

## **Dataset Link:**

The dataset used in the project can be found on Kaggle:

<https://www.kaggle.com/datasets/jsphyg/weather-dataset-rattle-package>

## **Description of Project:**

This code uses Artificial Neural Networks (ANN) to predict the possibility of rain based on various weather factors like temperature, humidity, wind speed, etc. The dataset used in this project contains historical weather data of Australia from 2007 to 2017.

## **A brief explanation of the outputs of the Predicting rainfall amount Project**

1. Data preprocessing: The code performs the following data preprocessing steps: it loads the dataset and handles missing values and categorical data. It also normalizes the numerical data and encodes the categorical data using OneHotEncoding.
2. Model building and training: The code builds and trains an Artificial Neural Network (ANN) model using the following steps: it splits the dataset into training and testing sets. It builds an ANN model using the Keras library. It compiles the model, sets the loss function and optimizer, and fits the model on the training data.
3. Model evaluation: The code evaluates the ANN model using the following steps: it calculates various performance metrics such as accuracy, precision, recall, and F1-score. It also plots the confusion matrix to check the classification performance.
4. Feature importance: The code analyzes the feature importance using the following steps: it calculates the correlation between the features and the target variable. It plots the correlation matrix and heatmap to visualize the correlations. It also calculates the feature importance using the SHAP (SHapley Additive exPlanations) library and plots the summary plot.
5. Prediction: The code performs the following steps for prediction: it makes predictions on the testing data using the final ANN model. It plots the actual and predicted values on a graph.
6. Conclusion: The code interprets the results and discusses the strengths and weaknesses of the model. It provides suggestions for future improvements.

## **Description of Output:**

The output of this code is a binary classification model that predicts whether or not it will rain on a given day based on the input features. The code also provides various evaluation metrics such as accuracy, precision, recall, F1 score, ROC curve, and confusion matrix to evaluate the performance of the model.

### **Instructions on How to Run the Code/Project/File:**

1. To run the code in the notebook, you will need to have Jupyter Notebook installed on your computer. Once you have installed the Jupyter Notebook, you can download the notebook from the Kaggle website and open it in the Jupyter Notebook.
2. Before running the code, you will need to make sure that you have downloaded the necessary data files and saved them in the correct directory. The notebook provides instructions on how to download the data files and where to save them.
3. Once you have downloaded the data files and opened the notebook in Jupyter Notebook, you can run each cell of the notebook by clicking on the cell and then clicking the "Run" button in the toolbar or by using the keyboard shortcut "Shift + Enter".
4. It is recommended that you run the code cells in order, as some cells depend on the output of earlier cells. The notebook also provides explanations and comments for each code cell, so beginners can follow along and understand the code.

**Note: Make sure to update the file paths in the code cells to match the location of the downloaded dataset and kernel files on your local machine.**