Machine-Learning-Model-for-Weather-Forecasting

The purpose of this project is to predict temperature using various algorithms, including linear regression, random forest regression, and decision tree regression. The output value will be a numerical prediction based on the latest time series data.

Use of Algorithms

predict temperature using various algorithms, including linear regression, random forest regression, and decision tree regression, based on the Global Weather Repository dataset, which provides four months of historical weather information. To train the models, 80% of the data is utilized, while the remaining 20% serves as the test set. For example, when predicting temperature using these machine learning algorithms, we leverage this four-month dataset. Unlike traditional weather forecasting, which primarily relies on physics-based simulations and differential equations, this approach harnesses artificial intelligence to improve predictive accuracy. In conclusion, machine learning has significantly transformed weather forecasting, enhancing precision and predictive capabilities. In the coming years, further advancements in these technologies are expected to enhance our ability to predict weather events and mitigate disasters such as hurricanes, tornadoes, and thunderstorms.

METHODOLOGY

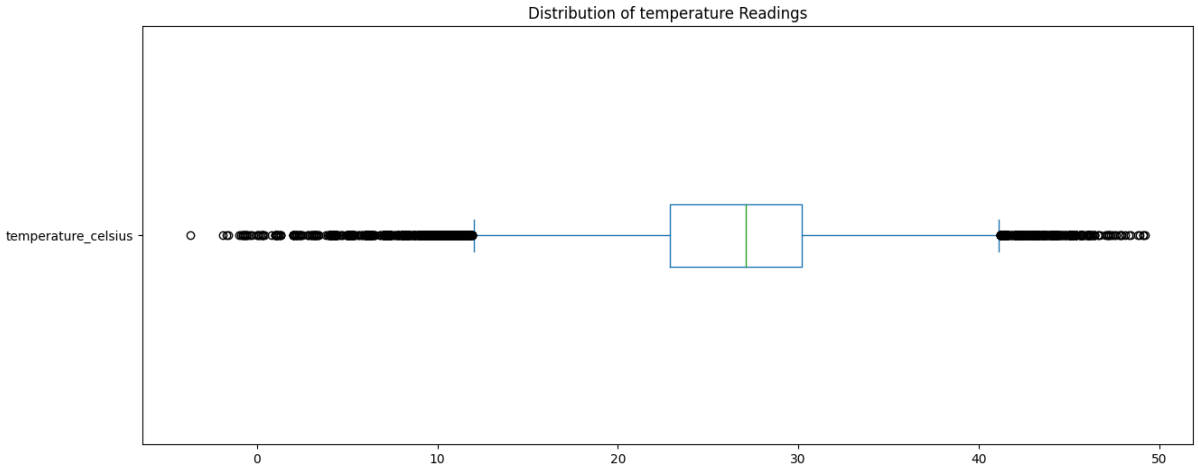
"The dataset used in this project was sourced from the 'Global Weather Repository' available on Kaggle. It includes comprehensive weather data from around the world, with four months of detailed historical weather information. This dataset was created to meet the growing need for global weather data and includes variables such as temperature, humidity, wind speed, and precipitation. The data is valuable for visualizing weather patterns and predicting future conditions over short or long periods, such as days, weeks, or months.

