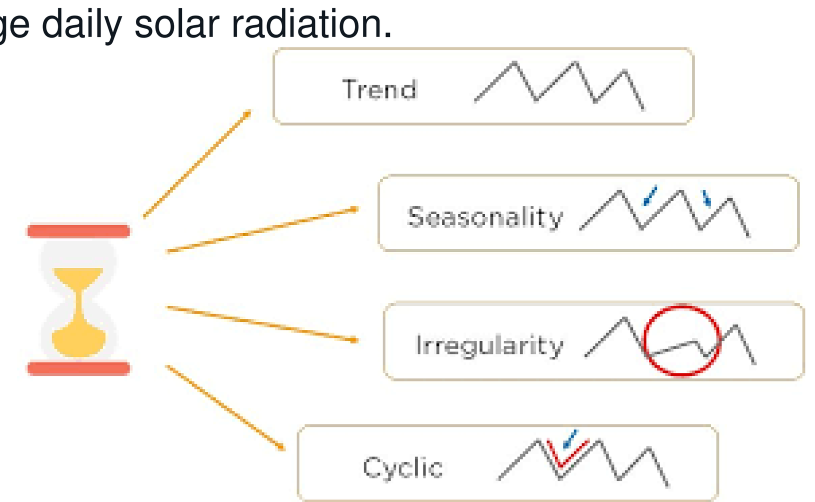
Comparative Evaluation of machine and deep learning algorithms for solar radiation Prediction

Mohammed El-Amine Azz, Raza Imam, Mohammed Talha Alam, Mohammed Al Zarooni

Mohamed bin Zayed University of Artificial Intelligence

Introduction

- Time Series Prediction has been an area of exploration with many applications in economy and weather.
- Good predictions and forecasting are essential for decision-making in a lot of fields.
- We explore different machine and deep learning models to make predictions about average daily solar radiation.

