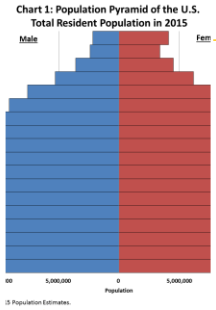
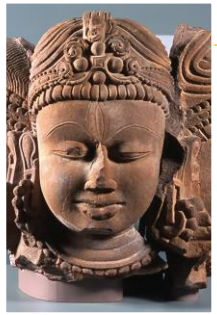
Developing a Deep Learning Model for Predicting Agricultural Weather Conditions and Crop Water Requirements in Hyderabad, India Using Time Series Data

India Overview



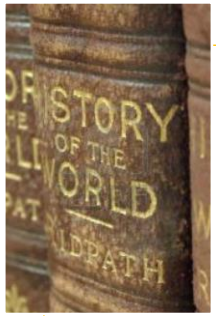
Demographics

- Constitutional republic with 28 states, 8 union territories. Capital: New Delhi.



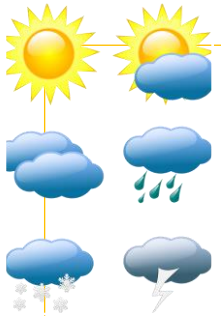
Cultural Heritage

- Home to diverse ethnic groups, hundreds of languages.



Historical Influence

- Islamic rule, European colonization (Portuguese, British). Independence in 1947, partitioned into India and Pakistan.



Geography and Climate:

- Tropical monsoon climate with diverse weather patterns.

Hyderabad Area

- Vibrant city in Telangana, semi-arid region.
- Diverse climatic influences affecting agriculture, economy, and daily life.

Problem Statement

Challenge
Accurate weather prediction in Hyderabad.

Issues
Existing models fail to combine time series and geospatial data.

Impact
Poor agricultural planning and water management.

Need
Advanced model for reliable forecasts.

Data Source and Parameters

Provider	• International Crops Research Institute for the Semi-Arid Tropics (ICRISAT)
Dataset	• Over 40 years of daily weather data (since 1978)
Parameters	• Max/Min temperatures, Relative humidity (morning/afternoon), Wind speed, Rainfall, Bright sunshine hours, Evaporation, Radiation, Reference crop evapotranspiration (FAO56-ET)
Geospatial Data	• Longitude • Latitude

Research Objectives

- Develop deep learning model for predicting future weather conditions.
- Identify significant seasonal and geographical variations in historical weather data
- Utilize K-means clustering for season classification in Hyderabad.

Research Questions

- How can a deep learning model be developed to predict future weather conditions, incorporating historical and geospatial data?
- What trends and correlations exist in the weather data over time and across different locations?
- How can seasonal variations in weather patterns in Hyderabad be effectively characterized and understood, considering the dynamic nature of climatic conditions?