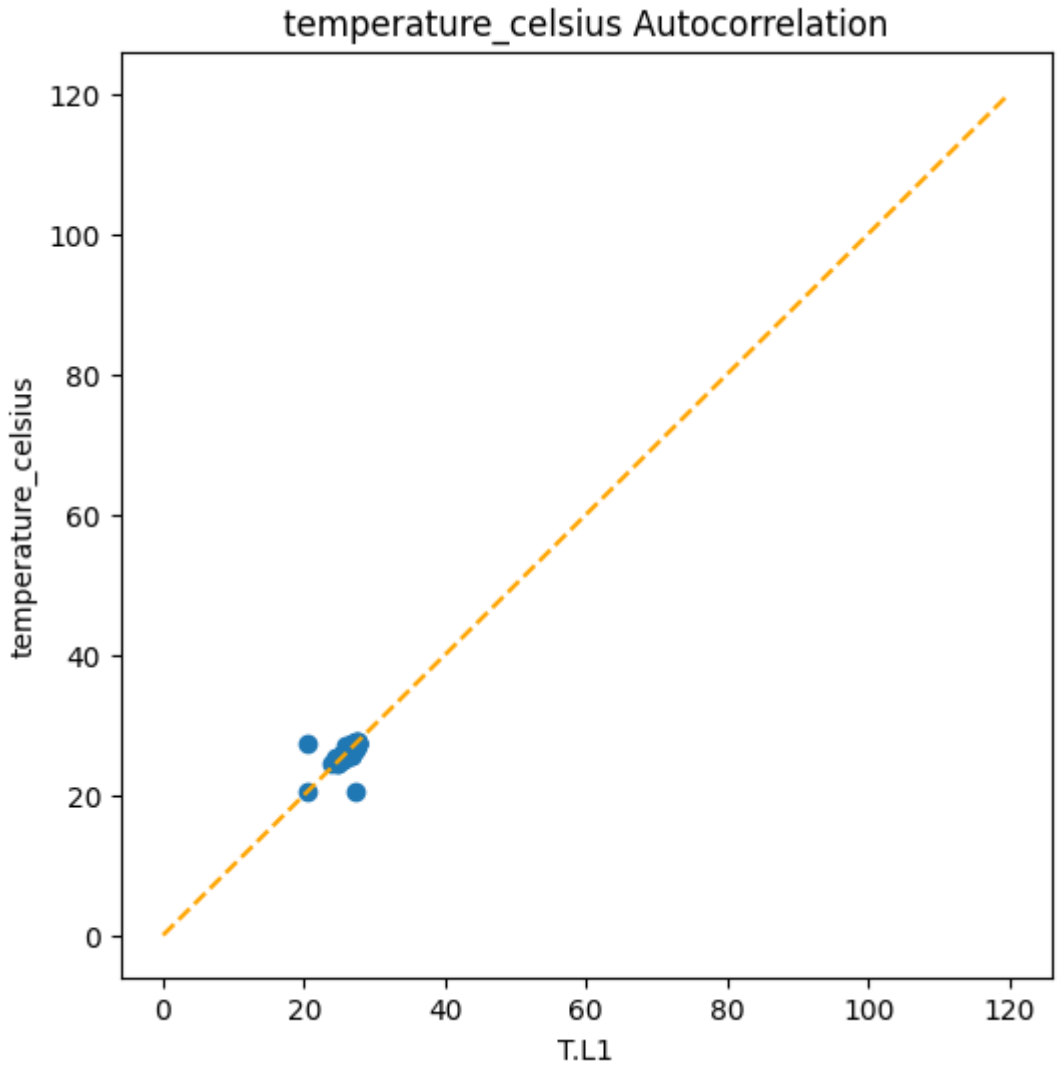
Note: The data was collected from various sources and aggregated into the repository, but its accuracy is not guaranteed.

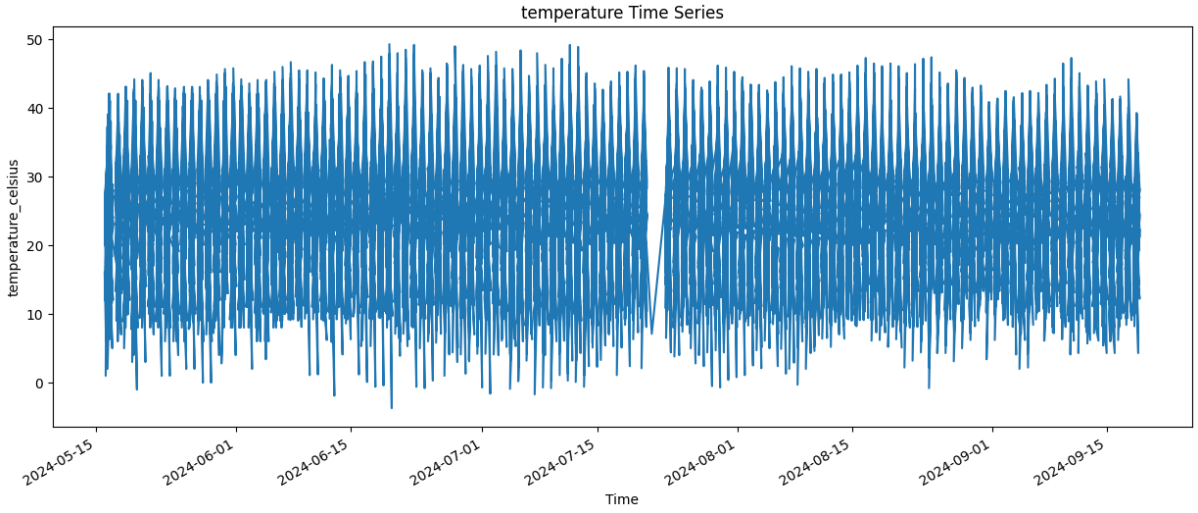
The primary objective of this dataset is to enable the prediction of weather parameters like temperature using the large volume of data provided. Additionally, it can be used to create visualizations that illustrate the impact of global climate phenomena, such as global warming, on weather variables.

