CHAPTER 1
SYNOPSIS

1. Synopsis

1.1 Project Title

Solar Power forecast using Machine Learning

Team Members:

1. Aniket
2. Mrigais
3. Dhananjay Jha
4. Harjit Singh

1.2 Internal Guide

Dr. Sunil Rathod

1.3 Technical Keywords

- Solar Power
- Machine Learning
- Prediction Model
- Data Analytics
- Algorithms

1.4 Problem Statement

- To develop a system for accurately predicting the power forecasting of solar power for PV systems using machine learning for a short period of time.
- The system comprises of machine learning modules which will predict the power generation by solar PV systems by taking into consideration different meteorological and weather related parameters.

1.5 Abstract

It discusses the theoretical assumptions and design aspects of developing a Model which will predict the solar power generation beforehand. The project aims at promoting the use of renewable source of energy by developing a model which will accurately predict the solar power generation. Climate change and energy crisis have motivated us to make use of renewable non conventional source of energy. So developing a model to predict the power generation using various Machine Learning Algorithms will be beneficial to both Industries and Residents.

1.6 Goals and Objectives

- The main objective is to benchmark different forecasting techniques of solar PV panel energy output. Towards this end, machine learning and statistical techniques can be used to dynamically learn the relationship between different weather conditions and the energy output of PV systems.
- This is being done to optimize the energy structure and improve the performance of a PV system.
- Accurate prediction of PV power output is required to make better generation plans, support the spatial and temporal compensation, and achieve coordinated power control, so that the need for energy storage capacity and operating costs can be reduced.
- Our aim is to investigate the future engineering methodologies, which can be used to increase the overall prediction accuracy.
- We will be using various techniques to train models on solar irradiance data and different meteorological parameters to forecast solar irradiance, and therefore power, for different forecasting horizons in the short-term future.

1.7 Relevant Mathematics related to the Project

LSTM Recurrent Neural Network Algorithm and Mathematics

The basic RNN that we have implemented has the structure below
Steps:

1. Implement the calculations needed for one time-step of the RNN.
2. Implement a loop over $\mathbf{T_x}$ time-steps in order to process all the inputs, one at a time Here $\mathbf{a^{(i)}}$, $\mathbf{x^{(i)}}$, $\mathbf{y^{(i)}}$ represents $\mathbf{i^{th}}$ activation function, training example input and target output respectively.

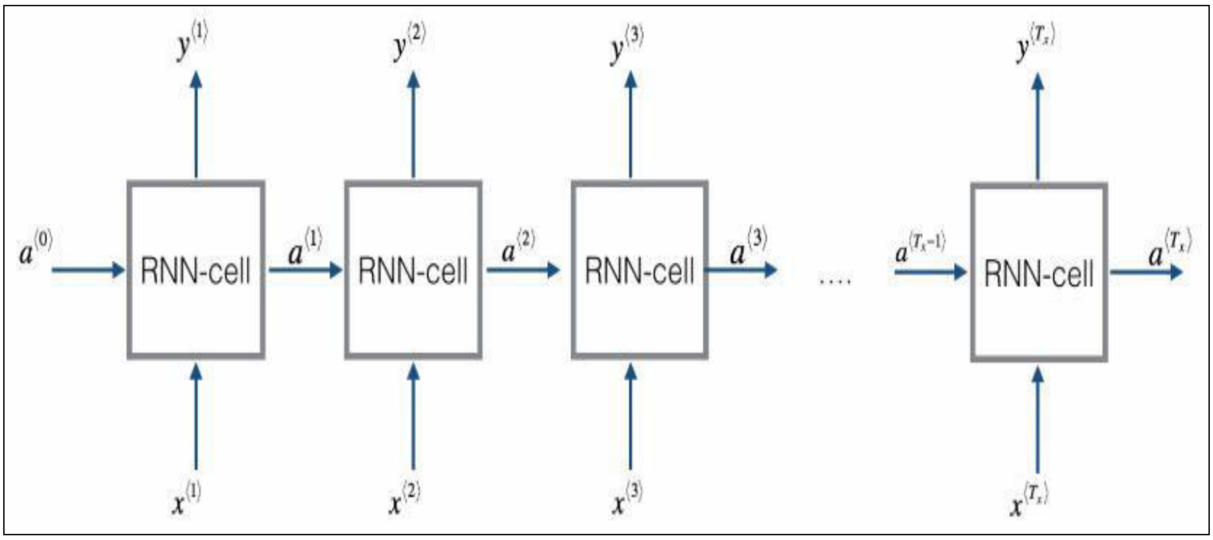


Figure 2: RNN Working

A Recurrent neural network can be seen as the repetition of a single cell. First we have implemented the computations for a single time-step. The following figure describes the operations for a single time-step of an RNN cell.

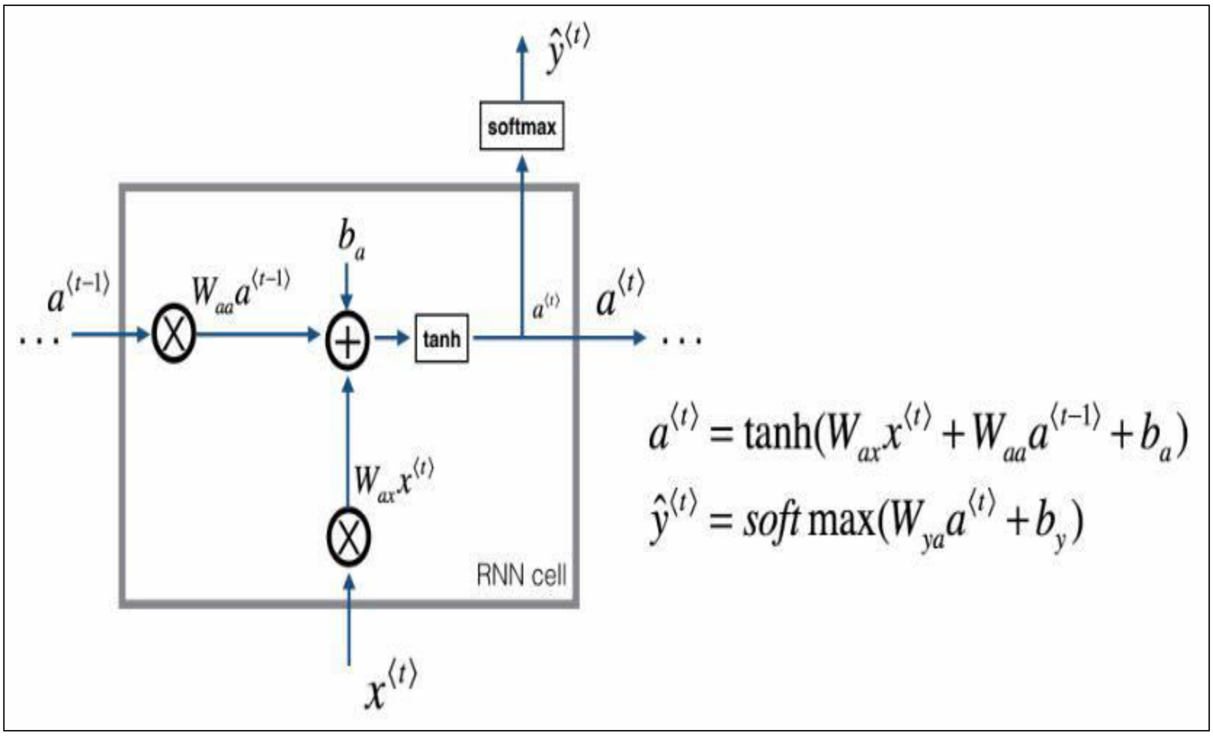


Figure 3: RNN Cell

1. Compute the hidden state with tanh activation: $\mathbf{a^{<t>}=tanh(W_{ax}x^{<t>}+W_{aa}a^{<t-1>}+b_a)}$
2. Using the new hidden state $\mathbf{a^{<t>}}$, compute the prediction $\mathbf{y^{<t>}=softmax(W_{ya}\ a^{<t>}+b_y)}$
3. Here $\mathbf{W_{ax}}$ is set of weather parameters governing the connection from x to the hidden layer.
4. $\mathbf{W_{aa}}$ is vectorized weather parameter for horizontal connection and $\mathbf{W_{ya}}$ governs the output prediction. What this notion notation means is to just take the two vectors and stack them together.
5. $\mathbf{b_a}$ on top indicates a bias used for computing activation output.
   softmax function outputs a vector that represents the probability distributions of a list of potential Solar Power outcomes.

Including an LSTM layer vastly improved performance, while the nonlinear hyperbolic tangent and sigmoid layers exhibited lower errors than standard linear activation functions.

The optimized neural network comprised three hidden layers (one LSTM) in addition to an input and output layer [6]. The introduction of a nonlinear hidden layer and an LSTM layer were each found to greatly increase the accuracy of test predictions [7]. Xavier-He initialization was utilized to select an ideally distributed initial value for the RNN weights. The ‘adam’ optimizer combined the benefits of both RMSProp and AdaGrad in adaptive moment estimation 20% dropout rate to effect regularization [5]. A mean-squared loss function was also used to train the RNN to maintain consistency.

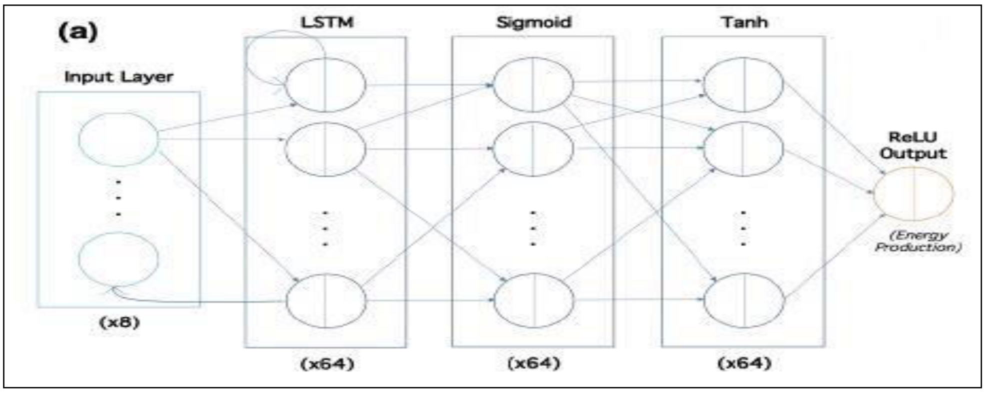
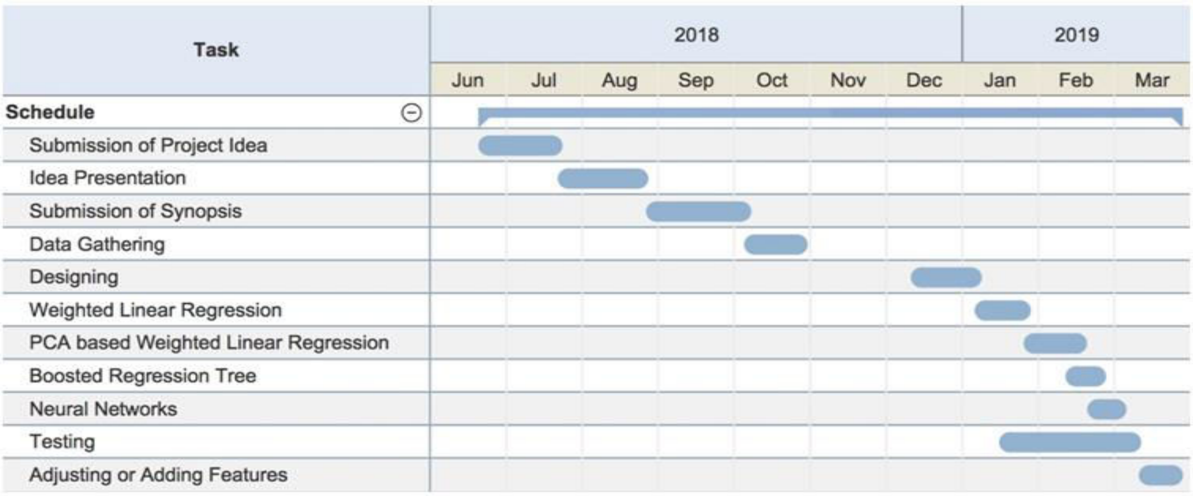


Figure 4: Depiction of flow of hidden layers in optimized neural network

1.8 Journals we used to publish our paper

1. IJSART
2. IJRASET

1.9 Plan Of Project Execution



1.10 Literature Survey

Predicting Solargeneration from weather forecasts

Authors: N. Sharma, P. Sharma, D. Irwin, and P. Shenoy

1. Deep Learning for Solar Power Forecasting
   Authors: Andr´e Gensler, JanoschHenze, Bernhard Sick

2. Short-term Power forecast of Solar PV Systems
   Authors: MayukhSamanta, Bharath K. Srikanth, J B Yerrapragada

1.11 Requirements

Software Requirements

- Anaconda 5.2
- Flask Server
- HTML, CSS
- Python 3
- OS- Windows7/MAC

Hardware Requirements

- RAM 8GB
- Processor i3 or Higher
- Hard Disk 500GB

CHAPTER 2

INTRODUCTION

2. Introduction

2.1 Project Idea

Climate change and energy crisis have led us to use renewable energy use and Solar Energy is one of the most appropriate option for use. It is renewable as well as non conventional source of energy and available in abundance. Power generated using Solar PV Panels depends on many external factors namely weather and meteorological factors. Factors such as Wind, Cloud and Rain also affect the rate of Power Generated. We will work on developing a model which will have high level of accuracy in predicting solar power. To do so we will compare various Machine Learning Algorithms and find the most accurate.

The Dataset will be divided into Training and Test Data after pre-processing and scaling and various Machine Learning Algorithms will be applied to find out the most accurate of them. The most suitable one will be applied in the model to predict the power.

The dataset used in this work is historical weather data from Amherst, MA, and is maintained by the University of Massachusetts, Amherst – Computer Science Weather Station.

