Weather Prediction

A MINI PROJECT REPORT

Submitted by

ARYAN CHAKRABORTY
(RA2011003011043)

AVIK KUMAR
(RA2011003011036)

SHIKHAR PANDEY
(RA2011003011018)

Under the guidance of
Dr . L Kavisankar
(Assistant Prof, CTECH)

In partial satisfaction of the requirements for the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING
With specialization in COMPUTER SCIENCE

SCHOOL OF COMPUTING
COLLEGE OF ENGINEERING AND TECHNOLOGY
SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
KATTANKULATHUR - 603203 , May 2023

COLLEGE OF ENGINEERING & TECHNOLOGY
SRM INSTITUTE OF SCIENCE & TECHNOLOGY
S.R.M. NAGAR, KATTANKULATHUR – 603 203

BONAFIDE CERTIFICATE

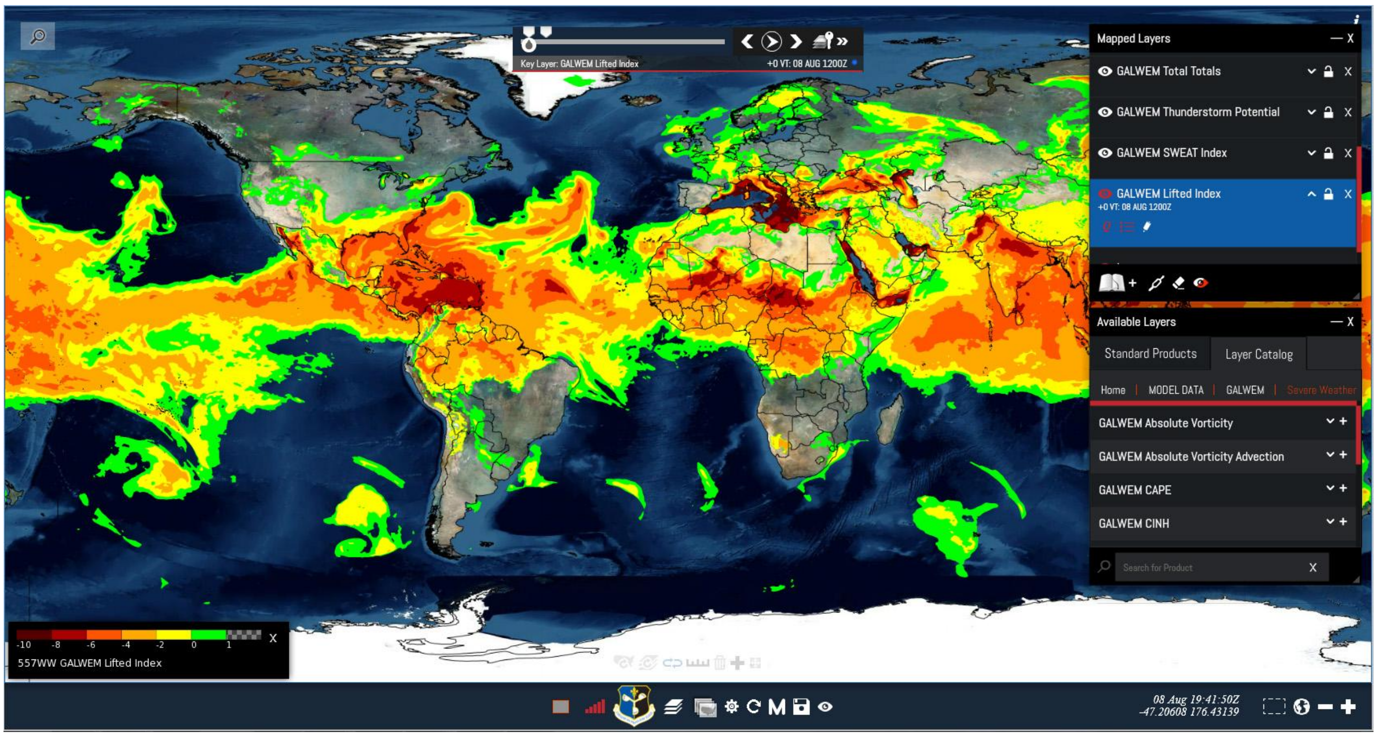
Certified that this project report **Weather Prediction model** is the bonafide work of **ARYAN CHAKRABORTY (RA2011003011043), AVIK KUMAR (RA2011003011036) , SHIKHAR PANDEY(RA2011003011018)** of III Year/VI Sem B.tech(CSE) who carried out the mini project work under my supervision for the course 18CSC305J- Artificial Intelligence in SRM Institute of Science and Technology during the academic year 2022-2023(Even sem).