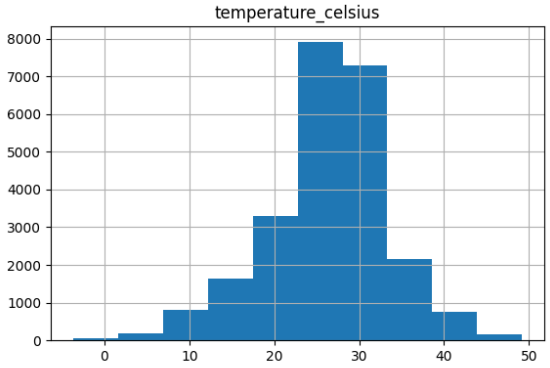
In this project, we focus on predicting temperature using the historical weather data from the repository, employing various machine learning algorithms. We apply multiple regression techniques, starting with linear regression, followed by decision tree regression, and finally random forest regression, to forecast future temperature

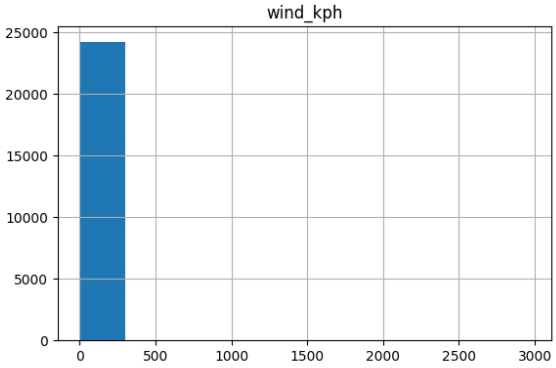
trends."

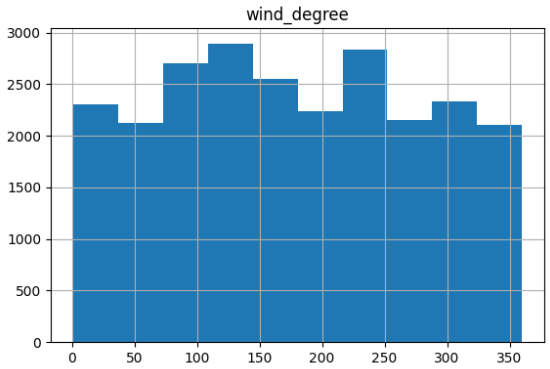

Distribution of temperature Readings
temperature_celsius
0
10
20
30
40
50

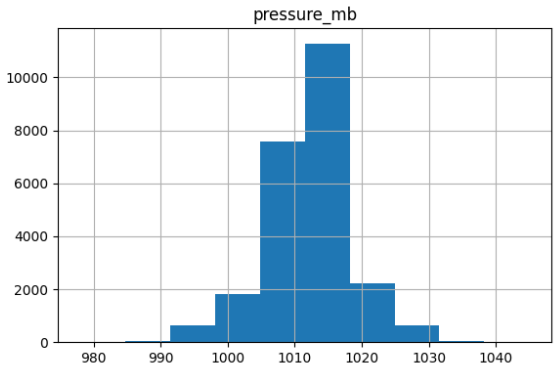

temperature_celsius Autocorrelation
temperature_celsius
T.L1
0
20
40
60
80
100
120

