Solar Prediction for Smart Homes

INTRODUCTION:

Solar generation capacity is experiencing a dramatic increase in India, with its peak capacity having risen 8 times from 2,650 MW on 26 May 2014 to over 20 GW as on 31 January 2018. The country's solar installed capacity reached 31.696 GW as of 31 October 2019.

In this project, we argue that predicting solar generation for small-to-medium-sized solar deployments raises a different set of challenges than predicting it for massive solar farms. Specifically, the location of massive solar deployments is carefully chosen to be in open spaces that minimize occlusions, which enables installers to maximize solar output by precisely tuning the orientation of the panels or employing “trackers” that continuously change the tilt of the panels to track the sun. Further, industrial solar farm operators routinely clean the panels to keep them free from dust or snow in order to maintain optimal solar output. At the same time, industrial operators also have the technical expertise and resources to carefully design and tune custom models to predict future solar output.

Since solar energy today only contributes a small fraction of grid energy, the lack of effective planning does not currently pose a significant problem. However, as solar penetration rises, effective planning by homes and utilities will be increasingly important to maintain grid stability.

In this project we aim to predict solar energy generation for the next 48 hours or any given time.

The popularity of rooftop solar for homes is rapidly growing. However, accurately forecasting solar generation is critical to fully exploiting the benefits of locally-generated solar energy. We present machine learning techniques to predict solar intensity from publicly-available weather forecasts.The predicted solar intensity values can be used to calculate solar energy generation under the assumption that solar intensity is directly proportional to the energy generated.

FEATURE SELECTION:

A variety of features were used for the purpose of prediction.
The interrelation between features was studied and the most important ones were selected. NSRDB datasets were for training and testing purpose. Accuracies(RMSE) of various models were analysed and the one with highest accuracy(lowest RMSE) was selected as the model to be used.

This model will be used to predict the value of solar intensity 48 hours into the future given input of various features at current time which is the overall goal of the project. Machine learning is used to achieve this particular goal.

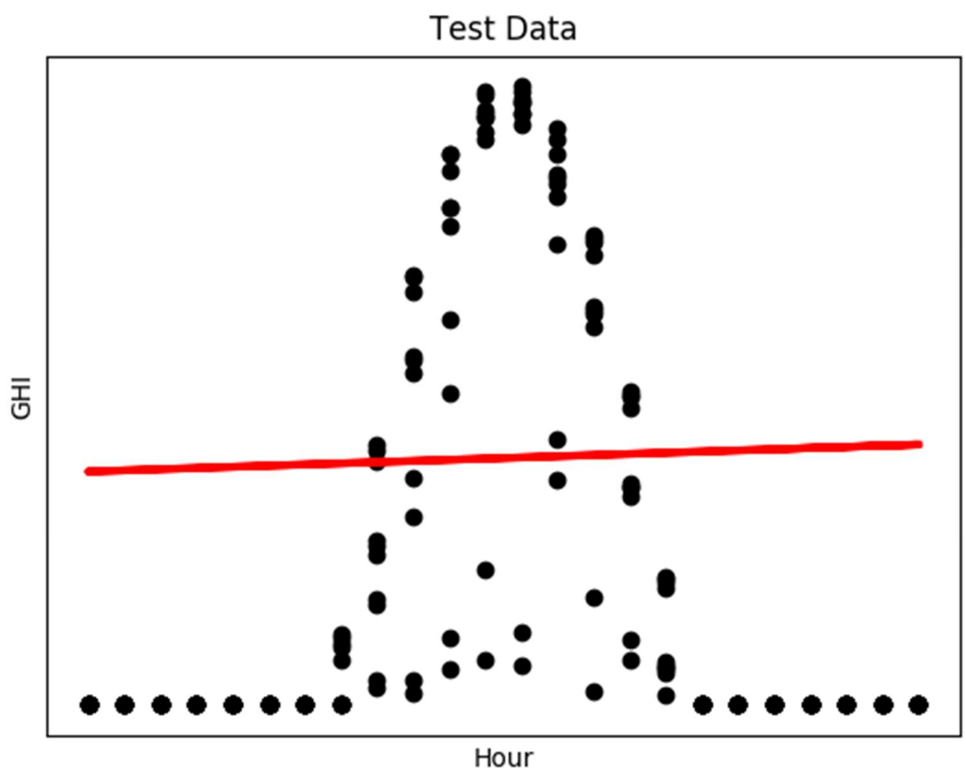
The terms solar intensity, GHI, solar power and solar energy are used interchangeably as solar energy generation is assumed to be directly proportional to the solar intensity.

