Weather Forecasting

SHIVAIBHAV DEWANGAN
DSAI
211020448
International Institute of Information Technology
Naya Raipur, India
shivaibhav21102@iiitnr.edu.in

SHIVAM KUSHWAHA
DSAI
211020449
International Institute of Information Technology
Naya Raipur, India
shivam21102@iiitnr.edu.in

Abstract - Weather forecasting is the prediction of the state of the atmosphere for a given location using the application of science and technology. This includes temperature, rain, cloudiness, wind speed, and humidity. Weather warnings are a special kind of short-range forecast carried out for the protection of human life. Weather warnings are issued by the governments throughout the world for all kinds of threatening weather events including tropical storms and tropical cyclones depending upon the location. The forecast may be short-range or Long-range. It is a very interesting and challenging task. While weather forecasting has come a long way over the years, it still remains a challenging task, as weather patterns can be unpredictable and complex. Additionally, the effects of climate change have added new complexities to forecasting, making it even more critical to have accurate and reliable weather predictions.

1. INTRODUCTION :

Weather forecasting is a critical component of modern society, as it provides valuable information for planning and decision-making in various sectors such as agriculture, transportation, aviation, and emergency management. Accurate weather forecasts are essential for preparing for severe weather events, reducing damage to property and infrastructure, and saving lives. With the advancements in technology and the growing understanding of atmospheric science, weather forecasting has become more accurate and reliable than ever before. This research paper aims to explore the history and development of weather forecasting, the various methods and technologies used in predicting weather, and the challenges and limitations that weather forecasters face. Through a comprehensive analysis of the current state of weather forecasting, this paper seeks to provide insights into the potential for further improvements in this critical field.

DATASET :

The data is obtained from Darksky.net which includes Ta hourly/daily summary for Szeged, Hungary area, between 2006 and 2016. The dataset contains various features related to temperature, wind speed, pressure and summary of the weather on that particular data. The dataset has over 96429 entries and 12 features in all.

METHODOLOGY :

Steps Performed:

Data Cleaning: Identifying and correcting or removing inaccurate, incomplete, or irrelevant data in a dataset. It involves detecting and correcting errors, filling missing values, and removing duplicates or outliers and reducing the skewness to improve the quality of the data.

Feature Selection: We have used correlation analysis, feature importance analysis, and domain knowledge to select the most relevant features for our model so that we can get the most accurate model.

Visualization: Visualization is an important toolin exploratory data analysis and helps to understand patterns and relationships in the data. Different types of visualizations, such as scatter plots, histograms, box plots and heat maps are used to visualize the data and relationships between variables and also help in removing the outliers. Visualization can provide insights into the data and guide the selection of features and models for machine learning tasks.

Model Training and Testing: In **Linear Regression,** the model is trained using a set of input features and a continuous target variable. The goal is to find a linear relationship between the features and the target variable. The model is then tested using a separate test dataset to evaluate the performance.

Performance Evaluation: Performance evaluation is an essential step in machine learning to assess the quality of the trained models. Multiple performance metrics are used to evaluate the model's accuracy, such as R2 Score, Mean Absolute Error, Mean Squared Error and Root Mean Squared Error. Evaluating the models on the test datahelps in selecting the best-performing modelfor deployment.

EVALUATION METRICS :

Evaluation metric refers to a measure that we use to evaluate different models.
We have used the following evaluation metrics:

- **R2 Score (R - Squared):** R-squared (R2) is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable in a regression model.

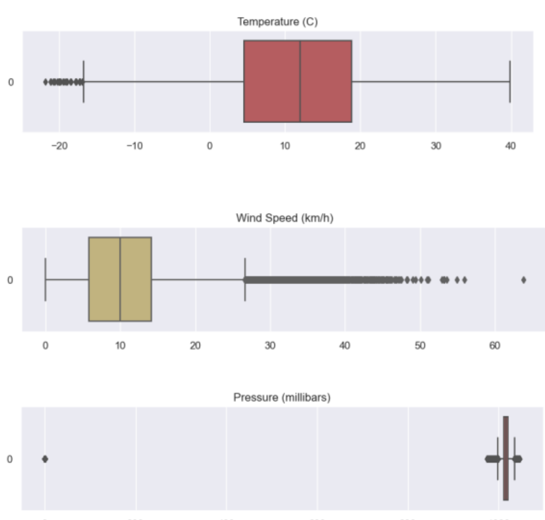
- **MSE(Mean Square Error):** MSE (Mean Squared Error) is a commonly used metric to evaluatethe performance of regression models. It calculates the average squared difference between the predicted and actual values, providing a measure of how wellthe model fits the data.

- **MAE(Mean absolute error):** Mean absolute error (MAE) is a measure of errors between paired observations expressing the same phenomenon. MAE is calculated as the sum of absolute errors divided by the sample size.