2.2 Motivation Of The Project

Roof-top mounted solar photovoltaic (PV) systems are becoming an increasingly popular means of incorporating clean energy into the consumption profile of its users. It is one of the most efficient renewable source of energy which can be used over non renewable sources of energy as Fossil Fuels. There are certain influencing factors which promote the use of Solar Energy such as environment friendly and safer than traditional electricity current. The motivation behind taking up this project was to implement a model which would help people manage the energy resources in an efficient and economic way. This model can help the user to pre-plan and use the power according to the prediction made by different machine learning and statistical techniques and avoid any sorts of loss due to sudden weather changes which are not in their control. Application of this model incurs low cost for installation (economical), safer and comparatively more available than other energy resources. Electric utilities often allow the inter-connection of such systems to the grid, compensating system owners for electricity production. As the systems grow in number and their contribution to the overall load profile becomes increasingly significant, it becomes imperative for utilities to accurately account for them while planning and forecasting generation.

2.3 Major Constraints

The user should have a Smartphone device which high speed internet is accessible.

2.4 Methodologies Of Problem Solving And Efficiency Issues

Step1: User will enter the name of the city whose weather data he/she will be using for solar power forecasting.
Step2: API will fetch the weather data from open source weather forecast website.
Step3: The weather data will be fed to the machine learning algorithms.
Step4: The algorithms will then use this data to forecast the solar power.
Step5: The prediction will be displayed to the user through graphical representation.

2.4.1 Scenario In Which Modular Approach Used

Project is a top down approach which uses modular approach. City name is given as the input to the application module and output of this module is shown as the forecasting through a graphical means of representation.

2.4.2 Planed Outcome

Develop an application which would help people to know the near future solar power through most accurate predictions done by the machine learning algorithms used in the application.

2.5 Literature Survey

A similar study has already been done previously. The comparative study is given below.

PAPER NO.	PAPER NAME	ADVANTAGES	LIMITATIONS
1	N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, “Predicting solar generation from weather forecasts using machine learning.”	• 27% more accurate than existing models. • 51% better than simple approaches that only use the past to predict the future.	• It does not incorporate information from multiple weather metrics and their impact on solar intensity.
2	Gensler-Janosch, A., et al. “Deep Learning for solar power forecasting - An approach using AutoEncoder and LSTM Neural Networks.”	• Performance achieved can also be transferred to other regenerative energy sources, e.g. forecasting of wind power output. • Feature Extraction Capability.	• It needs to take into account if an overestimation or an underestimation is preferred.
3	Mayukh Samanta, Bharath Srikanth, Jayesh Yerrapragada, “Short Term Power Forecasting Of Solar PV Systems Using Machine Learning Techniques.”	• High Accuracy using Hybrid Model.	• Predicts a Non Zero Solar irradiance during period of day when there is completely no sunlight.

Table 1: Comparative Study of Previous Research

CHAPTER 3

ANALYSIS

3. Analysis

3.1 Project Estimates

3.1.1 Time Estimates

Time estimate is about 9 months

Sr. No.	Estimates	Time Taken
1	Literature Survey	2
2	Design	1
3	Presentation	1
4	Coding	3
5	Testing	1
6	Reporting	1
7	TOTAL	9

3.2 PROJECT SCHEDULE

3.2.1 Project Task Set

Major Tasks in the Project stages are:
1: Research about solar power forecasting.
2: Research about selected topic.
3: Literature survey.
4: Deciding the flow of project plan.
5: Determine the Requirement.
6: Dividing the task.
7: Formulate the code.
8: Testing.
9: Project Complete Demonstration.

3.3 Risk Management

P-class problem: P is set of all decision problems which can be solved in polynomial time by a deterministic. Since it can be solved in polynomial time, it can be verified in polynomial time. Therefore P is a subset of NP. NP-class problem:‖ NP‖ means ― we can solve it in polynomial time if we can break the normal rules of step-by-step computing‖ . NP-Hard problem: A problem is NP-hard if an algorithm for solving it can be translated into one for solving any NP-problem (nondeterministic polynomial time) problem. NP-hard therefore means ― at least as hard as any NP-problem,‖ although it might, in fact, be harder. NP-Complete problem: Since this amazing‖ computer can also do anything a normal computer can, we know that‖ P‖ problems are also in‖ NP‖ . So, the easy problems are in‖ P‖ (and‖ NP‖), but the really hard ones are only in ‖ NP‖ , and they are called ‖ NP-complete‖ . It is like saying there are things that People can do (‖ P‖), there are things that Super People can do (‖ SP‖), and there are things *only* Super People can do (‖ SP-complete‖). Our project, Nature Trail Location Based Discovery Application can be realized in P (Polynomial time).

3.3.1 Timeline Chart

3.4 Requirement Analysis

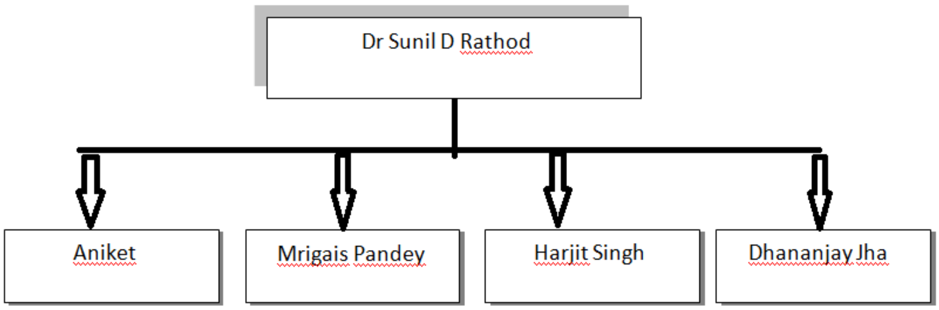
3.4.1 Hardware Requirements

Sr.No	Parameters	Recommended requirements	Justification
1	Android Mobile Phone	Kitkat version	Must not be less than this
2	CPU Speed	1.2 GHz	Speed must be equal or more than this
3	RAM	1 GB	Must not be less than this

3.4.2 Software Requirements

- Anaconda 5.2
- Python 3.6
- HTML, CSS
- Flask
- Open source API key of *openweathermap.org* and *darksky.net*

3.5 Team Organization



CHAPTER 4

DETAILED DESIGN DOCUMENT

4. Detailed Design Document

4.1 Introduction

This topic specifies the interpreted and expected implementation of the project that should facilitate efficiency and responsiveness of the application to the user.

4.2 Architectural Design

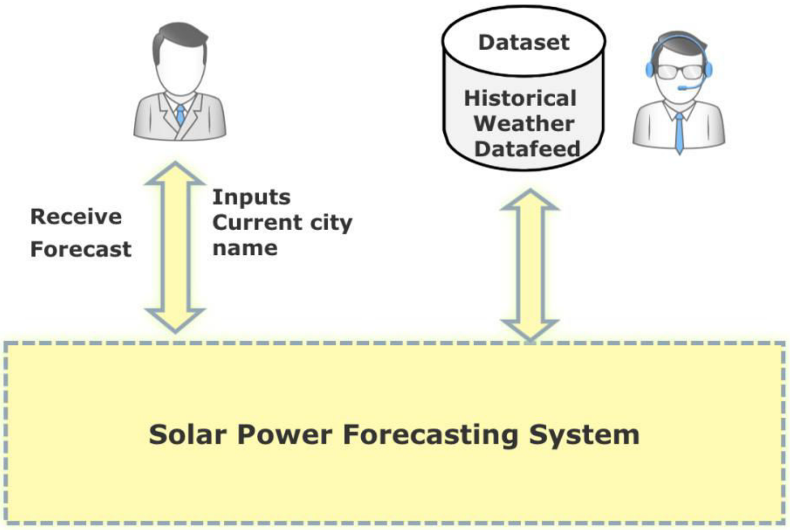


Fig:- Architectural Design of Project

The user inputs the name of the city he wants weather details of in the solar power forecasting system. This system includes the machine learning algorithms that will be used for forecasting. When the API fetches the data from historic weather data, this data is fed to the solar power forecasting system .Now the user receives the forecast from the solar power forecasting system in a graphical form.

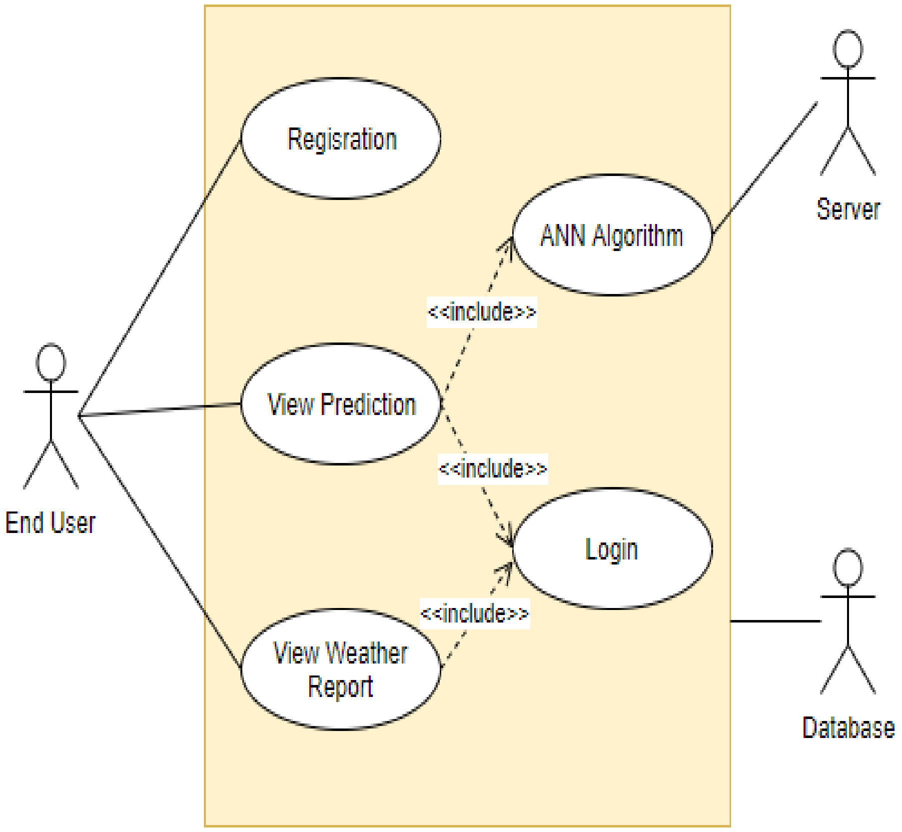
CHAPTER 5

MODELING

5. Modeling

5.1 Usage Scenario

5.1.1 Use Case Diagram

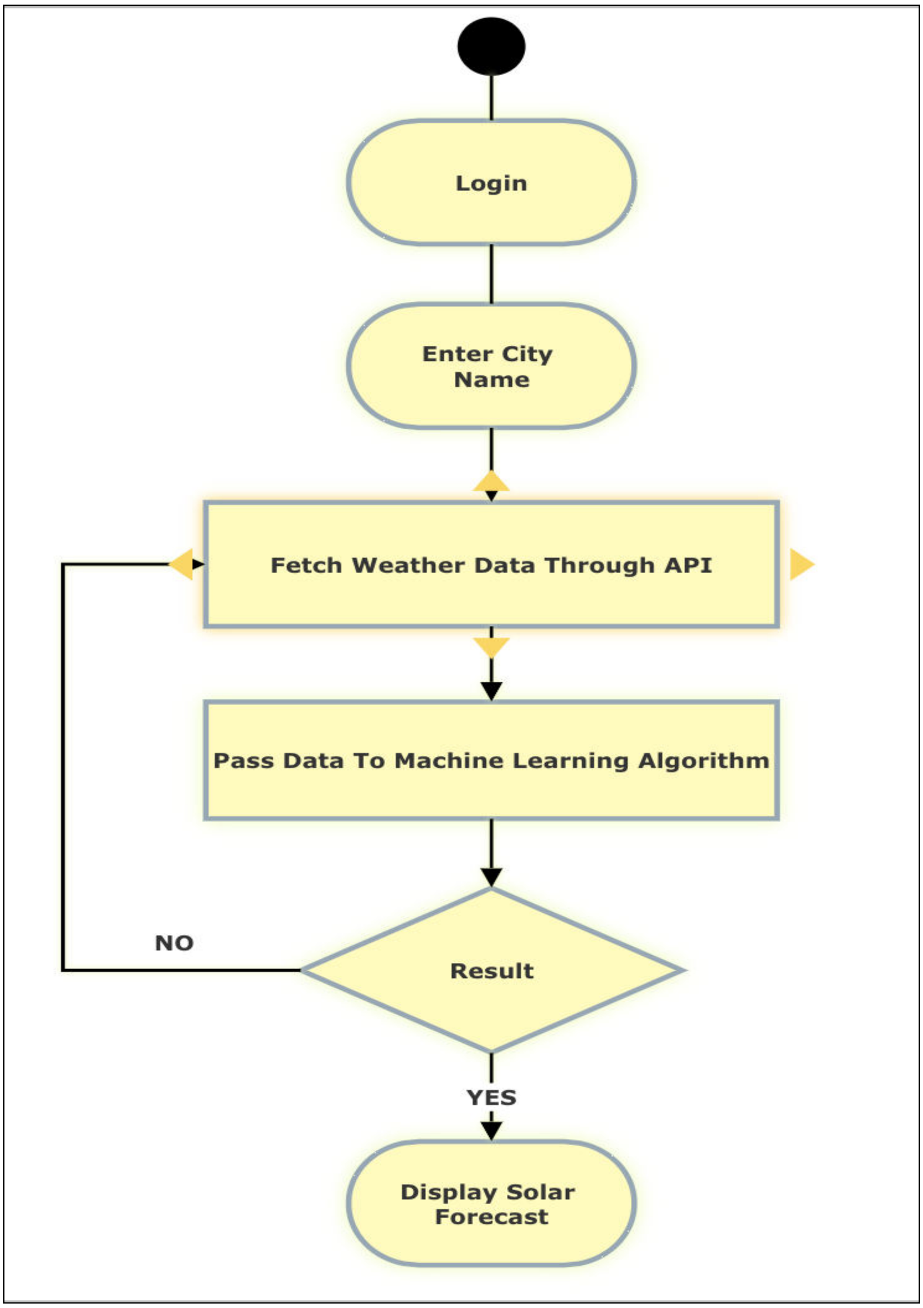


The use-case diagram clearly depicts three actors in the system. Two human actors and one non-human actor. The two human actors are the Admin and the End user. The Admin feeds the dataset to the machine learning which then prepares, cleans, and analyzes the data. All this data preprocessing techniques include data correction, correction of errors and application of neural network algorithms.