The NSRDB dataset was used for building the machine learning model.
The dataset had to cleaned and preprocessed before giving it to the model for training.
The dataset was split in the ratio of 1:5 (train/test).
The features that are present in the dataset are hour, dew point, solar zenith angle, wind speed, precipitable water, wind direction, relative humidity, temperature, pressure.
The interrelation between the features was studied and the most important ones were selected. Heat Map was used for the same.
The value which is to be predicted is present as ‘GHI’ column in the dataset.
The most accurate model was selected as the main model.

DEPENDENCE OF GHI ON VARIOUS FEATURES:

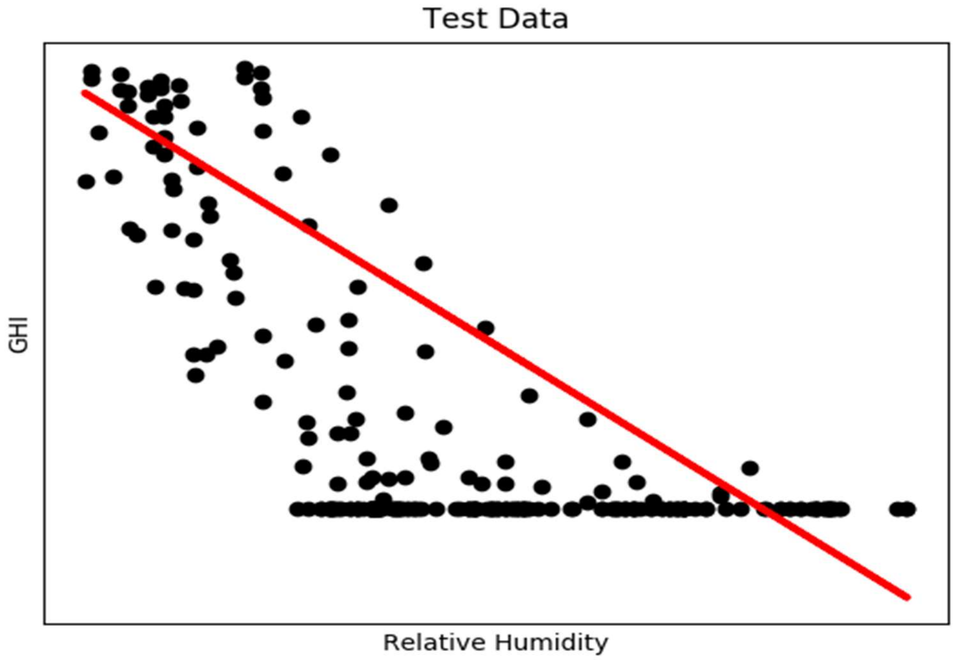
Although various features were considered, it was observed that GHI of a particular location has different degree of dependence on these features.

Hour:



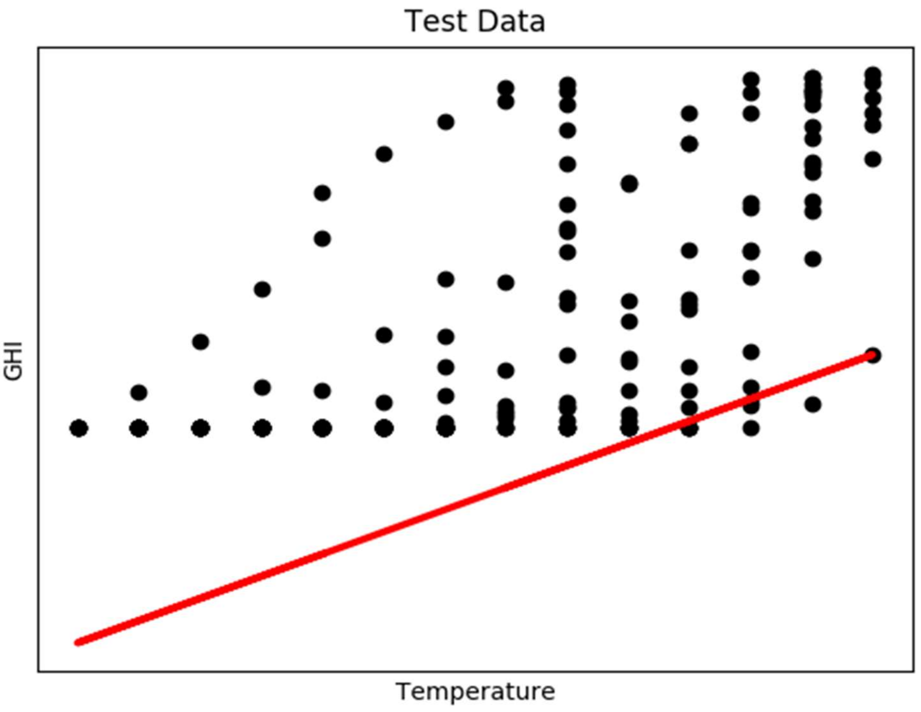
The hours are measured in a 24h clock format ranging from 0-24. GHI is the most in the afternoon hours and shows a symmetric downward parabola with axis of symmetry at Hour=12.

