CADT
IDRI Institute of Digital Research & Innovation

Project proposal

Topic: Weather Forecasting time series

I. Problem statement

Weather forecasting has gained researchers from worldwide societies over decades due to its substantial impact on global human life from agriculture, air traffic control to public health and safety. Anyway agriculture is the backbone of the economy.In a country like Cambodia, which has ever increasing demand for food due to rising population, advances in agriculture sectors are required to meet the needs and weather forecasting is in high demand for several applications in agriculture, air traffic services, foods ,energy and environment control. Weather forecasting is an expectation of what the weather will resemble in the next 1h, tomorrow , or 1 week from now. It will enable the farmers to maximise their crops yield and minimise the input cost as well as the losses due to reasons like uncertain rainfall, droughts etc.

II. Objective

This project aims to explore the time series data using some machine techniques. Most of the work related to machine learning for agriculture either solves the purpose of cultivating a crop and suggests weather data based on the statistical information. to the ecosystem. Most of the work does not handle the planting of crops based on the climate.

In this research aims:

a. To understand the trend of data based on each location and each month by using statistical information.

b. Weather forecasting is an expectation of what the weather will resemble in the next 1h, tomorrow , or 1 week from now.

c. We will use some machine learning techniques such as AR,MA, ARIMA, ARIMAX, SARIMA, SARIMAX

III. Dataset

We have joined MOU with RUA university . CE SAIN in collaboration with the Faculty of Land Management and Land Administration of the Royal University of Agriculture and the University of California, Davis, has installed five weather stations located in Phnom Penh, Kampong Cham, Kampong Thom ,Siem Reap, Battambang and Banteay meanchey. The weather station provides weather data to facilitate environmental agriculture research in order to increase agricultural modernization and facilitate students and researchers in agricultural development. We can say it is a public dataset and generated once a month since 2018 till today. There are 11 features such as date ,time ,water content ,solar radiation ,rainfall, temperature, RH, wind speed, wind direction, gust speed and dew points. We are going to use forecasting problems by using machine learning techniques to solve this problem based on time series data.

IV. Method

Data generated from 6 stations in Cambodia. It is a kind of structured data, and it is a forecasting problem. Machine learning technique is the most robust technique for predicting weather forecasting. Most of the work related to machine learning for agriculture either solves the purpose of cultivating a crop and suggests weather data based on the statistical information. Based on the time series data it allows us to either do analysis or forecasting . Time series analysis comprises methods for analysing time series data in order to extract meaningful statistics and other characteristics of the data. We will use some machine learning techniques such as AR,MA, ARIMA, ARIMAX, SARIMA, SARIMAX etc..

The advantage of weather forecasting in agriculture

- Help farmers better understand making an informed decision on
- Crop Growth/Irrigation
- Fertiliser Timing and Delivery
- Pest and Disease Control

when to work efficiently in their day-to-day operation.

The challenge of weather forecasting in agriculture

-It has many parameter

-It difficult to forecast correctly

-The weather forecasters get blamed if the weather is different from the forecast

V Project Timeline

Tasks	Weather forecasting				
	August	September	September	October	November
A. Define Problem					
B. Data Collection					
C. Study about data					
D. Data preparation					
E. EDA					
F. Feature Engineering					
G. Feature Selection					
H. Model creation					
I. Hyper parameter Tuning					
J. Conclusion					
K. Final Thesis					

References

-WW2 Temperatures w/ Linear Regression
https://www.kaggle.com/code/caesarmario/ww2-temperatures-w-linear-regression/notebook
-Time series: Delhi Weather Forecasting (ARIMA)
https://www.kaggle.com/code/amar09/time-series-delhi-weather-forecasting-arima/notebook
-Intro to Time Series Forecasting
https://www.kaggle.com/code/iamleonie/intro-to-time-series-forecasting/notebook#Feature-Engineering
-Summary of Weather_LinearRegression
https://www.kaggle.com/code/zahrahasanzadehh/summary-of-weather-linearregression/notebook
-Weather Conditions in World War 2
https://www.kaggle.com/code/ronikdedhia/weather-conditions-in-world-war-2