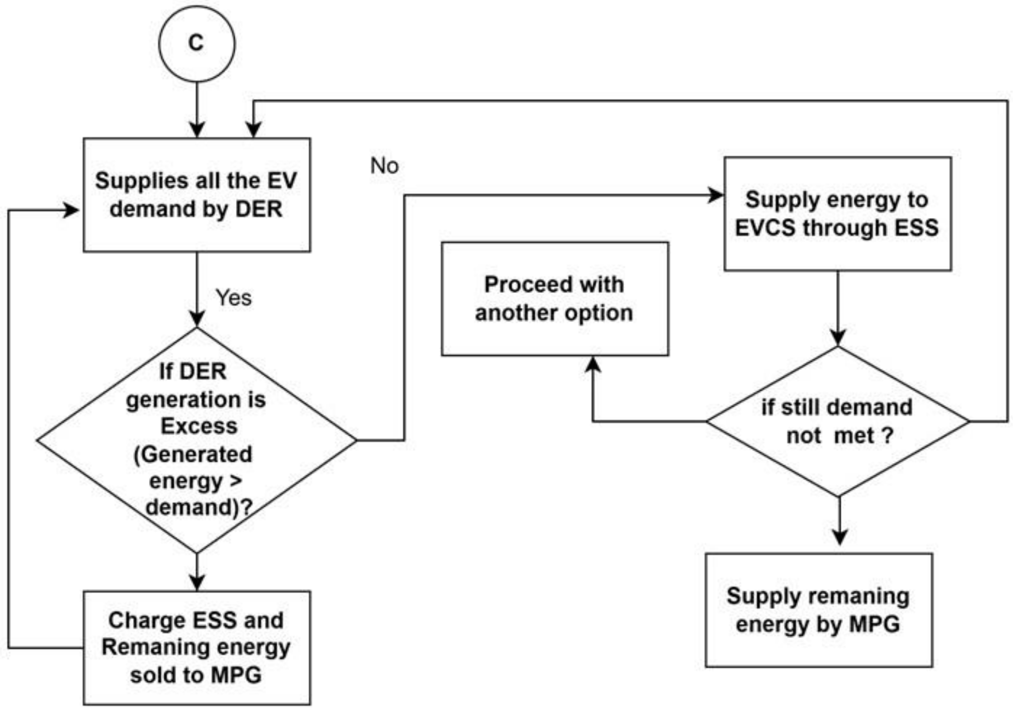

Figure 6: Flowchart of methodology C

4.3.3 Flowchart for DDQN algorithm

Initialization Phase

- Start process
- Initialization of Episode, runs from 1 to M episodes
- Define initial state at beginning of each episode [8]

Action selection and Environment Interaction

- Main neural network considers St as input and then predict the value function Q (S_t, a)
- Exploitation and Exploration proceeds based on epsilon greedy policy [8]
- Action is executed in the environment
- The transition is stored in Experience replay memory or the buffer. (S_t, a_t, r_t, S_{t+1})

Experience Replay and Target calculation

- The mini batch sampled from the replay memory

- S_{t+1}, i.e. next state is passed through target network
- the target value is computed

Loss computation and update

- The loss function is computed
- main network is updated
- Also, the target network is updated [8]

repeat till completes

- Follow the same process for next step
- continues till the end of episode
- proceed with the next episode

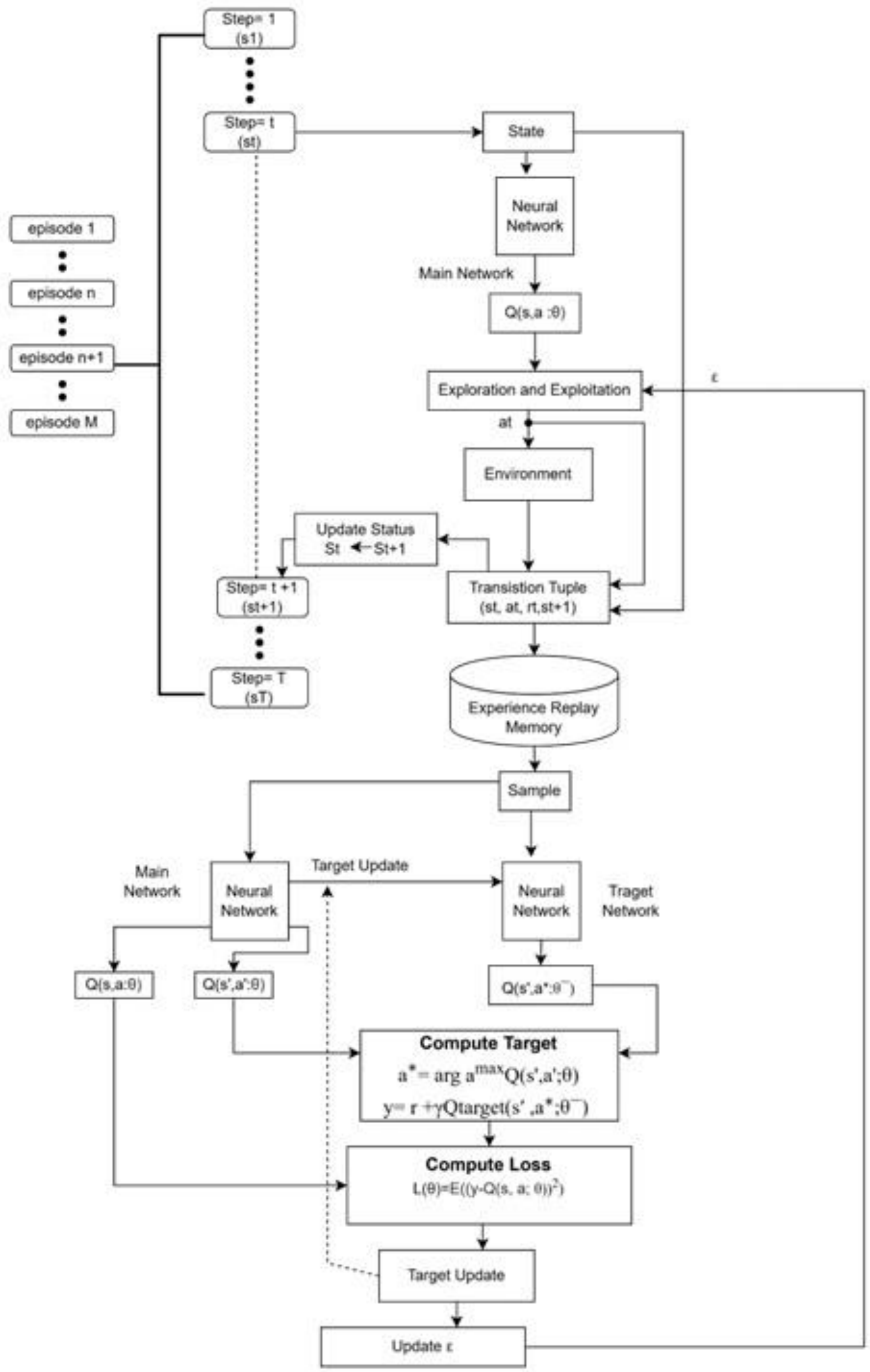

Figure 7:Flowchart of DDQN algorithm

4.4 Tools and Software

4.4.1 Python Programming language

- Python programming language is utilized to write the code and create the microgrid model.
- Python is one of the most user-friendly programming languages for its features like readability, versatility etc.
- It is widely utilized in AI, data science, ML, scientific computing and many more.
- Moreover, it has wide range of libraries so that models and research could be carried in an efficient way.

4.4.2 Libraries used

- NumPy: Fundamental library utilized to perform numerical computing in python. It supports a wide range of arrays and matrices to operate in data structure with greater ease. Typically, this library is used in scientific computing, data analysis and in ML.
- Matplotlib: This is the library famous called as plotting library. it helps to create plots of bar chart, line graph, histograms etc. Typically, it is utilized in data visualization, data analysis and ML models for evaluation.
- Torch: This library is open-source deep learning library. It is typically applicable for Reinforcement learning algorithms, Research, prototyping in AI, CNNs, RNNs, NLP, Transformers etc.
- Pandas: This library is known for data analysis and manipulation. It is responsible to provide data frame for easier analysis. It also conducts data filtering, grouping, merging, time series analysis etc.
- Optuna: It is an automatic hyperparameter optimization library that is utilized in deep learning models in ML. It is concerned to fine tune the hyper parameters of RL algorithms.
- Tensor flow: it is also open-source deep learning library utilized in ML and AI application

4.4.3 Compiler

- Anaconda compiler is the free and open-source distribution of Python and R programming language. It comes with preinstalled libraries and necessary tools. This compiler is mostly used in data science, Deep learning, Scientific computing and big data analysis.
- Jupyter Notebook is web-based coding environment that allows the user to write, run and visualize the code in python and R programming language.
- Google Collab is the cloud based Jupyter Notebook environment which allows the user to write, run and visualize the code in python.

4.4.4 Setup

The Model is implemented using following specifications

- Python, PyTorch library
- Similarly, the Computation is carried using laptop PC with AMDRyzen7435HSprocessor running at 3.1 GHz, and 16GB RAM was used.
- Besides, the training sequence for 1 month worth of simulation took 120 minutes.

4.4.6 Overleaf

Overleaf is used to prepare the report, it is the online latex editor that helps in writing scientific papers, research, reports and other academic documents. It also supports mathematical equations, citations and easy formatting.

4.4.7 Microsoft Office

Microsoft Word and Microsoft PowerPoint is utilized to prepare the report and presentation slides. Microsoft office is a multipurpose software, Word helps to prepare the report as in both Doc as well as PDF format. Similarly, PowerPoint is utilized to prepare the slides as it has multiple functions to prepare the slides in an interactive way.

CHAPTER FIVE: RESULTS

The outcomes for the methodologies described (Standalone Charging Station, Microgrid-Based Modelling, and DER-Based Modelling) focus on optimizing energy management and increasing Revenue. Below are the specific outcomes for each methodology.

5.1 Economic Metrics

The revenue, profit, ESS State of Charge for each model is trained and obtained using the DDQN-based reinforcement learning algorithm.

5.1.1 Standalone Model

The trained plot of Revenue for standalone model is provided below. Revenue for this model is evaluated as per the real- world data. These data are used to evaluate revenue and profit, and these parameters are trained using DDQN algorithm for about 1500 episodes. From the bill information, the bill amount i.e. the cost is Rs.164720. After training model with DDQN algorithm the Revenue is obtained as: Rs. 160,000. Hence the Profit is Rs. -4000. This the standalone model that is existing, is not showing profit rather has Rs. -4000.