SIGNATURE
Dr. L Kavisankar
Assistant Professor
Department of Computing Technologies

SIGNATURE
Dr. M. PUSHPALATHA
HEAD OF THE DEPARTMENT
Department of Computing Technologies

TABLE OF CONTENTS

WEATHER PREDICTION MODEL



ABSTRACT

A weather prediction model is a computational tool that uses mathematical algorithms and statistical techniques to analyze current and historical weather data, and to make predictions about future weather conditions. These models typically use large amounts of data from weather stations, satellites, and other sources to create a detailed picture of atmospheric conditions, which is then used to generate weather forecasts.

The basic idea behind a weather prediction model is to use mathematical equations to describe how weather variables like temperature, pressure, and humidity interact with each other over time. These equations are then solved using a computer, which allows meteorologists to generate predictions about future weather conditions.

One of the main advantages of weather prediction models is that they allow forecasters to make more accurate and detailed predictions than would be possible using traditional weather observation methods alone. By using data from a wide range of sources, and by incorporating complex mathematical algorithms, these models are able to generate forecasts with a high degree of accuracy.

However, it is important to note that no weather prediction model is perfect, and there will always be some degree of uncertainty associated with any forecast. Nevertheless, these models are a valuable tool for meteorologists and other weather professionals, as they can help to improve the accuracy of weather forecasts, and to provide important information to people who need to make decisions based on weather conditions, such as farmers, transportation providers, and emergency responders.

Traditionally, climate assessment has been performed reliably by treating the environment as a liquid. This indicates an insufficient understanding of environmental variations, so it limits climate forecasts to 10-day periods because climate projections are essentially unreliable. But machine learning is moderately hearty for most barometric destabilizing effects compared to traditional techniques. Another favorable position of machine learning is that it does not depend on the physical laws of environmental processes.

Background

Given the limitations of traditional weather forecasting methods in India, the project aims to increase accuracy and predict future weather using machine learning techniques. By incorporating machine learning algorithms into weather prediction, it has the potential to improve accuracy and extend the forecasting period.
Machine learning algorithms can analyze large volumes of historical weather data and identify patterns and relationships that are difficult for humans to discern. This analysis can lead to more accurate predictions, especially when dealing with complex atmospheric conditions and interdependencies.

To apply machine learning techniques to weather forecasting, we need a dataset consisting of historical weather data, including various atmospheric variables such as temperature, wind speed, pressure systems, rainfall, and other relevant conditions. This dataset should cover a sufficiently long period to capture different weather patterns and variations.

The training process involves feeding the model with past weather data and their corresponding outcomes (i.e., the weather conditions observed in the future). The model will learn the relationships between the input variables like

- temperature
- wind speed
- humidity
- previous precipitation information

These will be used to predict the target variable (e.g., weather conditions). After training, the model can make predictions for future weather based on new input data.

To increase the accuracy of the predictions, it's crucial to ensure the quality and reliability of the input data. This includes

- data cleaning
- handling missing values
- addressing any biases or anomalies present in the dataset

Additionally, we can incorporate external factors such as satellite imagery, oceanic data, or regional climatic patterns to enhance the predictive capabilities of your model.

Machine learning models are not infallible, and they rely on the quality and representativeness of the training data. Regular monitoring and evaluation of the model's performance are necessary to identify any biases, errors, or limitations. Continuously updating the model with new data and refining the algorithms can further improve its accuracy over time.

In conclusion, by leveraging machine learning techniques and applying them to weather forecasting in India, you have the potential to overcome the limitations of traditional methods. With accurate predictions extending beyond the current month and incorporating various atmospheric conditions, your project can contribute to more reliable weather forecasts, enabling better preparedness and decision-making in various sectors that rely on weather information.

OBJECTIVE

The purpose of the project is to predict temperature using various algorithms such as linear regression, random forest regression, and decision tree regression.
The output of these algorithms will be numerical values based on multiple additional factors including :

1. maximum temperature
2. minimum temperature
3. cloud cover
4. humidity
5. sun hours in a day
6. precipitation
7. pressure
8. wind speed.

To accomplish this, we need a dataset that includes historical weather data with the variables mentioned (maximum temperature, minimum temperature, cloud cover, humidity, sun hours, precipitation, pressure, and wind speed) as input features, as well as the corresponding temperature values as the target variable.

Here's an overview of the approach to this project:

1. **Data Collection**: Gather historical weather data that includes the desired variables for the time period and location of interest. This data can be obtained from various sources, such as meteorological organizations, weather APIs, or publicly available datasets.

2. **Data Preprocessing**: Clean the data and handle any missing values or outliers. Ensure that all variables are in the appropriate format and scale for the algorithms you plan to use. Split the data into training and testing sets to evaluate the performance of your models.

3. **Feature Selection**: Analyze the importance of each input variable in relation to the target variable (temperature). You can use techniques such as correlation analysis or

feature importance from random forest models to determine the most relevant features for temperature prediction.

4. **Algorithm Selection**: Choose suitable regression algorithms for your task, such as linear regression, random forest regression, and decision tree regression. These algorithms have different strengths and weaknesses, so it's beneficial to experiment with multiple models to see which performs best for your specific dataset.

