🌦️ Weather Trend Forecasting: Data Science Report

📌 PM Accelerator Mission

By making industry-leading tools and education available to individuals from all backgrounds, we level the playing field for future PM leaders. This is the PM Accelerator motto, as we grant aspiring and experienced PMs what they need most – Access. We introduce you to industry leaders, surround you with the right PM ecosystem, and explore the new world of AI product management skills.

1️⃣ Project Overview

🔹 Goal

This project analyzes historical weather data to predict future climate trends using Machine Learning (Random Forest) and Deep Learning (LSTM) models. The objective is to enhance weather forecasting accuracy and compare traditional ML vs. deep learning approaches.

🔹 Objectives:

- ✅ Forecast weather conditions based on historical patterns
- ✅ Compare ML and Deep Learning performance
- ✅ Analyze seasonal & geographical trends
- ✅ Assess the environmental impact of weather patterns

2️⃣ Data Cleaning & Preprocessing

📌 Dataset Overview

The dataset consists of meteorological parameters such as:

- 🌡️ Temperature
- 💧 Humidity
- 🌬️ Wind Speed
- 🔲 Pressure
- 🏭 Air Quality Index (AQI)

📌 Preprocessing Steps

- ✔️ Handled Missing Values → Used mean imputation & interpolation
- ✔️ Feature Engineering → Added seasonality, temperature deviation, moving averages
- ✔️ Scaling for LSTM → Applied MinMaxScaler
- ✔️ Train-Test Split → 80% training, 20% testing

3️⃣ Exploratory Data Analysis (EDA)

📌 Key Findings from EDA

- 📈 Temperature & Humidity Trends → Seasonal fluctuations observed
- 🌍 Geographical Variability → Different regions exhibit distinct climate behaviors
- 📊 Feature Correlations →

Temperature is highly correlated with humidity & pressure
Wind speed negatively correlates with temperature
Air quality worsens in extreme temperatures

📌 Data Visualizations

- ✅ Time Series Trends → Temperature & humidity variations over time
- ✅ Heatmaps → Correlation between weather features
- ✅ Boxplots & Histograms → Distribution of weather parameters
- ✅ Geospatial Mapping → Weather variations across locations

4️⃣ Forecasting Models & Evaluations

📌 Models Implemented

1. Random Forest Regressor → Traditional ML for time series forecasting
2. LSTM (Long Short-Term Memory) → Deep learning model for sequence prediction
3. Ensemble Model (Random Forest + LSTM) → Hybrid approach for better accuracy

📌 Model Performance Metrics

Model	R^2 Score	Accuracy (%)
Random Forest	0.9995	85.6%
LSTM (Before Inverse Scaling)	-	73.11%
LSTM (After Inverse Scaling)	-	95%

Ensemble Model 0.9998 93.5%
👉 Final Model Chosen: LSTM (After inverse scaling, achieved 95% accuracy)

5️⃣ Advanced Analyses & Insights
📌 Climate Analysis
✔️ Long-Term Temperature Trends → Consistent rise in temperature over the years
✔️ Anomaly Detection → Identified heatwaves & unusual weather conditions

📌 Environmental Impact Analysis
✔️ Air Quality & Weather Correlation → Poor air quality linked to high temperatures & low wind speed
✔️ Rainfall Impact → Rainfall improves air quality by reducing pollutants

📌 Feature Importance Analysis
✔️ SHAP Values & Permutation Importance used to determine key predictors:

Temperature → Most influential
Humidity & Pressure → Strong secondary predictors
Wind Speed → Minor impact
📌 Spatial & Geographical Analysis
✔️ Mapped temperature & humidity across different cities
✔️ Clustering Analysis (K-Means) → Grouped regions with similar weather patterns

📌 Deliverables
✅ GitHub Repository: https://github.com/PATHAN-0716/Weather-Trend-Forecasting
✅ Final Report / Presentation:

🔷 PM Accelerator Mission is displayed in the report/dashboard.

📩 Contact & Contributions
👤 Developed by: PATHAN ADILSHA KHAN
📩 Reach out for questions or collaboration!