Then the use-case ML Model forecasts the solar power to the use-case End User. This forecast is provided to the use-case End User through graphical representation.

5.2 Functional Model and Description

5.2.1 Data Flow Diagram



CHAPTER 6

CODING

6. Coding

6.1 Implementation

The project demonstrates a web based Solar Power Predictor Application hosted on Flask server. Firstly we trained our model using Long Short Term Memory, Recurrent Neural Network ML Algorithm. The activation functions used in the hidden layer of RNN are Sigmoid and Tanh. The model was trained using the dataset obtained from repository mentioned in Section III of this paper. After training the algorithm we dumped our model in Pickle. To do the prediction of Solar Power user can login into web based app and select the city name for which prediction has to be done. After selecting the city, the API key provided by openweathermap.org and darksky.net will fetch weather data in JSON format. In the background this fetched data will be scaled using the scale factor obtained while training the algorithm. After scaling the data will be passed to the pickle and the output predicted by the model will again be scaled to KWh and displayed as the amount of Solar Polar generated for the queried city.

6.2 Tools and Technologies Used

6.2.1 Software's Used

- Anaconda Navigator for Machine Learning Implementation
- PyCharm for User Interface Implementation
- Flask as a Server
- DB Browser for SQLite
- Web Browser for User Interface Interaction

6.2.2 Hardware Specification

- The Hardware used to code the Project was in Macbook Pro

6.2.3 Programming Languages

- Python for Machine Learning
- HTML and CSS for Web Interface

6.3 Algorithms

6.3.1 Machine Learning Development

RNN.py

The class contains code for the execution of Machine Learning Algorithm. It has following functions:

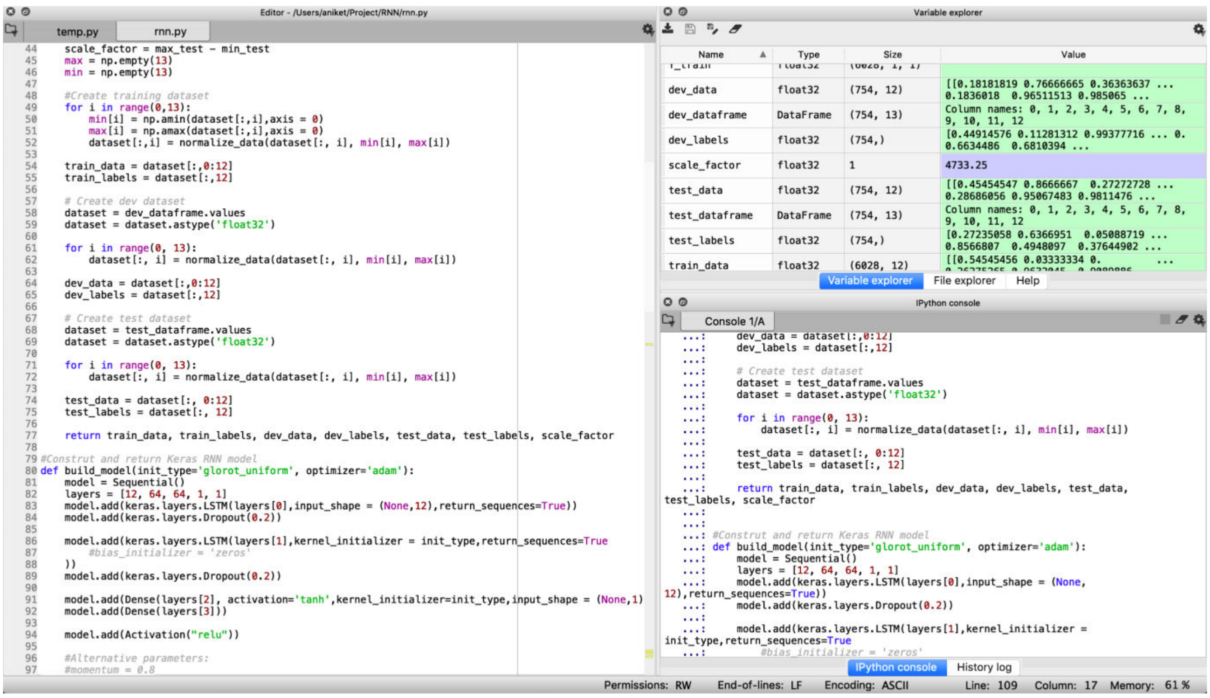
I. import_data()
This function imports the Training, Development and Test Dataset which are in the ratio of 8:1:1 respectively.

II. normalize_data()
This function normalizes the Dataset using Min Max Normalization and Scales them using the Scaling factor.

III. build_model()
This function creates a LSTM RNN Model using the Training Dataset. It also uses the activation functions ReLu, Tanh, Sigmoid.

IV. write_to_csv()
Writes the predicted the Solar Power Output to a CSV File.

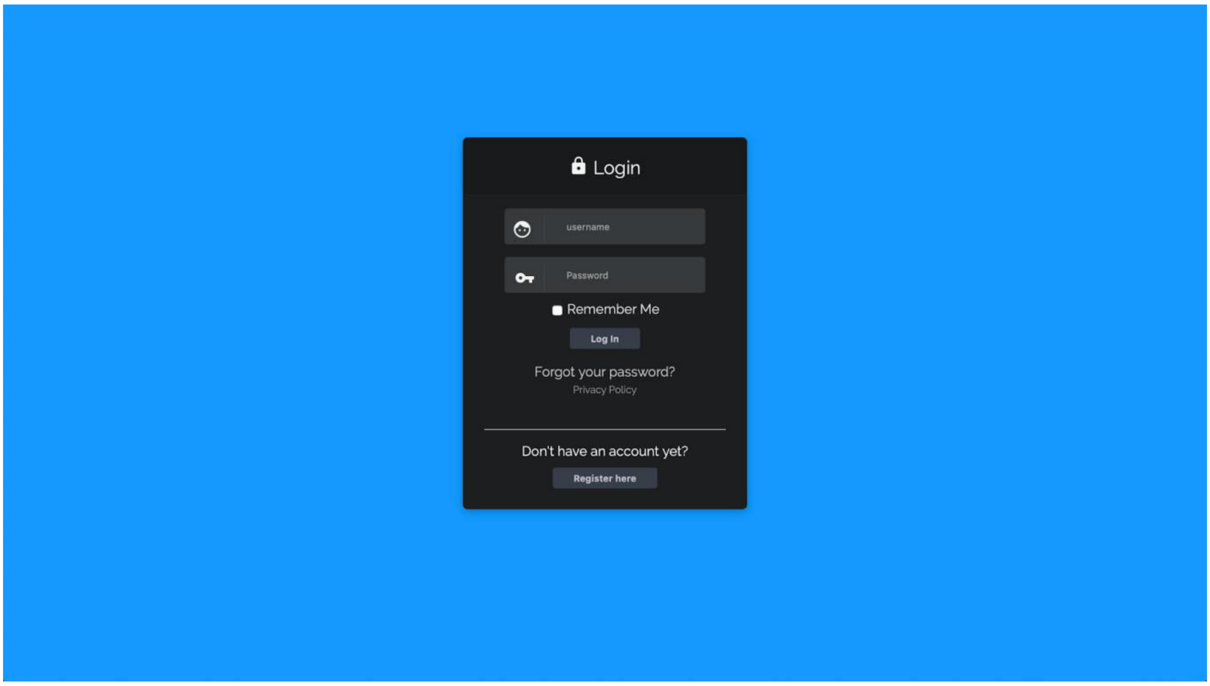
V. mse()
Calculates the Mean Square Error for Accuracy calculation.



6.3.2 User Interface Development

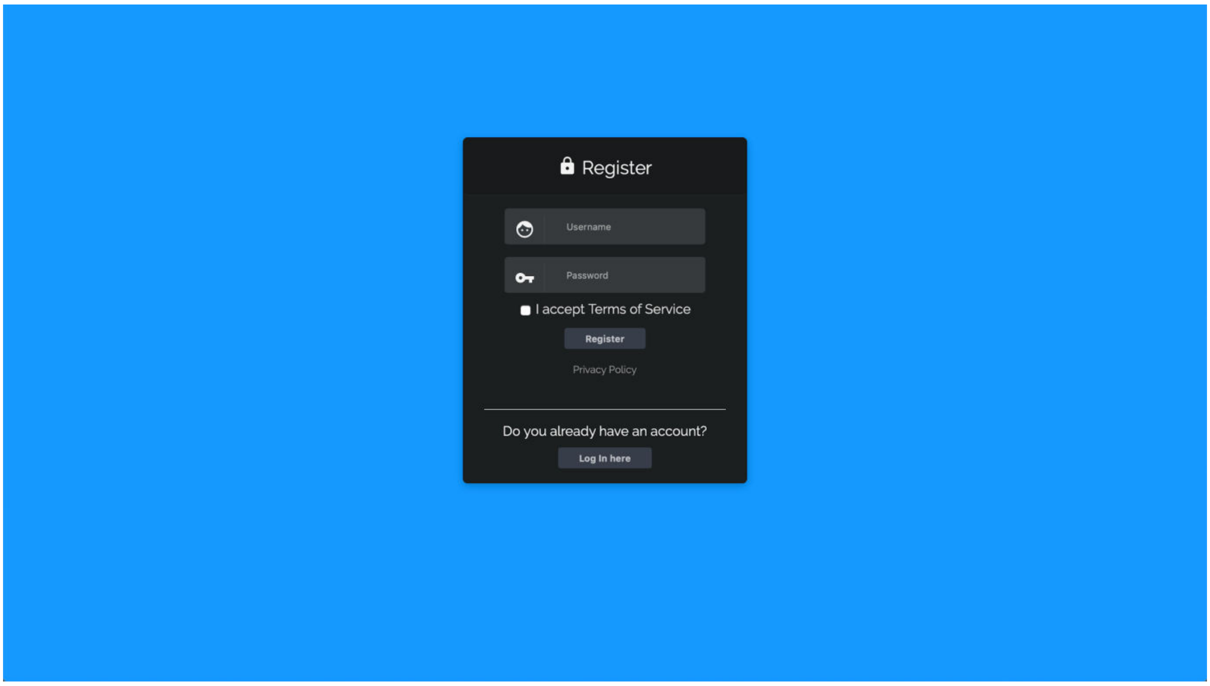
Login.html

This Page allows user to enter details to Log into the System.



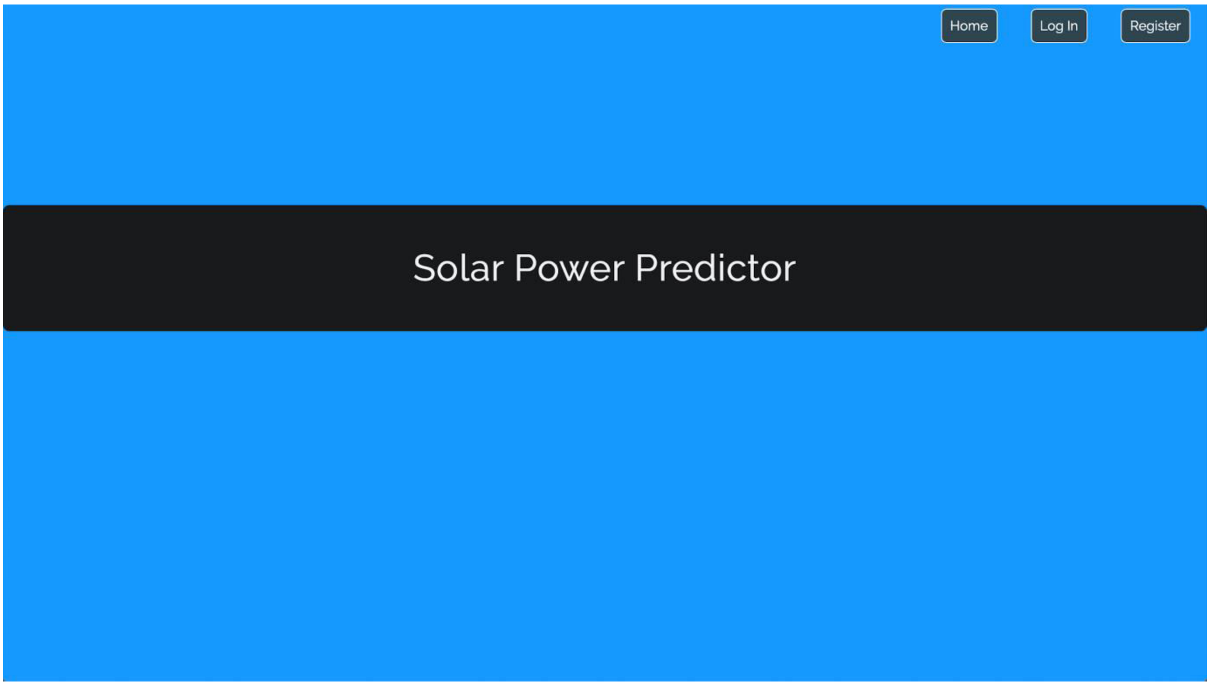
Registration.html

This page allows user to register for a new account.