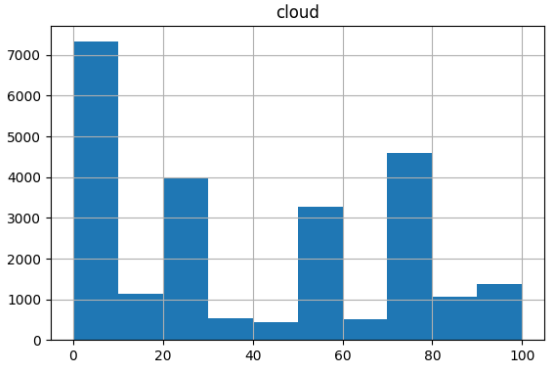


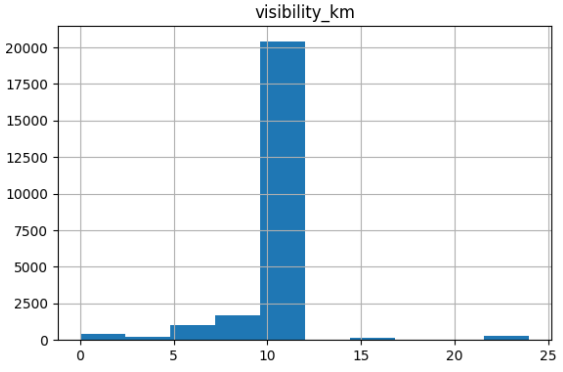










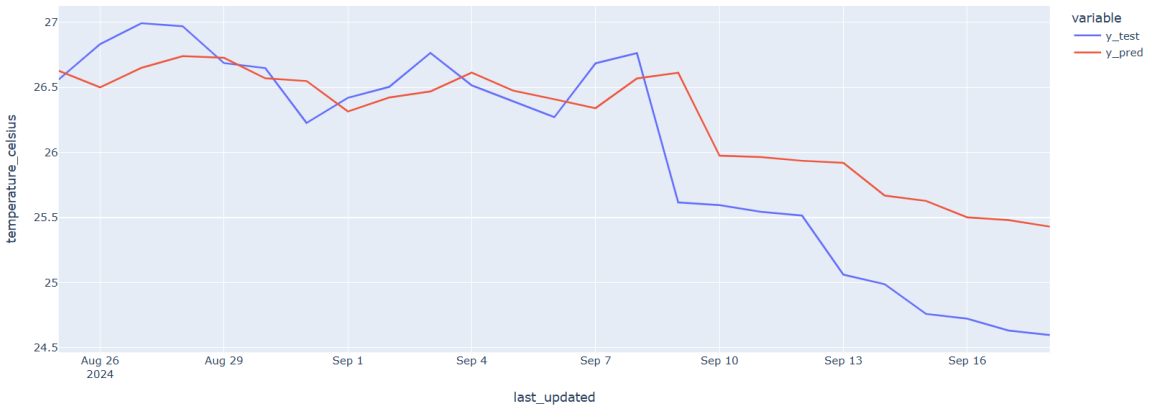


Multiple LinearRegression

This regression model has high mean absolute error, hence turned out to be the least accurate model. Given below is the plot of y_test and y_pred:

2]:

last_updated	y_test	y_pred	Diff
2024-08-25	26.561538	27.089231	0.527692
2024-08-26	26.832308	26.723077	-0.109231
2024-08-27	26.993333	27.059162	0.065829
2024-08-28	26.970769	26.863131	-0.107638
2024-08-29	26.687179	26.509231	-0.177949



Decision Tree Regression:

This regression model has medium mean absolute error, hence turned out to be the little accurate model.

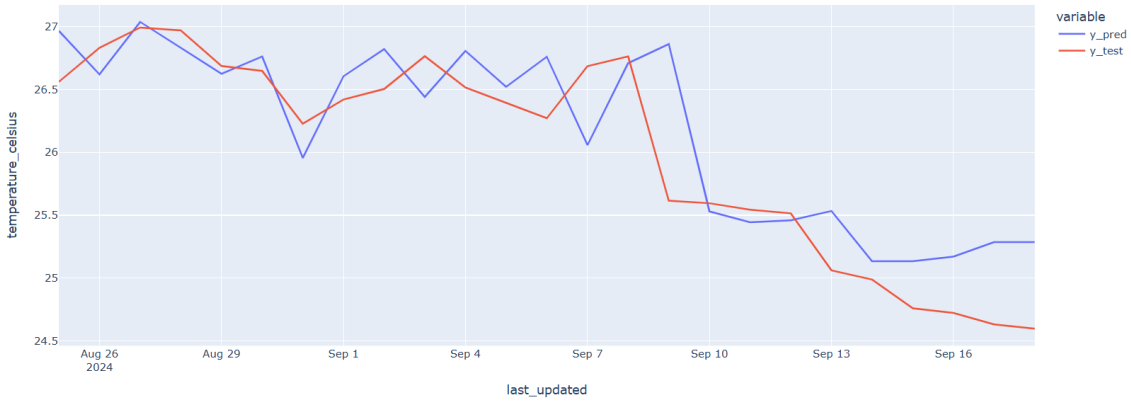
[22]:

last_updated	y_test	y_pred	Diff
2024-08-25	26.561538	26.967436	0.405898
2024-08-26	26.832308	26.620723	-0.211585
2024-08-27	26.993333	27.038038	0.044704
2024-08-28	26.970769	26.834098	-0.136671
2024-08-29	26.687179	26.624589	-0.062591

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 for random forest regression:

last_updated	y_test	y_pred	Diff
2024-08-25	26.561538	26.967436	0.405898
2024-08-26	26.832308	26.620723	-0.211585
2024-08-27	26.993333	27.038038	0.044704
2024-08-28	26.970769	26.834098	-0.136671
2024-08-29	26.687179	26.624589	-0.062591


variable
y_pred
y_test
temperature_celsius
27
26.5
26
25.5
25
24.5
Aug 26
2024
Aug 29
Sep 1
Sep 4
Sep 7
Sep 10
Sep 13
Sep 16
last_updated