5. **Model Training and Evaluation**: Train each regression model on the training dataset and evaluate their performance using appropriate evaluation metrics (e.g., mean squared error, mean absolute error, or R-squared). Adjust the model parameters as needed to optimize performance.

6. **Model Comparison**: Compare the performance of the different algorithms to determine which one provides the most accurate temperature predictions. Consider factors such as prediction accuracy, computational efficiency, and interpretability.

7. **Prediction**: Once you have selected the best-performing model, you can use it to make temperature predictions for future time periods based on the input features you have available. Monitor the model's performance over time and retrain it periodically with updated data to ensure its accuracy.

SCOPE AND APPLICATIONS

The scope and applications of weather prediction models are extensive and can benefit various industries and sectors. The scope of weather prediction models is to provide accurate and reliable predictions of future weather conditions. This information can help people and organizations make informed decisions regarding their activities and operations.

One of the primary applications of weather prediction models is in the agricultural sector. Farmers can use weather forecasts to plan their crop cycles, irrigation schedules, and other farming practices. This helps to optimize yields and minimize the risk of crop failure due to extreme weather conditions.

Another significant application of weather prediction models is in the transportation sector. Accurate weather forecasts can help airlines and shipping companies plan their routes, schedule departures and arrivals, and avoid potential weather-related disruptions. This can improve operational efficiency, reduce costs, and improve customer satisfaction.

The energy sector is another industry that relies on weather prediction models. The production of renewable energy, such as wind and solar power, is heavily influenced by weather conditions. Therefore, accurate weather forecasts can help energy companies plan and optimize their production schedules.

Finally, weather prediction models can also benefit the general public by providing timely and accurate warnings of severe weather conditions such as hurricanes, tornadoes, and floods. This can help people to take appropriate safety measures and minimize the risk of injury or damage to property.

Overall, the scope and applications of weather prediction models are broad, and their importance in various industries and sectors cannot be overstated.

INTRODUCTION

Weather prediction is the task of predicting the atmosphere at a future time and a given area. This has been done through physical equations in the early days in which the atmosphere is considered fluid. The current state of the environment is inspected, and the future state is predicted by solving those equations numerically, but we cannot determine very accurate weather for more than 10 days and this can be improved with the help of science and technology.

Machine learning can be used to process immediate comparisons between historical weather forecasts and observations. With the use of machine learning, weather models can better account for prediction inaccuracies, such as overestimated rainfall, and produce more accurate predictions. Temperature prediction is of major importance in a large number of applications, including climate-related studies, energy, agricultural, medical, or etc.

There are numerous kinds of machine learning calculations, which are Linear Regression, Polynomial Regression, Random Forest Regression, Artificial Neural Network, and Recurrent Neural Network. These models are prepared depending on the authentic information given of any area. Contribution to these models is given, for example, if anticipating temperature, least temperature, mean air weight, greatest temperature, mean dampness, and order for 2 days. In light of this Minimum Temperature and Maximum Temperature of 7 days will be accomplished.

Machine Learning

Machine learning has the potential to improve weather forecasting by leveraging its robustness to perturbations and its ability to analyze large volumes of data. The advancement of technology, including the use of the Internet of Things (IoT), data science, and artificial intelligence, has indeed contributed to increased accuracy and predictability in weather forecasting.

In the past, weather forecasting relied on limited data sources and less accurate atmospheric models. However, with the availability of more advanced technologies, such as satellites and weather monitoring systems, we now have access to a vast amount of real-time data on various atmospheric variables. This data, combined with historical records, provides a rich source of information for training machine learning models.

Machine learning algorithms can effectively analyze and identify complex patterns in weather data. They can learn from historical observations and capture the relationships between different variables, allowing for more accurate predictions. These algorithms are capable of considering a wide range of factors simultaneously, including temperature, pressure systems, wind patterns, humidity, and more, to generate forecasts.

Additionally, the integration of artificial intelligence techniques, such as deep learning, has further improved the accuracy of weather forecasting. Deep learning models, such as convolutional neural networks (CNNs) or generative adversarial networks (GANs), can extract meaningful features from meteorological data and generate highly accurate predictions.

The advancements in data science have also played a significant role in weather forecasting. Data scientists can clean and preprocess large datasets, handle missing values, and address data biases, ensuring the quality and reliability of input data for machine learning models. They can also employ advanced statistical techniques to validate and evaluate the performance of these models.

Furthermore, the use of ensemble methods, where multiple machine learning models are combined, has become common in weather forecasting. By aggregating predictions from different models, scientists can reduce uncertainties and increase the overall accuracy of the forecasts. It's important to note that while machine learning and advanced technologies have greatly improved weather forecasting, they still have limitations. Weather is a highly complex and chaotic system, influenced by numerous factors, and there are inherent uncertainties involved in predicting its behavior. Therefore, continuous research, model refinement, and collaboration between meteorologists, data scientists, and domain experts are crucial to further enhance the accuracy and reliability of weather forecasts.