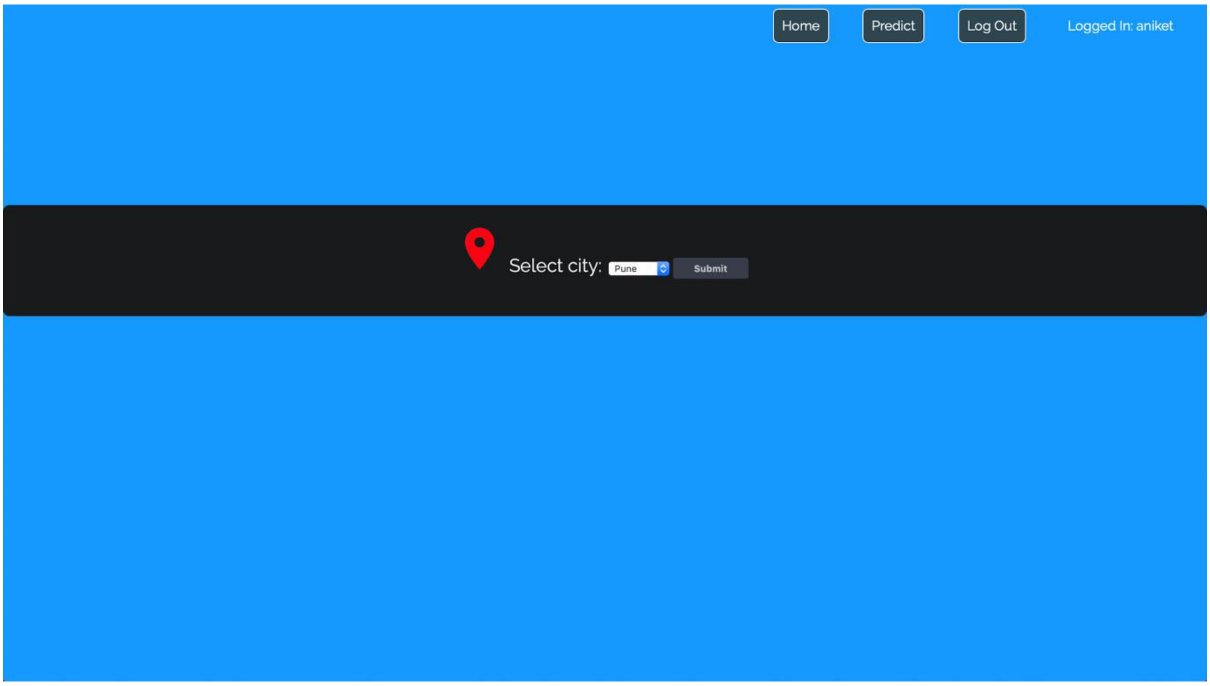
Home.html

This is the Home Page of the Application



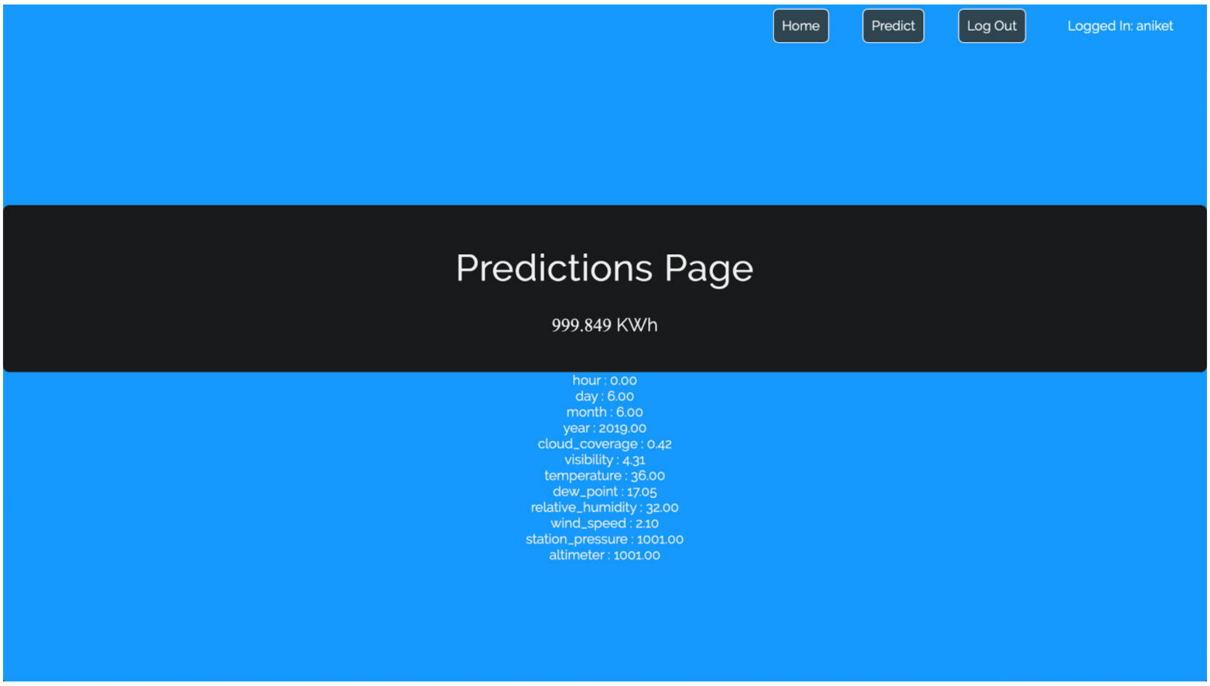
Index.html

This Page allows user to select a City Name for Solar Power Prediction.



Predict.html

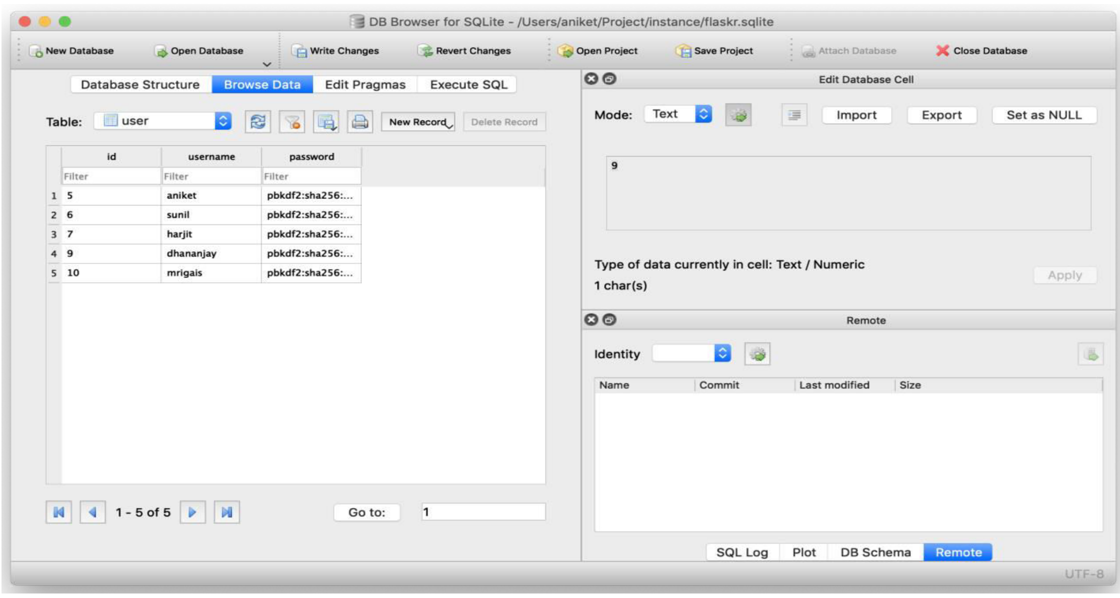
This Page displays the Solar Power Prediction and Weather Data fetched



6.3.3 Database

db.py

Before we get information from the database the database must be connected to first. It must be checked for its version and must be updated if required. This class extends __init__.py in order to connect with the database provided. The name of the database is stored in the database_name variable.The version plays an important role to check and update the version of the database that is being used. The Password here is in encrypted form because Flask initializes its database in that format.



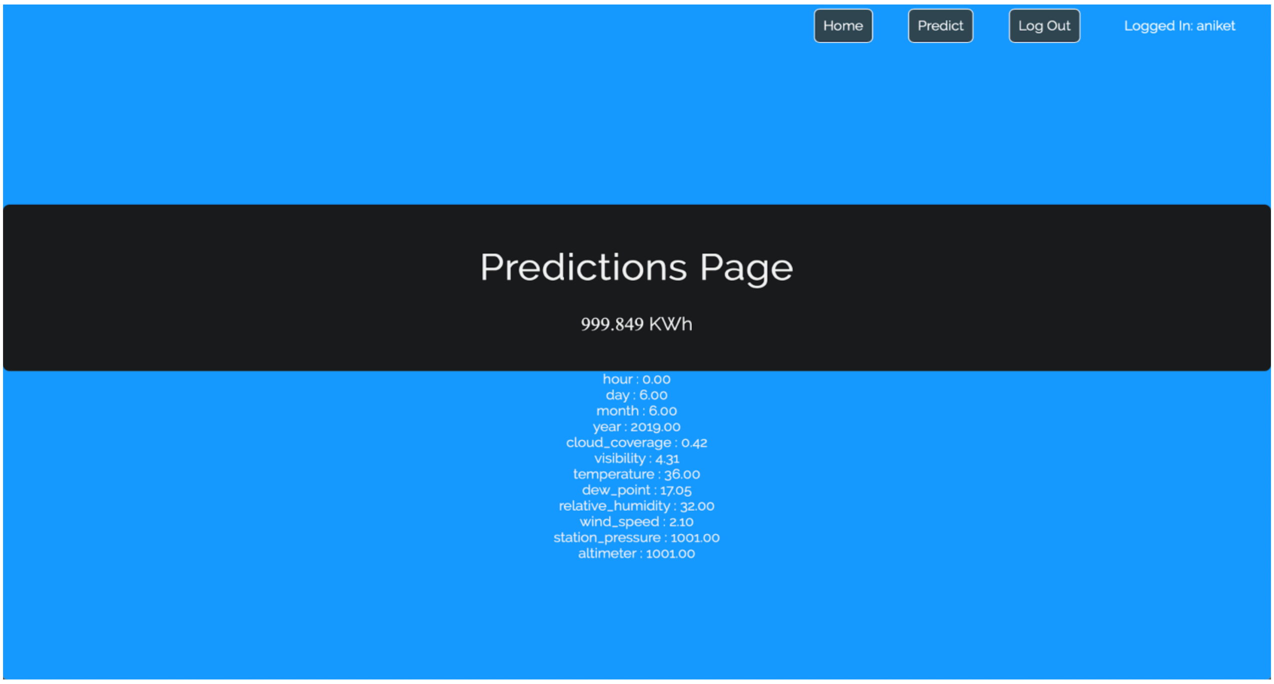
CHAPTER 7

RESULTS

7. Results

7.1 Outcome

This page contains the Predicted Solar Power and Real Time Weather data of the queried city.



CHAPTER 8

TESTING

8. Testing

8.1 Introduction

- When a person is unfamiliar of the weather condition of a particular area and wants to know and manage the solar power they'll be having for upcoming days, it is important for them to have the knowledge of solar power in near future so that they can keep a track of power consumption and manage it. This forecast enhances the way a person will manage the solar power usage in their household while keeping a track of the weather condition for upcoming days. Roof-top mounted solar photovoltaic (PV) systems are becoming an increasingly popular means of incorporating clean energy into the consumption profile of its users. It is one of the most efficient renewable sources of energy which can be used over non renewable sources of energy like Fossil Fuels.

- Initially the screen will show the user information and by entering the city name, the API key will fetch the weather data from ***openweathermap.org.***This data will be fed to the ML algorithm which in return will forecast the solar power that will be generated in upcoming days.

8.2 **Test Cases**

Test ID	1
Test Case	Building RNN Model using Training Model
Unit Test	Machine Learning
Assumptions	i. Dataset is imported ii. Dataset is normalized and scaled
Test Data	i. Weather Data Parameters and Solar Power
Steps To Be Executed	i. Building the RNN Model ii. Smoothing the Parameters using Development Dataset
Expected Result	Trained Solar Power Predictor RNN Model
Actual Result	Trained Solar Power Predictor RNN Model
Verdict	Test Case Passed.

Test ID	2
Test Case	Testing the Model for accuracy using Test Dataset
Unit Test	Machine Learning
Assumptions	i. Dataset is already imported ii. Dataset is normalized and scaled iii. Model is already Trained
Test Data	i. Weather Data Parameters
Steps To Be Executed	i. Passing Test Weather Dataset to the Model ii. Calculating the Mean Square Error for Accuracy
Expected Result	Accurate Test Prediction
Actual Result	Accurate Test Prediction
Verdict	Test Case Passed.

Test ID	3
Test Case	Loading RNN Pickle Model on Server for prediction
Unit Test	Flask Server Service
Assumptions	i. Flask Server is initiated ii. Model is dumped in Pickle
Test Data	i. Weather Data Parameters
Steps To Be Executed	i. Reshape the Weather Data in 3D Array ii. Normalize the Array using Scale Factor
Expected Result	The Solar Power generated prediction from the model
Actual Result	The Solar Power generated prediction from the model
Verdict	Test Case Passed.

Test ID	4
Test Case	To verify clicking the Registration button creates new User Account
Unit Test	Authorization System
Assumptions	Client is connected to the server
Test Data	i. Username ii. Password
Steps To Be Executed	i. Enter the Username ii. Enter the Password iii. Click on Registration
Expected Result	i. User should be able to Login ii. User should be able to view the Predictions
Actual Result	i. User should be able to Login ii. User should be able to view the Predictions
Verdict	Test Case Passed.

Test ID	5
Test Case	To verify clicking the Login button logs into the system
Unit Test	Authorization System
Assumptions	Client is connected to the server
Test Data	i. Username ii. Password
Steps To Be Executed	i. Enter the Username ii. Enter the Password iii. Click on Login
Expected Result	User should be able to access Prediction Page
Actual Result	User should be able to access Prediction Page
Verdict	Test Case Passed.

Test ID	6
Test Case	Fetching Real Time Weather Data using API
Unit Test	API Service
Assumptions	i. Flask Server is initiated ii. Internet Connection is available
Test Data	i. Weather Data Parameters
Steps To Be Executed	i. Requesting Weather Data using City Name ii. Dumping the JSON File Data into Array
Expected Result	Displaying the Real Time Weather Data on Prediction Page
Actual Result	Displaying the Real Time Weather Data on Prediction Page
Verdict	Test Case Passed.

Test ID	7
Test Case	Predicting Solar Power
Unit Test	API and Flask Service
Assumptions	i. Flask Server is initiated ii. Internet Connection is available
Test Data	i. Real Time Weather Data Parameters
Steps To Be Executed	i. Requesting Weather Data using City Name ii. Dumping the JSON File Data into Array iii. Passing the Array to the model for Prediction
Expected Result	Displaying the Solar Power on Prediction Page
Actual Result	Displaying the Solar Power on Prediction Page
Verdict	Test Case Passed.

Test ID	8
Test Case	Validating creation of Database
Unit Test	Query and Database implementation
Assumptions	Database Browser of SQLite to view Table after creation
Test Data	i. Query to create Database ii. Name of the Database
Steps To Be Executed	Initiate the db.py using Flask
Expected Result	Database created.
Actual Result	Database created.
Verdict	Test Case Passed.