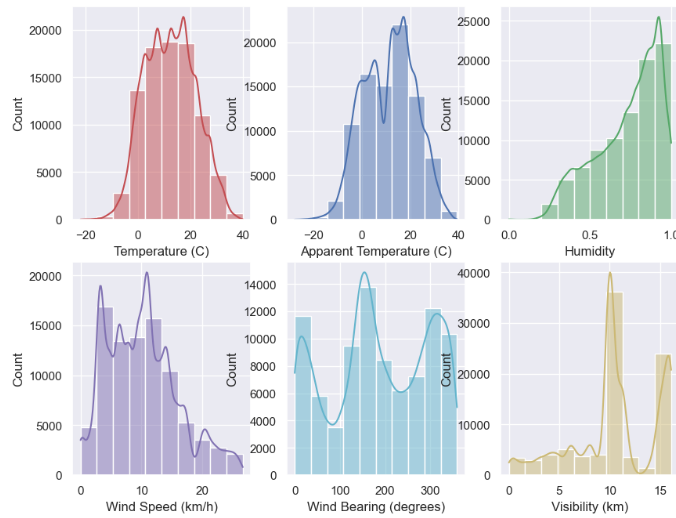
- **RMSE(Root Mean Squared Error):** Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are; RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data is around the line of best fit.

OBSERVATIONS:

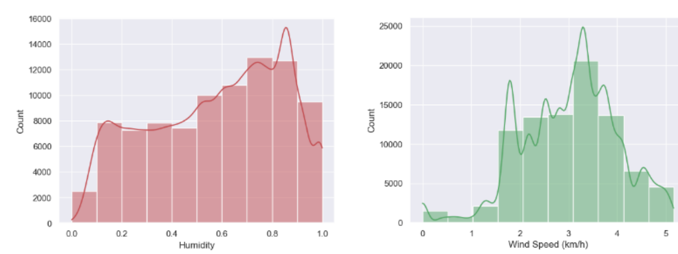
Box plot : used to detect the outliers using the IQR range.



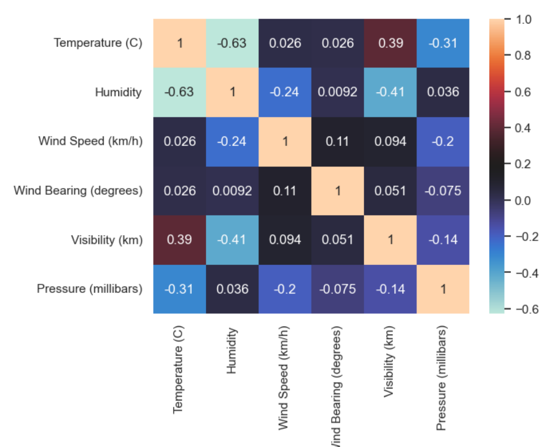
Data Distribution – checking for the skewness in the data.



In left skewed data, applying square transformation and in right skewed data, applying square root transformation in order to reduce the skewness.



Heatmap :



RESULTS :

R2 Score for Linear Regression :

We achieved a R2 score of 0.99129 which is excellent.

Other Metrics used :

	Metric	Value
0	R2 Score	0.991129
1	MSE	0.008815
2	RMSE	0.093888
3	MAE	0.074437

Final Coefficient of Attributes :

	Feature	Coefficient
0	Precip Type	0.052700
1	Temperature (C)	0.998936
2	Humidity	0.014347
3	Wind Speed (km/h)	-0.057215
4	Wind Bearing (degrees)	0.006672
5	Visibility (km)	0.003706
6	Pressure (millibars)	0.008455
7	Drizzle	0.025058
8	Dry	-0.103961
9	Dry and Mostly Cloudy	-0.074148
10	Dry and Partly Cloudy	-0.086801
11	Foggy	0.030150
12	Humid and Mostly Cloudy	0.001121
13	Humid and Overcast	0.017314
14	Humid and Partly Cloudy	-0.000111
15	Light Rain	0.026035
16	Mostly Cloudy	0.015246
17	Overcast	0.007238
18	Partly Cloudy	0.003788
19	Rain	0.018158

Intercept:

```
Intercept : -0.056664581020808506
```

CONCLUSION :

We have used various approaches to find the best optimum model. At the start, we tried to find out the pattern hidden in the data using data distribution and scatter plot. After that we tried simple linear regression whose results were dissatisfying. Hence, we tried multiple linear regression to and found the results to be very good. Hence it was the best fit for the dataset we have.

FUTURE SCORE

1. **Real-time Monitoring** : Another future scope for a weather forecasting project is real-time monitoring. This can be done by using advanced sensors and weather monitoring systems that can provide real-time updates on weather conditions, allowing for more accurate and timely predictions.

2. **Integration with other systems:** Weather forecasting can be integrated with other systems to improve the overall efficiency and effectiveness of the project. For example, weather data can be integrated with transportation systems to help drivers plan their routes more efficiently, or with agricultural systems to help

farmers make better decisions about planting and harvesting.

3. **Customized weather alerts :** Customized weather alerts are another future scope for a weather forecasting project. This can be achieved by using artificial intelligence to analyze weather data and provide personalized alerts based on the user's location and specific weather preferences.

4. **Climate Change Analysis :** Finally, climate change analysis is an increasingly important future scope for weather forecasting projects. By analyzing historical weather data and predicting future trends, weather forecasting projects can help us better understand and prepare for the impacts of climate change.

REFRENCES :

1. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3350281
2. https://ieeexplore.ieee.org/document/8448971
3. https://edubirdie.com/examples/weather-forecasting-using-machine-learning/