Data Cleaning Process

Step 1: Data Loading and inspection

Data: ICRISAT Weather 1978 to 2018

Handling Missing Values

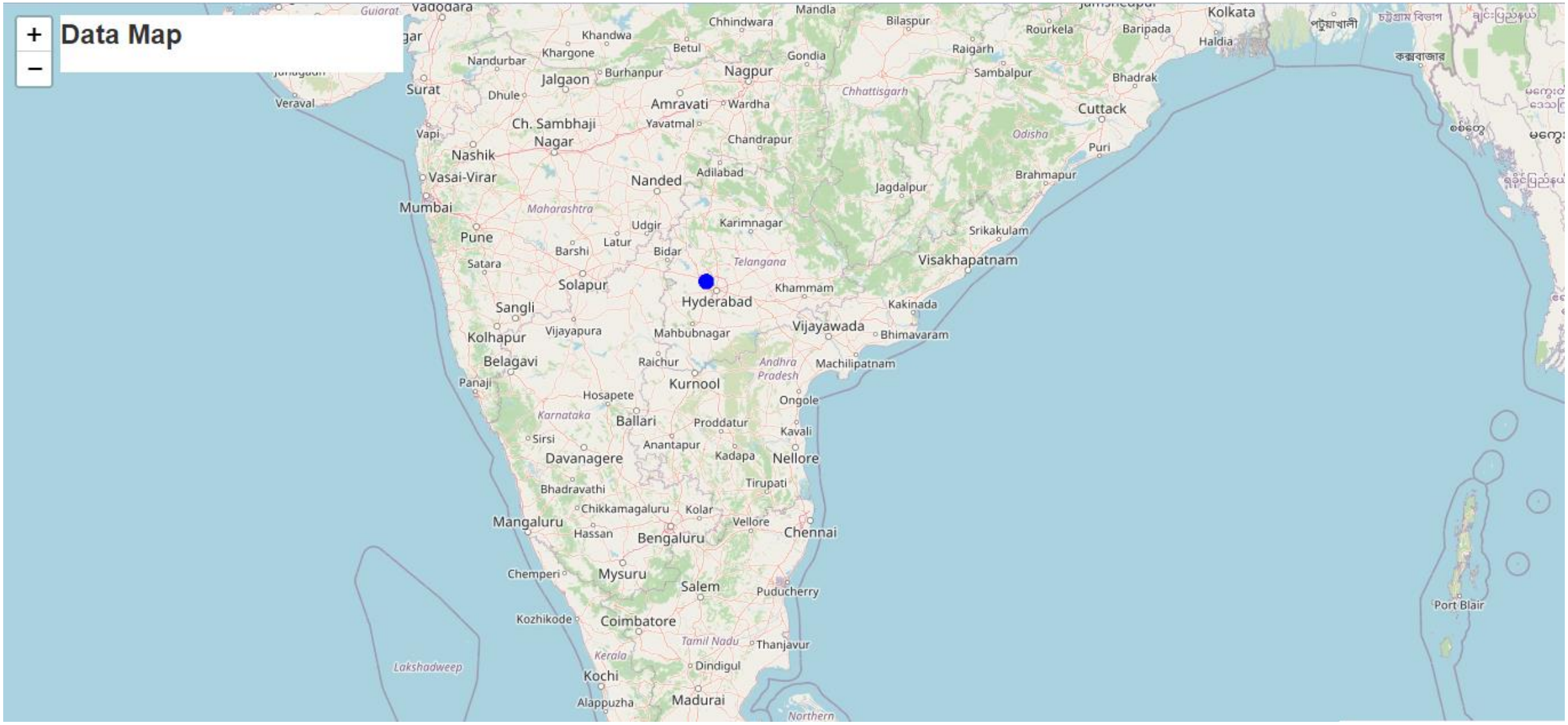
Handling Duplicates

Step 2: Deleting Unnecessary Columns

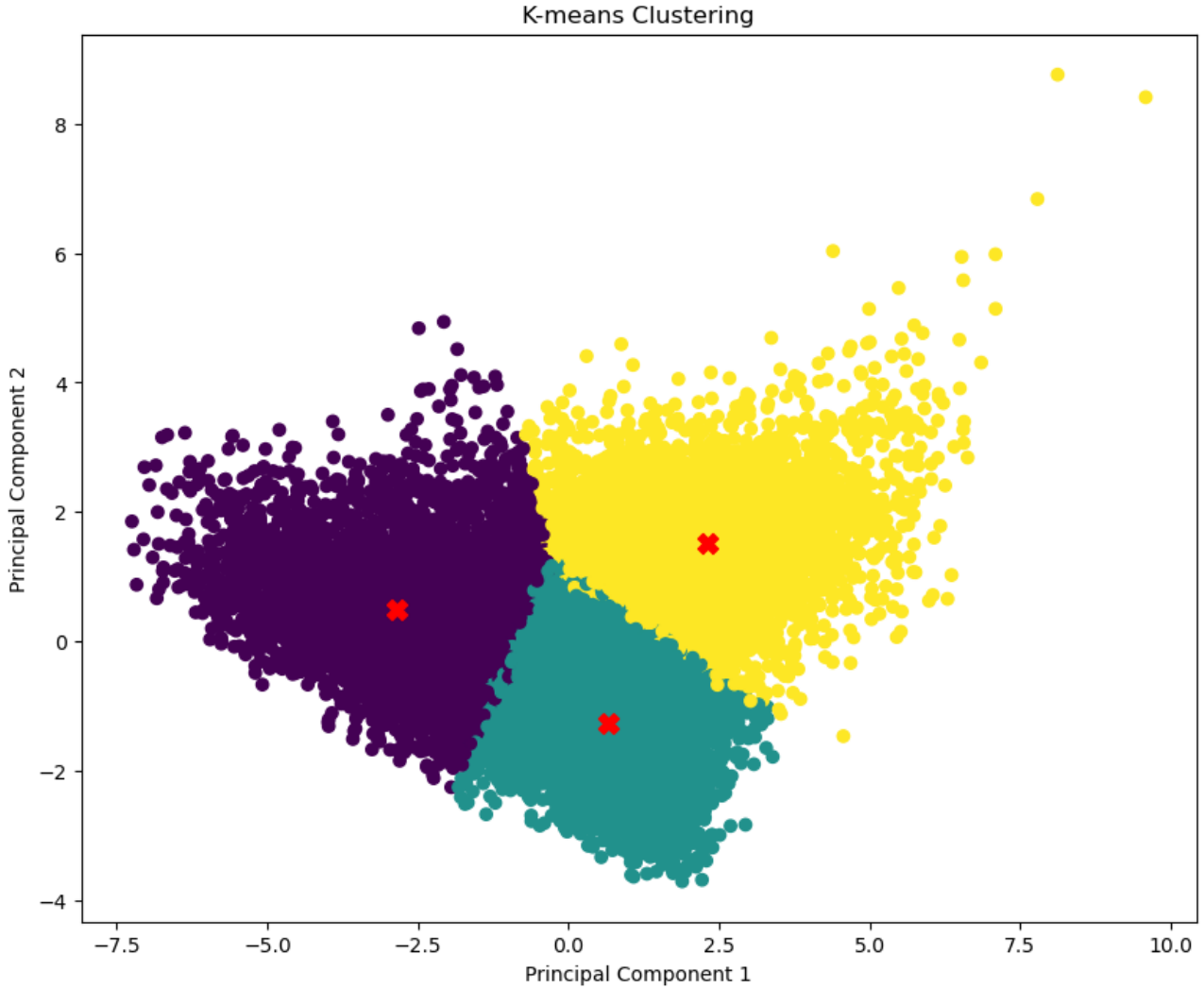
Dealing With outliers

Clean Dataset for EDA

Data Point


Data Map
Vadodara
Surat
Veraval
Khandwa
Khargone
Nandurbar
Burhanpur
Jalgaon
Dhule
Amravati
Wardha
Nagpur
Chhindwara
Mandla
Betul
Gondia
Bilaspur
Rourkela
Raigarh
Sambalpur
Baripada
Haldia
Kolkata
Bhadrak
Cuttack
Puri
Brahmapur
Ch. Sambhaji Nagar
Yavatmal
Chandrapur
Nashik
Vasai-Virar
Mumbai
Maharashtra
Nanded
Adilabad
Jagdalpur
Udgir
Karimnagar
Pune
Latur
Barshi
Bidar
Telangana
Satara
Solapur
Hyderabad
Khammam
Srikakulam
Visakhapatnam
Sangli
Kakinada
Kolhapur
Vijayapura
Mahbubnagar
Vijayawada
Bhimavaram
Belagavi
Raichur
Andhra Pradesh
Machilipatnam
Panaji
Kurnool
Hosapete
Ongole
Karnataka
Ballari
Proddatur
Kavali
Sirsi
Anantapur
Davanagere
Kadapa
Nellore
Tirupati
Bhadravathi
Chikkamagaluru
Kolar
Mangaluru
Vellore
Hassan
Bengaluru
Chennai
Chemperi
Mysuru
Salem
Puducherry
Kozhikode
Coimbatore
Tamil Nadu
Thanjavur
Kerala
Dindigul
Lakshadweep
Kochi
Alappuzha
Madurai
Northern
Port Blair

