Solar Irradiance Forecasting

Project Overview

Business Understanding

Irradiance is the amount of solar energy that reaches the earth's surface per unit area. This irradiance is the basis for the calculation of overall power production capacity; forecasting irradiance allows us to forecast production. Accurate irradiance forecasting is essential for renewable energy generators, especially solar power plants, to optimize their energy production and meet regulatory requirements.

In Australia, the grid rewards solar power plants that can predict their expected irradiance (really their power generation capability) for the next 5 minutes. This helps the grid balance the supply and demand of electricity and prevent blackouts or surges. Other countries and regions may soon adopt similar or stricter rules for solar power plants.

Objective

The aim of this project is to <u>accurately forecast irradiance (GHI) with 30 minutes horizon and one minute granularity</u> (IE, predicting t+1, t+2, t+3…, t+30) using provided historical data.

Dataset

Used public dataset from Folsom, California.

- Day ahead weather forecasts for closest weather stations (North American Mesoscale Forecast System (NAM))
- Satellite imagery from GOES-15
- Imagery from Whole Sky Cameras
- RGB Ratios from Whole Sky Camera Imagery
- Target dataset with irradiance measures at a minutely granularity

Exploratory data analysis (EDA)

1. **Day ahead weather forecasts for closest weather stations**: a public North American Mesoscale Forecast System (NAM) Hourly dataset from Folsom, California (Latitude 38.579454, Longitude -121.260320) includes forecast time period, **GHI Downward short-wave radiation flux (W/m2), Cloud cover, precipitation, Pressure, components of wind, temperature, Relative humidity** with 14980 records.

2. **Satellite imagery from GOES-15 dataset** used to extract **cloud cover ratio** using following equation 1 and **Average, Standard Derivation and Entropy** features

   Cloud Cover Ratio = Number of Cloud Pixels / Total Number of Cloud and Sky Pixels (Eq. 1)

   Average $\mu = \frac{1}{N}\sum_{i=1}^{N} v_i.$ (Eq. 2) **Standard Derivation** $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(v_i - \mu)^2},$ (Eq. 3) **Entropy** $e = -\sum_{\substack{i=1 \\ p_i \neq 0}}^{N_B} p_i \log_2(p_i),$ (Eq. 4)

3. **Imagery from Whole Sky Cameras** used to extract **cloud cover ratio** using equation 1.

   Preprocessing Images:

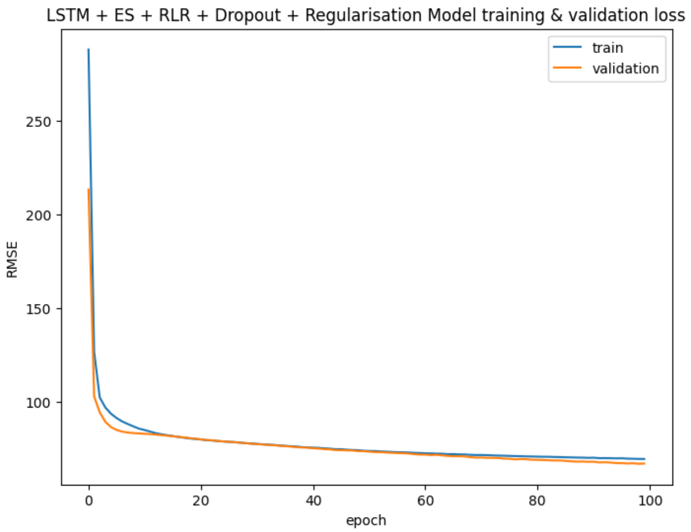
   1. Convert Fish-Eye Image to flat Image
   2. Detect Solar Position
   3. Convert the image into Color segmentation
   4. Classify the cloud and sky pixels using threshold pixel value
   5. Convert the cloud pixel to cloud cover ratio using Eq.1