Humidity:



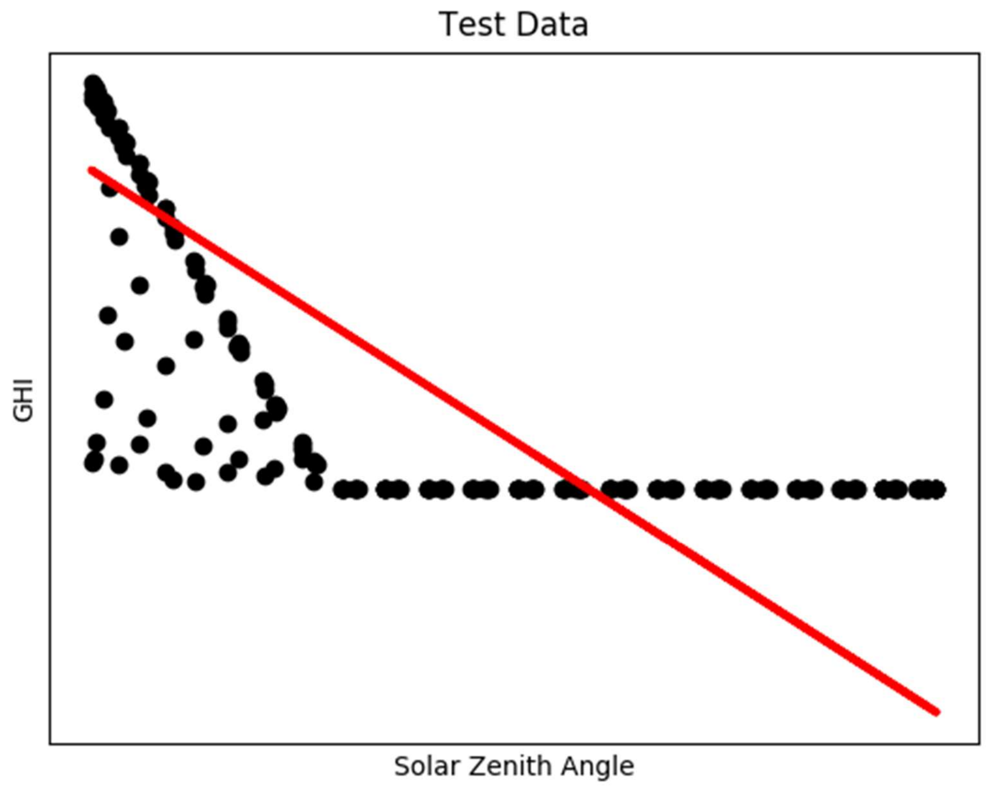
It can be observed from the graph that GHI is inversely proportional to the Relative Humidity of a particular location.

Temperature:



The GHI and temperature are directly proportional to each other and show a linear relationship.

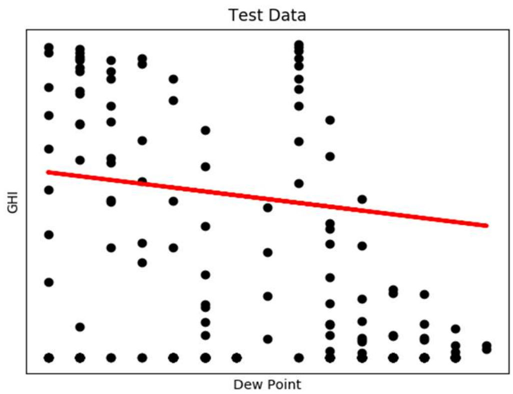
Solar Zenith Angle:

