RESEARCH ARTICLE OPEN ACCESS

Weather Forecasting

Dr. Imtiyaz Khan [1], Muhammad Umar Khan [2], Muaaz Mulla [3], Dr.M.Upendra Kumar [4]

[1] Assistant Professor CS and Ai dept. MJCET OU Hyderabad TS India
[2] BE CSE AI DS MJCET OU Hyderabad, [3] BE CSE AI DS MJCET OU Hyderabad
[4] Professor and Associate Head CS and Ai dept. MJCET OU Hyderabad TS India

ABSTRACT

Weather forecast is more helpful for people as it predicts how the future weather is going to be and people may plan accordingly. The activities of many primary sectors depend on the weather for production, e.g. farming.

The climate is changing at a drastic rate nowadays, which makes the old weather prediction methods less effective and more hectic. To overcome these difficulties, the improved and reliable weather prediction methods are required. The weather forecast can be done in many ways like using the previous data or analyzing the current clouds.

The success of machine learning in weather forecasting depends on the quality and quantity of input data, the choice of algorithm, and the design of the prediction system. This abstract highlights the potential of machine learning for weather forecasting and the ongoing efforts to improve its accuracy and reliability.

Keywords: - Machine Learning, Prediction, Forecasting

I. INTRODUCTION

Weather forecasting using machine learning is a method that employs artificial intelligence algorithms to predict future weather conditions. The approach involves training machine learning models on historical weather data to learn the relationship between atmospheric variables and weather patterns. The problem of weather forecasting using machine learning is to develop a system that can accurately predict future weather conditions based on past data and current conditions. The models are then used to make predictions based on current weather conditions and trends. The advantages of using machine learning in weather forecasting include improved accuracy, faster processing times, and the ability to handle large and complex datasets.

In this paper, we will first collect a large dataset of historical weather data from a trusted source. The data will then be pre-processed and cleaned to remove any missing or incorrect values. Next, the Prophet model will be trained on this historical data, using a suitable training-validation split to

ensure the model is able to make accurate predictions on unseen data.

Finally, the trained model will be used to make weather predictions for a given location and time frame, taking into account any known holidays or events that may impact the weather. The results of these predictions will be evaluated against actual weather observations, to determine the accuracy of the model and identify any areas for improvement.

Overall, this paper will demonstrate the power of using time series analysis and the Prophet model for weather forecasting, and provide valuable insights into the patterns and trends in weather data.

II. EXISTING SYSTEM

Currently, the National Oceanic and Atmospheric Administration (NOAA) collects around 100 terabytes of data per day. This data is fed into supercomputers that provide 1 to 10 day forecasts through numerical computation of several physical processes such as atmospheric dynamics, thermal radiation, vegetation, lake and ocean effects, etc. Because there are so many numbers to crunch, these numerical computations take several hours to run. For example, if a numerical computation takes six hours to compute a forecast, it can only run three or four times per day and when the forecast is finally made, it is based on data that is already six hours old.

How Google Is Using Machine Learning to Predict the Weather:

Using radar images, Google treats this as a computer vision problem. They use a "data-driven *physics-free* approach," which means they are not using atmospheric conditions and physics to predict the weather. Instead, they treat weather prediction as an image-to-image translation problem. One where image analysis of radar images and the use of convolutional neural networks (CNNs) can be utilized to predict the weather.

III. PROPOSED SYSTEM

The proposed system for weather forecasting using the Prophet model would leverage the latest advancements in time series forecasting to provide accurate and reliable weather predictions.

The system would gather data from various sources such as weather stations, satellites, and weather forecast agencies. The data would include temperature, precipitation, wind speed and direction, pressure, and other relevant weather parameters.

The collected data would be processed and cleaned to remove any missing values, outliers, or inconsistencies. The data would then be transformed and aggregated as necessary to ensure it is in the right format for modeling.

The Prophet model would be trained on the processed data using advanced machine learning algorithms. The model would be optimized to account for various factors such as seasonality, trends, and fluctuations in the weather data.

The trained Prophet model would then be used to generate weather forecasts for future periods. The forecasts would be based on the past data and would take into account the trends and patterns observed in the data.

The system would continuously evaluate and refine the forecasts by comparing the actual weather data with the predicted values. The model would be updated and retrained periodically to ensure the accuracy of the forecasts.

The proposed system would be highly scalable and would be able to handle large amounts of weather data. It would provide real-time weather forecasts and would be accessible to a wide range of users, including weather forecasters, meteorologists, and the general public.