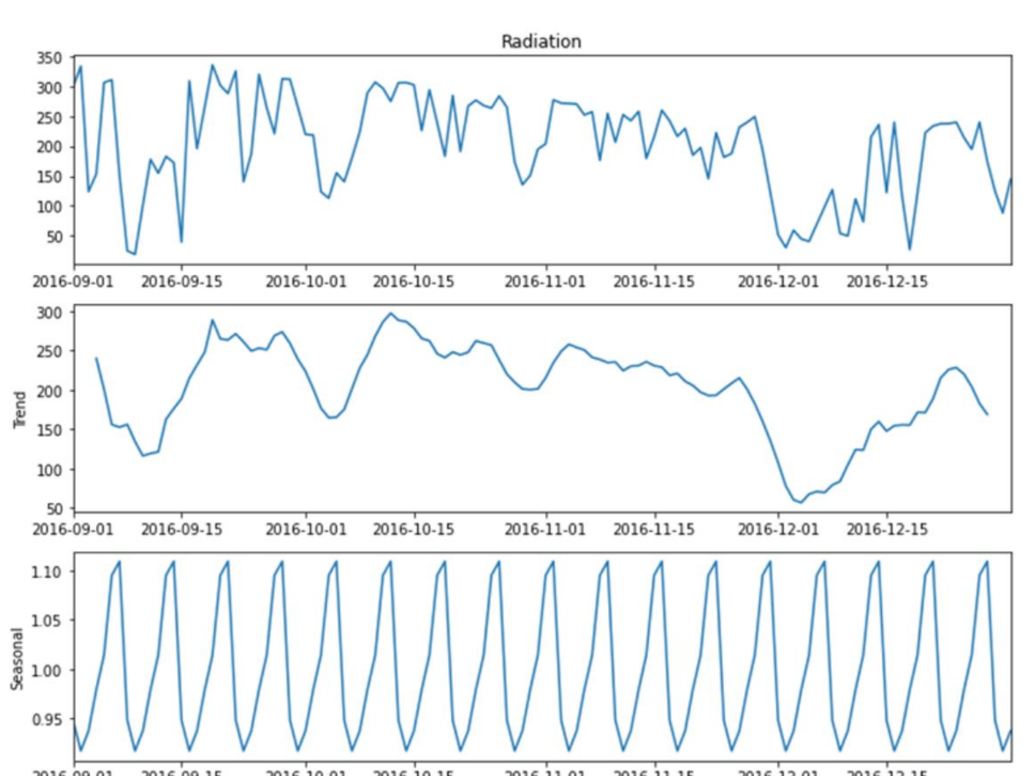


Research Objectives

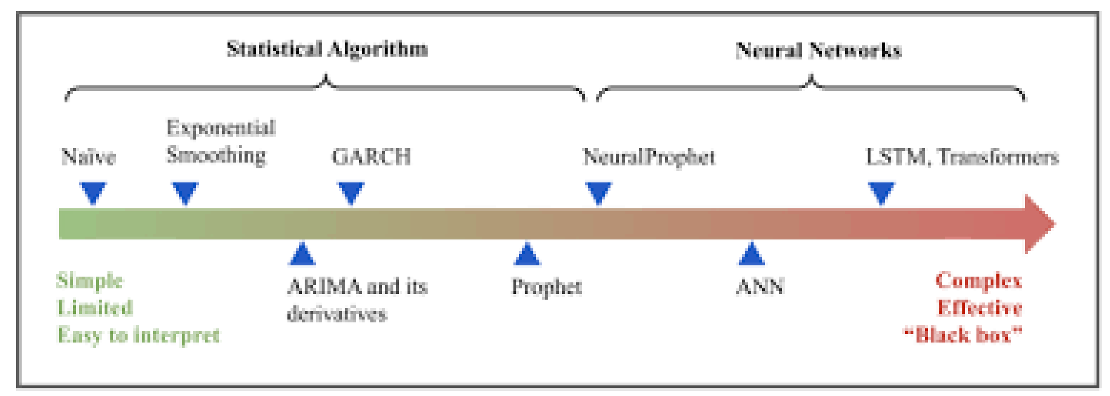
- We employ 4 time series prediction models, 2 of which are machine learning methods, and 2 are deep learning methods.
- We experiment with both predictions with ground truths and without ground truths to test the models' capabilities to tackle in-sample and out-of-sample predictions.
- Our empirical evaluations of the models under the different approaches are assessed on metrics like MAPE and RMSE.

Dataset

- The dataset contains 6 climate features, from which we will focus only on Radiation. The observations are daily from 1 September 2016 to 30 December 2016.
- To tackle missing data, we interpolated the few missing days.
- We split the dataset into training and testing with a 2:1 ratio. Unlike the usual splitting, in time series data the split is not random.
- The Testing dataset is used to evaluate all of the accuracy claims made about the models in the Results and Comparisons section.



Methods



Machine Learning Methods

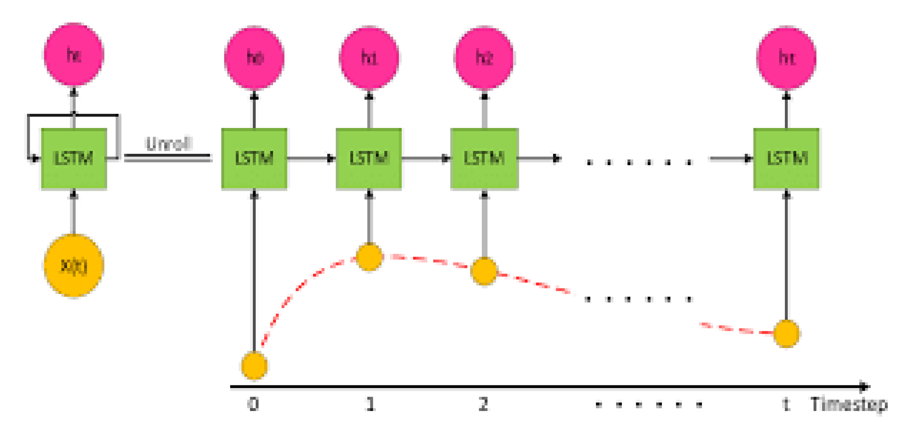
- **ARIMA** (Auto Regressive Integrated Moving Average): is a class of models that 'explains' a given time series based on its own past values with the following equation:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + .. + \beta_p Y_{t-p} \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + .. + \phi_q \epsilon_{t-q}$$

- **Exponential Smoothing**: Similar to ARIMA, except the weight decrease exponentially the further we go back.

Deep Learning Methods

- **LSTM**: an RNN based architecture that is widely used in natural language processing and time series forecasting.