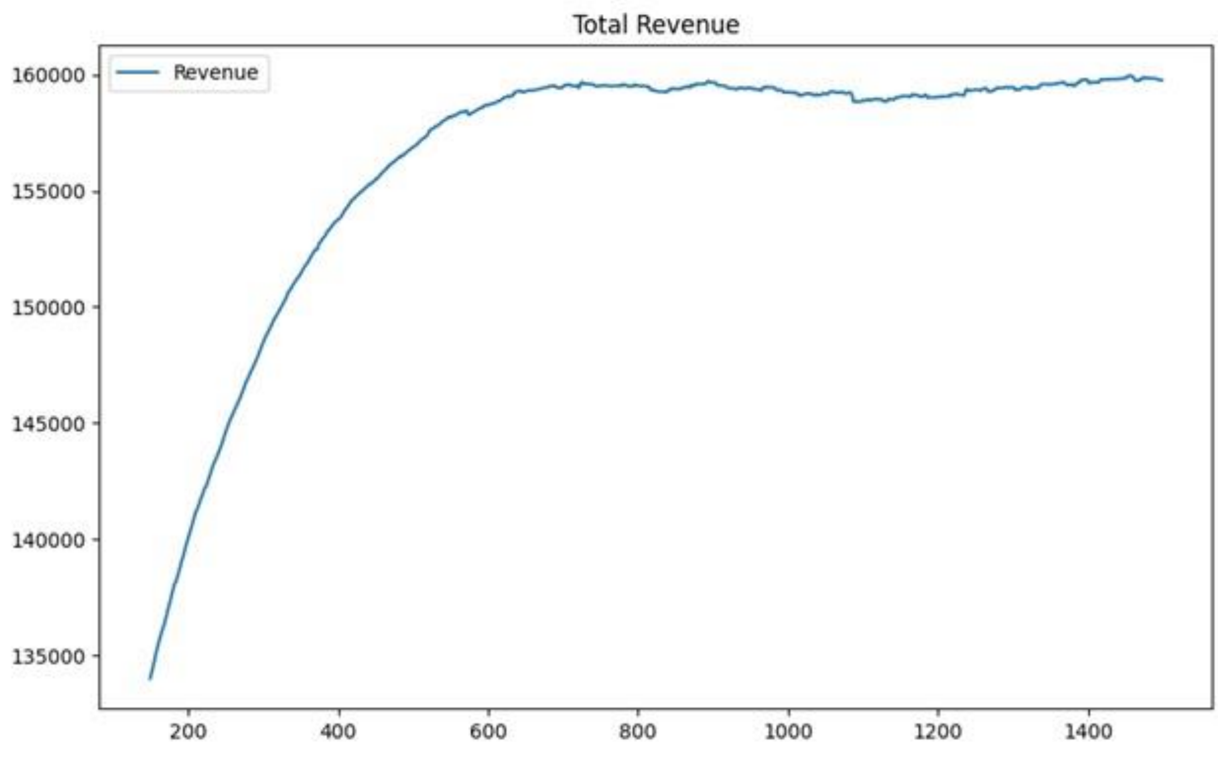

Figure 8: Revenue vs episode plot for standalone model

5.1.2 Micro Grid based model

The trained plot of Revenue for the Micro Grid based model is provided below. Revenue for this model are evaluated as per the real-world data. These data are used to evaluate revenue and profit, and these parameters are trained using DDQN algorithm for about 1500 episodes. From the bill information, the bill amount i.e. the cost is Rs. 164720. After training the model with DDQN algorithm, the Revenue for this model is obtained as: Rs. 178,500. Hence the Profit is Rs. 14500. Hence, the microgrid model shows a profit of about Rs. 14,500. The peaks in the graph at 1000 episode show how epsilon greedy policy has been utilized to perform exploration and exploitation.

Exploration: It is the process of trying the new actions that agent has not taken at all to study about the environment. It might not give the highest reward but leads to better understanding at the future. Example, if the agent has two options, one with reward 5 and other not certain. Here, the choice of uncertain action is exploration which might have reward 2 or 10 as agent does not know unless it tires.

Exploitation: This is the process by which agent utilizes the knowledge to find the best-known action to find the highest expected reward. This process concerns maximization of reward based on what is already known. Example, upon trying multiple action the agent knows that the action gives reward of 9 and the agent will stick with this reward utilizing knowledge of exploitation.

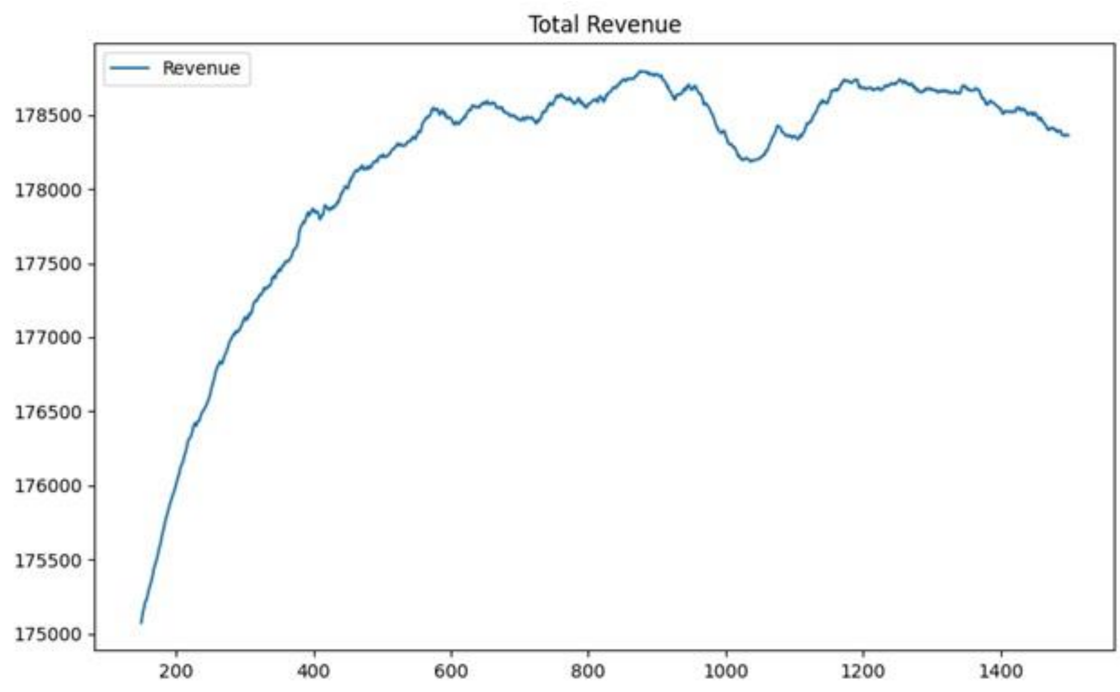

Figure 9: Revenue vs Episode plot of Microgrid based model

Similarly, Confusion matrix for the model is also plotted. This confusion matrix shows the overall accuracy of 78 percent. Here, actual represents the optimal choice that had to be taken. Similarly, predicted is the choice taken by the agent.

The accuracy percent is calculated using the formulae as:

$$= \left\{\frac{TP+TN}{TP+TN+FP+FN}\right\} * 100 \text{.............(9)}$$

Where,

TP: True Positive is the value of optimal 0 row and action 0 column= 15495

TN: True Negative is the value of optimal 0 row and action 1 column = 208658

FP: False Positive is the value of optimal 1 row and action 0 column= 3601

FN: False Negative is the value of optimal 1 row and action 1 column = 756246

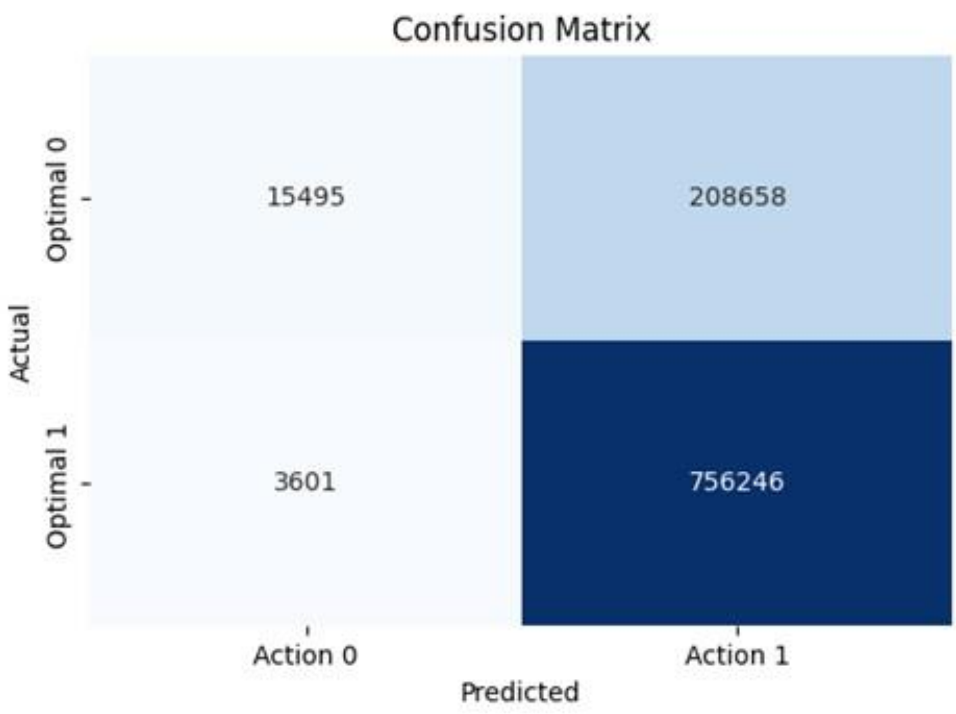

Figure 10: Confusion matrix for micro grid-based model

5.1.3 Solar based model

The trained plot of Revenue for the solar-based model is provided below. Revenue for this model are evaluated as per the real-world data. These data are used to evaluate revenue and profit, and these parameters are trained using DDQN algorithm for about 2500 episodes. From the bill information, the bill amount, i.e. the cost is Rs. 164720. After training the model with the DDQN algorithm, the revenue for this model was obtained at Rs. 166,500. Hence the Profit is 2500. The lower peaks in the graph show how epsilon greedy policy have been utilized to perform exploration and exploitation.

Exploration: It is the process of trying the new actions that agent has not taken at all to study about the environment. It might not give the highest reward but leads to better understanding in the future. Example, if the agent has two options, one with reward 5 and other not certain. Here, the choice of uncertain action is exploration which might have reward 2 or 10 as agent does not know unless it tires.

Exploitation: This is the process by which agent utilizes the knowledge to find the best-known action to find the highest expected reward. This process concerns maximization of reward based on what is already known. Example, upon trying multiple action the agent knows that the action gives reward of 9 and the agent will stick with this reward utilizing knowledge of exploitation.

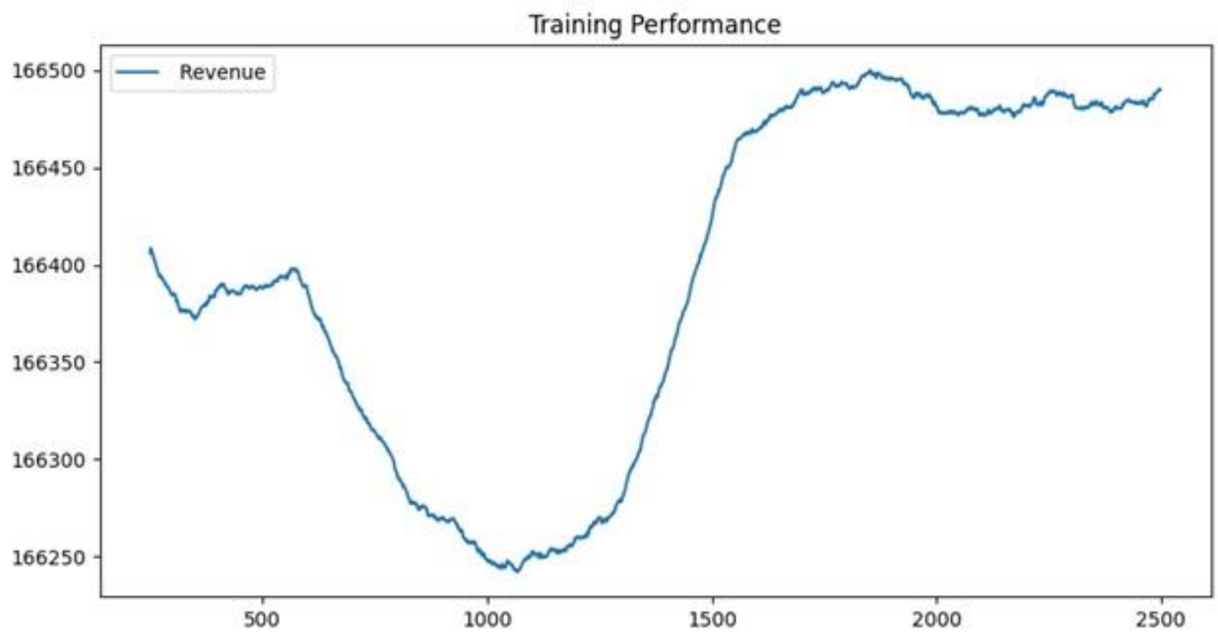

Figure 11:Revenue Vs Episode plot for DER based modelling

Similarly, Confusion matrix for the model is also plotted. This confusion matrix shows the overall accuracy of 71 percent. Here, actual represents the optimal choice that had to be taken. Similarly, predicted is the choice taken by the agent.