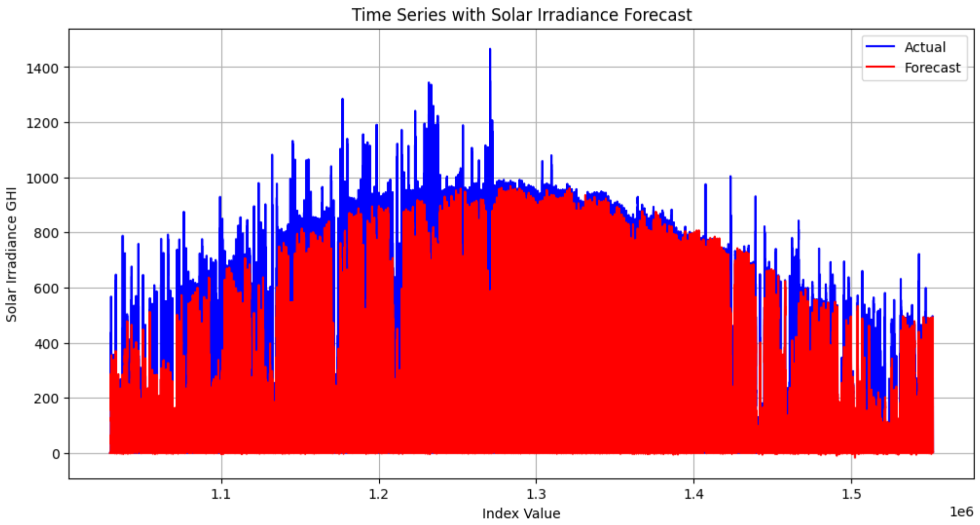
4. **RGB Ratios from Whole Sky Camera Imagery** includes the red-to-blue components with **Average, Standard Derivation and Entropy** features.

5. Target dataset with **irradiance measures** at a minutely granularity.

Model Implementation

- **Data Split** Use 2014 and 2015 for training data, and the data from 2016 to test
- **Data Normalized** by Min Max Standardization using `MinMaxScaler`
- **Data Lags** for forecast irradiance (GHI) with 30 minutes horizon and one minute granularity (IE, predicting t+1, t+2, t+3…, t+30).
- LSTM Model Building using Tensorflow
- Model Trained on 100 epochs and 3432 (int(trainX.shape[0]/epochs) * 0.10) batch size and calculate loss using Root Mean Square Error.