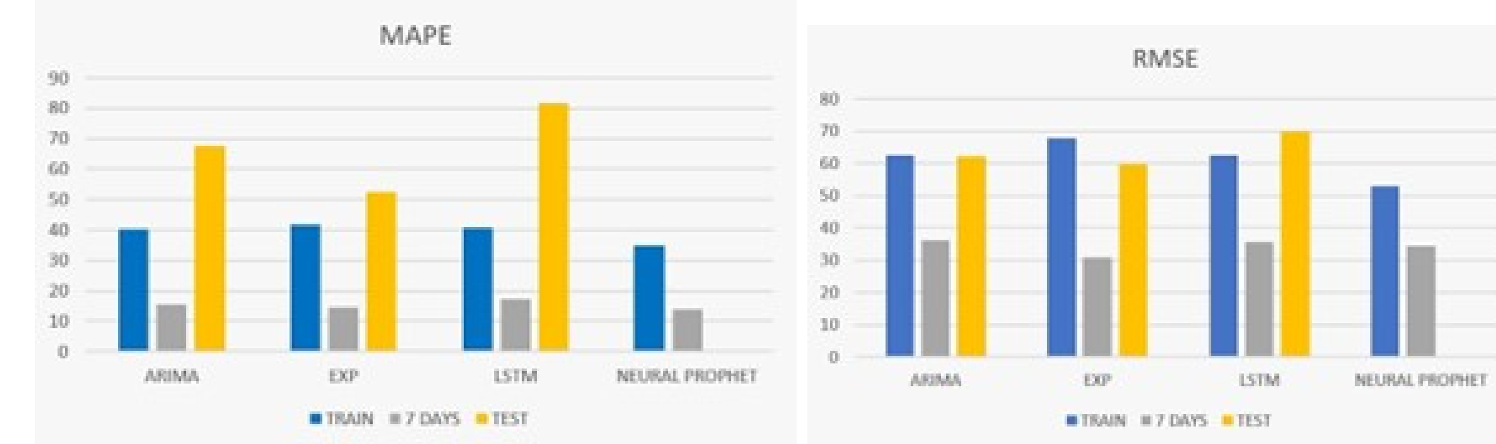
- **Neural Prophet**: Facebook-derived algorithm that uses neural networks.

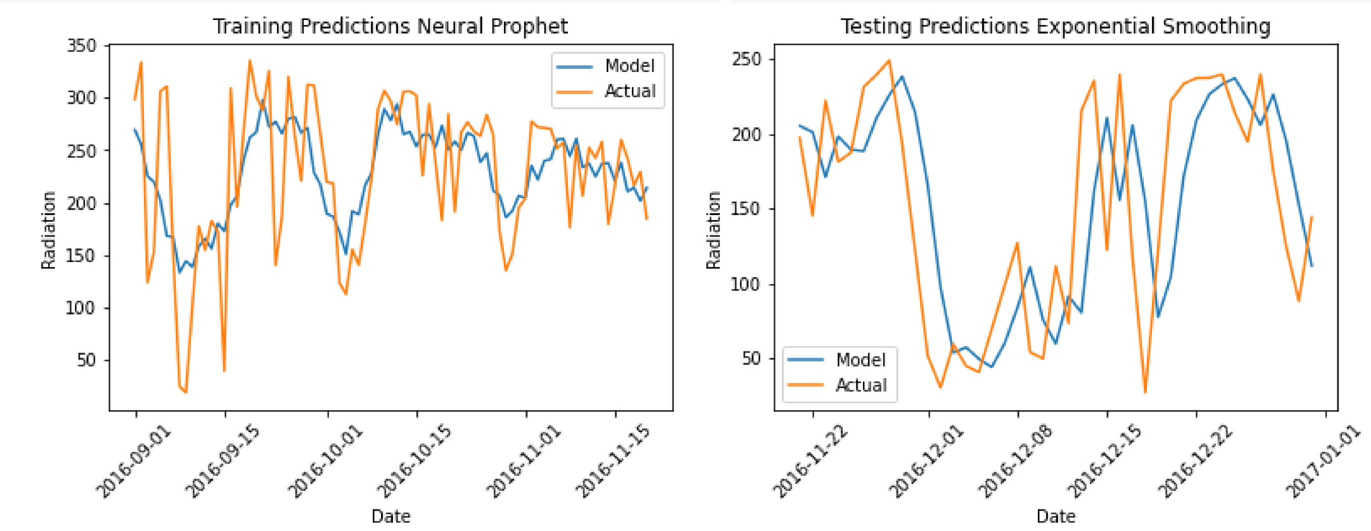
$$y(t) = g(t) + s(t) + h(t) + \epsilon(t)$$