The accuracy percent is calculated using the formulae as:

$$= \left\{\frac{TP+TN}{TP+TN+FP+FN}\right\} * 100 \text{............}(10)$$

Where,

TP: True Positive is the value of optimal 0 row and action 0 column= 361138

TN: True Negative is the value of optimal 0 row and action 1 column = 98788

FP: False Positive is the value of optimal 1 row and action 0 column= 189212

FN: False Negative is the value of optimal 1 row and action 1 column = 334862

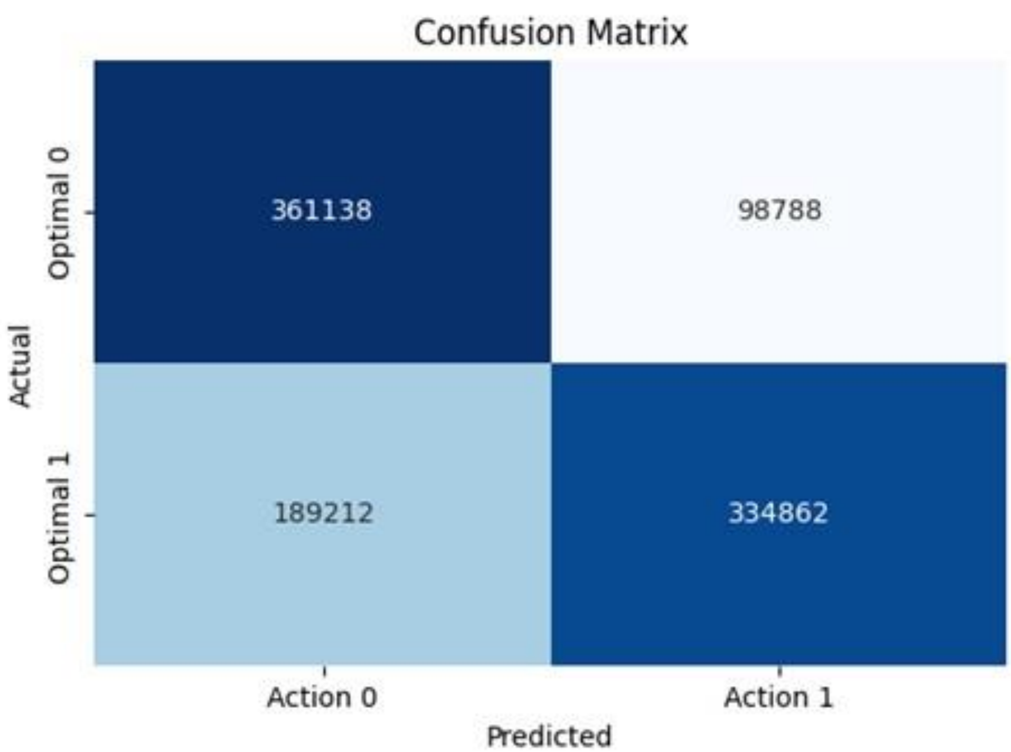

Figure 12: Confusion matrix for micro grid-based model

5.2 State of charge of ESS

5.2.1 SOC percent for Microgrid based model

The SOC percentage is evaluated using the battery capacity data. Battery capacity of 150Kwh is assumed while training the model. The SOC percent versus time step graph for microgrid based model is shown below. while obtaining the soc, each model is trained till 1500 episodes and then final episodes' soc percent vs each time steps is plotted. This graph shows that, ESS is charged and discharged continuously. Hence, in Microgrid based model ESS supplies energy to EV that is shown by the discharging pattern. Similarly, when the rising pattern shows that ESS is charged. In the graph of SOC % the minimum value is 20% of the capacity of ESS. As, the graph shows zero line at certain instances, here SOC% of ESS is 20% of the ESS capacity.

The graphical plot of electrical energy generated from solar plant is shown below. This plot shows that solar energy is generated only at daytime and at other time there is no generation shown by the zero line.

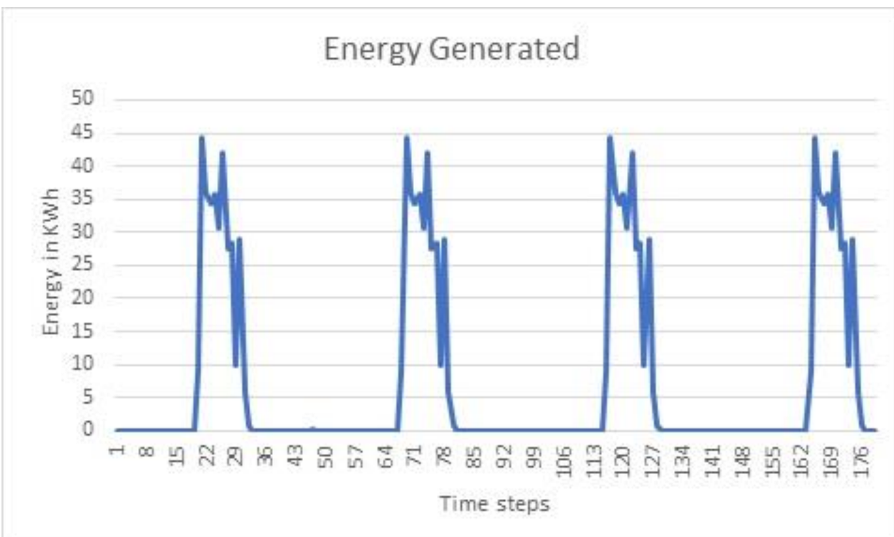

Figure 13: Electrical energy Generated from solar plant

Similarly, energy demanded by charging station is only seen during the daytime from morning 8 to evening 8. At, this time only the graph peaks are observed. Also, the bill information of Bhaktapur charging station is obtained by from the owner. The latest month's bill is observed as Rs. 1,64,720.

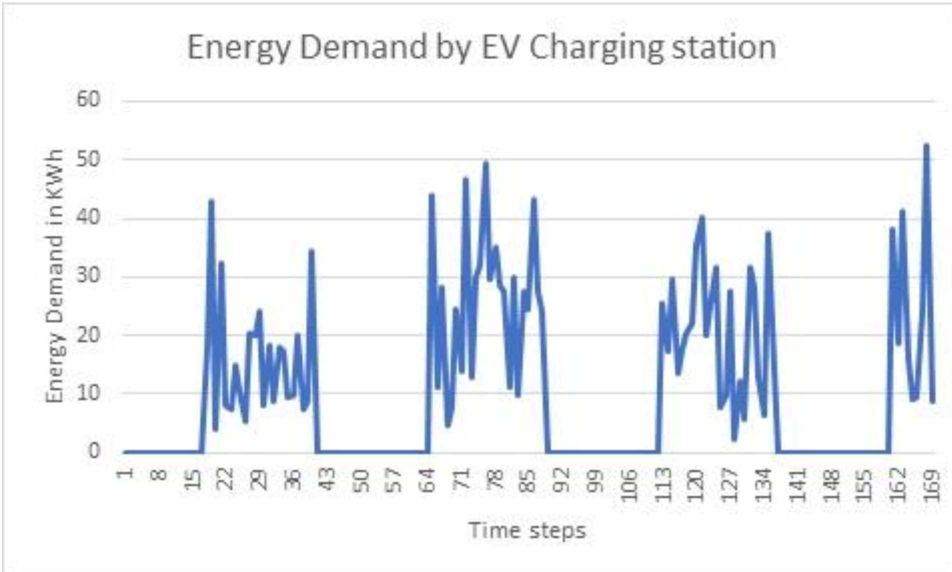

Figure 14: Electrical energy demanded by Charging station

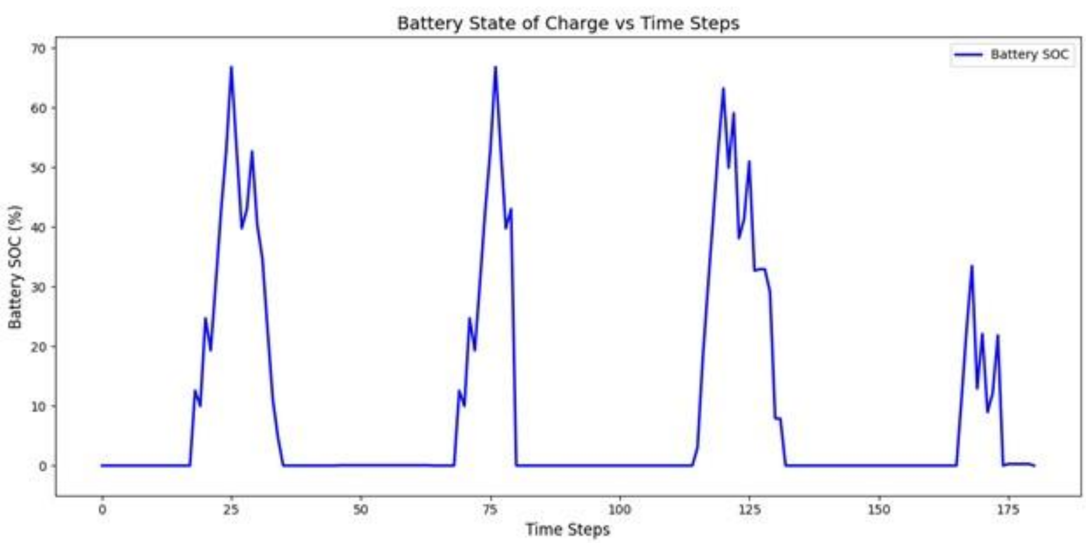

Figure 15:SOC Percent Vs Time step for Microgrid based model

5.2.2 Soc percent for Solar based model

The SOC percentage is evaluated using the battery capacity data. Battery capacity of 150Kwh is assumed while training the model. The SOC percent versus time step graph for solar-based model is shown below. This graph shows that ESS is charged but no discharging pattern is observed. As the discharging pattern is not seen, this shows that solar energy is sufficient to satisfy demand for electric vehicles. The extra amount of solar energy is used to charge the battery, seen by the increasing trend. In the graph of SOC % the minimum value is 20% of the capacity of ESS.

The graphical plot of electrical energy generated from solar plant is shown below. This plot shows that solar energy is generated only at daytime and at other time there is no generation shown by the zero line.

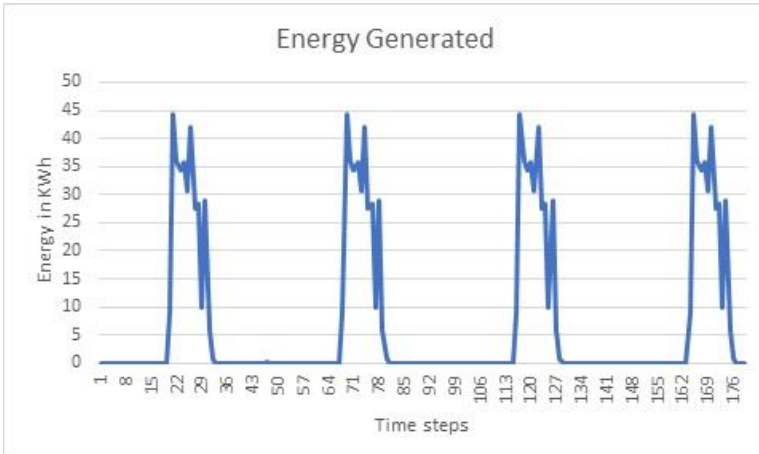

Figure 16: Electrical energy Generated from solar plant

Similarly, energy demanded by charging station is only seen during the daytime from morning 8 to evening 8. At, this time only the graph peaks are observed. Also, the bill information of Bhaktapur charging station is obtained by from the owner. The latest month’s bill is observed as Rs. 1,64,720.