In summary, the advancement of technology, including the use of machine learning, data science, and artificial intelligence, has significantly improved weather forecasting over the years. These technologies have allowed for the analysis of large volumes of data, capturing complex patterns, and generating predictions. While challenges and uncertainties remain, the integration of advanced technologies offers a promising opportunity for further advancements in weather forecasting.

USE OF ALGORITHMS:

The project relies on a dataset comprising historical weather data, including the aforementioned input variables and corresponding temperature values. The data is divided into training and testing sets, with 80% of the data used for training the algorithms and 20% reserved for testing the models' performance. For instance, if the goal is to predict the temperature in Kanpur, India, a training dataset spanning 8 years of historical weather data would be utilized, while 2 years of data would be allocated for testing.

Regression algorithms, such as linear regression, decision tree regression, and random forest regression, are employed to build predictive models. These algorithms analyze the relationships between the input variables and the target variable (temperature) to generate accurate predictions. Each algorithm is trained using the training dataset and fine-tuned to optimize its performance.

To evaluate the models' effectiveness, various evaluation metrics such as mean squared error, mean absolute error, and R-squared are utilized. The models are compared based on their prediction accuracy, computational efficiency, and interpretability. The best-performing model is selected for temperature prediction.

1. Multiple Linear Regression:

Multiple linear regression is a powerful statistical method used in weather prediction to analyze the relationships between multiple predictor variables and a target variable, such as precipitation or temperature. By considering multiple factors simultaneously, this technique helps estimate the impact of each predictor variable on the target variable while accounting for the influence of other variables.

In the context of weather prediction, the predictor variables can include various atmospheric conditions such as temperature, humidity, wind speed, cloud cover, air pressure, and more. These variables are selected based on their relevance and potential impact on the target variable, which could be temperature or any other weather parameter of interest.

The multiple linear regression model assumes a linear relationship between the predictor variables and the target variable. It estimates the coefficients for each predictor, representing the magnitude and direction of their influence on the target variable. The model takes into account the combined effect of all the predictor variables to make predictions or forecast the target variable. Multiple linear regression provides valuable insights into the relationships between predictor variables and the target variable, helping meteorologists and weather forecasters make more accurate predictions. It allows for the consideration of multiple factors simultaneously, capturing the complex interactions between different atmospheric conditions and their impact on weather patterns Overall, multiple linear regression serves as a useful tool in weather prediction by leveraging the relationships between multiple predictor variables and a target variable. It contributes to improving the accuracy and understanding of weather forecasts by incorporating various atmospheric conditions into the prediction process.

1. Decision Tree Regression:

Decision tree algorithms are commonly used for weather prediction due to their ability to handle both numerical and categorical data and their interpretability. Decision trees recursively partition the data based on the values of predictor variables and make predictions based on the majority class or mean value within each partitionDecision trees offer several advantages for weather prediction tasks. They are interpretable, allowing meteorologists to understand the decision-making process of the model and gain insights into the relationships between predictor variables and the target variable. Decision trees can handle both numerical and categorical data, making them suitable for weather prediction tasks that involve a mix of different types of variables. They are also robust to outliers and can capture nonlinear relationships between variables.Decision tree algorithms are widely used in weather prediction tasks due to their interpretability and ability to handle diverse types of variables. By recursively partitioning the data based on predictor variables, decision trees can provide accurate predictions for weather parameters such as precipitation or temperature categories.

2. Random Forest Regression:

Random Forest Regression is a powerful machine learning algorithm that can be effectively applied to weather prediction tasks. It is an ensemble method that combines multiple decision trees to create a robust and accurate predictive model. Random Forest Regression offers several advantages for weather prediction due to its ability to handle complex relationships, handle both numerical and categorical data, and mitigate overfitting.
Random Forest Regression offers several benefits for weather prediction tasks. It can handle nonlinear relationships between predictor variables and the target variable, capturing complex patterns in the data. The ensemble nature of Random Forest helps improve generalization and reduces the risk of overfitting. Additionally, it can handle both numerical and categorical data, making it suitable for weather prediction tasks that involve various types of variables

Random Forest Regression is a powerful algorithm for weather prediction tasks. By combining multiple decision trees and leveraging their collective predictions, Random Forest Regression provides accurate and robust predictions for weather parameters. It handles complex relationships, can handle various types of variables, and offers improved generalization capabilities.

DATASET DETAILS

The dataset utilized in this arrangement has been gathered from Kaggle which is "Historical Weather Data for Indian Cities" from which we have chosen the data for "Kanpur City". The dataset was created by keeping in mind the necessity of such historical weather data in the community. The datasets for the top 8 Indian cities as per the population. The dataset was used with the help of the worldweatheronline.com API and the wwo_hist package. The datasets contain hourly weather data from 01-01-2009 to 01-01-2020. The data of each city is for more than 10 years. This data can be used to visualize the change in data due to global warming or can be used to predict the weather for upcoming days, weeks, months, seasons, etc.
Note: The data was extracted with the help of worldweatheronline.com API and we cannot guarantee the accuracy of the data.
The main target of this dataset can be used to predict the weather for the next day or week with huge amounts of data provided in the dataset. Furthermore, this data can also be used to make visualizations which would help to understand the impact of global warming over the various aspects of the weather like precipitation, humidity, temperature, etc.