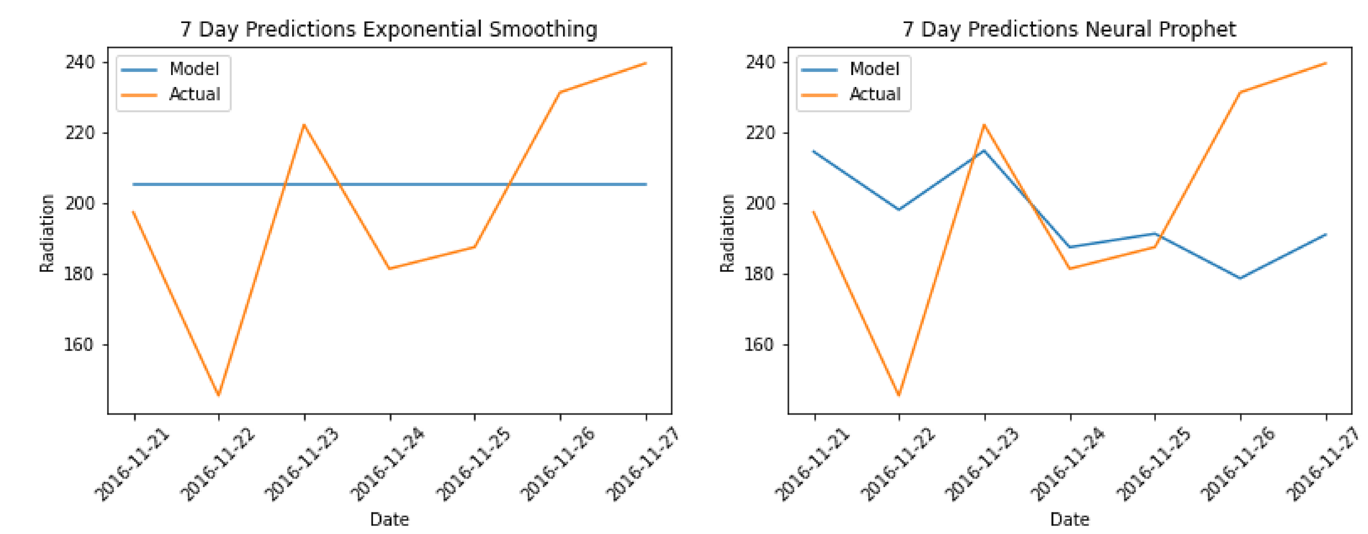
Results & Comparisons

Metric\Model	ARIMA	Exponential Smoothing	LSTM	Neural Prophet
Training MAPE (%)	40.50	41.58	40.60	**34.93**
Testing MAPE (%)	67.53	**52.43**	81.65	-
7 Days MAPE (%)	15.44	14.44	17.15	**13.81**
Training RMSE	62.49	67.96	62.37	**52.94**
Testing RMSE	62.09	**59.82**	69.81	-
7 Days RMSE	36.19	**30.89**	35.69	34.44

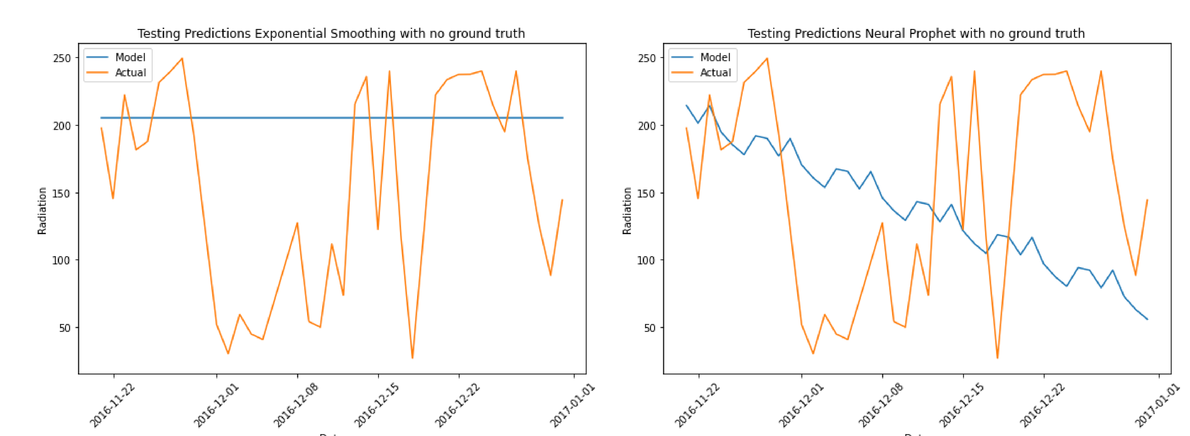
- For the 4 models used, we evaluated both MAPE and RMSE for the training predictions, testing with ground truth predictions and 7-day testing predictions with no use of the test set ground truth.
- In terms of MAPE, the best performing models in training, testing with ground truth, and 7-day testing are Neural Prophet, Exponential Smoothing, and Neural Prophet respectively. The worst in the same order are Exponential Smoothing, LSTM and ARIMA Respectively.







- Although Exponential smoothing achieved a higher RMSE over the 7-day period, over a longer period, Neural Prophet is better.



Conclusion & Future Works

- For 1-day predictions, Exponential Smoothing was the best, while Neural Prophet was the best for multi-day predictions with no ground truth.
- Improve the accuracy of LSTM via Hyperparameter tuning.
- Explore multivariate time-series analysis to improve future predictions.

References