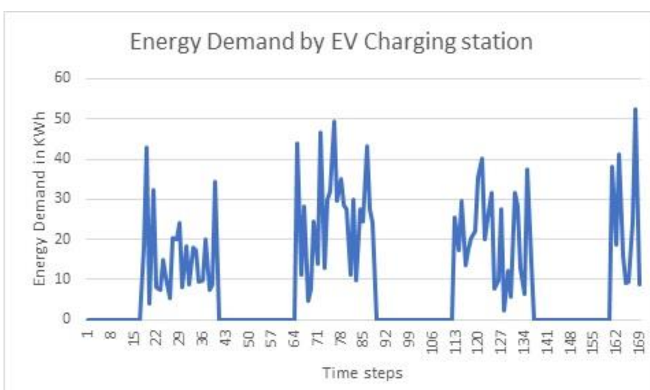

Figure 17: Electrical energy demanded by Charging station

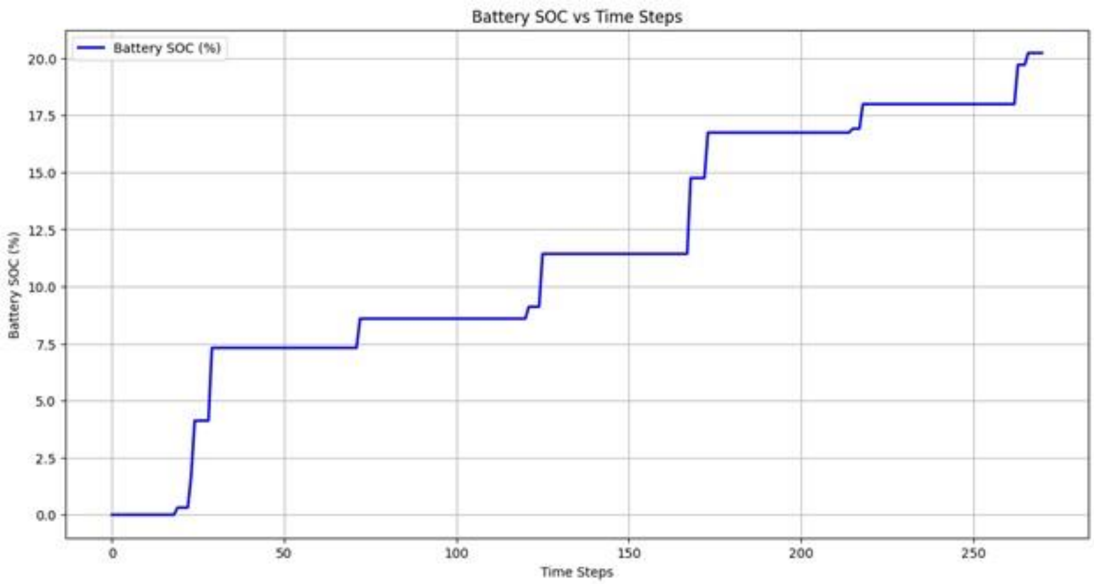

Figure 18:SOC Percent Vs Time step for solar based model

CHAPTER SIX: FORECAST OF THE EV CHARGING STATION

6.1 EV Import of last 12 year

While observing the data of import of Vehicles from Department of Custom of Nepal, it is observed that the trend of Electric Vehicle has increased. The data of Vehicle import for past 12 years is observed. The trend line shown in the figure below explains that users of EV have increased from year 2077-2078. Moreover, the Diesel/Petrol based vehicle have also reduced its number as EV have increased its number from year 2077. Similarly, following the policy dictated by government as:

- Downpayment of both EV and Diesel/Petrol vehicle will be increased by 40 percentage [6].
- Dictated in May 2019, the Prime Minister's Committee on Climate Change e-mobility by 2030, 90 percent of private passenger vehicle sales will be electric, and 60 percent of public vehicle will be electric [5].
- As per budget speech of 2022/23 NEA will operate charging station at 50 and private charging station will be promoted.[5]

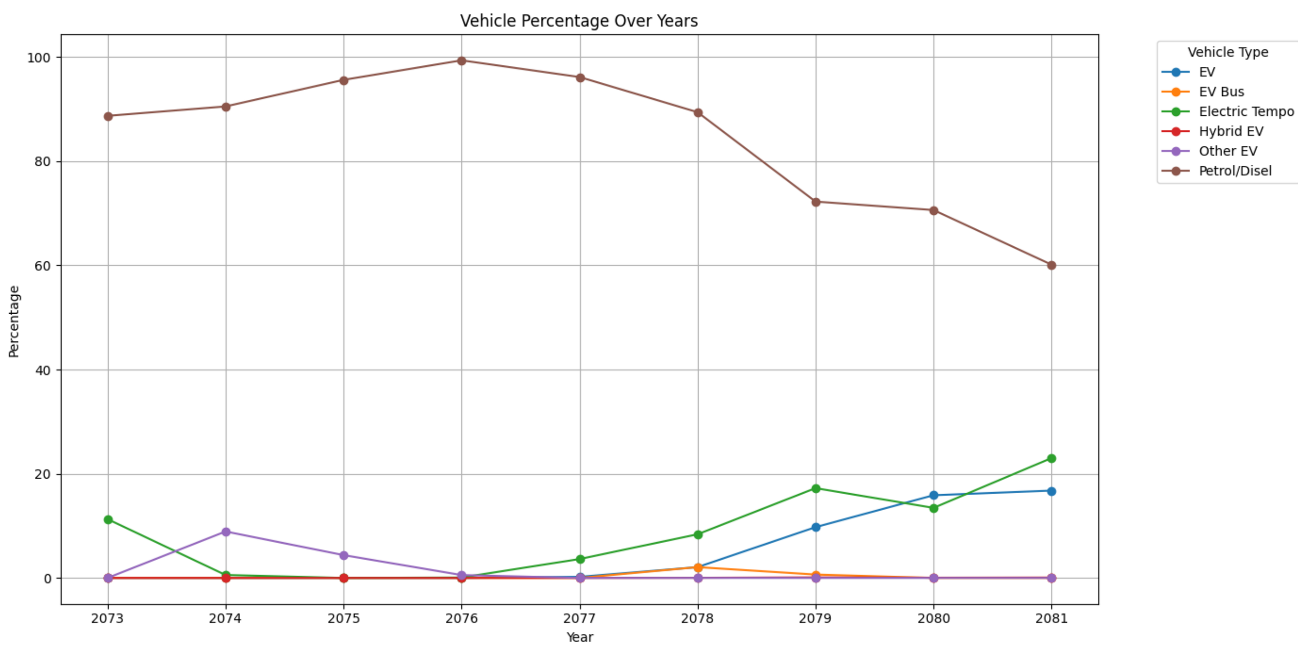

Figure 19: 12-year trend of vehicle number

As per the policies, regression analysis is performed and 5-year trend of vehicle number for each type of vehicle is fore casted. This forecast trend line shows that the EV number will increase, and the number of Diesel/Petrol Vehicle will decrease. Moreover, during the past year the increase on Electric Tempo was only seen but in future EV vehicle such as 'car', 'jeep', 'van', motorbike' etc will also consequently increase. Besides, certain government policies should be revised like the revenue of private charging station is very low so policy related to charging station should be made.

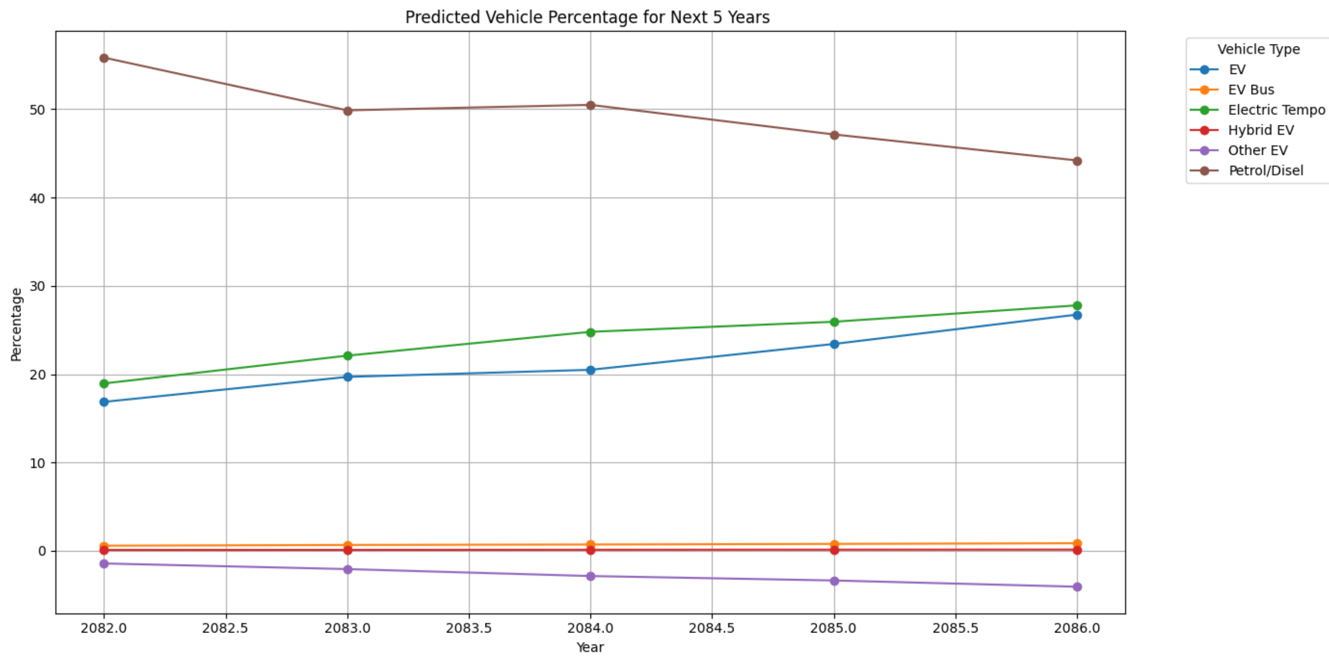

Figure 20:Forecast of 5-year vehicle number in percentage

6.2 Scenario of Bhaktapur Charging station in 5 Year

The data of yearly revenue, yearly charging volume, yearly electricity bill of Bhaktapur charging station is studied. Charging station provided a year data of this parameters as enlisted in the Table below:

Table 5:1 year data of charging station

Station Name	Total Charging Volume(kWh)	Number of Total Charges(times)	Total number of connectors	Total revenue(Rs)	Electricity Bill(Rs)
Bhaktapur Charging station	8520	284	3	66,981	66981
Bhaktapur Charging station	13860	462	3	108,847	108847.8
Bhaktapur Charging station	21031	701	3	160000	164720.88
Bhaktapur Charging station	26,981	900	3	205,500	210,885
Bhaktapur Charging station	24,236	810	3	184,500	189,750
Bhaktapur Charging station	25,492	850	3	195,000	199,950
Bhaktapur Charging station	39,747	1325	4	305,000	310,320
Bhaktapur Charging station	38,899	1300	4	298,500	303,520
Bhaktapur Charging station	27,345	910	4	208,500	213,885
Bhaktapur Charging station	28,698	955	4	218,500	224,520
Bhaktapur Charging station	24,678	825	4	187,500	193,200
Bhaktapur Charging station	24,240	812	4	147,500	153,700

Similarly, Forecasting analysis of charging station is done using regression analysis. This analysis incorporates real world impact as well as the policies:

1. 40% downpayment in every Electric Vehicle considering 15% reduction in EV adoption.
2. With the increase in demand the charging station increases that would reduce load of 10% in each charging station
3. Revenue status if charging station is lower so 5% reduction in investment of charging station
4. As hybrid EVs have also started increasing so this impact a reduction of 10% dependency on charging station.
5. Many EV users are considering charging at home so a 15% reduction in demand for public charging station is considered
6. Charging cost structure of private charging station ranges is higher than public charging station so increase of demand in public charging station is considered by 15%
7. Nepal electrical energy is dependent on hydro so seasonal variation is seen that impact the demand/supply of electrical energy.

All these impacts are combinedly considered and applied to each parameter of charging station and a 5-year prediction is made.

- Also, some random fluctuation of +-5% is considered for variability.
- Besides considering economic recession, policy change and increased competition in year 2 and 4 predicted values are reduced by 15%.