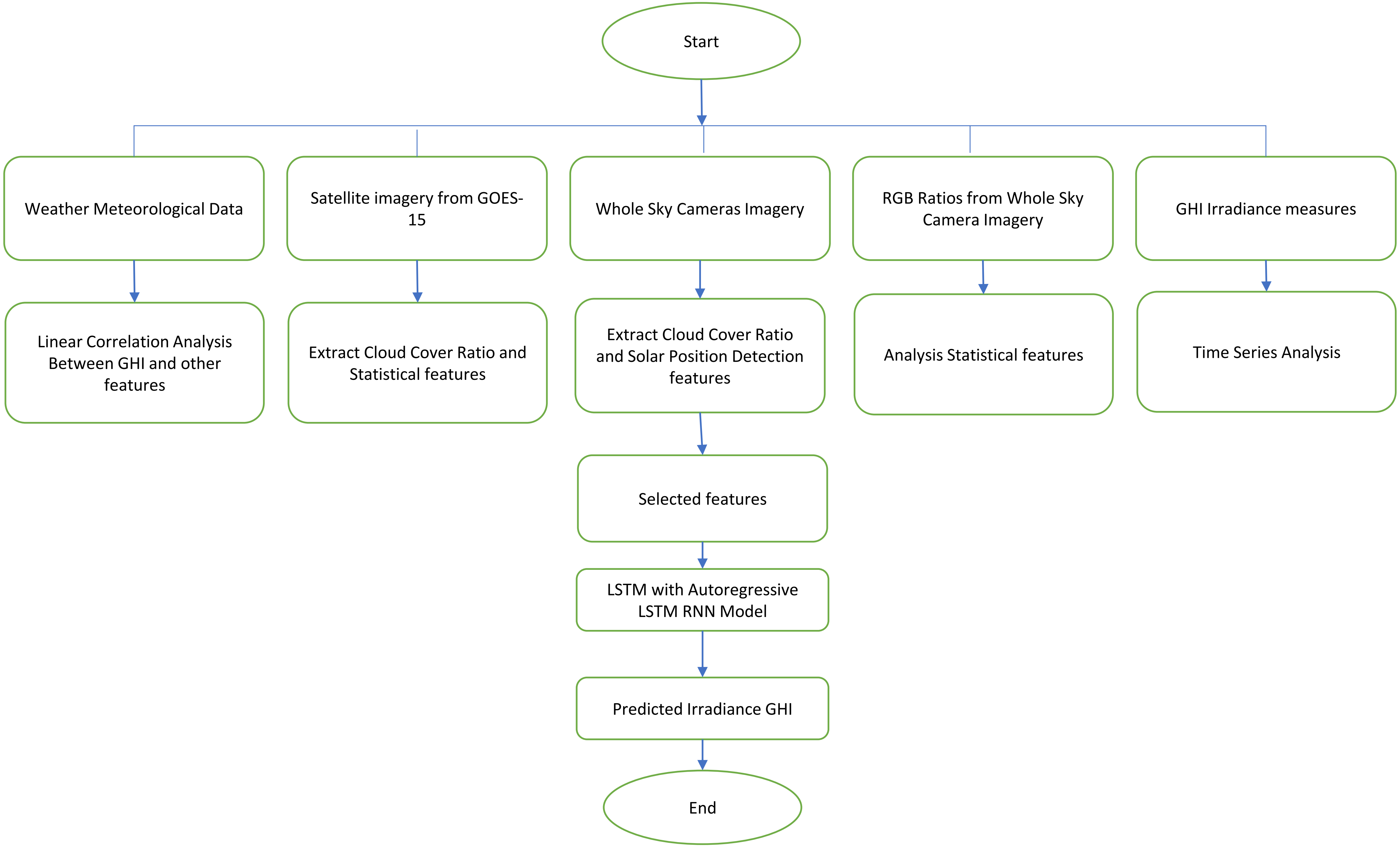


Model Training

Conclusion:

Based on the results of the all plots, we can conclude that the **LSTM model with Early Stopping, Reduce Learning Rate, Dropout and Regularization** has the best performance. This model has the highest test accuracy and the lowest test loss.

Model Name	Test Accuracy	Test Loss
LSTM + Early Stopping	0.34600067138671875	34.138160705566406
LSTM + ES + RLR + Dropout + Regularisation	0.32450488209724426	39.718406677246094
LSTM + ES + RLR + Regularization	0.3236911296844824	37.5727297973633
LSTM + ES + RLR + Dropout	0.29803672432899475	38.60310745239258
Base LSTM	0.27373212575912476	37.25522994995117
LSTM + ES + RLR + Dropout + Regularisation + Batch Normalization	0.1579645723104477	43.317256927490234
LSTM + ES + RLR + Dropout + Batch Normalization	0.1566721796989441	52.868812561035156
LSTM + Early Stopping + Reduce Learning Rate	0.13069225847721 1	31.88323402404785
LSTM + ES + RLR + Regularisation + Batch Normalization	0.01757076568901539	51.620784759521484


Start
Weather Meteorological Data
Satellite imagery from GOES-15
Whole Sky Cameras Imagery
RGB Ratios from Whole Sky Camera Imagery
GHI Irradiance measures
Linear Correlation Analysis Between GHI and other features
Extract Cloud Cover Ratio and Statistical features
Extract Cloud Cover Ratio and Solar Position Detection features
Analysis Statistical features
Time Series Analysis
Selected features
LSTM with Autoregressive LSTM RNN Model
Predicted Irradiance GHI
End