Test ID	9
Test Case	Validating creation of Table
Unit Test	Query and Database implementation
Assumptions	Database Browser of SQLite to view Table after creation
Test Data	For User Account Table i. ID ii. Username Iii. Password
Steps To Be Executed	Initiate the schema.sql using Flask
Expected Result	Table created.
Actual Result	Table created.
Verdict	Test Case Passed.

CHAPTER 9

CONCLUSION

9. Conclusion

This model will help user predict the Solar Power Generation. It will guide the user through unfamiliar situation which can occur so that he could save power prior itself.

Currently there is not much use of Solar Power in India but once the people start realizing the importance of renewable sources of energy, they will eventually adopt this prediction model to conserve the Solar Energy.

It will also help in promoting use of renewable source of energy.

CHAPTER 10

REFERENCES

References

[1] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning," in Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on, pp. 528–533, IEEE, 2011.

[2] Gensler-Janosch, A., et al. "Deep Learning for solar power forecasting — An approach using AutoEncoder and LSTM Neural Networks." 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016.

[3] Mayukh Samanta, Bharath Srikanth, Jayesh Yerrapragada, "Short Term Power Forecasting Of Solar PV Systems Using Machine Learning Techniques".

[4] http://s35695.mini.alsoenergy.com/Dashboard/2a5669735065572f4a42454b772b714d3d

[5] https://machinelearningmastery.com/use-keras-deep-learning-models-scikit-learn-python/.

[6] Adele Kuzmiakova, Gael Colas, Alex McKeehan, "Short-term Memory Solar Energy Forecasting at University of Illinois"

[7] Sunil Rathod, Aniket, Mrigais Pandey, Dhananjay Jha, Harjit Singh, "Power forecast of Solar Panels using Machine Learning Techniques: A Survey" in IJSART- Volume 4 Issue 10 – October 2018

[8] https://github.com/sborgeson/local-weather.

[9] NOAA website : https://www.ncdc.noaa.gov/cdo-web/datatools/l cd

[10] www.openweathermap.org

[11] www.darksky.net

PART I ANNEXURE A

LABORATORY ASSIGNMENTS ON PROJECT ANALYSIS OF ALGORITHMIC DESIGN

1. What is IDEA Matrix?

Ideas are delicate. To bring them to life, one needs to have a strategy that Develops them well. Ideas need to be transformed from concept to reality through a brooding and development process. We need an IDEA matrix for transforming Concepts into reality. The processes outlined below interact at so many levels and the lines are not clear-cut between them. It is possible to go from the I stage to the E stage, then back to the D stage then the A stage and vice-versa. That‘s the beauty of nurturing ideas. In this project, we are using IDEA Matrix for:

I	**D**	**E**	**A**
Increase	Define	Experiment	Accelerate
Improve	Deliver	Evaluate	Associate
Ignore	Decrease	Eliminate	Avoid

Table A1: IDEA Matrix

IDEA Matrix(Elaborated):

Increase: Efficiency, Reliability, Convenience	Define: Protocol Procedure Methodologies	Experiment: Different ways to deploy application locally.	Accelerate: Updating the application.
Improve: Connectivity between application and database.	Deliver: Location specific data and relevant data.	Evaluate : User experience and client feedback on application.	Associate: With the problem faced by the user and improve on it.
Ignore: Inconvenient and abnormal suggestions which divert from requirement.	Decrease: Effort of the system administrator and deploy locally.	Eliminate: Dependencies of internet and deploy locally.	Avoid: Unauthorized access of database.

Table A2: Project IDEA Matrix

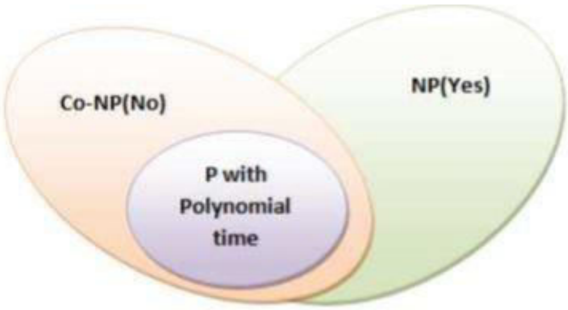
P-Complete problems:

Definition: P is the set of all decision problems solvable by deterministic algorithms in polynomial time. NP is the set of all decision problems solvable by nondeterministic algorithms in polynomial time. In complexity theory, the notion of P-complete decision problems is useful in the analysis of both: which problems are difficult to parallelize effectively • which problems are difficult to solve in limited space Formally, a decision problem is P-complete (complete for the complexity class P) if it is in P and that every problem in P can be reduced to it by using an appropriate reduction.

- What is P? • P is set of all decision problems which can be solved in polynomial time by a deterministic.
- Since it can be solved in polynomial time, it can be verified in polynomial time.
- Therefore, P is a subset of NP.

P:

We have developed a mobile system that can measure the calories of the food from that food‘s image taken by the user‘s smartphone. Once the user captures the image of the food item on the plate, the image is sent to the cloud for food recognition and calorie computation. Food recognition is done by deep learning running in the cloud: the image is recognized and the calorie details matching the image are fetched from the database that also exists in the cloud. The result is then prompted back to the user‘s phone.



What is NP?

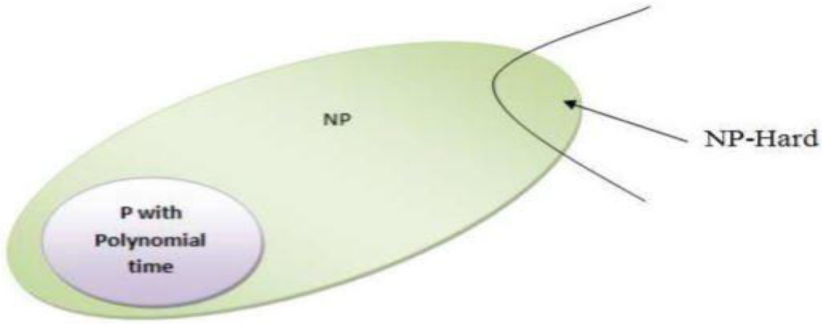
"NP" means "we can solve it in polynomial time if we can break the normal rules of step-by-step computing".

What is NP Hard?

A problem is NP-hard if an algorithm for solving it can be translated into one for solving any NP-problem (nondeterministic polynomial time) problem. NP-hard therefore means "at least as hard as any NP-problem," although it might, in fact, be harder.

NP-Hard:

The issue of calibration has been bothersome, because it is difficult for the user to take the photo with one hand on the phone and other hand‘s thumb near the plate. • Required more time.



What is NP-Complete?

Since this amazing "N" computer can also do anything a normal computer can, we know that "P" problems are also in "NP". So, the easy problems are in "P" (and "NP"), but the hard ones are *only* in "NP", and they are called "NP-complete". It is like saying there are things that People can do ("P"), there are things that Super People can do ("SP"), and there are things *only* Super People can do ("SP-complete").

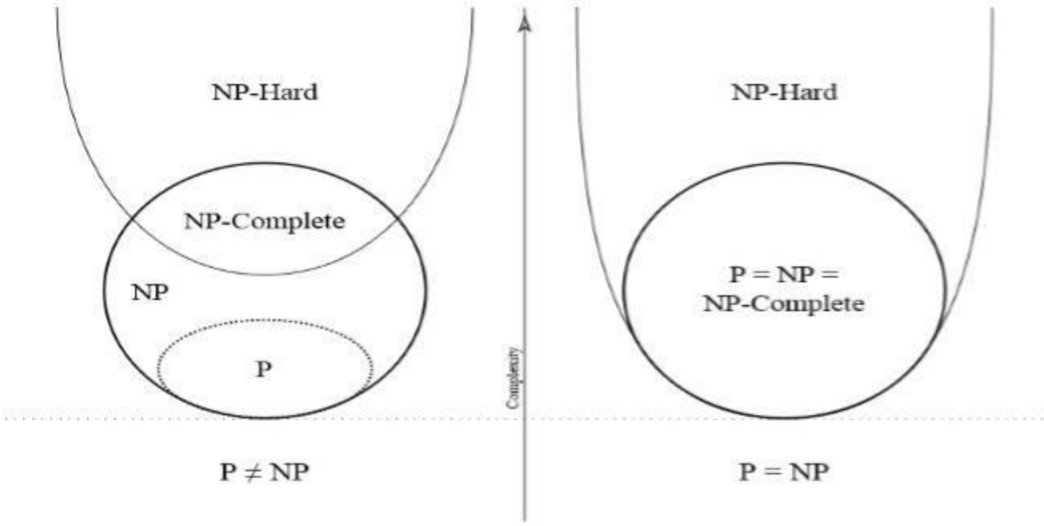
NP-Complete:

We have used Support Vector Machine (SVM) for image processing, Map Reduce for the cloud model, Volume calculation and finger calibration for calorie calculation. While our system has achieved excellent results, the issue of calibration has been bothersome, because it is difficult for the user to take the photo with one hand on the phone and other hand‘s thumb near the plate. In this work, we have addressed this issue by proposing a distance calculation method between the object on the plate and the person taking the image. By using distance calculation method, the user now does not have to keep the finger in the plate for calibration. We use deep learning to accurately train and classify the food object to its corresponding label. The system will then record this value and send it to the cloud along with the food photo captured. The image will then be processed in the cloud with the help of virtualization, and the results (including the calorie value and the food object label) will be sent back to the user on the mobile device. We have also proposed a new method whereby the application will assist the user in real time for determining the ideal distance from which the user must capture the photo.

Theory of NP-completeness:

The group of problems is further subdivided into two classes: **NP-complete:** A problem that is NP-complete can be solved in polynomial time if and only if all other NP-complete problems can also be solved in polynomial time**. NP-hard:** If an NP-hard problem can be solved in polynomial time then all NP-complete problems can also be solved in polynomial time All NP-complete problems are NP-hard but some NP-hard problems are known not to be NP complete.

NP-Complete NP-hard



Euler diagram for P, NP, NP-complete, and NP-hard set of problems. The left side is valid under the assumption that P≠NP, while the right side is valid under the assumption that P=NP (except that the empty language and its complement are never NP-complete) About the theory above, we have encountered that feasibility of project in question comes under the category of P-Complete problems for the most part. The outputs/outcomes that have been theoretic zed are expected to occur within polynomial time. The outcomes are found to be deterministic and are traceable to an extent.

Part II ANNEXURE B

LABORATORY ASSIGNMENTS ON PROJECT QUALITY AND RELIABILITY TESTING OF PROJECT DESIGN

Divide and Conquer Strategy

In computer science, divide and conquer (D&C) is an algorithm design paradigm based on multi-branched recursion. A divide and conquer algorithm works by recursively breaking down a problem into two or more sub-problems of the same (or related) type (divide), until these become simple enough to be solved directly (conquer). The solutions to the sub-problems are then combined to give a solution to the original problem

.

• Divide:

The problem i.e. problem statement P into a number of sub problems (p1, p2 . . . pn) that are themselves smaller instances of the same problem statement P we have defined. In the project, the problem of uploading a large data on VMs and crunching data after uploading data into VMs. It takes to much time and keep VMs waiting and also cost of data loading and processing is increase

• Conquer:

Recursively solving those sub problems. To solve problem of large data on VMs of cloud we provide a big data provisioning service that incorporates hierarchical and peer-to-peer data distribution techniques to speed-up data loading into the VMs used for data processing. Also some tricks over data distribution. System will dynamically mutates source of data for speed up data loading. After exploiting the above divide and conquer strategies the system found in concurrent functional dependency processing and identified following objects:

1. P2P connections:

The downside of the hierarchical approach is that it provides no fault tolerance during the transfer. If one of the VM deployments fails or the VM gets stuck after the transfers have been initiated, it is not easy to recover from failure and reschedule transfers (all the branches from the failing point need to be re-created and transfers re-started). Failure of one of the upstream leaves in the hierarchy dries the flow of data to the nodes that were supposed to be fed from there. This also implies more synchronization is required. To deal with this issue, we adopted an approach that also takes advantage of the fact that the data centre environment presents low-latency access to VMs, no NAT or Firewall issues, and no ISP traffic shaping to deliver a P2P (Bit Torrent) delivery approach for big data in the data Centre.

2. Hierarchical approach:

Semi-centralized approaches are hard to maintain, especially if new data are continuously added; centralized approaches do not scale well once you get past a few hundred VMs (in our experiments we observed that the server containing the Data starts dropping connections and overall throughput Decreases by 2-3 orders of magnitude). A next logical step would be to benefit from the knowledge IaaS providers have on the underlying network topology of the data centre . Building a relay tree where VMs get data not from the original store, but from their parent node in the hierarchy, which ideally is in the same rack. This way N VMs will access the central server to fetch data, and as soon as some blocks are downloaded by these N VMs, they will provide the blocks to N additional VMs (ideally in their same racks), and so on. This way we also confine most of the traffic within top of the rack switches and avoid more utilised routers. The VMs need to be finely configured to download the data from the right location at the right time.

Advantages

Solving difficult problems Divide and conquer is a powerful tool for solving conceptually difficult problems: all it requires is a way of breaking the problem into sub-problems, of solving the trivial cases and of combining sub-problems to the original problem.

Algorithm efficiency

The divide-and-conquer paradigm often helps in the discovery of efficient algorithms.

Parallelism

Divide and conquer algorithms are naturally adapted for execution in multi-processor machines, especially shared-memory systems where the communication of data between processors does not need to be planned in advance, because distinct sub-problems can be executed on different processors. **Memory access** Divide-and-conquer algorithms naturally tend to make efficient use of memory caches. The reason is that once a sub-problem is small enough, it and all its sub-problems can, in principle, be solved within the cache, without accessing the slow.

Memory access

Divide-and-conquer algorithms naturally tend to make efficient use of memory caches. The reason is that once a sub-problem is small enough, it and all its sub-problems can, in principle, be solved within the cache, without accessing the slow.