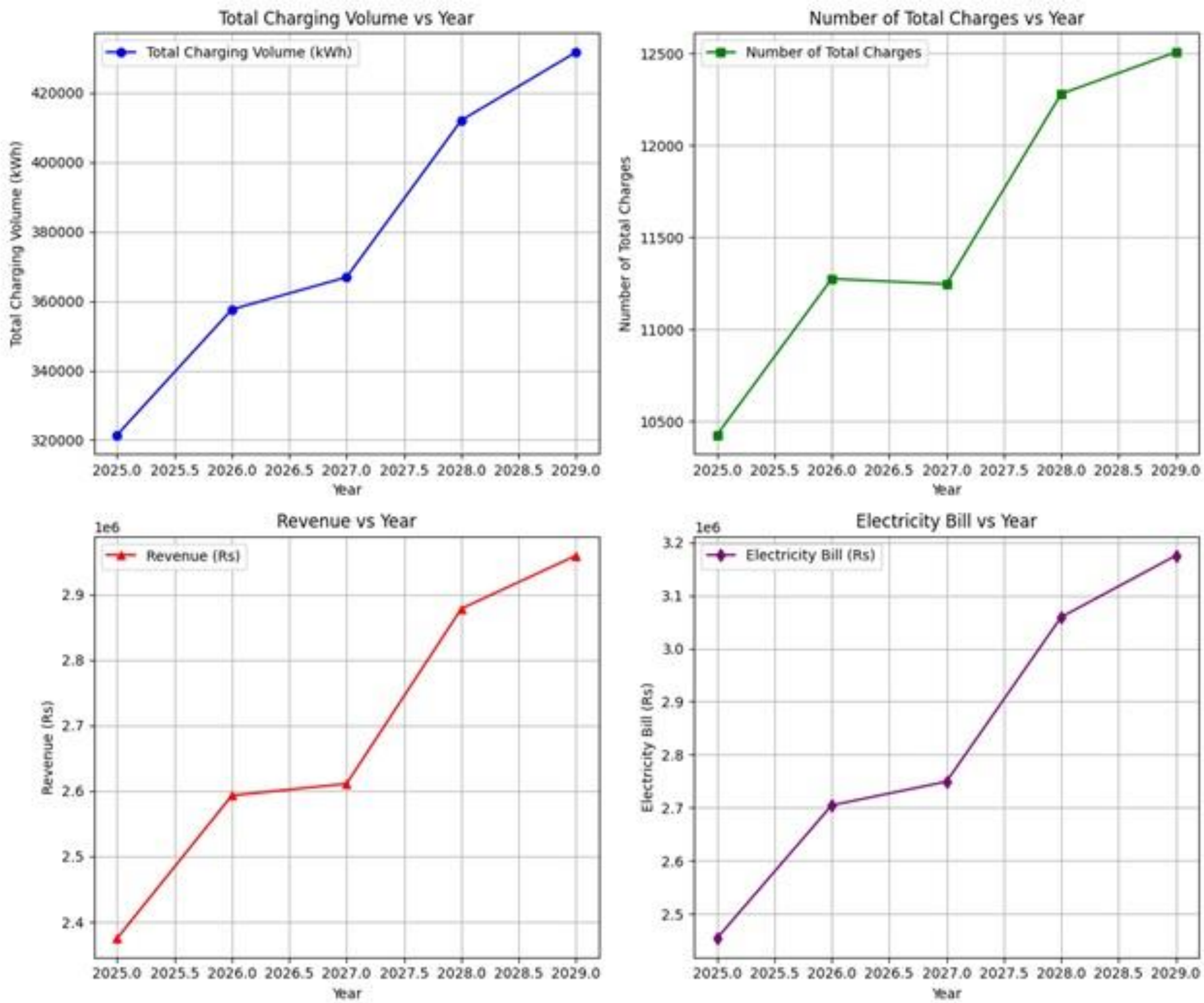

Figure 21: Forecast of Bhaktapur charging station

CHAPTER SEVEN: CONCLUSION

The present energy scenario of charging station and supply /demand status shows that there is very low profit. Soon, if charging station-based scheduling methodology are not bought to practice then the charging station might be closed. In contrary, while observing the bill and revenue status of Bhaktapur charging station, it is found that the bill amount is around Rs. 164000. But the Revenue is less, so there is negative profit.

Hence, the research has presented models of charging station explaining three different scenarios with the application of DDQN Reinforcement learning algorithm. First scenario is the present scenario, only MPG is the provider of energy requirement by charging station. In this scenario both practical and DDQN based model of Bhaktapur charging station is observed. The profit is obtained negative. Similarly, another model is Microgrid based model which has solar, ESS and MPG as the provider of energy source. In this model there is highest profit. Furthermore, the last model is solar based model, it has profit but is less compared to microgrid based model. Therefore, the microgrid model have maximum revenue and profit but with the solar based model the load of MPG can be shifted. With the application of DDQN based Reinforcement learning algorithm, the data training is complete in 30min for 2500 episodes.

Similarly, the Data of Import vehicle count is studied from Department of Custom of Nepal to understand the trend of vehicle. While observing the 12-year data of vehicle, it is found that EV number have been increasing from 2077. Moreover, regarding government policy, EV 2018 act policy the future five-year Vehicle count is fore casted using regression analysis. While observing the Forecast result, it is seen that the EV number is bound to increase and consequently the Diesel/Petrol vehicle number will gradually decrease. Besides, the Bhaktapur charging station's five-year forecast is also carried and its charging volume in KWh, Revenue. Electricity Bill, Total EV number is observed.

CHAPTER EIGHT: SUGGESTION FOR FURTHER RESEARCH

Based on the above results and conclusion obtained from the research, multiple suggestions can be given so that performance of EV charging station can be improved. Some suggestions are:

- **Adoption of Intelligent Scheduling Techniques:** As, DDQN algorithm has been utilized to model the charging station in this research. There are variety of other RL algorithms that could be employed like Actor Critic, Policy based algorithm.

- **Policy Support and Incentives**: Government based incentives is necessary to consider the application of smart energy management systems possible. Policymakers should be involved in future efforts to suggest incentives for charging stations powered by green energy and smart scheduling to reduce grid load and emissions.

- **Expanding and Scaling Microgrid Model:** The successful result of the microgrid based scheduling model at Bhaktapur station suggests its application to other semi-urban and urban locations.

- **Integration of Forecasting and Demand Analysis:** The import trends analysis and the analysis of forecasting reveals tremendous growth in EV uptake. Future vehicles must have advanced forecasting modules to factor in EV growth, volatility in energy price, and patterns in charging in planning infrastructure augmentation and resource optimization allocation.

- **Economic Feasibility and Life Cycle Cost Analysis:** Besides technical efficiency, future studies should also consider overall economic viability analysis, including life cycle cost (LCC), analysis of DER devices (solar panels, batteries), maintenance cost, and degradation effects over time.

- **User Behaviour and Dynamic Pricing:** Considering user behaviour in the scheduling model would further optimize energy consumption and reduce peak demand. Future studies can also explore dynamic pricing strategies to promote off-peak charging and improve load management.

REFERENCE

[1] A. S. KIRASEYA PREUSSER, “Energy Scheduling for a DER and EV ChargingStation Connected Microgrid with Energy Storage,” *IEEE Access,* January 2023.

[2] A. G. L.-B. Heba Abdullah, “Reinforcement Learning Based EV Charging Management Systems–A Review,” *IEEE Access,* 2021.

[3] Y. He, B. Venkatesh and L. Guan, “Optimal Scheduling for Charging and Discharging of Electric Vehicles,” *IEEE Transaction on smart Grid,* 2012.

[4] S. Wang, S. Bi and Y. A. Zhang, “Reinforcement Learning for Real-Time Pricing and Scheduling Control in EV Charging Stations,” *IEEE Transactions on Industrial Informatics,* 2021.

[5] “Nepal E-Mobility Country Profile,” Feburary 2024. [Online]. Available: https://asiantransportobservatory.org/documents/160/Nepal_emobility.pdf.

[6] “Nepal Rastra Bank,” [Online]. Available: https://www.nrb.org.np/.

[7] “Department of Customs,” [Online]. Available: https://www.customs.gov.np/.

[8] “Deep Learning Bible - 5. Reinforcement Learning - Eng,” [Online]. Available: https://wikidocs.net/174609.

[9] J. E. Anderson, N. Bottcher and T. Kunimhof, “Approach to Determine Charging Infrastructure for 1 Million Electric Vehicles in Germany,” in *Transportation Research Board 2016 Annual Meeting*, Washington, DC, January 2016.

[10] N. Z. A. Matanov, “Developments and Challenges for Electric Vehicle Charging Infrastructure,” *BulEF Conference,* 2020.

[11] IEA, “Global EV Outlook 2023,” International Energy Agency, 2023.

[12] Q. L. X. D. J. K. F. Wang, “Smart Charging for Electric Vehicles: A Survey From the Algorithmic Perspective,” *IEEE Communications Surveys & Tutorials,* 2016.

[13] X. L. J. D. F. K. Qinglong Wang, “Smart Charging for Electric Vehicles: A Survey From the Algorithmic Perspective,” in *IEEE CommunicationSurveys 7 Tutorials*, 2016.

[14] B. V. L. G. Yifeng He, “Optimal Scheduling for Charging and Discharging of Electric Vehicles,” in *IEEE Transaction on smart Grid*, 2012.

[15] J. L. a. V. K. A. Chis, “Reinforcement learning-based,” in *IEEE Trans. Veh. Technol.*, 2017.

[16] A. Y. H. W. D. P. a. F. S. V. Moghaddam, “An online reinforcement learning approach for dynamic pricing of electric vehicle charging station,” in *IEEE Access*, 2020.

[17] A. Ö. S. M. M. A. V. V. a. I. A. Y. Yoldaş, “Enhancing smart grid with microgrids: Challenges and opportunities,” in *Renew.Sustain. Energy Rev.*, 2017.

[18] T. A. N. a. P. Toivanen, “Deep reinforcement learning for energy management in a microgrid with flexible demand,” in *Sustain. Energy, Grids Netw.*, 2021.

[19] S. M. J. Y. C. W. a. J. Y. Y. L. Hou, “‘Reinforcement mechanism design for electric vehicle demand response in microgrid charging stations,” in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2020.

[20] E. O. A. a. K. A. Folly, “Reinforcement learning techniques for optimal power control in grid-connected microgrids: A comprehensive review,” in *IEEE Access*, 2020.

APPENDIX A: PUBLICATION

[IOEGC16] Editor Decision Inbox ×

Suwarna Lingden Wed, Apr 2, 9:15 AM (2 days ago)
to me

Aryasupurna Timalsina:

We are pleased to inform you that your manuscript titled "Impact of DER on Electric Vehicle with the application of Reinforcement Learning" submitted to 16th IOE Graduate Conference is **Accepted** for presentation in the Conference as well as inclusion in the Peer-Reviewed Proceedings. Please note that inclusion in hard copy proceedings is contingent upon your timely response to further edits, if any, during the publication process.

With Warm Regards,
IOEGC-16 Editorial Team

Optimizing EV Charging station with RL based Solar DER integration

Aryasupurna Timalsina [a], Basanta Kumar Gautam[b],

Corresponding Email: [a] timarya123@gmail.com , [b] basanta.gautam@pcampus.edu.np

Abstract

Electric vehicles in Nepal are seen to be increasing, both private and public sector transportation are seen adopting electric vehicles as a means of transport. While analyzing the present energy demand and energy supply scenario of charging station, the profit is found very low. If similar supply demand scenario persists the profit will be negative and there might be situation of shutdown of private charging station. But with the introduction of a Micro grid-based model, Charging stations could be scheduled properly. If along with Main Power Grid other energy sources like solar, Battery storage system are added to the charging station then the demand can be fulfilled without extra load to MPG.This paper presents model of charging station, there are three different models of EV charging stations. The first model is a stand-alone model, where the EV receives energy from the main power grid. Similarly, other models are microgrid-based models and solar-based models. In the microgrid-based model players such as MPG, ESS, and Solar operate in unison to fulfill the demand for charging stations. Besides, the solar-based model has solar as the prime energy source to fulfill the demand for charging stations. All these models are operated and trained using Reinforcement Learning. With application of DDQN algorithm for scheduling the charging station, Microgrid based model showed the highest revenue of all three models.