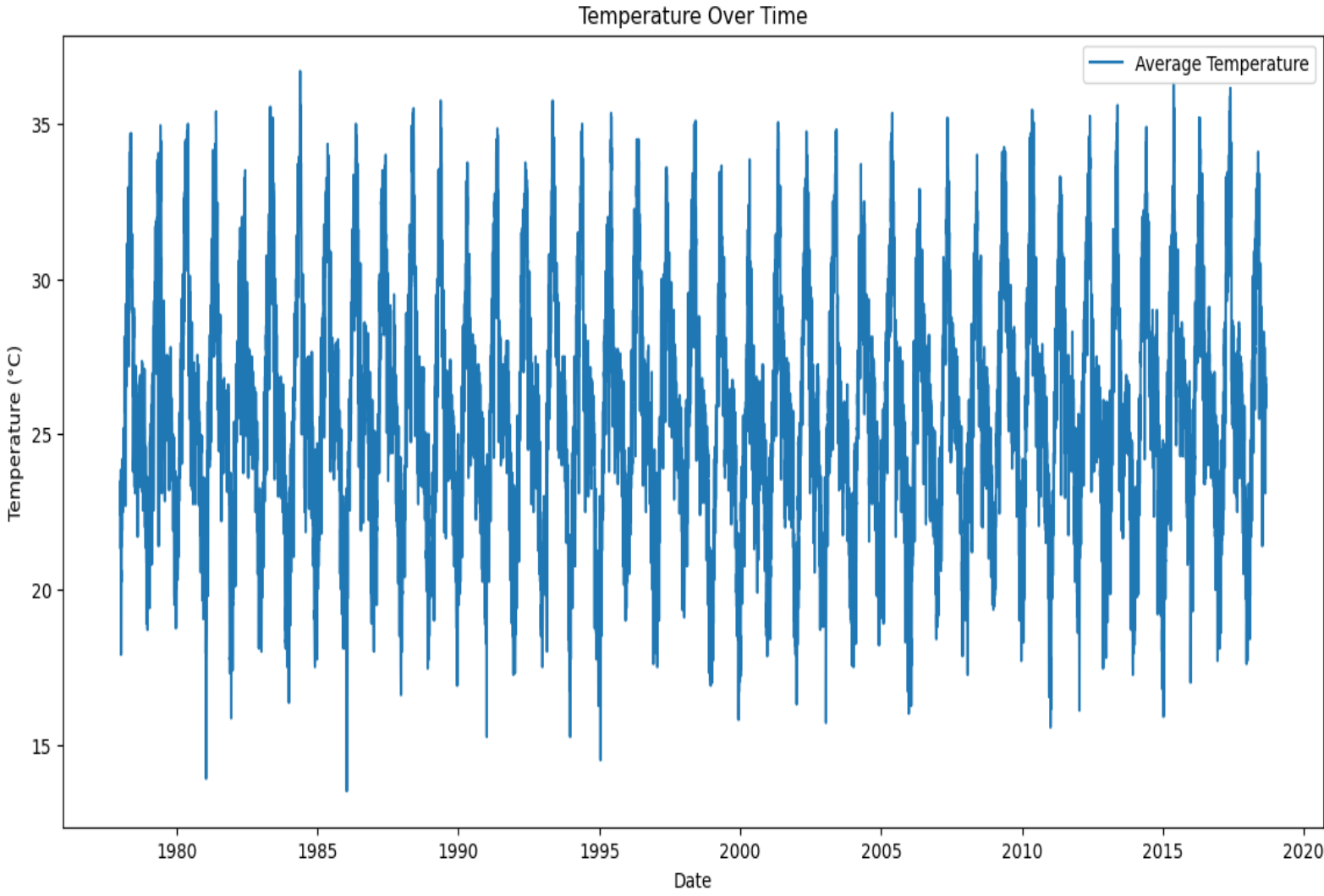
Seasons Clustering



- There are three weather seasons in Hyderabad

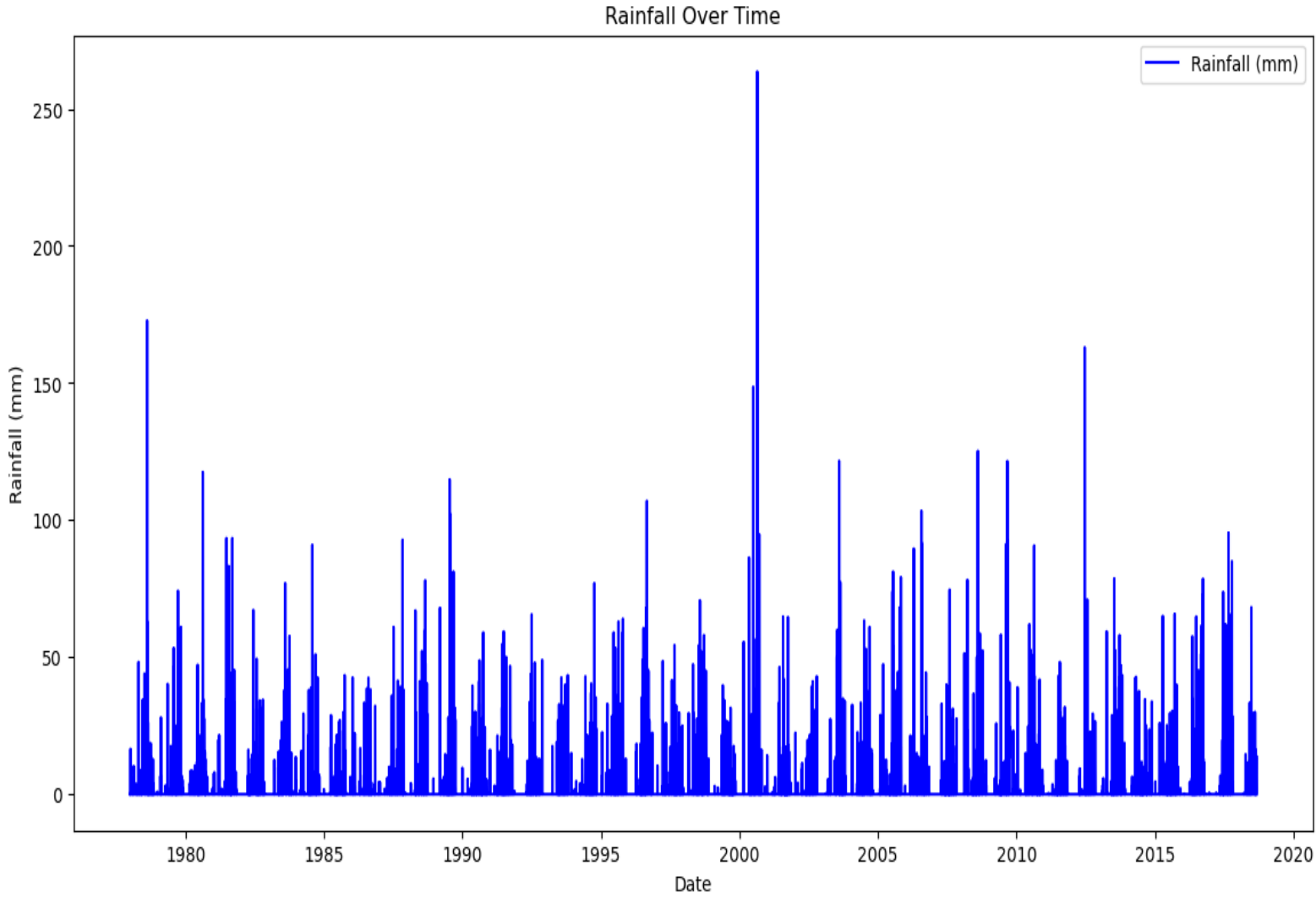
Time series- Temperature

- Average Temperature shows seasonal trends
  - Apply SARIMA for machine learning analysis analysis.