• Software Testing:

Testing is "the process of questioning a product in order to evaluate it", where the "questions" are things the tester tries to do with the product, and the product answers with its behaviour in reaction to the probing of the tester. Although most of the intellectual processes of testing are nearly identical to that of review or inspection, the word testing is connoted to mean the dynamic analysis of the product—putting the product through its paces. The quality of the application can and normally does vary widely from system to system but some of the common quality attributes include reliability, stability, portability, maintainability and usability. Refer to the ISO standard ISO 9126 for a more complete list of attributes and criteria. The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

Testing is a process rather than a single activity. This process starts from test planning then designing test cases, preparing for execution and evaluating status till the test closure. So, we can divide the activities within the fundamental test process into the following basic steps:

1)Planning and Control.
2) Analysis and Design.
3) Implementation and Execution.
4) Evaluating exit criteria and Reporting.
5) Test Closure activities.

1) Planning and Control:

Test planning has following major tasks:

i. To determine the scope and risks and identify the objectives of testing.

ii. To determine the test approach.

iii. To implement the test policy and/or the test strategy. (Test strategy is an outline that describes the testing portion of the software development cycle. It is created to inform PM, testers and developers about some key issues of the testing process. This includes the testing objectives, method of testing, total time and resources required for the project and the testing environments.).

iv. To determine the required test resources like people, test environments, PCs, etc.

v. To schedule test analysis and design tasks, test implementation, execution and evaluation.

vi. To determine the Exit criteria we need to set criteria such as Coverage criteria. (Coverage criteria are the percentage of statements in the software that must be executed during testing. This will help us track whether we are completing test activities correctly. They will show us which tasks and checks we must complete for a particular level of testing before we can say that testing is finished.)

1) Analysis and Design:

Test analysis and Test Design has the following major tasks:

i. To review the test basis. (The test basis is the information we need in order to start the test analysis and create our own test cases. Basically it‘s a documentation on which test cases are based, such as requirements, design specifications, product risk analysis, architecture and interfaces. We can use the test basis documents to understand what the system should do once built.)

ii. To identify test conditions.

iii. To design the tests.

iv. To evaluate testability of the requirements and system.

v. To design the test environment set-up and identify and required infrastructure and tools.

1) Implementation and Execution:

During test implementation and execution, we take the test conditions into test cases and procedures and other test ware such as scripts for automation, the test environment and any other test infrastructure. Test implementation has the following major task:

i. To develop and prioritize our test cases by using techniques and create test data for those tests. (In order to test a software application you need to enter some data for testing most of the features. Any such specifically identified data which is used in tests is known as test data.) We also write some instructions for carrying out the tests which is known as test procedures. We may also need to automate some tests using test harness and automated tests scripts. (A test harness is a collection of software and test data for testing a program unit by running it under different conditions and monitoring its behavior and outputs.)

ii. To create test suites from the test cases for efficient test execution. (Test suite is a collection of test cases that are used to test a software program to show that it has some specified set of behaviors. A test suite often contains detailed instructions and information for each collection of test cases on the system configuration to be used during testing. Test suites are used to group similar test cases together.)

iii. To implement and verify the environment.

Test execution has the following major task:

i. To execute test suites and individual test cases following the test procedures.

ii. To re-execute the tests that previously failed in order to confirm a fix. This is known as confirmation testing or re-testing.

iii. To log the outcome of the test execution and record the identities and versions of the software under tests. The test log is used for the audit trial. (A test log is nothing but, what are the test cases that we executed, in what order we executed, who executed that test cases and what is the status of the test case (pass/fail). These descriptions are documented and called as test log.). iv. To Compare actual results with expected results. v. Where there are differences between actual and expected results, it report discrepancies as Incidents.

4) Evaluating Exit criteria and Reporting: ‘

Based on the risk assessment of the project we will set the criteria for each test level against which we will measure the ―enough testing‖. These criteria vary from project to project and are known as exit criteria. Exit criteria come into picture, when: — Maximum test cases are executed with certain pass percentage. — Bug rate falls below certain level. — When achieved the deadlines.

4) Test Closure activities:

Test closure activities are done when software is delivered. The testing can be closed for the other reasons also like: When all the information has been gathered which are needed for the testing. When a project is cancelled. When some target is achieved. When a maintenance release or update is done.

TYPES OF TESTS

• **Unit testing**:

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

• Integration testing :

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures : interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify

Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

• System Test

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

• White Box Testing

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

• Black Box Testing

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box.you cannot ―see‖ into it. The test provides inputs and responds to outputs without considering how the software works.

Part III ANNEXURE C

PROJECT PLANNER

1 Project Estimates

1.1 Time Estimates

Time estimate is about 9 months

Sr. No.	Estimates	Time Taken
1	Literature Survey	2
2	Design	1
3	Presentation	1
4	Coding	3
5	Testing	1
6	Reporting	1
7	TOTAL	9

1.2 PROJECT SCHEDULE

1.2.1 Project Task Set

Major Tasks in the Project stages are:
1: Research about solar power forecasting.
2: Research about selected topic.
3: Literature survey.
4: Deciding the flow of project plan.
5: Determine the Requirement.
6: Dividing the task.
7: Formulate the code.
8: Testing.
9: Project Complete Demonstration

1.3 Risk Management

P-class problem: P is set of all decision problems which can be solved in polynomial time by a deterministic. Since it can be solved in polynomial time, it can be verified in polynomial time. Therefore P is a subset of NP. NP-class problem:‖ NP‖ means — we can solve it in polynomial time if we can break the normal rules of step-by-step computing‖ . NP-Hard problem: A problem is NP-hard if an algorithm for solving it can be translated into one for solving any NP-problem (nondeterministic polynomial time) problem. NP-hard therefore means — at least as hard as any NP-problem,‖ although it might, in fact, be harder. NP-Complete problem: Since this amazing‖ computer can also do anything a normal computer can, we know that‖ P‖ problems are also in‖ NP‖ . So, the easy problems are in‖ P‖ (and‖ NP‖), but the really hard ones are only in ‖ NP‖ , and they are called ‖ NP-complete‖ . It is like saying there are things that People can do (‖ P‖), there are things that Super People can do (‖ SP‖), and there are things *only* Super People can do (‖ SP-complete‖

1.3.1 Risk Identification

This section discusses the differences between identifying generic risks and product specific risks. Generic risks can be listed on a checklist to examine for every software product. Examining the project plan and the software statement of scope identifies product-specific risks. Students may need to be shown examples of software project risk checklists. The risk assessment table shown in this section provides students with a good to begin quantifying the impact of many types of risk.

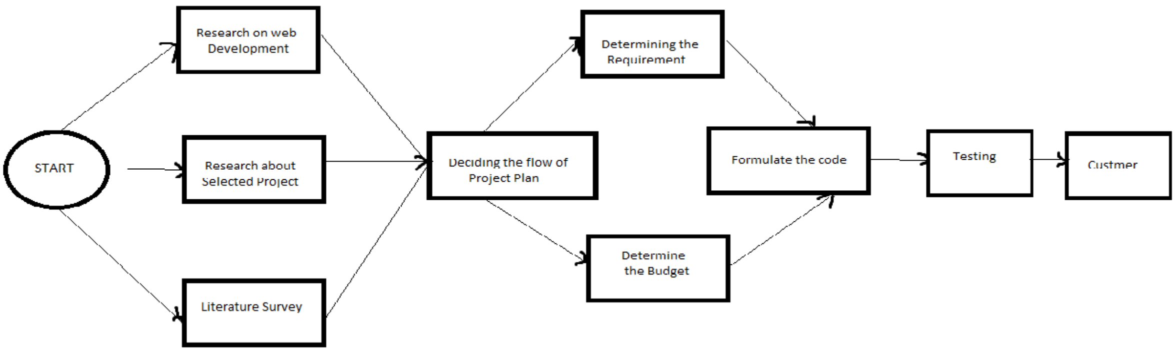


Figure C.1 Task Network

C.4 Timeline Chart

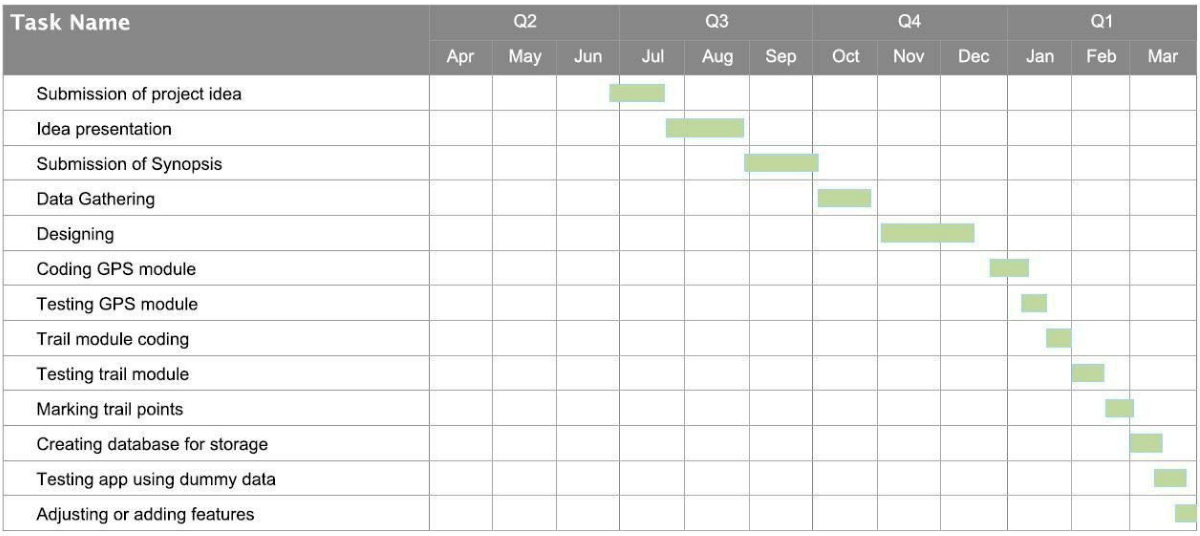


Figure C.2: Timeline

3.3 Team Organization

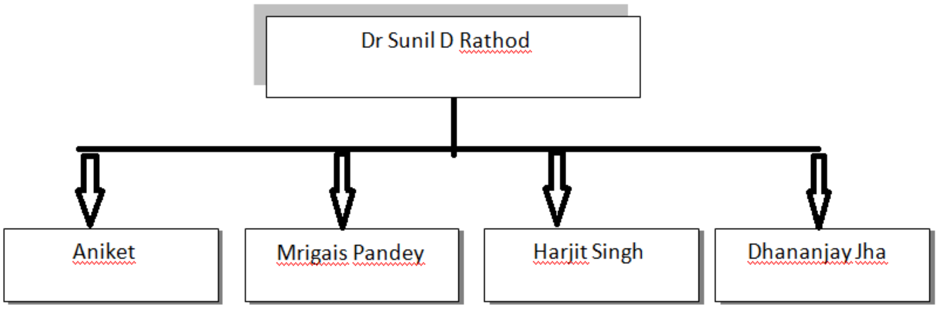


Figure C.3: Team Organization

Part IV

PUBLISHED PAPER

SEM-1 Published Paper

IJSART - *Volume 4 Issue 10 –OCTOBER 2018* *ISSN* [ONLINE]: 2395-1052

Power Forecast of Solar Panels Using Machine Learning Techniques: A Survey

Aniket Pandey[1], Mrigais Pandey[2], Dhananjay Jha[3], Harjit Singh[4], Dr. Sunil D Rathod[5]
Department of Computer Engineering
Dr. D Y Patil School of Engineering, Pune

Abstract- The non-renewable sources of energy are limited and will get exhausted eventually. Looking at the current need of electric power and its fulfilment, the non-conventional way of generating this energy has become essential. Climate change and energy crisis have motivated us to make use of renewable non-conventional source of energy. This paper discusses the theoretical assumptions and design aspects of developing a Model which will predict the solar power generation beforehand. The paper aims at promoting the use of renewable source of energy by developing a model which will accurately predict the solar power generation. The suggested model uses various Machine Learning Algorithms to predict the power generation which will be beneficial to both Industries and Residents

Keywords- Solar Power, Machine Learning, Prediction Model

I. INTRODUCTION

Climate change and energy crisis have led us to use renewable energy use and Solar Energy is one of the most appropriate option for use. It is renewable as well as non-conventional source of energy and available in abundance. Power generated using Solar PV Panels depends on many external factors namely weather and meteorological factors. Factors such as Wind, Cloud and Rain also affect the rate of Power Generated. A proposed model uses the methodology which will have high level of accuracy in predicting solar power.[1] To do so we will compare various Machine Learning Algorithms and find the most accurate to be used to get desired result.

The Dataset will be divided into Training and Test Data after pre-processing and scaling and various Machine Learning Algorithms will be applied to find out the most accurate of them. The most suitable one will be applied in the model to predict the power.

The dataset[4] used in this work is historical weather data from Amherst, MA, and is maintained by the University of Massachusetts, Amherst – Computer Science Weather Station.

II. MOTIVATION

Roof-top mounted solar photovoltaic (PV) systems are becoming an increasingly popular means of incorporating clean energy into the consumption profile of its users. It is one of the most efficient renewable sources of energy which can be used over non-renewable sources of energy such as Fossil Fuels. There are certain influencing factors such as environment friendly which promote the use of Solar Energy and it is also safer than traditional electricity current.

The motivation behind taking up this project was to implement a model which would help people manage the energy resources in an efficient and economical way. This model can help the user to pre-plan and use the power according to the prediction made by different machine learning and statistical techniques and avoid any sorts of loss due to sudden weather changes which are not in their control. Application of this model incurs low cost for installation (economical), safer and comparatively more available than other energy resources. Electric utilities often allow the inter-connection of such systems to the grid, compensating system owners for electricity production. As the systems grow in number and their contribution to the overall load profile becomes increasingly significant, it becomes imperative for utilities to accurately account for them while planning and forecasting generation.

III. LITERATURE REVIEW

A similar study has already been done previously. The comparative study is given below. The Advantages and Limitations of the Papers are discussed which will be overcome in our proposed model.

