CSE4005 - Data Warehousing and Data Mining

Faculty: S. BHARATH BHUSHAN

Slot: F2+SF1

DWDM REPORT-1

Project Title: Weather Analysis and Prediction

MEMBERS:

B. Harshavardhan Reddy-22BCE20157

T. Narasimha-22BCE20152

A. Mohit-22BCE8268

M. Vishruth Reddy-22BCE7584

Project Title:

Weather Analysis and Prediction

This project focuses on exploring weather data to identify patterns in temperature, humidity, precipitation, and other meteorological elements. The aim is to visualize weather trends, forecast possible weather conditions, and analyse seasonal changes that could impact various sectors, such as agriculture, transportation, or energy consumption.

Software Requirements:

The tools that will be needed to interact with this dataset in doing exploratory data analysis, and probably some predictive modelling are as follows:

1. **Programming Language:**

Python: Selected because of its extensive libraries and community support for data science and machine learning.

2. **Libraries:**

- **pandas:** Manipulating data, cleaning, handling missing values.
- matplotlib, seaborn: Plotting, charting, and creating heatmaps to visualize patterns and trends.
- **NumPy:** Performing numerical operations on the dataset.
- **scikit-learn:** Any application of machine learning relevant to antibiotic use, such as clustering or predictive analysis.
- **stats models:** Statistical tests, modelling, and regression analysis.

3. **Tools:**

- **Kaggle API:** To download datasets from Kaggle directly into this working environment
- **Google Collab:** Development environment.

Abstract:

Weather plays a critical role in daily life, influencing transportation, agriculture, and energy sectors. This project aims to explore a weather dataset, examining variables such as temperature, humidity, wind speed, and precipitation to identify patterns across regions and seasons. The project will leverage Python and data science techniques to visualize these patterns and perform predictive modelling to forecast future weather conditions. Insights from this study can assist in creating better preparedness strategies for weather-related challenges and developing sustainable solutions for various industries.

Base Research Paper:

Several key research topics and papers can serve as a starting point for this project:

- **Name:** "Predictive Analysis for Weather Prediction using Data Mining with ANN: A Study"
- Reference : Predictive Analysis for Weather Prediction using Data Mining
- Introduction about paper :
  - Weather forecasting is important because changing environmental factors like temperature and rain can affect our daily lives. It's a challenging task for meteorologists, who use tools like satellites and radar to predict the weather.
  - In the past, people used simple methods like observing the wind and clouds. Today, we use advanced technology, including observations from ships and aircraft, and Doppler radar.
  - This paper is organized into:
    - Section 1: Background on weather forecasting, data mining, and artificial neural networks (ANNs).
    - Section 2: Review of current weather prediction methods.
    - Section 3: Conclusions and future research directions.

Dataset:

Source: Data Set

A typical weather dataset might include the following variables:

- District: Administrative district.
- Mandal: Subdivision within the district.
- date: Date of observation for the weather data.
- rainfall: Cumulative rainfall recorded up to the observation date.
- temp_min: Minimum temperature recorded on the given date.
- temp_max: Maximum temperature recorded on the given date.
- humidity_min: Minimum humidity level recorded on the given date.
- humidity_max: Maximum humidity level recorded on the given date.
- wind_speed_min: Minimum wind speed recorded on the given date.
- wind_speed_max: Maximum wind speed recorded on the given date.

district Administrative district	mandal Subdivision within the district	date Date of observation for the weather data.	# rainfall Cumulative rainfall recorded up to the observation date.	# temp_min Minimum temperature recorded on the given date	# temp_max Maximum temperature recorded on the given date	# humidity_min Minimum humidity level recorded on the given date	# humidity_max Maximum humidity level recorded on the given date	# wind_speed_min Minimum wind speed recorded on the given date	# wind_speed_max Maximum wind speed recorded on the given date
Nalgonda 6% Rangareddy 5% Other (966035) 89%	713 unique values	1551 unique values	0 391	2.7 87.6	-1 47.9	-1 100	-1 100	-1 31.3	-1 234
Medchal-Malkajgiri	Uppal	01-01-2018	0.0	12.1	32.6	23.8	100.0	0.0	6.6
Medchal-Malkajgiri	Uppal	02-01-2018	0.0	11.6	32.6	23.2	100.0	0.0	4.7
Medchal-Malkajgiri	Uppal	03-01-2018	0.0	13.0	33.0	31.5	100.0	0.0	6.3
Medchal-Malkajgiri	Uppal	04-01-2018	0.0	9.7	31.7	27.4	100.0	0.0	5.2
Medchal-Malkajgiri	Uppal	05-01-2018	0.0	8.8	31.0	28.6	100.0	0.0	7.1
Medchal-Malkajgiri	Uppal	06-01-2018	0.0	8.7	31.8	22.8	100.0	0.0	4.5
Medchal-Malkajgiri	Uppal	07-01-2018	0.0	9.2	31.6	21.4	100.0	0.0	3.7
Medchal-Malkajgiri	Uppal	08-01-2018	0.0	9.6	31.59	23.8	100.0	0.0	7.1
Medchal-Malkajgiri	Uppal	09-01-2018	0.0	10.5	30.4	26.5	100.0	0.0	6.1
Medchal-Malkajgiri	Uppal	10-01-2018	0.0	10.8	30.0	35.7	100.0	0.0	6.4
Medchal-Malkajgiri	Uppal	11-01-2018	0.0	11.8	32.4	34.0	100.0	0.0	6.2
Medchal-Malkajgiri	Uppal	12-01-2018	0.0	11.8	34.5	30.7	100.0	0.0	6.8
Medchal-Malkajgiri	Uppal	13-01-2018	0.0	11.7	34.6	21.1	100.0	0.0	7.6
Medchal-Malkajgiri	Uppal	14-01-2018	0.0	11.4	34.9	18.6	92.8	0.0	4.5
Medchal-Malkajgiri	Uppal	15-01-2018	0.0	14.3	34.4	25.9	100.0	0.0	5.4
Medchal-Malkajgiri	Uppal	16-01-2018	0.0	13.9	35.5	24.0	100.0	0.0	4.0
Medchal-Malkajgiri	Uppal	17-01-2018	0.0	12.0	35.3	18.0	100.0	0.0	8.5
Medchal-Malkajgiri	Uppal	18-01-2018	0.0	11.2	33.7	24.7	100.0	0.0	6.1
Medchal-Malkajgiri	Uppal	19-01-2018	0.0	9.6	34.1	21.6	100.0	0.0	5.5
Medchal-Malkajgiri	Uppal	20-01-2018	0.0	7.1	34.0	14.4	100.0	0.0	5.5
Medchal-Malkajgiri	Uppal	21-01-2018	0.0	8.7	32.6	14.3	100.0	0.0	7.4
Medchal-Malkajgiri	Uppal	22-01-2018	0.0	9.5	33.1	19.0	100.0	0.0	7.7