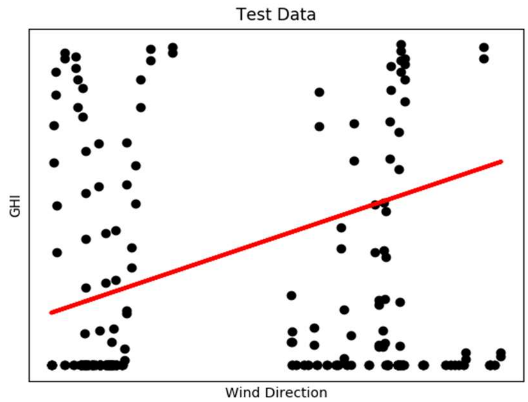


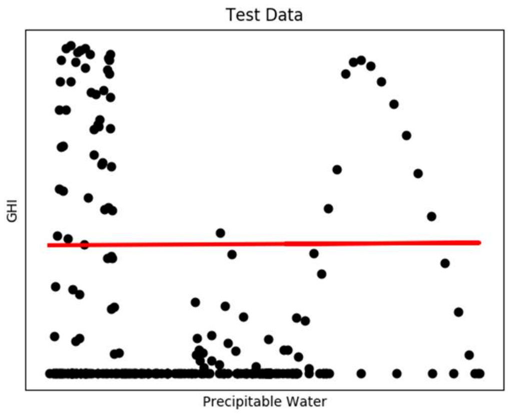
The solar zenith angle is the angle that zenith (surface normal) makes with the center of the sun’s disc. The plot of Zenith angle vs GHI shows a inverse relationship but it is more to that. It can be observed that the first half of plot along X axis, where solar zenith angle is between 0 to 90, the GHI is inversely related to Solar zenith angle with a linear relationship, but the next half of the plot (90<Solar Zenith Angle<180) GHI is constant irrespective of the Solar Zenith angle.

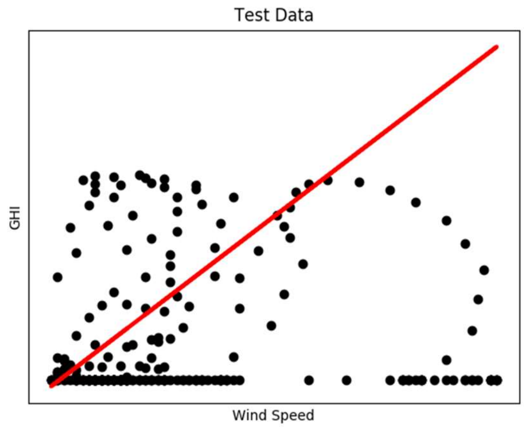
These were the features that influence GHI of a location drastically.

Other features like dew point, wind speed, wind direction and percipitable water showed a random relationship with GHI, the same can be observed in the following plots. These features were further tested and it was concluded that these do not make a significant impact on the GHI of a location.