In this project, we are concentrating on the temperature prediction of Kanpur city with the help of various machine learning algorithms and various regressions. By applying various regressions on the historical weather dataset of Kanpur city we are predicting the temperature like first we are applying Multiple Linear regression, then Decision Tree regression, and after that, we are applying Random Forest Regression.

Table 2.1: Historical Weather Dataset of Kanpur City

	maxtempC	mintempC	cloudcover	humidity	tempC	sunHour	precipMM	pressure	windspeedKmph
date_time									
2009-01-01 00:00:00	24	10	17	50	11	8.7	0.0	1015	10
2009-01-01 01:00:00	24	10	11	52	11	8.7	0.0	1015	11
2009-01-01 02:00:00	24	10	6	55	11	8.7	0.0	1015	11
2009-01-01 03:00:00	24	10	0	57	10	8.7	0.0	1015	12
2009-01-01 04:00:00	24	10	0	54	11	8.7	0.0	1016	11

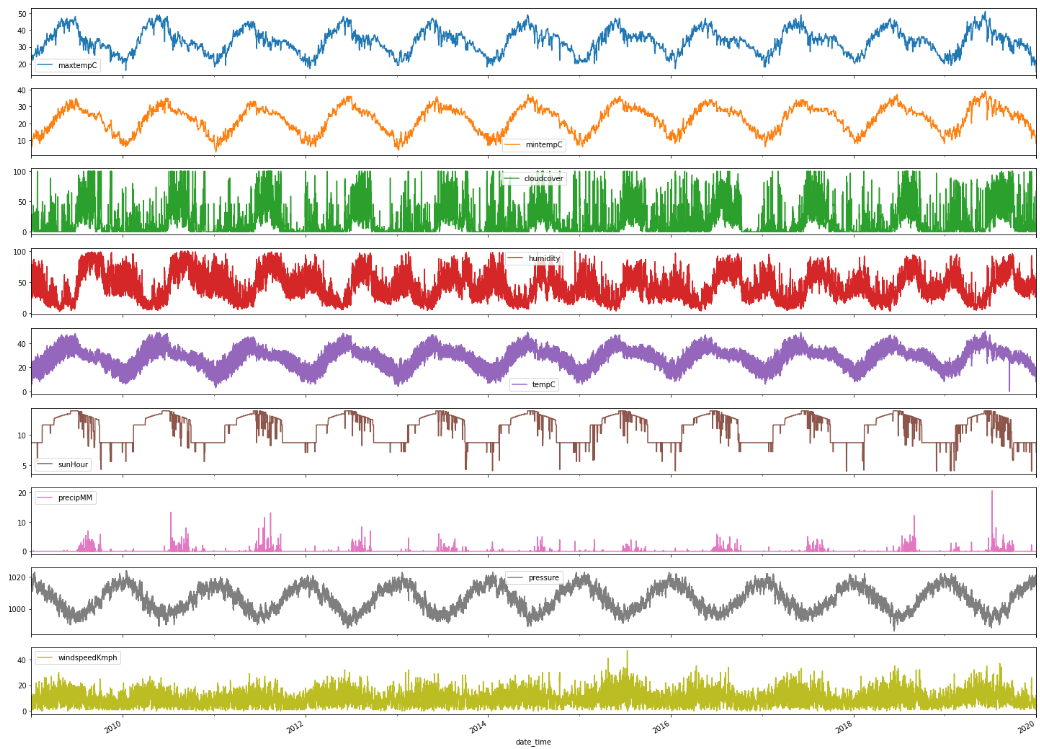


Figure 2.1: Plot for each factor for 10 years

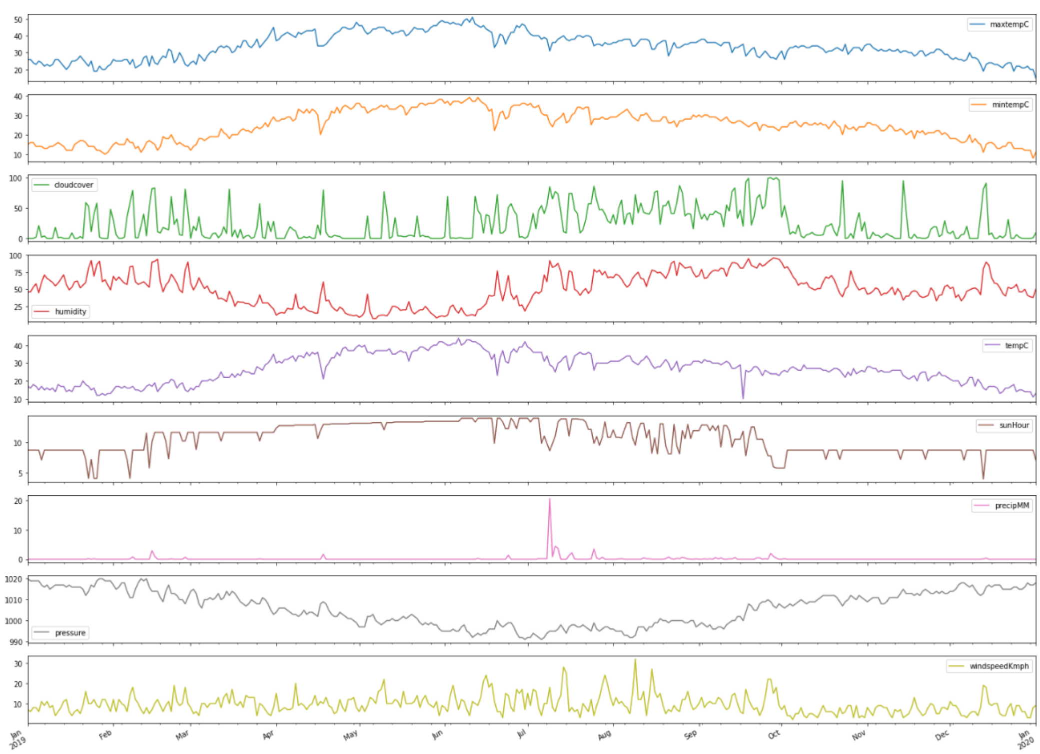


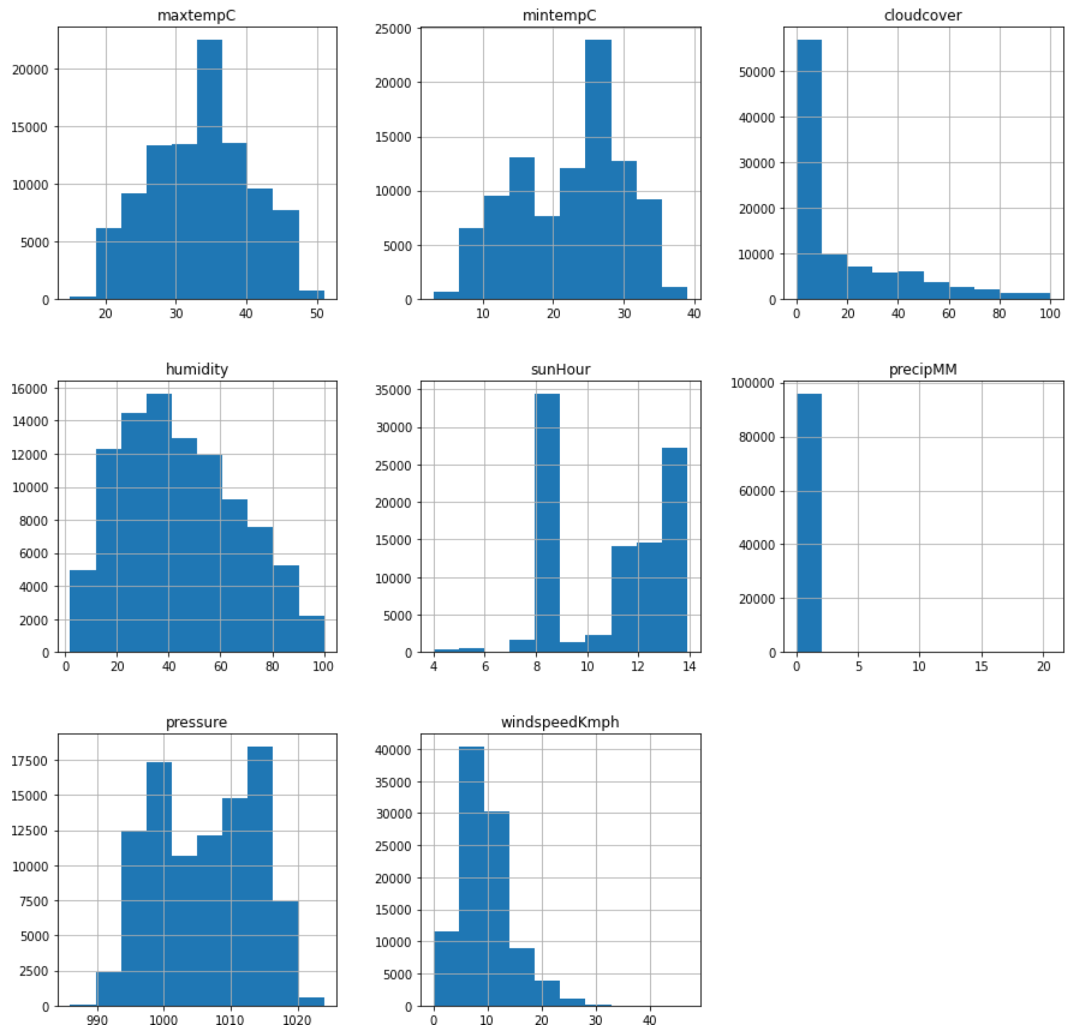
Figure 2.2: Plot for each factor for 1 year

EXPERIMENTATION

The record has just been separated into a train set and a test set. Each information has just been labeled. First, we take the train set organizer. We will train our model with the help of histograms and plots. The feature so extracted is stored in a histogram. This process is done for every data in the train set. Now we will build the model of our classifiers.