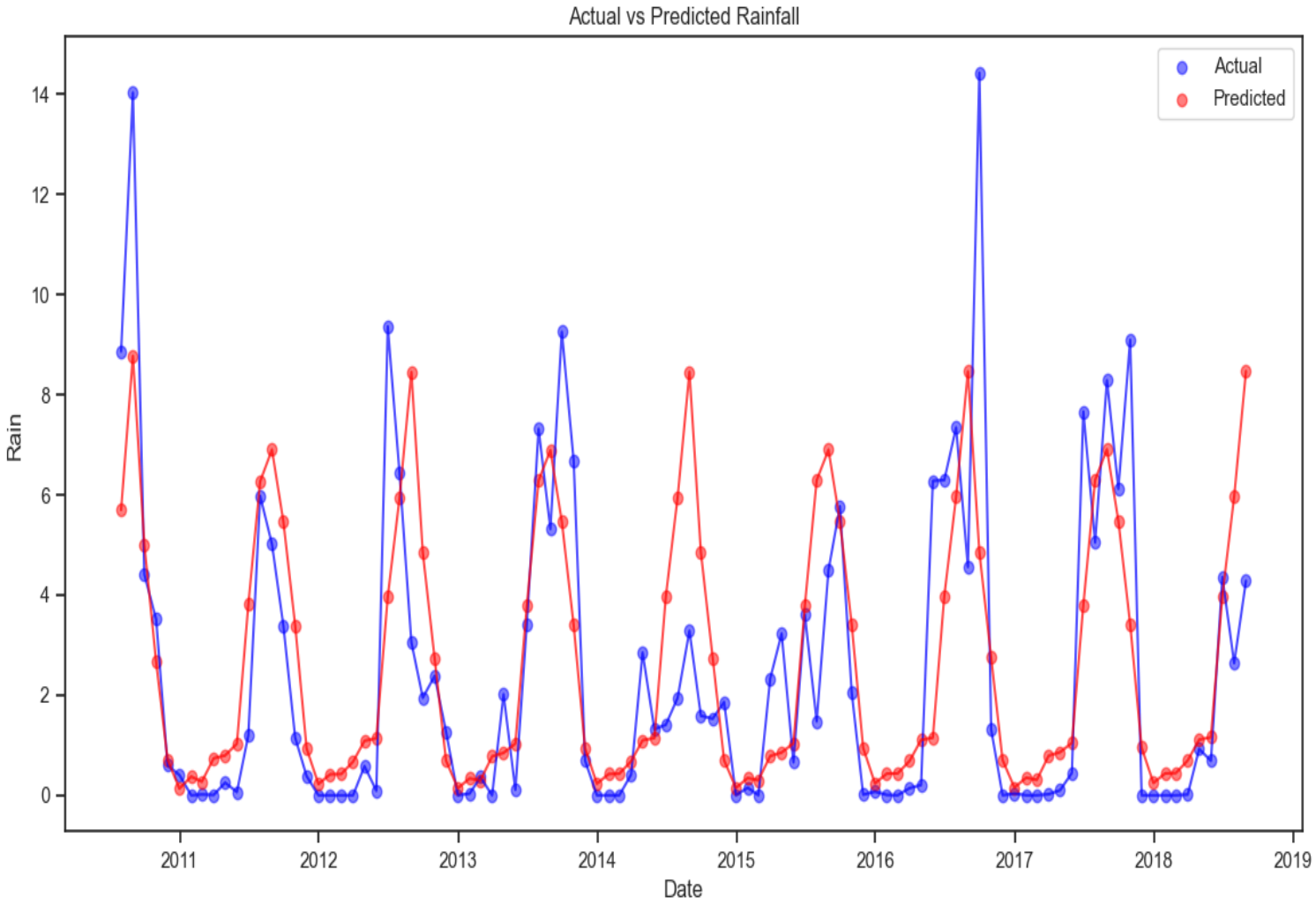
Time series- Rain

- Rainfalls shows seasonal trends
  - Apply SARIMA for machine learning analysis .