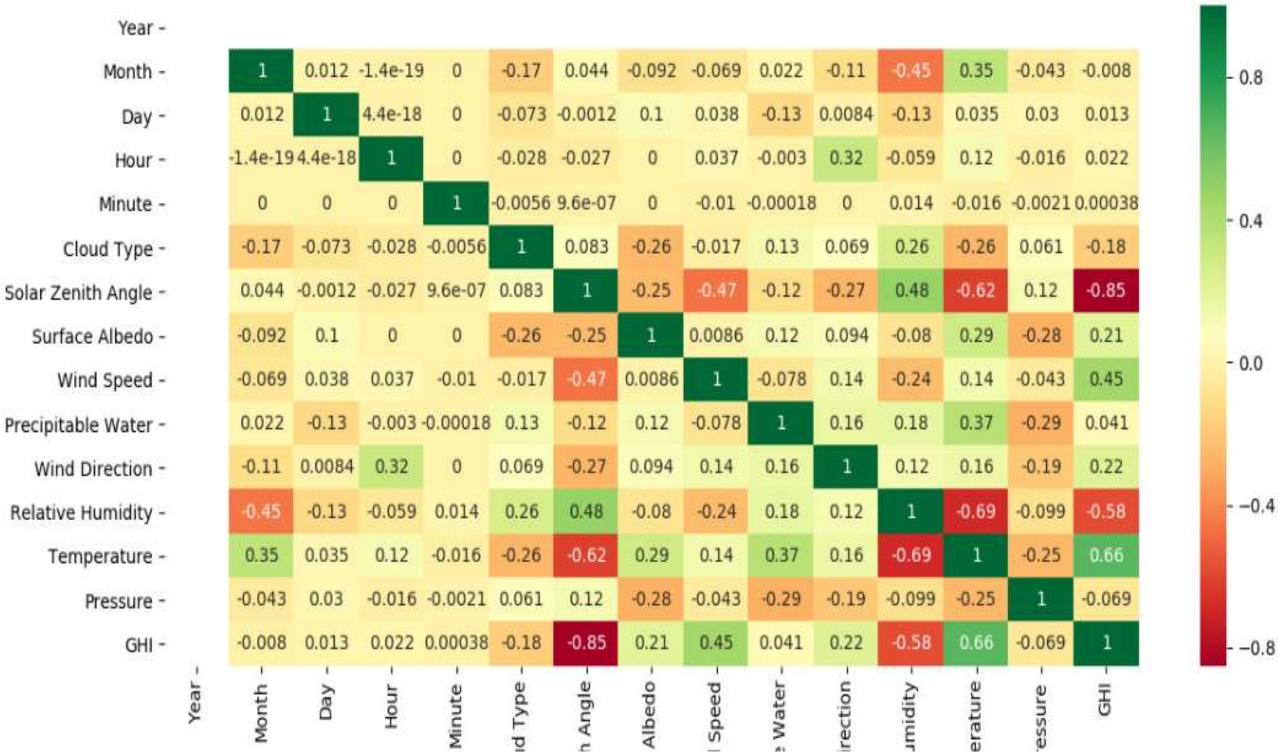






HEAT MAP:

The final heat map to show the relationship between the various features can be observed as follows



MODELS AND RESULTS:

We combined four years of data to predict solar intensity. We used 70 percent of data for training and 30 percent for testing purpose.We have considered 17 features to predict solar intensity.

1.Linear Model

We implement the basic linear model and achieved following results:
R2 Score : 0.906
RMSE : 95.80

2.Quadratic model

Expansion to linear model,we applied quadratic feature expansion to predict intensity. We found that Quadratic model reduces the error in testing data.

R2 Score : 0.99

RMSE : 8.74

3.KNN

We also applied nearest neighbour regressor in which we considered the value of neighbour = 7.

R2 Score : 0.982

RMSE : 41.24

4.Gradient Boosting Regression

We applied GBM regressor and got following results:

R2 Score : 0.995

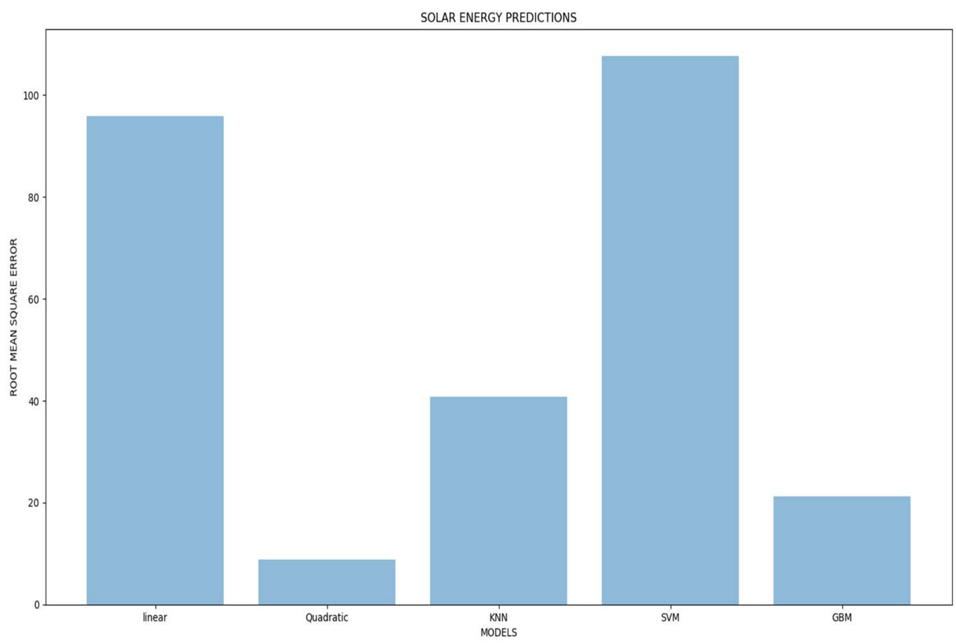
RMSE : 21.94

5.Suppot Vector Regression

We achieved following result using svr:

R2 Score : 0.883

RMSE : 107.37



CONCLUSION:

We were able to predict the value of solar intensity(GHI) given the input feature values.
The Model that was the most accurate was found out to be quadratic model giving RMSE value of 8.xx.
The problems that we'll be able to solve after getting the predicted value of the solar power are:
1.) Less dependence on the conventional energy resources
2.) Lessen down the skeptical views people have about solar generation by giving near accurate predictions.
3.) Efficient Energy management.