**Steps:**

1. **Data Collection and Exploration:**
   * I searched for the weather data on Kaggle.com and downloaded the nepali data.
2. **Data Cleaning and Transformation:**
   * For the data cleaning, my data didn't have any missing or Nan values.
   * If there were any, I would have used different approaches depending upon the type of missing values.
   * Example- If months were missing, i would have used the date to calculate the month; if districts were messing, i would have used a conditional function using the latitude and longitude.
   * If the Precipitation or other similar data were missing, I would have used a random number generator in between Q1 and Q3, and assigned the value. Or even leave them, cause we would have no idea about those values.
   * I changed date into DateTime format, and selected the top 10 wettest districts, because the data was too big, and including the driest as well as wettest districts would have skewed our prediction model way too much.
   * I log transformed the precipitation data to get less number and less range of extreme values. I couldn't ignore them because there were a very high number of outliers and they seemed genuine.
3. **Exploratory Data Analysis (EDA):**
   * For EDA, I used scatter plots and histograms to see the distribution of data. Most of the data were normally distributed with a couple of exceptions.
   * They had a pretty low linear relationship with Precipitation.
4. **Feature Selection:**
   * I used the Spearman's correlation as the main measure because of the right skewed precipitation data.
5. **Model Development:**
   * I used a random forest regressor to make a model. There were many outliers and the Linear relationship was less defined.
   * I took 80% of the data as training data and the remaining 20% as test data.
6. **Model Evaluation and Hyperparameter Tuning: ( Changing the parameters to make Learning better)**
   * I used 3 parameters for hypertuning. And cross validated each data 5 times.
   * I used rmse as an accuracy indicator and used the mean for the base rsme calculator.
   * The parameters i used are: 1. Depth 2. Minimum samples split. 3. n\_estimators