**WEATHER DATA ANALYSIS**

**A PROJECT REPORT**

**by**

**Shashank Singh (202410116100194)**

**Tarun Kumar (202410116100223)**

**Vishal (202410116100238)**

**Vishal Dagar (202410116100249)**

**Session:2024-2025 (II Semester)**

Under the supervision of

**Mrs. Komal Salgotra**

**Assistant Professor**

### KIET Group of Institutions, Delhi-NCR, Ghaziabad



### Department Of Computer Applications

**KIET GROUP OF INSTITUTIONS, DELHI-NCR, GHAZIABAD-201206**

( 2024- 2025)

**ABSTRACT**

Weather forecasting has always been a critical aspect of daily life, agriculture, aviation, and disaster management. With the advancement of Artificial Intelligence (AI), traditional weather prediction methods are being enhanced using machine learning and data analysis techniques. This project, **"Weather Data Analysis,"** aims to analyze historical weather data using AI-based methodologies to identify trends and predict future conditions.

In this project, a weather dataset containing attributes such as **temperature, humidity, wind speed, and precipitation** is analyzed. The data is first preprocessed to handle missing values and normalize variables. Various AI techniques, including **Exploratory Data Analysis (EDA)** and **machine learning models** like **Linear Regression**, are applied to identify correlations and predict temperature variations.

The implementation is carried out using **Python** and libraries such as **Pandas, NumPy, Matplotlib, Seaborn, and Scikit-learn** in a Jupyter Notebook environment. The project also includes visualizations such as correlation heatmaps and actual vs. predicted temperature plots to better understand weather patterns.

The results demonstrate that AI-based weather analysis can efficiently predict trends, thereby improving forecasting accuracy. Future improvements can include **deep learning models, real-time weather prediction, and IoT-based data collection** for enhanced precision. This project highlights the growing importance of AI in meteorology and its potential to revolutionize climate-related predictions.

**TABLE OF CONTENTS**

Page Number

1. Introduction 4
2. Methodology 5
3. Code Typed 6
4. Screenshots 12

**1. INTRODUCTION**

**Overview of Weather Data Analysis**

Weather data analysis plays a crucial role in predicting atmospheric conditions and understanding climate trends. Artificial Intelligence (AI) enhances weather forecasting accuracy by processing large datasets and identifying patterns that humans may miss. This project focuses on utilizing AI techniques for analyzing historical weather data, identifying trends, and predicting possible weather conditions.

**Importance of AI in Weather Forecasting**

* Accuracy Improvement: AI can process vast amounts of data quickly, leading to more precise forecasts.
* Pattern Recognition: Machine learning algorithms detect complex patterns in weather conditions.
* Climate Change Monitoring: AI helps in long-term climate trend analysis.
* Disaster Prediction: AI models can provide early warnings for hurricanes, floods, and heatwaves.

**Objective of the Project**

The primary goal of this project is to analyze historical weather data using AI techniques to derive insights and visualize trends. Specifically, we will:

* Collect weather datasets.
* Process and clean the data.
* Apply AI techniques such as machine learning for analysis.
* Visualize the results.

.

**2. Methodology**

**Data Collection**

The dataset used in this project consists of historical weather data, including temperature, humidity, wind speed, and precipitation. The data was collected from publicly available sources such as:

* Kaggle
* OpenWeatherMap API
* NOAA (National Oceanic and Atmospheric Administration)

**Data Preprocessing**

Before analysis, raw data requires cleaning and transformation. The following steps were performed:

* **Handling Missing Values:** Missing temperature or humidity values were filled using interpolation methods.
* **Data Normalization:** The dataset was normalized to improve machine learning model efficiency.
* **Feature Selection:** Unnecessary columns were removed to focus on essential weather attributes.

**AI Techniques Used**

* Exploratory Data Analysis (EDA): Statistical methods were used to summarize trends in the dataset.
* Machine Learning Models: Algorithms like Linear Regression and Decision Trees were implemented to predict temperature trends.
* Visualization: Graphs and charts were used for trend analysis.

**Tools and Technologies**

* **Programming Language:** Python
* **Libraries Used:** Pandas, NumPy, Matplotlib, Scikit-learn, Seaborn
* **Development Environment:** Google Collab

1. **Code Typed**

# Import necessary libraries

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.model\_selection import train\_test\_split

from sklearn.linear\_model import LinearRegression

# Load dataset

df = pd.read\_csv("weather\_data.csv")

# Data Cleaning - Handling missing values

df.fillna(method='ffill', inplace=True)

# Feature Selection

df = df[['Temperature', 'Humidity', 'Wind Speed', 'Precipitation']]

# Exploratory Data Analysis

plt.figure(figsize=(8,5))

sns.heatmap(df.corr(), annot=True, cmap="coolwarm")

plt.title("Correlation between Weather Variables")

plt.show()

# Splitting Data

X = df[['Humidity', 'Wind Speed', 'Precipitation']]

y = df['Temperature']

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Model Training

model = LinearRegression()

model.fit(X\_train, y\_train)

# Prediction

y\_pred = model.predict(X\_test)

# Visualization of Predictions

plt.scatter(y\_test, y\_pred, alpha=0.5)

plt.xlabel("Actual Temperature")

plt.ylabel("Predicted Temperature")

plt.title("Actual vs Predicted Temperature")

plt.show()

1. **Screenshots**