Keywords

Main Power Grid(MPG)–Energy Storage System(ESS)–Electric vehicle Charging Station (EVCS) –Distributed Energy Resources (DER)–Reinforcement Learning–Double Deep Q Learning (DDQN)

1. Introduction

The transportation sector of Nepal is one of the prominent contributors of greenhouse gases, about 44 percent [1] of greenhouse gases and 99 percent of CO2 emission occurs from the road infrastructure. Due to such a scenario, Electric Vehicles have gained more popularity as these are toxic gas-free fuels. Also, the Electric Vehicle Association of Nepal has estimated that in 2018, the EVs number reached about 45000[5].In addition, the EV Act 2019 policy has said that 90 percent private and 60 percent public of transport will be converted to electric vehicles by 2030.[2] Similarly, the number of electric vehicle charging stations has also increased. At present there are at least 350 charging stations in Nepal. From May 2019, it is mentioned that 25 percent of private and 25 percent of public sales will be made electric by 2025.

Hence, such a growing number of EVs demand more electrical energy. To fulfill the demand, the present scenario of the Grid supply is not enough. In addition, while visiting the charging station, it is confirmed that the current profit status of the charging station is very low. If the same topology of energy supply from only the Main Grid exists, then the status of the private charging station in the future will be in danger. Hence, there is a need for a Micro Grid and Distributed Energy sources interaction with the Main Power Grid. The combination of DER, ESS, along with the Main Power Grid will not only boost the environment of the charging station but will also help to manage the energy supply and shift the load from the Main Power Grid. For the interaction of multiple players, i.e. MPG, ESS, Charging station, DER, there needs an appropriate scheduling methodology. Scheduling of charging station is very important because it concerns energy management. Machine learning based algorithms such as Reinforcement Learning technique[3] could be employed to perform the scheduling of the charging station as energy management tool. These algorithms not only help in scheduling, but also help to obtain the optimal scenario at which the revenue and profit could be maximized. Moreover, instead of using model-based data, real-world data of energy generation and vehicle count can bring more optimized charging station scheduling. In terms of DER, the number of solar plants has increased, so the solar plant could be utilized in the microgrid model to facilitate the scheduling of charging stations.

Reinforcement learning is of a variety of types,[4] but among those, the DDQN based algorithm is a popular one. It has lots of advantages over other forms of reinforcement learning. DDQN algorithm is an advanced form of DQN algorithm that has addressed the issue of overestimation bias. Unlike other forms of learning, it uses Q-learning methods that high-dimensional state spaces could be handled easily. It is a sample efficient learning tool compared to policy gradient RL technique. Besides, it is not so complex to understand compared to actor critic methods. Also, this algorithm has a strong baseline for other extended RL based algorithms[5]. The practical application of DDQN is seen in robotics, autonomous driving, and many other real-world applications.

2. Methodology

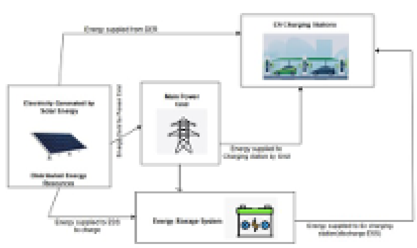

Figure 1: General block diagram

The model of the charging station[6] with all the players in action is shown. This diagram has main power grid, solar plant as DER, EV charging station block and ESS block. The entire model works in three different ways. Individual ways are explained:

2.1 Stand alone charging station

In this scenario charging station receives energy from only MPG. This form of scheduling action is termed as stand alone charging scheme. The mathematical model for this scheme is

$$\text{Minimize } Z = \sum_{t=1}^{T} P_{\text{buy}} \cdot e_b^t \tag{1}$$

here, the constraints for this model is energy supplied by grid must meet the energy demand by charging station, $e_b^t = e_{EVCS}^t$

2.2 Micro Grid based Modeling

In this scheduling scenario there exist two actions.
Action 1

- ESS Charges charging station
- DER Sells Energy to the Grid
- Grid Supplies Additional Energy if Needed

The objective function for this action is:

$$\text{Maximize } Z_1 = \sum_{t=1}^{T} \left(P_{\text{sell}} \cdot e_s^t - P_{\text{buy}} \cdot e_b^t \right) \tag{2}$$

Constraints:

- Energy Demand by charging station, $e_c^t + e_b^t = e_{EVCS}^t$
- ESS energy balance, $0 <= e_c^t <= min(s_t, r_{max})$
- Energy trade by DER, $e_s^t = e_g^t$
- Grid Supplies Additional Energy, $e_b^t = e_{EVCS}^t - e_c^t$

Action 2

- DER energy is first used to Charge the ESS
- Excess DER Energy Sold to the Grid
- Charging station is Powered by the Grid

The objective function for this action is:

$$\text{Maximize } Z_2 = \sum_{t=1}^{T} (P_{\text{sell}} \cdot e_s^t - P_{\text{buy}} \cdot e_b^t) \tag{3}$$

Constraints are

- ESS energy balance, $e_r^t <= r_{max})$
- Energy trade DER, $e_s^t = e_g^t - e_r^t$
- Grid Supplies charging station, $e_b^t = e_{EVCS}^t$

2.3 DER based modeling

In this model, the EVs are directly charged from the solar energy generated by the DER. There are two cases in this scenario, excess energy and deficit energy. The objective function for this scenario is:

$$\text{Maximize } Z_3 = \sum_{t=1}^{T} \left(P_{\text{sell}} \cdot e_s^t - P_{\text{buy}} \cdot e_b^t \right) \quad (4)$$

2.4 Set Up

The scheduling action is carried with the use of DDQN algorithm. In order to perform the computation and utilize the logic of methodology explained necessary parameters are defined in the table, These parameters are the setup defined during

Table 1: Coding setup

Parameter	Value	Description
ESS capacity	150KW	Total capacity of ESS
Max rate	100KW	Charging discharging rate
Energy Price	{17<h<23:7,23<h or h<5:3.7, else: 5.5}	Energy price based on time of day
State Size	4	No of state variables, State variables as: battery soc, solar energy, energy price, EV KWh demand
Action size	2 or 3	Depending on the model
Episodes	1500 or 2500	Number of episodes for training
Discount factor gamma	0.9 to 0.99	Discount the rewards
Epsilon min	0.01 to 0.1	Min exploration rate
Epsilon decay	0.99 to 0.9999	Decay rate of exploration
Batch size	32 to 256	Sample of experience used during training
Memory size	1000 to 10000	Maximum size of replay memory
EV KW usage	Data from excel	Energy demand of EVs
EV revenue	Data from excel	Revenue from EVs
Solar energy	Data from excel	Solar energy at each time step
timesteps	30min	time at which data are taken

coding. There are major hyperparameter explained in the table. These hyperparameters are fine tuned with help of optuna library and then defined in the code. Besides, the data of solar generation, EV revenue and KWh consumption are merged at same time instant. Normalization of these data are also carried before performing the coding.

2.4.1 Neural Network Architecture

Reinforcement learning[7] is employed to build the scheduling model of charging station. DDQN is a modified version of DQN type of algorithm used to perform the training on data set. The neural network structure for this algorithm has 4 input layer as the number of states. 2 hidden layer activated by Rectified linear unit activation function and 2 or 3 output layer depending on the model type. In DDQN algorithm two type of Q network exist, to solves issue of overestimation bias. These are online network and Target network. Online network is utilized during

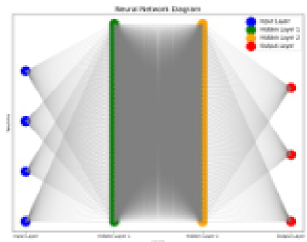

Figure 2: Neural Network Architecture

action selection to select best action and Target network is utilized during action evaluation which is delayed copy of online network. In this algorithm the target Q value is evaluated using Bellman equation.

$$\text{Target} = r + yQ_{target}(s', arga'_{max} Q_{onine}(s', a')) \quad (5)$$

similarly, the loss function is evaluated using following relation,

$$\text{Loss} = 1/N \sum_{i=1}^{N} (Q_{online}(s_i, a_i) - Target_i)^2 \quad (6)$$

2.4.2 Flowchart of DDQN algorithm

The overall flow of DDQN algorithm[5] is explained below by the flow diagram. This diagram explains about initialization phase, action selection, environment interaction, experience replay, target calculation, loss computation and update.

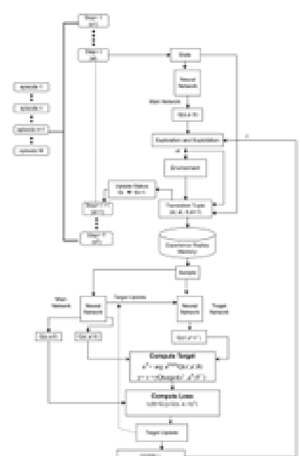

Figure 3: Neural Network Architecture

3. Results

The outcomes for each methodology focus on optimizing energy management and increasing Revenue. Below are the specific outcomes for each methodology.The revenue, profit, ESS State of Charge for each model is trained and obtained using the DDQN algorithm.

3.1 Standalone Model

The trained plot of Revenue for standalone model is provided below. Revenue for this model is evaluated as per the real world data. These data are used to evaluate revenue and profit, and these parameters are trained using DDQN algorithm for about 1500 episodes. From the bill information, the bill amount

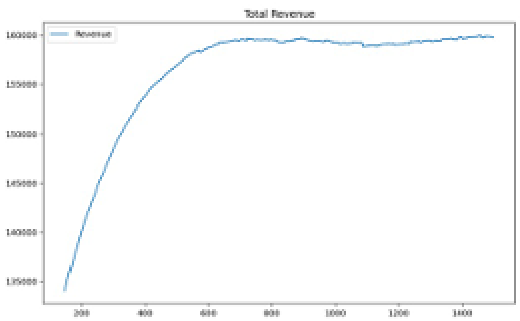

Figure 4: Revenue Plot of Stand alone model

i.e. the cost is 164720. After training the model with DDQN algorithm the Revenue is obtained as: 160,000. Hence the Profit is -4000. This the standalone model that is existing, is not showing profit rather has -4000.

3.2 Micro Grid based model

The trained plot of Revenue for the Micro Grid based model is provided below. Revenue for this model is evaluated as per the real world data. These data are used to evaluate revenue and profit, and these parameters are trained using DDQN algorithm for about 1500 episodes. From the bill information, the bill amount i.e. the cost is 164720. After training the model with DDQN algorithm, the Revenue for this model is obtained as: 178,500. Hence the Profit is 14500.Hence, the microgrid model shows a profit of about 14,500. Similarly, Confusion matrix for the model is also plotted. This confusion matrix shows the overall accuracy of 78 percent. Here, actual represents the optimal choice that had to be taken and

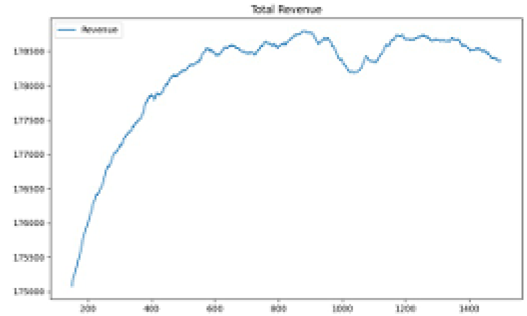

Figure 5: Revenue Plot of Micro based model

predicted is the choice taken by the agent.

Figure 6: Confusion matrix for micro grid based model

3.3 Solar based model

The trained plot of Revenue for the solar-based model is provided below. Revenue for model is evaluated as per the real world data. These data are used to evaluate revenue and profit, and these parameters are trained using DDQN algorithm for about 2500 episodes. From the bill information, the bill amount, i.e. the cost is 164720. After training the model with the DDQN algorithm, the revenue for this model was obtained at 166,500. Hence the Profit is 2500. Similarly, Confusion matrix for the model is also plotted. This confusion matrix shows the overall accuracy of 71 percent.