The classifiers which we will take into account are Linear Regression, Decision Tree Regression, and Random Forest Regression. With the help of our histogram, we will train our model. The most important thing in this process is to tune these parameters accordingly, such that we get the most accurate results. Once the training is complete, we will take the test set.

Now for each data variable of the test set, we will extract the features using feature extraction techniques and then compare its values with the values present in the histogram formed by the train set. The output is then predicted for each test day. Now in order to calculate accuracy, we will compare the predicted value with the labeled value. The different metrics that we will use are the confusion matrix, R2 score, etc.


maxtempC
mintempC
cloudcover
humidity
sunHour
precipMM
pressure
windspeedKmph

CODING AND TESTING

```
import warnings
warnings.filterwarnings('ignore')
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt


import sklearn
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.linear_model import LinearRegression
from sklearn import preprocessing


%matplotlib inline


weather_df = pd.read_csv('kanpur.csv', parse_dates=['date_time'], index_col='date_time')
weather_df.head(5)
weather_df.columns
weather_df.shape
weather_df.describe()
weather_df.isnull().any()


weather_df_num=weather_df.loc[:,['maxtempC','mintempC','cloudcover','humidity','tempC'
, 'sunHour','HeatIndexC', 'precipMM', 'pressure','wind speed mph']]
weather_df_num.head()
weather_df_num.plot(subplots=True, figsize=(25,20))
weather_df_num['2019':'2020'].resample('D').fillna(method='pad').plot(subplots=True,
figsize=(25,20))
weather_df_num.hist(bins=10,figsize=(15,15))
```

```
train_X,test_X,train_y,test_y=train_test_split(weather_x,weather_y,test_size=0.2,random_s
tate=4)
train_X.shape
train_y.shape
train_y.head()
```

```
model=LinearRegression()
model.fit(train_X,train_y)
prediction = model.predict(test_X)
np.mean(np.absolute(prediction-test_y))
print('Variance score: %.2f' % model.score(test_X, test_y))
for i in range(len(prediction)):
  prediction[i]=round(prediction[i],2)
pd.DataFrame({'Actual':test_y,'Prediction':prediction,'diff':(test_y-prediction)})
```

```
from sklearn.tree import DecisionTreeRegressor
regressor=DecisionTreeRegressor(random_state=0)
```

```
regressor.fit(train_X,train_y)
prediction=regression.predict(test_X)
np.mean(np.absolute(prediction2-test_y))
```

```
print('Variance score: %.2f' % regressor.score(test_X, test_y))
for i in range(len(prediction2)):
  prediction2[i]=round(prediction2[i],2)
pd.DataFrame({'Actual':test_y,'Prediction':prediction2,'diff':(test_y-prediction2)})
```

```
from sklearn.ensemble import RandomForestRegressor
regr=RandomForestRegressor(max_depth=90,random_state=0,n_estimators=100)
regr.fit(train_X,train_y)
```

```
prediction3=regr.predict(test_X)
np.mean(np.absolute(prediction3-test_y))
print('Variance score: %.2f' % regr.score(test_X, test_y))
for i in range(len(prediction3)):
  prediction3[i]=round(prediction3[i],2)
pd.DataFrame({'Actual':test_y,'Prediction':prediction3,'diff':(test_y-prediction3)})
```

RESULT AND DISCUSSION

The results of the implementation of the project are demonstrated below.

Multiple Linear Regression:

This regression model has high mean absolute error, hence turned out to be the least accurate model. Given below is a snapshot of the actual result from the project implementation of multiple linear regression.

date_time	Actual	Prediction	diff
2013-07-10 08:00:00	34	33.209030	0.790970
2015-11-04 20:00:00	25	25.275755	-0.275755
2015-09-21 09:00:00	34	31.975338	2.024662
2017-02-16 11:00:00	28	20.496727	7.503273
2012-07-21 01:00:00	28	28.401085	-0.401085
...	...	...	...
2019-03-30 09:00:00	37	33.187428	3.812572
2015-11-12 12:00:00	32	28.483724	3.516276
2019-12-31 05:00:00	8	15.177361	-7.177361
2019-08-02 17:00:00	35	35.363251	-0.363251
2019-10-22 08:00:00	26	27.890691	-1.890691

19287 rows × 3 columns

Decision Tree Regression:

This regression model has medium mean absolute error, hence turned out to be the little accurate model. Given below is a snapshot of the actual result from the project implementation of multiple linear regression.

