Weather Prediction Developed by Yen-Chen Shih

(a) Definition of the Problem

- This problem involves predicting temperature (Temp_C) using weather data.
- The feature variables include all columns except for date/time, which will be used to predict the target variable.
- By constructing regression models, our goal is to predict temperature based on given weather features.
- This is a typical supervised learning problem, as the model uses labeled data during training and attempts to predict a continuous numerical target (temperature).
- The importance of the task lies in accurately predicting future temperatures based on past meteorological data, which is crucial for fields such as agriculture, transportation, and construction.

(b) Rational of target variable selection

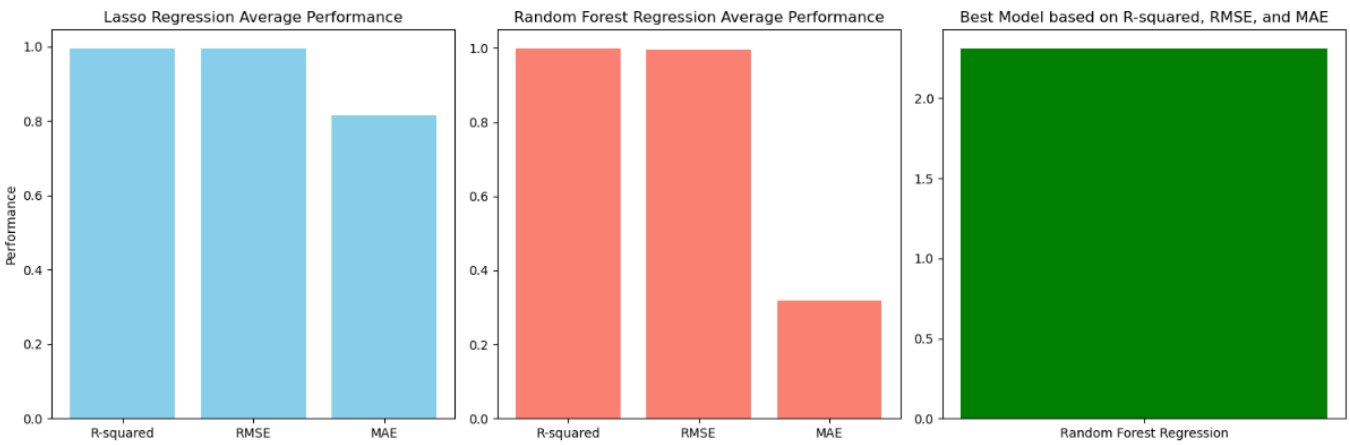
In this problem, I chose temperature (Temp_C) as the target variable because temperature is a key meteorological variable.

Understanding and predicting temperature fluctuations play a crucial role in various applications such as weather forecasting apps, outdoor activity planning apps, travel planning apps, and energy demand prediction.

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

- This problem involves predicting temperature based on meteorological variables, which is a regression task.
- - Therefore, supervised learning algorithms such as random forest regression, and lasso regression are suitable for this problem.

(d) Conclusion with a at least one visual



Based on the evaluation of multiple regression models, including Lasso Regression and Random Forest Regression, it was found that both models performed well in predicting temperature. However, Random Forest Regression exhibited slightly better performance in terms of R-squared, RMSE, and MAE.

Therefore, the Random Forest Regression model is recommended for predicting temperature based on meteorological variables.

(e) Scope for future work

- **Feature Engineering:** Further exploration of feature engineering techniques could improve model performance.
- **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 for the frontend.
- The frontend utilizes asynchronous JavaScript (fetch API) to send data to the backend and handle prediction responses without reloading the page. This enhances user experience by providing a seamless and responsive interface.
- The integration of machine learning with web development in this solution demonstrates a comprehensive understanding of both domains and effectively showcases the application of predictive modeling in a real-world context.
- Hence, it deserves the extra 5 points.