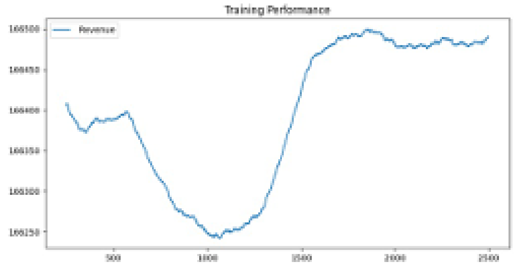

Figure 7: Revenue Plot of Solar based model

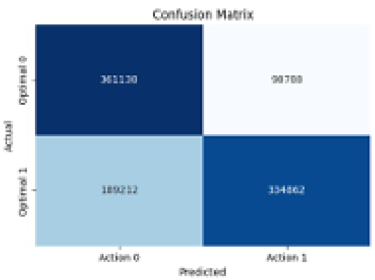

Figure 8: Confusion matrix for micro grid based model

3.4 Soc percent of ESS for Microgrid based model

The SOC percentage is evaluated using the battery capacity data. Battery capacity of 150Kwh is assumed while training the model. The SOC percent versus time step graph for microgrid-based model is shown below. while obtaining the soc, each model is trained till 1500 episodes and then final episodes' soc percent vs each time steps is plotted. This graph shows that, ESS is charged and discharged continuously. Hence, in Micro-grid based model ESS supplies energy to EV that is shown by the discharging pattern. Similarly, when the rising pattern shows that ESS is charged.

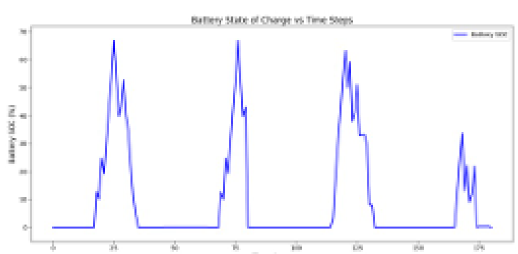

Figure 9: SOC Percent Vs Time step for Microgrid based model

3.5 Soc percent of ESS for Solar based model

The SOC percentage is evaluated using the battery capacity data. Battery capacity of 150Kwh is assumed while training the model. The SOC percent versus time step graph for solar-based model is shown below. This graph shows that ESS is charged but no discharging pattern is observed. As the discharging pattern is not seen, this shows that solar energy is sufficient to satisfy demand for electric vehicles. The extra amount of solar energy is used to charge the battery, seen by the increasing trend.

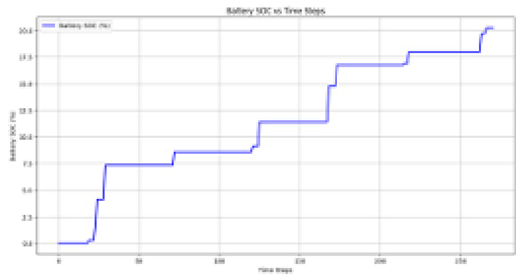

Figure 10: SOC Percent Vs Time step for solar based model

4. Conclusion

The present energy scenario of charging station and supply /demand status shows that there is very low profit. Soon, if charging station based scheduling methodology are not bought to practice then the charging station might be closed. In contrary, while observing the bill and revenue status of Bhaktpur charging station, it is found that the bill amount is around 164000. But the Revenue is less, so there is negative profit.Hence, the research has presented models of charging station explaining three different scenario with the application of DDQN Reinforcement learning algorithm. First scenario is the present scenario, only MPG is the solo provider of energy. In this scenario both practical and DDQN based model of Bhaktpur charging station is observed. The profit is obtained negative i.e.-4000. Similarly, other model is Microgrid based model which has solar, ESS and MPG as the provider of energy source. In this model there is profit of 14500. Furthermore, the last model is solar based model, it has profit but is less compared to microgrid based model. Therefore, the microgrid model have maximum revenue and profit but with the solar based model the load of MPG can be shifted. All three model implemented using Python, PyTorch library. The computation is carried using laptop with AMD Ryzen 7435HS processor running at 3.1 GHz, and 16GB RAM. Besides,training sequence for 1 month worth of simulation took 120 minutes.

References

[1] Nepal E-mobility country profile, 2024. Accessed: Feb. 2024.

[2] Nepal Rastra Bank, 2024. Accessed: 2024.

[3] Shuoyao Wang, Suzhi Bi, and Yingjun Angela Zhang. Reinforcement learning for real-time pricing and scheduling control in ev charging stations. *IEEE Transactions on Industrial Informatics*, 17(2):849–859, 2021.

[4] Heba M. Abdullah, Adel Gastli, and Lazhar Ben-Brahim. Reinforcement learning based ev charging management systems–a review. *IEEE Access*, 9:41506–41531, 2021.

[5] *Deep Learning Bible- 5. Reinforcement Learning*. Dec. 28 2024.

[6] Kiraseya Preusser and Anke Schmeink. Energy scheduling for a der and ev charging station connected microgrid with energy storage. *IEEE Access*, 11:73435–73447, 2023.

[7] Yifeng He, Bala Venkatesh, and Ling Guan. Optimal scheduling for charging and discharging of electric vehicles. *IEEE Transactions on Smart Grid*, 3(3):1095–1105, 2012.

APPENDIX B: PLAGIARISM TEST REPORT

Aryasupurna Timalsina

final_report.pdf

Tribhuvan University

Document Details

Submission ID

trn:oid:::3117:450639466

Submission Date

Apr 20, 2025, 8:49 AM GMT+5:45

Download Date

Apr 20, 2025, 8:51 AM GMT+5:45

File Name

final_report.pdf

File Size

1.0 MB

41 Pages

9,252 Words

47,465 Characters

5% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Match Groups

- **13** Not Cited or Quoted 2%
  Matches with neither in-text citation nor quotation marks
- **0** Missing Quotations 0%
  Matches that are still very similar to source material
- **30** Missing Citation 3%
  Matches that have quotation marks, but no in-text citation
- **1** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

3%	Internet sources
3%	Publications
0%	Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **13** Not Cited or Quoted 2%
  Matches with neither in-text citation nor quotation marks
- **0** Missing Quotations 0%
  Matches that are still very similar to source material
- **30** Missing Citation 3%
  Matches that have quotation marks, but no in-text citation
- **1** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

3%	Internet sources
3%	Publications
0%	Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	**conference.ioe.edu.np**	**1%**
2	Publication	**"The 37th Annual Conference on Power System and Automation in Chinese Unive...**	**<1%**
3	Internet	**medium.com**	**<1%**
4	Publication	**Marco Grassi, Michele Nucci, Francesco Piazza. "Towards a semantically-enabled ...**	**<1%**
5	Internet	**fastercapital.com**	**<1%**
6	Publication	**Parham, Zahra. "Explaining the Behavior of Reinforcement Learning Agents Usin...**	**<1%**
7	Internet	**core.ac.uk**	**<1%**
8	Internet	**www.researchsquare.com**	**<1%**
9	Publication	**P. Lazzeroni, S. Olivero, M. Repetto. "Economic perspective for PV under new Itali...**	**<1%**
10	Publication	**Qilong Zhang, Xiangping Chen, Guangming Li, Junjie Feng, Anqian Yang. "Model P...**	**<1%**

APPENDEX C: DATA ANALYSIS IN EXCEL

	A	B
1	Data Time	Active Energy Export (-A)(kWh)
2	2081-07-21 00:00:00	0
3	2081-07-21 00:30:00	0
4	2081-07-21 01:00:00	0
5	2081-07-21 01:30:00	0
6	2081-07-21 02:00:00	0
7	2081-07-21 02:30:00	0
8	2081-07-21 03:00:00	0
9	2081-07-21 03:30:00	0
10	2081-07-21 04:00:00	0
11	2081-07-21 04:30:00	0
12	2081-07-21 05:00:00	0
13	2081-07-21 05:30:00	0
14	2081-07-21 06:00:00	0
15	2081-07-21 06:30:00	0
16	2081-07-21 07:00:00	0
17	2081-07-21 07:30:00	0
18	2081-07-21 08:00:00	0
19	2081-07-21 08:30:00	0
20	2081-07-21 09:00:00	0
21	2081-07-21 09:30:00	8.96
22	2081-07-21 10:00:00	44.16
23	2081-07-21 10:30:00	36.16
24	2081-07-21 11:00:00	34.24
25	2081-07-21 11:30:00	35.68
26	2081-07-21 12:00:00	30.56
27	2081-07-21 12:30:00	42.08

Figure 22:Normalized data set of Solar generation

Date	SOC(initial)%	SOC(Final)%	type	cost	percentage difference	battery pack	KWh	Cost per kw	Revenue
2024-01-27	25	98	car	10	73	60	43.8	16.66666667	730
2024-01-27	49	95	Taxi	6	46	24	11.04	25	276
2024-01-27	45	98	kinglong	7	53	53.58	28.3974	13.06457633	371
2024-01-27	61	80	taxi	6	19	24	4.56	25	114
2024-01-27	74	100	DFSR	6	26	30	7.8	20	156
2024-01-27	36	82	kinglong	7	46	53.58	24.6468	13.06457633	322
2024-01-27	57	85	kyc	6	28	50	14	12	168
2024-01-27	17	95		10	78	60	46.8	16.66666667	780
2024-01-27	37	90	taxi	6	53	24	12.72	25	318
2024-01-27	44	100	kinglong	7	56	53.58	30.0048	13.06457633	392
2024-01-27	36	100	kyc	6	64	50	32	12	384
2024-01-27	31	100	byd	9	69	71.8	49.542	12.53481894	621
2024-01-27	45	100	king long	7	55	53.58	29.469	13.06457633	385
2024-01-27	30	100	kyc	6	70	50	35	12	420
2024-01-27	43	100	kyc	6	57	50	28.5	12	342
2024-01-27	45	100	kyc	6	55	50	27.5	12	330
2024-01-27	53	100	taxi	6	47	24	11.28	25	282
2024-01-27	40	100	kyc	6	60	50	30	12	360
2024-01-28	82	100	kinglong	7	18	53.58	9.6444	13.06457633	126
2024-01-28	45	100	kyc	6	55	50	27.5	12	330
2024-01-28	51	100	kyc	6	49	50	24.5	12	294
2024-01-28	38	100	nexon	9	62	70	43.4	12.85714286	558
2024-01-28	45	100	kyc	6	55	50	27.5	12	330
2024-01-28	35	80		7	45	54	24.3	12.96296296	315

Figure 23: Normalized EV data

A	B	C	D	E	F	G
Data Time	Solar Generation	EV KWh	EV Revenue			
2081-07-21 00:00:00	0	0	0			
2081-07-21 00:30:00	0	0	0			
2081-07-21 01:00:00	0	0	0			
2081-07-21 01:30:00	0	0	0			
2081-07-21 02:00:00	0	0	0			
2081-07-21 02:30:00	0	0	0			
2081-07-21 03:00:00	0	0	0			
2081-07-21 03:30:00	0	0	0			
2081-07-21 04:00:00	0	0	0			
2081-07-21 04:30:00	0	0	0			
2081-07-21 05:00:00	0	0	0			
2081-07-21 05:30:00	0	0	0			
2081-07-21 06:00:00	0	0	0			
2081-07-21 06:30:00	0	0	0			
2081-07-21 07:00:00	0	0	0			
2081-07-21 07:30:00	0	0	0			
2081-07-21 08:00:00	0	0	0			
2081-07-21 08:30:00	0	18.2172	272			
2081-07-21 09:00:00	0	42.75	627			
2081-07-21 09:30:00	8.96	3.864	140			
2081-07-21 10:00:00	44.16	32.31	540			
2081-07-21 10:30:00	36.16	8.1	216			
2081-07-21 11:00:00	34.24	7.44	372			
2081-07-21 11:30:00	35.68	15	125			
2081-07-21 12:00:00	30.56	10.5	105			
2081-07-21 12:30:00	42.08	5.4096	210			
2081-07-21 13:00:00	27.52	20.5	205			
2081-07-21 13:30:00	28.48	20.0928	312			
2081-07-21 14:00:00	9.76	24.111	630			

Figure 24: Merged Data set

Bill No: [illegible]

Consumer Number: [illegible] **Consumer ID :** [illegible]