The methodology for weather forecasting using machine learning typically involves the following steps:

1. Data collection: Collect historical weather data from various sources such as satellites, weather balloons, radar, and ground-based instruments.
2. Data pre-processing: Clean and pre-process the data to remove missing values, outliers, and other anomalies.
3. Feature extraction: Extract relevant features from the data that can be used to train machine learning models. This typically involves calculating various statistics such as mean, standard deviation, and correlation.
4. Model selection: Choose a suitable machine learning algorithm that is appropriate for the problem at hand. This typically involves evaluating different algorithms and choosing the one with the highest accuracy.
5. Model training: Train the chosen machine learning model on the pre-processed data using an optimization algorithm.
6. Model validation: Evaluate the performance of the trained model on a validation dataset to ensure its accuracy and reliability.
7. Model deployment: Deploy the trained model in a real-world weather forecasting system, integrating it with other weather forecasting tools and data sources.

Parameters:

- Temperature
- Wind speed
- Relative Humidity
- Precipitation

Installation:

To follow this project, please install the following locally:

- Python 3.8+
- The packages defined in requirements.txt

It's recommended to use Jupyter Lab

Machine learning model Facebook prophet algorithm:

We'll predict the weather using the Facebook prophet algorithm. Prophet uses an additive model to add up seasonal effects and trends to make a prediction.

The advantage of prophet is that it automatically identifies seasonality in the data -and weather data has strong seasonal effects. So without any feature engineering, you can get good baseline accuracy. It can also scale to multiple time series (think data from adjacent weather stations) easily.

Prophet is an open source time series forecasting library developed by Facebook. It is based on decomposable time series models and uses a Bayesian approach to model non-linear growth and seasonality trends in time series data.

Prophet is designed to work well with time series data that has clear patterns, such as daily, weekly or yearly seasonality, and growth trends that are not too complex. It can handle missing data, outliers and large fluctuations, making it a good choice for weather forecasting.

IV. DATASET

DATE	Temp_max	Temp_min	R_Humid	Precip	Wind_speed
01-01-1982	26.55	12.96	73.38	0.00	2.48
02-01-1982	27.60	13.61	69.19	0.00	2.15
03-01-1982	27.74	13.77	73.00	0.00	2.98
04-01-1982	27.42	14.15	76.00	0.09	3.69
05-01-1982	25.90	14.81	78.81	0.20	2.93
...	...	...	...	...	...
27-12-2022	27.42	18.94	87.81	0.68	2.50
28-12-2022	28.30	18.58	85.38	1.46	1.69
29-12-2022	27.46	18.40	81.38	0.19	1.59
30-12-2022	27.69	17.58	74.19	0.01	1.27
31-12-2022	27.87	16.76	74.56	0.00	1.45

- The dataset is taken from Data Acesss Viewer which is developed by NASA.
- These data include long-term climatologically averaged estimates of meteorological quantities and surface solar energy fluxes.
- The dataset ranges from the year 1982 to 2022.

V. IMPLEMENTATION

- Load in and clean data
- Define targets and predictors
- Train model
- Scale model to entire dataset using cv

Make future predictions

```
# Check invalid columns
weather.apply(lambda x: x == -999.00).sum()
```

```
Temp_max      3
Temp_min      3
R_Humid       3
Precip        3
Wind_speed    3
dtype: int64
```

```
# Clean invalid columns
weather[weather==-999.00]=np.nan
weather = weather.ffill()
```

```
weather.apply(pd.isnull).sum()
```

```
Temp_max      0
Temp_min      0
R_Humid       0
Precip        0
Wind_speed    0
dtype: int64
```

```
weather.index = pd.to_datetime(weather.index)
```

```
# Setup time series for prophet
weather["y"] = weather.shift(-1)["Temp_max"]
weather = weather.ffill()
```

Fig 4.1 Setting up time series for prophet

```
weather["ds"] = weather.index
```

```
weather.shape
```

```
(14975, 7)
```

```
predictors = weather.columns[~weather.columns.isin(["y","ds"])]
```

```
train = weather[:"2021-12-31"]
test = weather["2021-12-31":]
```

```
# Fit initial prophet model
from prophet import Prophet

def fit_prophet(train):
    m = Prophet()
    for p in predictors:
        m.add_regressor(p)
    m.fit(train)
    return m

m = fit_prophet(train)
```

```
23:07:11 - cmdstanpy - INFO - Chain [1] start processing
23:07:14 - cmdstanpy - INFO - Chain [1] done processing
```

```
predictions = m.predict(test)
```

```
predictions
```

