E&ICT Academy, IIT Kanpur
A Joint Initiative of MeitY & IIT Kanpur

Online Data Science Hackathon by EICT Academy, IIT Kanpur in association with The Ikigai Lab

You might ask yourself: Is it possible to predict the weather? What will it be like tomorrow in Nagaland? What clothing do I need? Weather can be predicted when looking at historical data and also with the help of data science. The idea behind data science is to create models that can predict the weather for other days, weeks or even months. For example, in the case of weather, you can train a model to predict the weather for tomorrow by looking at data from the past. This model will show the best weather for tomorrow. It's not always correct, but it's a pretty good option.

Northeastern states of India are considered as a place of varied weather conditions ranging from snow, rain, storms, etc. A group of researchers from Guwahati tried to acquire the daily weather data from various sensors that are managed by a research organization. The researchers were interested in prediction of tremendous change in the nature of weather and their impact on the environment.

A snip cut of the whole data had been provided to you with various factors as follows:

Formatted Date: Includes the date and time of data acquisition.
Summary: Nature of current weather
Precip Type: Precipitation is rain, snow, sleet, or hail — any kind of weather condition where something's falling from the sky.
Temperature: acquired temperature in degree centigrade.
Apparent Temperature: Apparent temperature is the temperature equivalent perceived by humans, caused by the combined effects of air temperature, relative humidity and wind speed.
Humidity: percentage of the moisture content in the atmosphere.
Wind Speed: speed of the wind measured in km/hr
Wind bearing: wind bearing measured in degrees. It indicates the direction toward which an object is moving.
Visibility: visible range in km. Visibility is a measure of the horizontal opacity of the atmosphere at the point of observation and is expressed in terms of the horizontal distance at which a person should be able to see and identify.
Cloud Cover: Cloud cover (also known as cloudiness, cloudage, or cloud amount) refers to the fraction of the sky obscured by clouds when observed from a particular location.
Pressure : Pressure of the environment in millibars.
Daily Summary: Daily forecast.

Q. Your task is to analyze the given scene. Find the reports of the weather in a graphical pattern, encode the necessary with different python tools and finally train with different machine learning tools to produce output which could serve as the target that is the same as daily summary when new data comes into play. Also find the accuracy metrics such that the number of false positives are less than the true positives and also the accuracy is optimized.