PAPER NO.	PAPER NAME	ADVANTAGES	LIMITATIONS
1	N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning."	• 27% more accurate than existing models • 51% better than simple approaches that only use the past to predict the future	• It does not incorporate information from multiple weather metrics and their impact on solar intensity
2	Gensler Andre, A., et al. "Deep Learning for solar power forecasting - An approach using AutoEncoder and LSTM Neural Networks."	• Performance achieved can also be transferred to other regenerative energy sources, e.g. forecasting of wind power output • Feature Extraction Capability	• It needs to take into account if an overestimation or an underestimation is preferred
3	Mayukh Samanta, Bharath Srikanth, Jayesh Yerrapragada, "Short Term Power Forecasting Of Solar PV Systems Using Machine Learning Techniques."	• High Accuracy using Hybrid Model	• Predicts a non zero solar irradiance during period of day when there is completely no sunlight

Table 1: Comparative Study of Previous Research

IJSART - *Volume 4 Issue 10 –OCTOBER 2018* ISSN [ONLINE]: 2395-1052

IV. PROPOSED SYSTEM

The obtained Dataset is in unprocessed format. So it will be pre-processed to fill the empty data and make it standardized. After pre-processing is done the Data will undergo scaling to bring data on a common scale. After this the data will be split into Training and Test Data. The Training Data will be used to Train the Model and after the Training is done Test Data will be passed to the model and Analysis will done to find out the accuracy of the model. While doing the Analysis various error factors will be considered to get accurate results. The algorithm with minimum error and maximum accuracy will be our model for prediction.

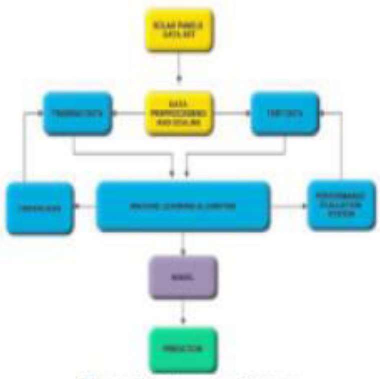


Figure 1 System Architecture

In our work we will be using various machine learning algorithms: weighted linear regression, PCA-based weighted linear regression, boosted regression trees, and neural networks.[2]

V. FUTURE USE AND SCOPE

The main objective is to benchmark different forecasting techniques of solar PV panel energy output.

Figure 2 Machine Learning Algorithms

Towards this end, machine learning and statistical techniques can be used to dynamically learn the relationship between different weather conditions and the energy output of PV systems. This is being done to optimize the energy structure and improve the performance of a PV system.[3]

Accurate prediction of PV power output is required to make better generation plans, support the spatial and temporal compensation, and achieve coordinated power control, so that the need for energy storage capacity and operating costs can be reduced. Our aim is to investigate the future engineering methodologies, which can be used to increase the overall prediction accuracy.[1] We will be using various techniques to train models on solar irradiance data and different meteorological parameters to forecast solar irradiance, and therefore power, for different forecasting horizons in the short-term future

VI. CONCLUSION

This model will help user predict the Solar Power Generation. It will guide the user through unfamiliar situation which can occur so that he could save power prior itself. It will also help in promoting use of renewable source of energy.

ACKNOWLEDGEMENT

It gives us great pleasure in presenting the paper on "Power Forecast of Solar Panels using Machine Learning Techniques". We would like to take this opportunity to thank our guide Dr. Sunil D Rathod for giving us all the help and guidance we need with indispensable support, suggestions and motivation during course of the paper writing work. We are really grateful to him. His valuable suggestions were insightful. We are also grateful to Prof. Soumitra Das, Head of Computer Engineering Department, DYPSOE, Pune. Our special thanks to Dr. E B Khedkar, Director DYPTC who motivated us and created a healthy environment for us to learn in the best possible way. We also thank all the staff members of our college for their support and guidance.

REFERENCES

[1] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning," in Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on, pp. 528–533, IEEE, 2011.

[2] Gensler-Janosch, A., et al. "Deep Learning for solar power forecasting — An approach using AutoEncoder and LSTM

IJSART - *Volume 4 Issue 10 –OCTOBER 2018* *ISSN* [ONLINE]: 2395-1052

Neural Networks." 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016

[3] Mayukh Samanta, Bharath Srikanth, Jayesh Yerrapragada, "Short Term Power Forecasting Of Solar PV Systems Using Machine Learning Techniques."

[4] http://s35695.mini.alsoenergy.com/Dashboard/2a5669735065572f4a42454b772b714d3d

[5] https://machinelearningmastery.com/use-keras-deep-learning-models-scikit-learn-python/

IJSART

ISSN [ONLINE] : 2395-1052

THE BOARD OF

IJSART

INTERNATIONAL JOURNAL FOR SCIENCE AND ADVANCE RESEARCH IN TECHNOLOGY

is here by awarding this certificate to

MRIGAIS PANDEY

In recognition of publication of the paper entitled

POWER FORECAST OF SOLAR PANELS USING MACHINE LEARNING TECHNIQUES: A SURVEY

Published in E-Journal

Volume 4,Issue 10 in October 2018

PAPER ID :IJSARTV4I1027796

EDITOR IN CHIEF

Email id : editor@ijsart.com | website : www.ijsart.com

IJSART

ISSN [ONLINE] : 2395-1052

THE BOARD OF

IJSART

INTERNATIONAL JOURNAL FOR SCIENCE AND ADVANCE RESEARCH IN TECHNOLOGY

is here by awarding this certificate to

ANIKET PANDEY

In recognition of publication of the paper entitled

POWER FORECAST OF SOLAR PANELS USING MACHINE LEARNING TECHNIQUES: A SURVEY

Published in E-Journal

Volume 4,Issue 10 in October 2018

PAPER ID :IJSARTV4I1027796

EDITOR IN CHIEF

Email id : editor@ijsart.com | website : www.ijsart.com

IJSART

ISSN [ONLINE] : 2395-1052

THE BOARD OF

IJSART

INTERNATIONAL JOURNAL FOR SCIENCE AND ADVANCE RESEARCH IN TECHNOLOGY

is here by awarding this certificate to

HARJIT SINGH

In recognition of publication of the paper entitled

POWER FORECAST OF SOLAR PANELS USING MACHINE LEARNING TECHNIQUES: A SURVEY

Published in E-Journal

Volume 4, Issue 10 in October 2018

PAPER ID :IJSARTV4I1027796

EDITOR IN CHIEF

Email id : editor@ijsart.com | website : www.ijsart.com

IJSART

ISSN [ONLINE] : 2395-1052

THE BOARD OF

IJSART

INTERNATIONAL JOURNAL FOR SCIENCE AND ADVANCE RESEARCH IN TECHNOLOGY

is here by awarding this certificate to

DHANANJAY JHA

In recognition of publication of the paper entitled

POWER FORECAST OF SOLAR PANELS USING MACHINE LEARNING TECHNIQUES: A SURVEY

Published in E-Journal

Volume 4, Issue 10 in October 2018

PAPER ID :IJSARTV4I1027796

EDITOR IN CHIEF

Email id : editor@ijsart.com | website : www.ijsart.com

SEM-2 Published Paper

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

Solar Power Prediction using Recurrent Neural Network

Sunil D Rathod[1], Aniket[2], Mrigais Pandey[3], Dhananjay Jha[4], Harjit Singh[5]
[1, 2, 3, 4, 5]*Computer Engineering Department, Dr. D Y Patil School of Engineering, Pune*

Abstract: The non-renewable sources of energy are limited and will get exhausted eventually. Looking at the current need of electric power and its fulfilment, the non-conventional way of generating this energy has become essential. Climate change and energy crisis have motivated us to make use of renewable non-conventional source of energy. This paper discusses the theoretical assumptions and design aspects of developing a Model which will predict the solar power generation beforehand. The paper aims at promoting the use of renewable source of energy by developing a model which will accurately predict the solar power generation. The suggested model uses Long Short Term Memory Recurrent Neural Network (LSTM RNN) Algorithms to predict the power generation which will be beneficial to both Industries and Residents.

Keywords: Solar Power, Long Short Term Memory (LSTM), Recurrent Neural Network (RNN), Machine Learning (ML)

I. INTRODUCTION

Climate change and energy crisis have led us to use renewable energy use and Solar Energy is one of the most appropriate option for use. It is renewable as well as non-conventional source of energy and available in abundance [1]. Power generated using Solar Photo Voltaic (PV) Panels depends on many external factors namely weather and meteorological factors. Factors such as Wind, Cloud and Rain also affect the rate of Power Generated. The proposed model uses the methodology which will have high level of accuracy in predicting solar power. To do so we have used LSTM RNN Machine Learning Algorithms after a brief study done by us in previous Survey Paper [7]. Using this trained ML model we will pass current weather data fetched through API to the model and the predicted solar power generation will be displayed.

II. MOTIVATION

Roof-top mounted solar photovoltaic (PV) systems are becoming an increasingly popular means of incorporating clean energy into the consumption profile of its users [2], [3]. It is one of the most efficient renewable sources of energy which can be used over non-renewable sources of energy such as Fossil Fuels. There are certain influencing factors such as environment friendly which promote the use of Solar Energy and it is also safer than traditional electricity current. The motivation behind taking up this project was to implement a model which would help people manage the energy resources in an efficient and economical way. This model can help the user to pre-plan and use the power according to the prediction made by RNN and statistical techniques and avoid any sorts of loss due to sudden weather changes which are not in their control. Application of this model incurs low cost for installation (economical), safer and comparatively more available than other energy resources. Electric utilities often allow the inter-connection of such systems to the grid, compensating system owners for electricity production [3]. As the systems grow in number and their contribution to the overall load profile becomes increasingly significant, it becomes imperative for utilities to accurately account for them while planning and forecasting generation [3].

III. DATASET AND FEATURES

The solar energy output necessary to power the campus of the University of Illinois in Urbana-Champaign was obtained from publicly-available repository [4]. The weather dataset used is historical weather data from Amherst, MA, and is maintained by the University of Massachusetts, Amherst – Computer Science Weather Station. It was obtained using the methodology detailed by National Oceanographic and Atmospheric Administration [8], [9]. We pre-processed these to obtain the numerical values for each feature and time-averaged them to obtain a consistent hourly resolution.

Weather Features	Unit	Weather Features	Unit
Cloud Coverage	% range	Relative Humidity	%
Visibility	Miles	Wind Speed	Mph
Temperature	°C	Station Pressure	inchHg
Dew Point	°C	Altimeter	inchHg

Table 1: A summary of selected meteorological parameters

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

IV. PROPOSED SYSTEM ARCHITECTURE

The obtained Dataset is in unprocessed format. So it will be pre-processed to fill the empty data and make it standardized. After pre-processing is done the Data will undergo scaling to bring data on a common scale. After this the data will be split into Training, Development and Test Data into 80%-10%-10% respectively. We use the hourly resolution, ranging from 6AM to 5PM. The total numbers of rows in our dataset are 7536. We had to take into account that our samples are not perfectly independent: in fact the solar energy output of an hour of a specified day is obviously correlated with the weather and the energy output of the previous hours of the same day. The Training Data will be used to Train the Model and after the Training is done Test Data will be passed to the model and Analysis will be done with this we can find out the accuracy of the model and Mean Squared Error. After completing the training process and testing it, the model will be bundled up into pickle. The Pickle now stores trained model. In the second stage we will fetch current weather data using API from Open Source Weather Data website which in our project happens to be openweathermap.org [10] and darksky.net [11]. The fetched data is in JSON format which will be fed to the Pickle model and the Solar Power generated will now be displayed. While doing the Analysis various error factors will be considered to get accurate results.

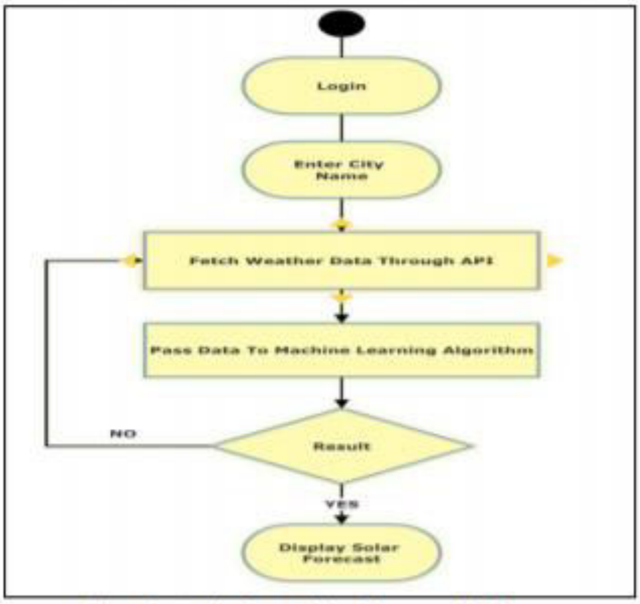


Figure 1: Application Activity Diagram with Flow

V. LSTM RECURRENT NEURAL NETWORK ALGORITHM AND METHOD

The prediction learning method implemented is an LSTM (Long Short Term Memory) recurrent neural network. We have assumed that a recurrent neural network is capable of capturing time-dependent trends in the data because feedback loops enable RNN's to exhibit memorization of temporal behaviour. Developing the Recurrent Neural Network involved sampling performance on the basis of a wide range of modifiable parameters which includes the size and number of hidden layers, types of activation functions, type of optimization and regularization, batch and epoch sizes, and cross-validation methods.

A. Forward Propagation for the basic Recurrent Neural Network

The basic RNN that we have implemented has the structure below

Steps:

1) Implement the calculations needed for one time-step of the RNN.
2) Implement a loop over T_x time-steps in order to process all the inputs, one at a time
3) Here $a^{(i)}$, $x^{(i)}$, $y^{(i)}$ represents i^{th} activation function, training example input and target output respectively.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

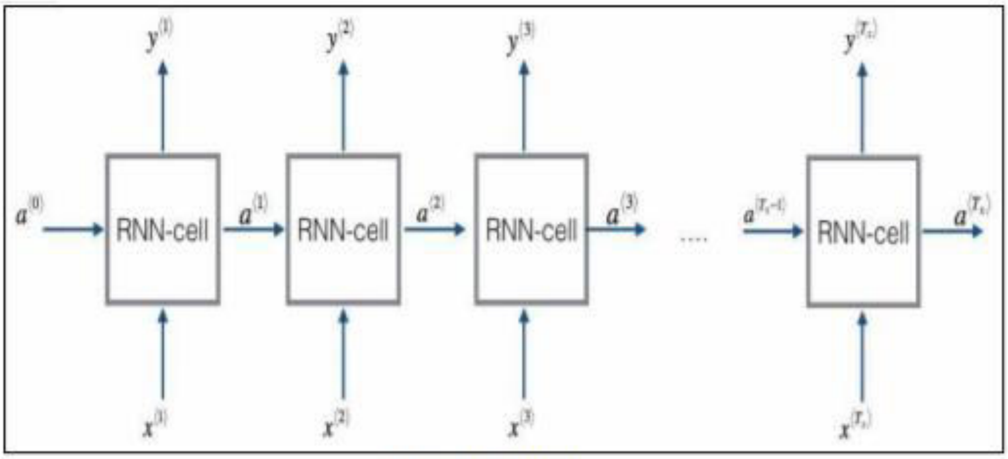


Figure 2: RNN Working

B. RNN Cell

A Recurrent neural network can be seen as the repetition of a single cell. First we have implemented the computations for a single time-step. The following figure describes the operations for a single time-step of an RNN cell.