	ds	trend	yhat_lower	yhat_upper	trend_lower	trend_upper	Precip	Precip_lower	Precip_upper	R_Humid	...	weekly	weekly_lower
0	2021-12-31	32.452690	24.708477	28.271243	32.452690	32.452690	-0.003841	-0.003841	-0.003841	-0.819212	...	-0.001389	-0.001389
1	2022-01-01	32.452691	24.510569	28.292467	32.452691	32.452691	-0.003841	-0.003841	-0.003841	-0.829809	...	0.001182	0.001182
2	2022-01-02	32.452691	28.875018	32.352478	32.452691	32.452691	-0.003841	-0.003841	-0.003841	0.552892	...	0.018281	0.018281
3	2022-01-03	32.452691	32.185623	35.614621	32.452691	32.452691	-0.003841	-0.003841	-0.003841	0.902595	...	-0.007633	-0.007633
4	2022-01-04	32.452692	37.948032	41.466762	32.452692	32.452692	-0.003841	-0.003841	-0.003841	0.796624	...	0.008182	0.008182
...	...	...	...	...	...	...	...	...	...	...	...	...	...
361	2022-12-27	32.452837	25.386283	28.949335	32.452334	32.453209	-0.002682	-0.002682	-0.002682	-1.158318	...	0.008182	0.008182
362	2022-12-28	32.452837	26.009877	29.452142	32.452329	32.453211	-0.001353	-0.001353	-0.001353	-1.055314	...	-0.057056	-0.057056
363	2022-12-29	32.452837	25.696122	29.236094	32.452323	32.453212	-0.003518	-0.003518	-0.003518	-0.885762	...	0.038433	0.038433
364	2022-12-30	32.452838	26.142815	29.677373	32.452316	32.453216	-0.003824	-0.003824	-0.003824	-0.580991	...	-0.001389	-0.001389
365	2022-12-31	32.452838	26.192095	29.622412	32.452313	32.453221	-0.003841	-0.003841	-0.003841	-0.596674	...	0.001182	0.001182

Figure 4.2 Fitting of prophet model and predicting

Visualization:

```
#Visualising Time Series(Temperature Over Time)
weather['2018-01-01':].plot('ds','y')
```

```
<AxesSubplot:xlabel='ds'>
```

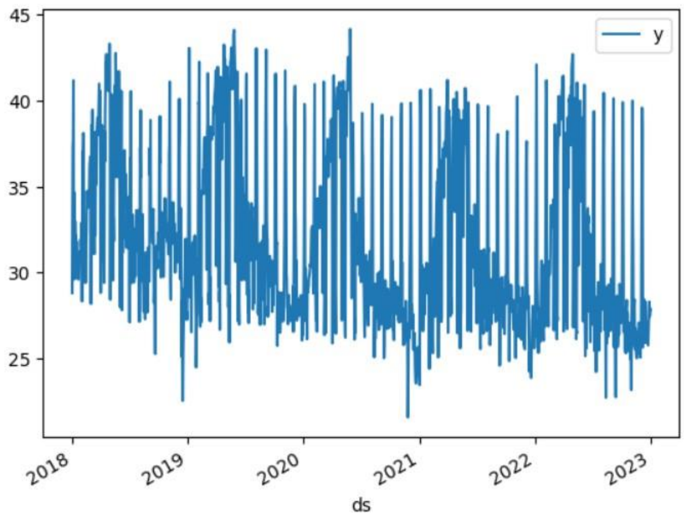

Fig4.3Visualizing temperature over years

Fig 4.4 Trends in Weather Data

```python
from prophet.utilities import regressor_coefficients

regressor_coefficients(m)
```

	regressor	regressor_mode	center	coef_lower	coef	coef_upper
0	Temp_max	additive	32.456864	0.716293	0.716293	0.716293
1	Temp_min	additive	20.556697	0.130366	0.130366	0.130366
2	R_Humid	additive	60.483559	-0.042388	-0.042388	-0.042388
3	Precip	additive	2.253426	0.001705	0.001705	0.001705
4	Wind_speed	additive	3.765791	0.013371	0.013371	0.013371

Fig 4.5 Regressor Coefficients

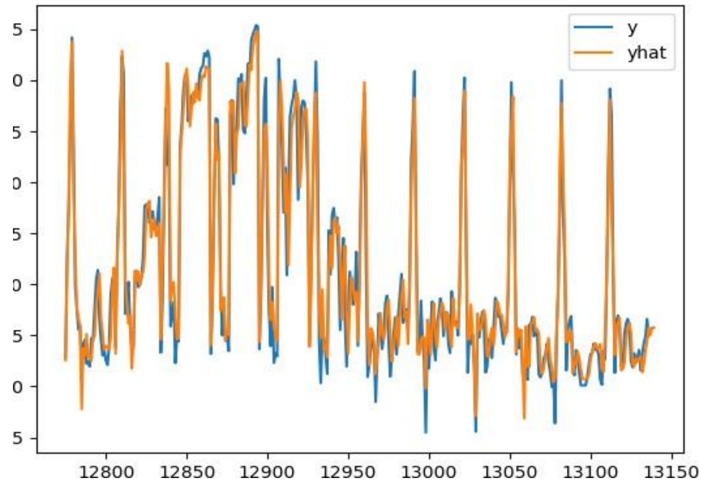

Fig 4.7actual (y) vs predicted (yhat) weather

Prediction:

```python
# Predict one day ahead with high accuracy
m = fit_prophet(weather)
m.predict(weather.iloc[-1:])
```

```
23:09:47 - cmdstanpy - INFO - Chain [1] start processing
23:09:49 - cmdstanpy - INFO - Chain [1] done processing
```