	Actual	Prediction	diff
date_time			
2013-07-10 08:00:00	34	34.0	0.0
2015-11-04 20:00:00	25	25.0	0.0
2015-09-21 09:00:00	34	34.0	0.0
2017-02-16 11:00:00	28	28.0	0.0
2012-07-21 01:00:00	28	28.0	0.0
...	...	...	...
2019-03-30 09:00:00	37	39.0	-2.0
2015-11-12 12:00:00	32	32.0	0.0
2019-12-31 05:00:00	8	9.0	-1.0
2019-08-02 17:00:00	35	36.0	-1.0
2019-10-22 08:00:00	26	27.0	-1.0

19287 rows × 3 columns

Random Forest Regression:

This regression model has low mean absolute error, hence turned out to be the more accurate model. Given below is a snapshot of the actual result from the project implementation of multiple linear regression.

	Actual	Prediction	diff
date_time			
2013-07-10 08:00:00	34	33.94	0.06
2015-11-04 20:00:00	25	24.43	0.57
2015-09-21 09:00:00	34	34.36	-0.36
2017-02-16 11:00:00	28	26.35	1.65
2012-07-21 01:00:00	28	28.17	-0.17
...	...	...	...
2019-03-30 09:00:00	37	32.99	4.01
2015-11-12 12:00:00	32	31.74	0.26
2019-12-31 05:00:00	8	10.62	-2.62
2019-08-02 17:00:00	35	35.72	-0.72
2019-10-22 08:00:00	26	26.85	-0.85

19287 rows × 3 columns

R2 Score:

Multiple Linear Regression

```
Mean absolute error: 1.20
Residual sum of squares (MSE): 2.51
R2-score: 0.96
```

Decision Tree Regression:

```
Mean absolute error: 0.56
Residual sum of squares (MSE): 1.12
R2-score: 0.98
```

Random Forest Regression

```
Mean absolute error: 0.47
Residual sum of squares (MSE): 0.63
R2-score: 0.99
```

FUTURE ENHANCEMENTS

There are several potential future enhancements that can be made to weather prediction models:

- Increased resolution: Weather prediction models could be enhanced by increasing their resolution, allowing for more accurate predictions of weather patterns in smaller geographic areas.

- Incorporation of new data sources: As new data sources become available, weather prediction models can be enhanced by incorporating this data into their algorithms. For example, the use of satellite imagery, ground-based sensors, and data from social media could all be useful additions to weather prediction models.

- Use of machine learning: Machine learning algorithms could be applied to weather prediction models to identify patterns and correlations in weather data that may not be immediately apparent to human analysts. This could help to improve the accuracy of predictions.

- Improved modeling of complex weather systems: Complex weather systems, such as hurricanes and tornadoes, are notoriously difficult to predict. Future enhancements to weather prediction models could focus on improving our understanding of these systems, allowing for more accurate predictions.

- Real-time updates: Real-time updates to weather prediction models could be used to provide more accurate and timely predictions of weather patterns. This could be particularly useful in situations where severe weather events are expected, such as hurricanes, tornadoes, and flash floods.

Overall, weather prediction models are constantly evolving and improving, and there are many potential future enhancements that could be made to these models to improve their accuracy and usefulness.

CONCLUSION

All the machine learning models: linear regression, various linear regression, decision tree regression, random forest regression were beaten by expert climate determining apparatuses, even though the error in their execution reduced significantly for later days, demonstrating that over longer timeframes, our models may beat genius professional ones.

Linear regression demonstrated to be a low predisposition, high fluctuation model though polynomial regression demonstrated to be a high predisposition, low difference model. Linear regression is naturally a high difference model as it is unsteady to outliers, so one approach to improve the linear regression model is by gathering more information. Practical regression, however, was high predisposition, demonstrating that the decision of the model was poor and that its predictions can't be improved by the further accumulation of information. This predisposition could be expected to the structure decision to estimate temperature dependent on the climate of the previous two days, which might be too short to even think about capturing slants in a climate that practical regression requires. On the off chance that the figure was rather founded on the climate of the past four or five days, the predisposition of the practical regression model could probably be decreased. In any case, this would require significantly more calculation time alongside retraining of the weight vector w, so this will be conceded to future work.

Talking about Random Forest Regression, it proves to be the most accurate regression model. Likely so, it is the most popular regression model used, since it is highly accurate and versatile. Below is a snapshot of the implementation of Random Forest in the project.

Weather Forecasting has a major test of foreseeing the precise outcomes which are utilized in numerous ongoing frameworks like power offices, air terminals, the travel industry focuses, and so forth. The trouble of this determination is the mind-boggling nature of parameters. Every parameter has an alternate arrangement of scopes of qualities.

REFERENCE

- https://www.researchgate.net/publication/329922758_Weather_Forecast_Prediction_An_Integrated_Approach_for_Analyzing_and_Measuring_Weather_Data
- http://cs229.stanford.edu/proj2016/report/HolmstromLiuVo-MachineLearningAppliedToWeatherForecasting-report.pdf
- https://dl.acm.org/doi/abs/10.1145/3292500.3330704
- https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3410
- https://arxiv.org/abs/2008.1078