Whole Sky Cameras Imagery Processing

Fish-Eye Image

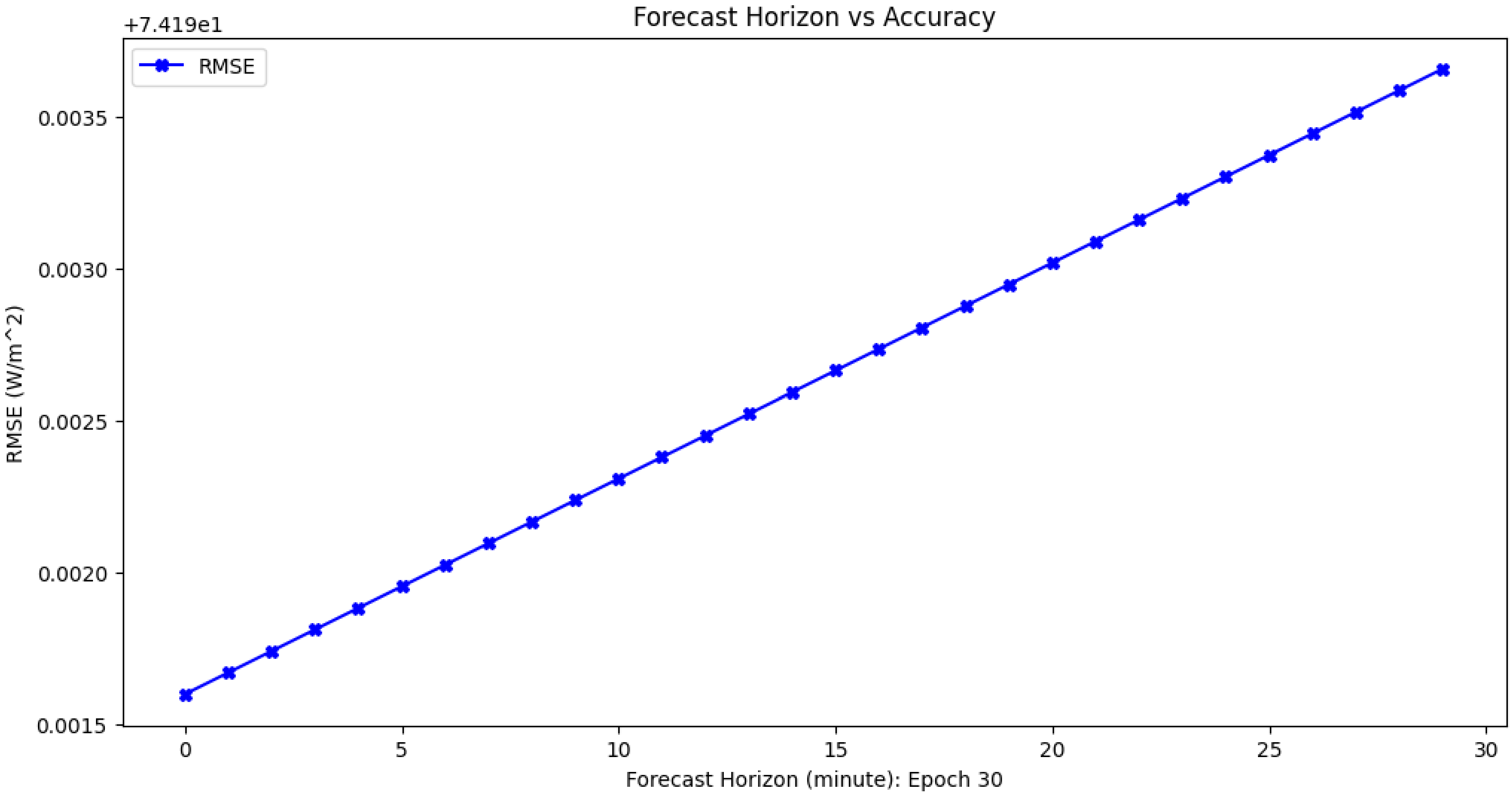
Flat Image

Colour segmentation

Image Binary Threshold

cloud_cover = (cloud_pixel/total_pixel)*100

Autoregressive LSTM RNN Model



Question?