	ds	trend	yhat_lower	yhat_upper	trend_lower	trend_upper	Precip	Precip_lower	Precip_upper	R_Humid	...	weekly	weekly_lower	weekly_up
0	2022-12-31	32.390398	26.086519	29.626329	32.390398	32.390398	-0.003664	-0.003664	-0.003664	-0.589287	...	0.00227	0.00227	0.002

1 rows × 37 columns

Fig 4.6 Predict one day ahead with high accuracy

Mean square error across data set:

```python
In [21]: mse(cv, actual_label="y")
```

```
1.9197731141332992
```

Fig 4.8 mean square error

VI. RESULT ANALYSIS

1. Predicting the weather for a given date:
2. Using Prophet model the we can predict weather for multiple days and evenyears into the future.
3. We can obtain the trend of the weather from this ML model.
4. This project can be used to forecast weather instead of existing system.

```python
# Predict multiple days ahead with lower accuracy
m = Prophet()
m.fit(weather)
future = m.make_future_dataframe(periods=3650)
forecast = m.predict(future)
```

```
23:09:51 - cmdstanpy - INFO - Chain [1] start processing
23:09:55 - cmdstanpy - INFO - Chain [1] done processing
```

Fig4.7 Predicting multiple days ahead

Performance analysis:

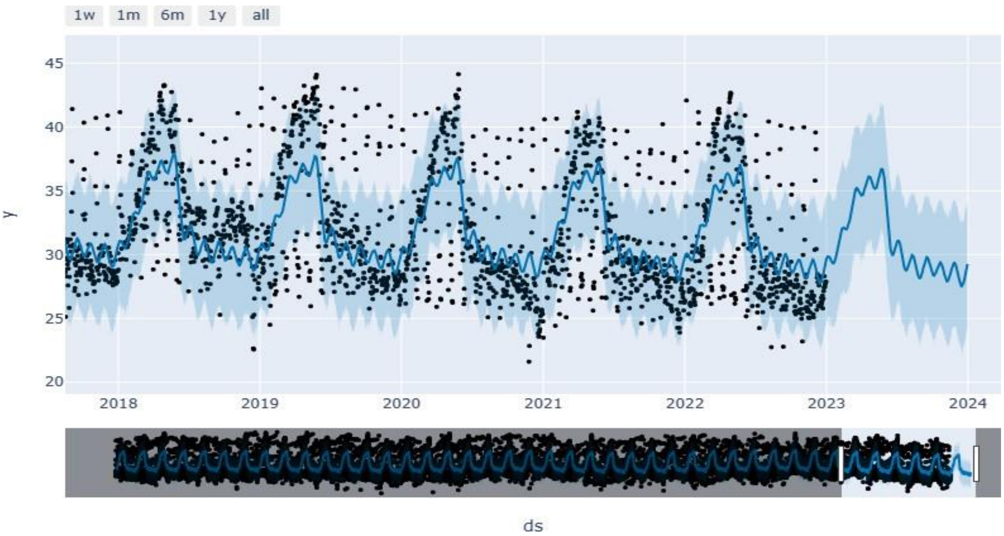

Fig 5.1 Weather prediction of past and for future

VIII. CONCLUSION

Machine learning has been applied in various forms to weather forecasting, and its results have been shown to be promising in some cases. However, weather forecasting is a complex problem and the accuracy of machine learning models can still be improved.

Additionally, machine learning models for weather forecasting typically require large amounts of data and computing resources, and their predictions can still be subject to errors and biases. Overall, machine learning has shown potential for improving weather forecasting, but there is still room for further research and development.

REFERENCES

[1] Dataset: The dataset is taken from NASA power access viewer. https://power.larc.nasa.gov/data-access-viewer

[2] "Time Series Analyzation and Prediction of Climate using Enhanced Multivariate Prophet" by J. Jagannathan and C. Divya : https://www.researchgate.net/publication/356053648_Time_Series_Analyzation_and_Prediction_of_Climate_using_Enhanced_Multivariate_Prophet

[3] "Time Series Prediction Based on Facebook Prophet: A Case Study, Temperature Forecasting in Myintkyina" by Sabai Phyu and Zar Zar Oo : https://doi.org/10.18100/ijamec.816894