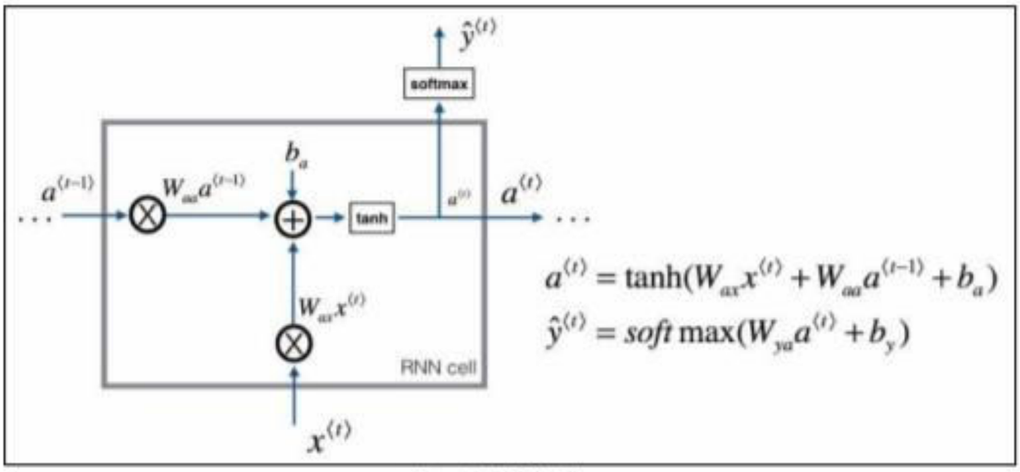


Figure 3: RNN Cell

1) Compute the hidden state with tanh activation: $\mathbf{a}^{\langle t \rangle}=\tanh(\mathbf{W_{ax}x}^{\langle t \rangle}+\mathbf{W_{aa}a}^{\langle t-1 \rangle}+\mathbf{b_a})$
2) Using the new hidden state $\mathbf{a}^{\langle t \rangle}$, compute the prediction $\mathbf{y}^{\langle t \rangle}=\mathbf{softmax}(\mathbf{W_{ya}}\ \mathbf{a}^{\langle t \rangle}+\mathbf{b_y})$
3) Here $\mathbf{W_{ax}}$ is set of weather parameters governing the connection from x to the hidden layer.
4) $\mathbf{W_{aa}}$ is vectorized weather parameter for horizontal connection and $\mathbf{W_{ya}}$ governs the output prediction. What this notion notation means is to just take the two vectors and stack them together.
5) $\mathbf{b_a}$ on top indicates a bias used for computing activation output.
6) **softmax** function outputs a vector that represents the probability distributions of a list of potential Solar Power outcomes.

Including an LSTM layer vastly improved performance, while the nonlinear hyperbolic tangent and sigmoid layers exhibited lower errors than standard linear activation functions.

The optimized neural network comprised three hidden layers (one LSTM) in addition to an input and output layer [6]. The introduction of a nonlinear hidden layer and an LSTM layer were each found to greatly increase the accuracy of test predictions [7].

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

Xavier-He initialization was utilized to select an ideally distributed initial value for the RNN weights. The 'adam' optimizer combined the benefits of both RMSProp and AdaGrad in adaptive moment estimation 20% dropout rate to effect regularization [5]. A mean-squared loss function was also used to train the RNN to maintain consistency.

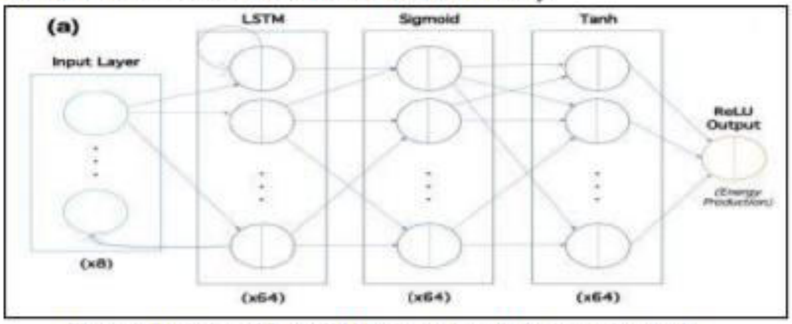


Figure 4: Depiction of flow of hidden layers in optimized neural network

VI. IMPLEMENTATION

The project demonstrates a web based Solar Power Predictor Application hosted on Flask server. Firstly we trained our model using Long Short Term Memory, Recurrent Neural Network ML Algorithm. The activation functions used in the hidden layer of RNN are Sigmoid and Tanh. The model was trained using the dataset obtained from repository mentioned in Section III of this paper. After training the algorithm we dumped our model in Pickle. To do the prediction of Solar Power user can login into web based app and select the city name for which prediction has to be done. After selecting the city, the API key provided by openweathermap.org and darksky.net will fetch weather data in JSON format. In the background this fetched data will be scaled using the scale factor obtained while training the algorithm. After scaling the data will be passed to the pickle and the output predicted by the model will again be scaled to KWh and displayed as the amount of Solar Polar generated for the queried city. An example of Solar Power predicted for Pune is displayed below which tells us the amount of Solar Power which will be generated by an individual Solar PV System in Pune City.

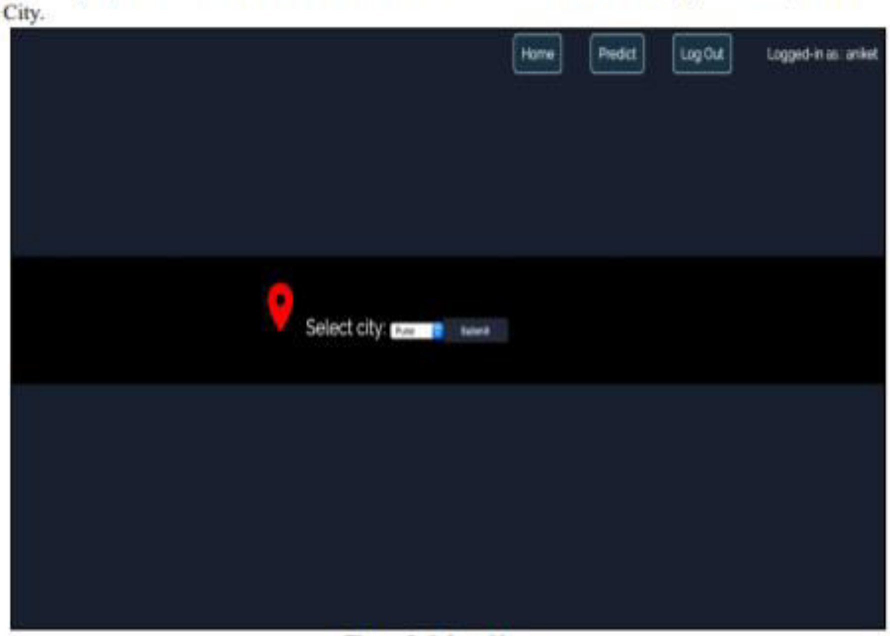


Figure 5: Select City

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

Figure 6: Predicted Solar Power

VII. FUTURE SCOPE

Towards this end, machine learning and statistical techniques can be used to dynamically learn the relationship between different weather conditions and the energy output of PV systems. This is being done to optimize the energy structure and improve the performance of a PV system. Accurate prediction of PV power output is required to make better generation plans, support the spatial and temporal compensation, and achieve coordinated power control, so that the need for energy storage capacity and operating costs can be reduced. Our aim is to investigate the future engineering methodologies, which can be used to increase the overall prediction accuracy. Further it can also be developed into an Application for better handy solution. Moreover if we connect it to IoT various other use such as automatic switching of the lights to save the power can be implemented.

VIII. CONCLUSIONS

This model will help user predict the Solar Power Generation. It will guide the user through unfamiliar situation which can occur so that he could save power prior itself. Currently there is not much use of Solar Power in India but once the people start realizing the importance of renewable sources of energy, they will eventually adopt this prediction model to conserve the Solar Energy. It will also help in promoting use of renewable source of energy.

IX. ACKNOWLEDGMENT

It gives us great pleasure in presenting the paper on "Power Forecast of Solar Panels using Machine Learning Techniques". We would like to take this opportunity to thank our guide Dr. Sunil D Rathod for giving us all the help and guidance we need with indispensable support, suggestions and motivation during course of the paper writing work. We are really grateful to him. His valuable suggestions were insightful. We are also grateful to Dr. Soumitra Das, Head of Computer Engineering Department, DYPSOE, Pune. Our special thanks to Dr. M Z Shaikh, Director DYPTC who motivated us and created a healthy environment for us to learn in the best possible way. We also thank all the staff members of our college for their support and guidance.

REFERENCES

[1] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning," in Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on, pp. 528–533, IEEE, 2011.
[2] Gensler-Janosch, A., et al. "Deep Learning for solar power forecasting — An approach using AutoEncoder and LSTM Neural Networks." 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016.
[3] Mayukh Samanta, Bharath Srikanth, Jayesh Yerrapragada, "Short Term Power Forecasting Of Solar PV Systems Using Machine Learning Techniques".
[4] http://s35695.mini.alsoenergy.com/Dashboard/2a5669735065572f4a42454b772b714d3d
[5] https://machinelearningmastery.com/use-keras-deep-learning-models-scikit-learn-python/.
[6] Adele Kuzmiakova, Gael Colas, Alex McKeehan, "Short-term Memory Solar Energy Forecasting at University of Illinois"
[7] Sunil Rathod, Aniket, Mrigais Pandey, Dhananjay Jha, Harjit Singh, "Power forecast of Solar Panels using Machine Learning Techniques: A Survey" in IJSART- Volume 4 Issue 10 – October 2018
[8] https://github.com/sborgeson/local-weather.
[9] NOAA website : https://www.ncdc.noaa.gov/cdo-web/datatools/l cd
[10] www.openweathermap.org
[11] www.darksky.net


ISSN No. : 2321-9653
iJRASET
International Journal for Research in Applied Science & Engineering Technology
IJRASET is indexed with Crossref for DOI-DOI : 10.22214
Website : www.ijraset.com, E-mail : ijraset@gmail.com
Certificate
It is here by certified that the paper ID : IJRASET23360, entitled
Solar Power Prediction using Recurrent Neural Network
by
Aniket
after review is found suitable and has been published in
Volume 7, Issue V, May 2019
in
International Journal for Research in Applied Science & Engineering Technology
Good luck for your future endeavors
ISRA Journal Impact Factor : 5.947
45.98
INDEX COPERNICUS
THOMSON REUTERS
doi
crossref
TOGETHER WE REACH THE GOAL
University Grants Commission
UGC Approved Journal
Editor in Chief, IJRASET


ISSN No. : 2321-9653
iJRASET
International Journal for Research in Applied Science & Engineering Technology
IJRASET is indexed with Crossref for DOI-DOI : 10.22214
Website : www.ijraset.com, E-mail : ijraset@gmail.com
Certificate
It is here by certified that the paper ID : IJRASET23360, entitled
Solar Power Prediction using Recurrent Neural Network
by
Mrigais Pandey
after review is found suitable and has been published in
Volume 7, Issue V, May 2019
in
International Journal for Research in Applied Science & Engineering Technology
Good luck for your future endeavors
ISRA Journal Impact Factor : 5.947
45.98
INDEX COPERNICUS
THOMSON REUTERS
doi
crossref
TOGETHER WE REACH THE GOAL
University Grants Commission
UGC Approved Journal
Editor in Chief, IJRASET

ISSN No. : 2321-9653

IJRASET

International Journal for Research in Applied Science & Engineering Technology

IJRASET is indexed with Crossref for DOI-DOI : 10.22214

Website : www.ijraset.com, E-mail : ijraset@gmail.com

Certificate

It is here by certified that the paper ID : IJRASET23360, entitled

Solar Power Prediction using Recurrent Neural Network

by

Harjit Singh

after review is found suitable and has been published in

Volume 7, Issue V, May 2019

in

International Journal for Research in Applied Science & Engineering Technology

Good luck for your future endeavors

University Grants Commission UGC Approved Journal

Editor in Chief, IJRASET

ISRA Journal Impact Factor : 5.947

45.98 INDEX COPERNICUS

THOMSON REUTERS

doi crossref

TOGETHER WE REACH THE GOAL IMPACT FACTOR : 7.177

ISSN No. : 2321-9653

IJRASET

International Journal for Research in Applied Science & Engineering Technology

IJRASET is indexed with Crossref for DOI-DOI : 10.22214

Website : www.ijraset.com, E-mail : ijraset@gmail.com

Certificate

It is here by certified that the paper ID : IJRASET23360, entitled

Solar Power Prediction using Recurrent Neural Network

by

Dhananjay Jha

after review is found suitable and has been published in

Volume 7, Issue V, May 2019

in

International Journal for Research in Applied Science & Engineering Technology

Good luck for your future endeavors

University Grants Commission UGC Approved Journal

Editor in Chief, IJRASET

ISRA Journal Impact Factor : 5.947

45.98 INDEX COPERNICUS

THOMSON REUTERS

doi crossref

TOGETHER WE REACH THE GOAL IMPACT FACTOR : 7.177

PART V

INFORMATION OF PROJECT GROUP MEMBERS

Name: Aniket
Seat No: B150884204
Email ID: aniket414@gmail.com
Phone no: 8690865419

Name: Mrigais Pandey
Seat No: B150884248
Email ID: mrigais@gmail.com
Phone no: 7045644720

Name: Harjit Singh
Seat No: B150884264
Email ID: hss2954@gmail.com
Phone no: 9730963697

Name: Dhananjay Jha
Seat No: B150884284
Email ID: jha.dhananjay1@gmail.com
Phone no: 9399860563