**Weather Prediction** Developed by Yen-Chen Shih

1. **Definition of the Problem**

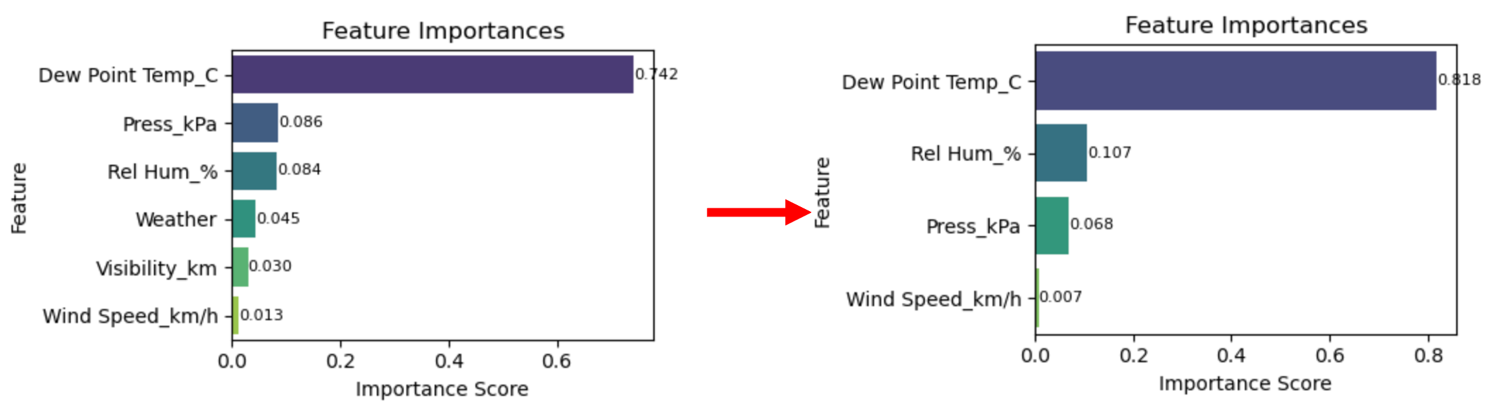
This task involves predicting temperature (Temp\_C) using weather data. I used regression models to forecast temperature based on various weather features, treating it as a supervised learning problem. Accurate temperature predictions are vital for sectors like agriculture, entertainment, and construction, relying on historical weather data to anticipate future conditions.

1. **Rational of target variable selection**

I chose temperature (Temp\_C) as the target variable because temperature is a key meteorological variable. Understanding and predicting temperature fluctuations play a crucial role in various applications such as weather-forecasting apps to help agriculture, outdoor enthusiasts, construction

**(c ) A short note on which kind of machine learning is suitable for this problem**

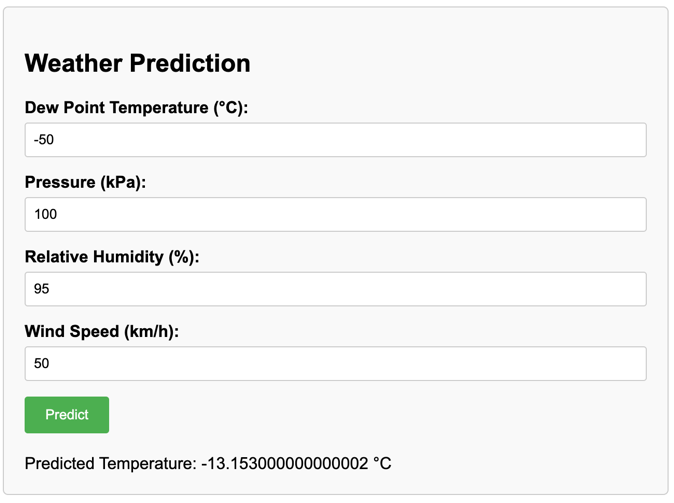
This task aims to predict temperature using meteorological variables, a regression problem. Hence, supervised learning algorithms like random forest regression and lasso regression are suitable. Despite similar R-squared and RMSE scores, the lower MAE score of Random Forest Regression makes it the preferred choice.

**(d) Conclusion with at least one visual**

Through feature importance analysis, the results show that the model's performance improves after removing the features "Weather" and "Visibility\_km" due to their non-normal distributions and low feature importance. The most important feature is "Dew Point Temperature (Dew Point Temp\_C)", with an importance score of 0.818. Other important features include "Relative Humidity (Rel Hum\_%)" (0.107), "Pressure (Press\_kPa)" (0.068), and "Wind Speed (Wind Speed\_km/h)" (0.007). This model can predict specified temperature values based on existing weather data.

**(e) Scope for future work**

**Model Ensembles**: Experimenting with ensemble methods, such as combining predictions from multiple models, could potentially enhance predictive accuracy. **Time-Series Analysis**: Considering the temporal aspect of meteorological data and incorporating time-series analysis techniques may lead to better predictions. **Hyperparameter Tuning**: Continued exploration of hyperparameter tuning methods to optimize model performance. **Integration with Other Data Sources**: Incorporating additional data sources, such as satellite imagery or geographical information, could provide supplementary insights for temperature prediction.

**(f) Did you go beyond the expectation and deserve the extra 5 points?**

The provided solution not only meets the expectations but also goes beyond by integrating a machine learning model for weather prediction into a web application using Flask for the backend and HTML/CSS/JavaScript(fetch API) for the front-end.

**Use Cases:**

* **Agriculture**: Helping Agricultural companies make informed decisions regarding planting, harvesting, and irrigation.
* **Outdoor** **enthusiasts**: People who enjoy outdoor activities can plan their activities accordingly.
* **Construction**: Construction companies can plan schedules, avoiding extreme weather conditions.

**Test Cases：**

|  |  |
| --- | --- |
| **Normal scenario test:**  Dew Point Temperature (°C): 20  Pressure (kPa): 90  Relative Humidity (%): 70  Wind Speed (km/h): 10  Predicted Temperature: 21.05 °C | **Extreme scenario test:**  Dew Point Temperature (°C): -50  Pressure (kPa): 100  Relative Humidity (%): 95  Wind Speed (km/h): 50  Predicted Temperature: -13.153000000000002 °C |