Consumer Name: **.BHAKTAPUR CHARGING STATION**

Address:

Tariff: LT [illegible]

Meter Number: [illegible]

Meter Pres Status: 001

Approved Load :

Instl Trans Capacity : 100 (PRIVATE)

MF KWH / MF KVA:1 / 1

Meter Prev Status : 001

Connection Status: Live

Trans Loss: N

Slot(Rate)	Curr.Rdg	Prev.Rdg	Units	TL Units	Bill Units	Rec.Demand	Voltage	Current
T1(8.7)	**22787**	11893	10894	0	10894	141.69	232	60
T2(6.9)	**16953**	9125	7828	0	7828	106.14	229	60
T3(6.9)	**3671**	1362	2309	0	2309	80.3	229	60

Prev. Demand: 0 **Curr. Demand:0** Excess Demand:0 PF:

Energy	Ex Loss	Dem Chg	Dem Loss	Ex Dem Loss	Total	In Rupees
164723.10	0.00	0.00	0.00	0.00	164723.10	
Coin Adjustment					0.00	
Penalty/Rebate Adj		0.00 / 0.00			0.00	
Other Adjustment					0.00	
Miscellaneous Adjustment					0.00	
Current Month Charge : (A)					164723.10	164723.10
Arrears : (B)					-2.22	-2.22
					Total Amt(A+B)	**164720.88**

Figure 25: Bill Status of Charging station

ANNEX

```
#third Model_final
import os
import numpy as np
import random
import torch
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F
from collections import deque
import pandas as pd
import matplotlib.pyplot as plt
import optuna

class DataPreprocessor:
    @staticmethod
    def preprocess_ev_data(ev_file_path):
        ev_data = pd.read_excel(ev_file_path)
        ev_data.fillna(0, inplace=True)

        ev_data['Normalized KWh'] = ev_data['KWh']
        ev_data['Normalized Cost'] = ev_data['cost']
        return ev_data

    @staticmethod
    def preprocess_wind_data(wind_file_path):
        wind_data = pd.read_excel(wind_file_path)
        wind_data.fillna(0, inplace=True)
        wind_power = wind_data['Active Energy Export
(A)(kWh)'].values.astype(float)
        return wind_power

class EVChargingEnv:
    def __init__(self, max_rate, ev_kw_usage, ev_revenue_collected,
wind_power,
                 grid_capacity=250, battery_capacity=200, with_ess=True,
with_der=True):
        self.max_rate = max_rate
        self.battery_capacity = battery_capacity
        self.ev_kw_usage = ev_kw_usage
        self.ev_revenue_collected = ev_revenue_collected
        self.wind_power = wind_power
        self.grid_capacity = grid_capacity
```

```
        self.with_ess = with_ess
        self.with_der = with_der

        self.time_steps = min(len(ev_kw_usage), len(wind_power))
        self.reset()

    def reset(self):
        self.current_time = 0
        self.battery_soc = 0
        self.total_reward = 0
        self.total_revenue = 0
        self.total_generation_cost = 0
        self.total_energy_supplied = 0
        self.total_energy_demanded = 0
        self.energy_supplied_list = []
        self.energy_demanded_list = []
        return self.get_state()

    def get_energy_price(self):
        hour = self.current_time % 24
        if 17 <= hour < 23: return 7
        elif 23 <= hour or hour < 5: return 3.7
        else: return 5.5

    def step(self, action):
        if self.current_time >= self.time_steps:
            return None, 0, True

        action_type = action
        ev_kw_usage = self.ev_kw_usage[self.current_time]
        ev_revenue = self.ev_revenue_collected[self.current_time]
        wind_energy = self.wind_power[self.current_time] if self.with_der
else 0
        energy_price = self.get_energy_price()
        energy_demanded = ev_kw_usage
        energy_supplied = 0
        revenue = ev_revenue
        cost = 0

        if wind_energy >= energy_demanded:
            excess = wind_energy - energy_demanded
            energy_supplied = energy_demanded
            if action_type == 0:
                revenue += excess * energy_price
            elif action_type == 1:
```

```
                charge = min(excess, self.max_rate, self.battery_capacity
- self.battery_soc)
                self.battery_soc += charge
                remaining = excess - charge
                revenue += remaining * energy_price if remaining > 0 else
0
        else:
            deficit = energy_demanded - wind_energy
            energy_supplied = wind_energy
            if action_type == 2 and self.with_ess and self.battery_soc >
0:
                discharge = min(deficit, self.max_rate, self.battery_soc)
                self.battery_soc -= discharge
                energy_supplied += discharge
                deficit -= discharge
            if deficit > 0:
                grid_energy = min(deficit, self.grid_capacity)
                energy_supplied += grid_energy
                cost += grid_energy * energy_price

        reward = revenue - cost
        self.total_reward += reward
        self.total_revenue += revenue
        self.total_generation_cost += cost
        self.energy_supplied_list.append(energy_supplied)
        self.energy_demanded_list.append(energy_demanded)
        self.total_energy_supplied += energy_supplied
        self.total_energy_demanded += energy_demanded

        self.current_time += 1
        done = self.current_time >= self.time_steps
        return self.get_state() if not done else None, reward, done

    def get_state(self):
        return [
            self.battery_soc / self.battery_capacity,
            self.ev_kw_usage[self.current_time],
            self.wind_power[self.current_time] if self.current_time <
len(self.wind_power) else 0,
            self.get_energy_price() / 10
        ]

class DQN(nn.Module):
    def __init__(self, state_size, action_size):
        super(DQN, self).__init__()
```

```python
        self.fc1 = nn.Linear(state_size, 64)
        self.fc2 = nn.Linear(64, 64)
        self.fc3 = nn.Linear(64, action_size)

    def forward(self, x):
        x = F.relu(self.fc1(x))
        x = F.relu(self.fc2(x))
        return self.fc3(x)

class DDQNAgent:
    def __init__(self, state_size, action_size,
                 gamma=0.95, epsilon_min=0.01,
                 epsilon_decay=0.995, batch_size=64,
                 lr=0.001, memory_size=2000):
        self.state_size = state_size
        self.action_size = action_size
        self.memory = deque(maxlen=memory_size)
        self.gamma = gamma
        self.epsilon = 1.0
        self.epsilon_min = epsilon_min
        self.epsilon_decay = epsilon_decay
        self.batch_size = batch_size
        self.model = DQN(state_size, action_size)
        self.target_model = DQN(state_size, action_size)
        self.update_target_model()
        self.optimizer = optim.Adam(self.model.parameters(), lr=lr)
        self.total_rewards = []
        self.total_revenues = []

    def update_target_model(self):
        self.target_model.load_state_dict(self.model.state_dict())

    def remember(self, state, action, reward, next_state, done):
        self.memory.append((state, action, reward, next_state, done))

    def act(self, state):
        if np.random.rand() <= self.epsilon:
            return random.randrange(self.action_size)
        with torch.no_grad():
            return
torch.argmax(self.model(torch.FloatTensor(state))).item()

    def replay(self):
        if len(self.memory) < self.batch_size:
            return
```

```
        minibatch = random.sample(self.memory, self.batch_size)
        states = torch.tensor([x[0] for x in minibatch],
dtype=torch.float32)
        actions = torch.tensor([x[1] for x in minibatch],
dtype=torch.long)
        rewards = torch.tensor([x[2] for x in minibatch],
dtype=torch.float32)
        non_final_mask = torch.tensor([x[3] is not None for x in
minibatch], dtype=torch.bool)
        non_final_next_states = torch.tensor([x[3] for x in minibatch if
x[3] is not None], dtype=torch.float32)
        dones = torch.tensor([x[4] for x in minibatch], dtype=torch.bool)

        current_q = self.model(states).gather(1, actions.unsqueeze(1))
        next_q = torch.zeros(self.batch_size)
        if len(non_final_next_states) > 0:
            next_q[non_final_mask] =
self.target_model(non_final_next_states).max(1)[0].detach()
        target_q = rewards + self.gamma * next_q
        loss = F.mse_loss(current_q.squeeze(), target_q)

        self.optimizer.zero_grad()
        loss.backward()
        self.optimizer.step()

        if self.epsilon > self.epsilon_min:
            self.epsilon *= self.epsilon_decay

    def train(self, env, episodes=3000):
        for episode in range(episodes):
            state = env.reset()
            done = False
            while not done:
                action = self.act(state)
                next_state, reward, done = env.step(action)
                self.remember(state, action, reward, next_state, done)
                state = next_state
                self.replay()
            self.update_target_model()
            self.total_rewards.append(env.total_reward)
            self.total_revenues.append(env.total_revenue)
            print(f"Episode {episode+1}: Reward={env.total_reward:.2f},
Revenue={env.total_revenue:.2f}")
```

```
    def save_model(self, path):
        torch.save(self.model.state_dict(), path)

    def load_model(self, path):
        self.model.load_state_dict(torch.load(path))
        self.model.eval()

def run_trained_simulation(env, agent):
    state = env.reset()
    done = False
    while not done:
        with torch.no_grad():
            action =
torch.argmax(agent.model(torch.FloatTensor(state))).item()
        next_state, _, done = env.step(action)
        state = next_state if not done else None
    print(f"Final Simulation Results:\n"
          f"Total Reward: {env.total_reward:.2f}\n"
          f"Total Revenue: {env.total_revenue:.2f}\n"
          f"Energy Supplied/Demanded Ratio:
{env.total_energy_supplied/env.total_energy_demanded:.2%}")

def objective(trial):
    params = {
        'gamma': trial.suggest_float("gamma", 0.9, 0.999),
        'epsilon_min': trial.suggest_float("epsilon_min", 0.01, 0.1),
        'epsilon_decay': trial.suggest_float("epsilon_decay", 0.99,
0.9999),
        'batch_size': trial.suggest_int("batch_size", 32, 256),
        'lr': trial.suggest_float("lr", 1e-5, 1e-2, log=True),
        'memory_size': trial.suggest_int("memory_size", 1000, 10000)
    }

    agent = DDQNAgent(state_size=4, action_size=3,**params)
    agent.train(env, episodes=5)
    return np.mean(agent.total_rewards)

    total_reward = 0
    for _ in range(5):  # Reduced episodes for faster trials
        state = env.reset()
        done = False
        while not done:
            action = agent.act(state)
            next_state, reward, done = env.step(action)
            agent.remember(state, action, reward, next_state, done)
```

```
            state = next_state
            agent.replay()
        total_reward += env.total_reward
    return total_reward / 5

if __name__ == '__main__':
    # Load and preprocess data
    ev_data =
DataPreprocessor.preprocess_ev_data("/content/drive/MyDrive/Colab
Notebooks/Book1122.xlsx")
    wind_power =
DataPreprocessor.preprocess_wind_data("/content/drive/MyDrive/Colab
Notebooks/refined1122.xlsx")

    # Create environment for optimization
    env = EVChargingEnv(
        max_rate=150,
        ev_kw_usage=ev_data['KWh'].values,
        ev_revenue_collected=ev_data['cost'].values,
        wind_power=wind_power,
        battery_capacity=200,
        grid_capacity=250
    )

    # Hyperparameter optimization
    study = optuna.create_study(direction='maximize')
    study.optimize(objective, n_trials=20)

    # Train with best params
    best_params = study.best_params
    print(f"Best hyperparameters: {best_params}")

    best_agent = DDQNAgent(
        state_size=4,
        action_size=3,
        **best_params
    )
    best_agent.train(env, episodes=3000)

    window = max(1, len(best_agent.total_rewards) // 10)  # Changed agent
-> best_agent
    moving_avg_rewards =
pd.Series(best_agent.total_rewards).rolling(window).mean()
    moving_avg_revenues =
pd.Series(best_agent.total_revenues).rolling(window).mean()
```

```python
plt.figure(figsize=(10, 5))
plt.plot(moving_avg_rewards, label='Moving Avg Reward')
plt.plot(moving_avg_revenues, label='Moving Avg Revenue')
plt.legend()
plt.title("Training Performance")
plt.show()

# Save and run simulation
best_agent.save_model("optimized_model.pth")
run_trained_simulation(env, best_agent)
```