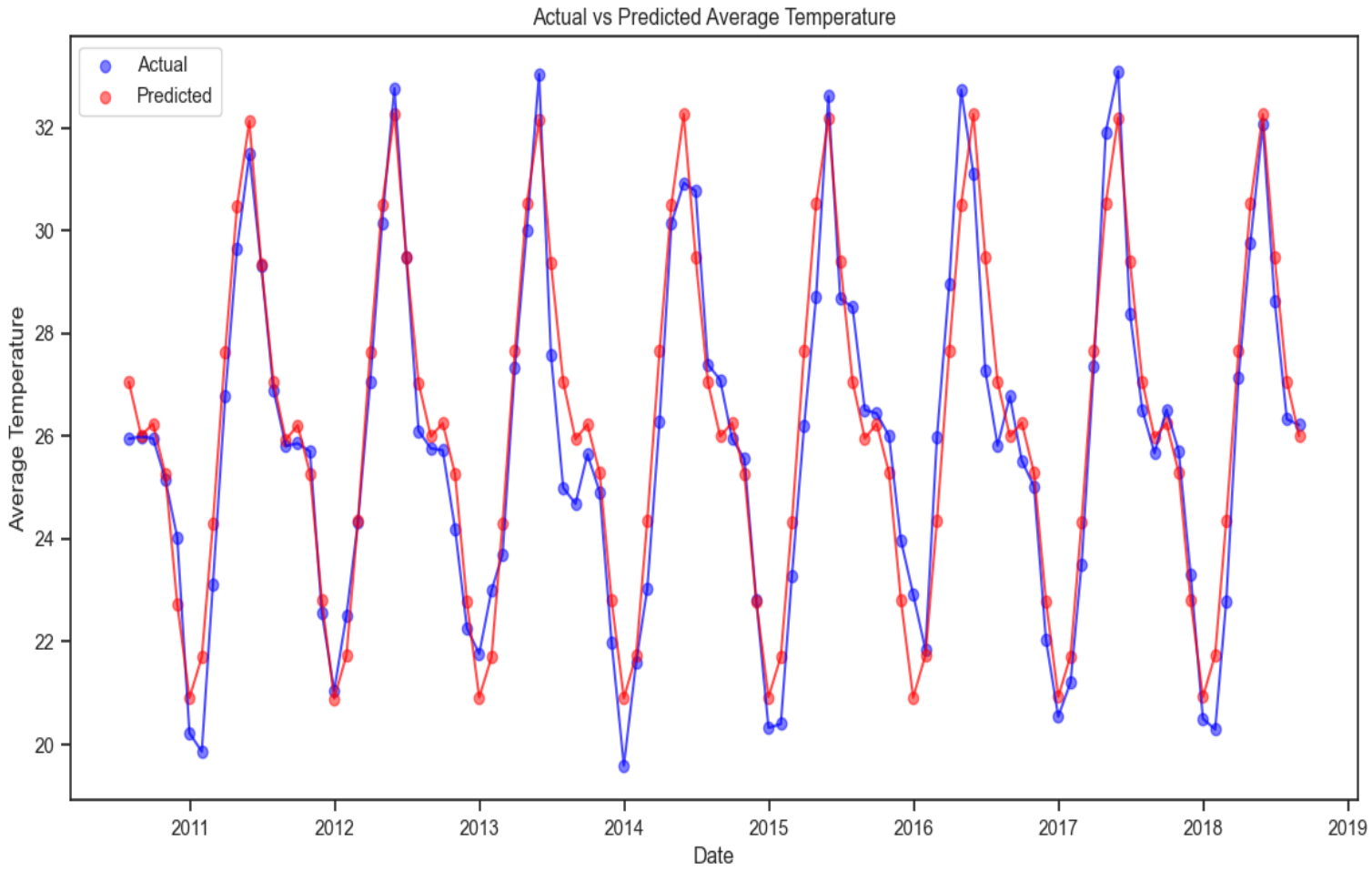
Best Machine learning Model-Rainfall

- Mean Squared Error: 5.068
- This shows average squared difference between predicted and actual rainfall values.



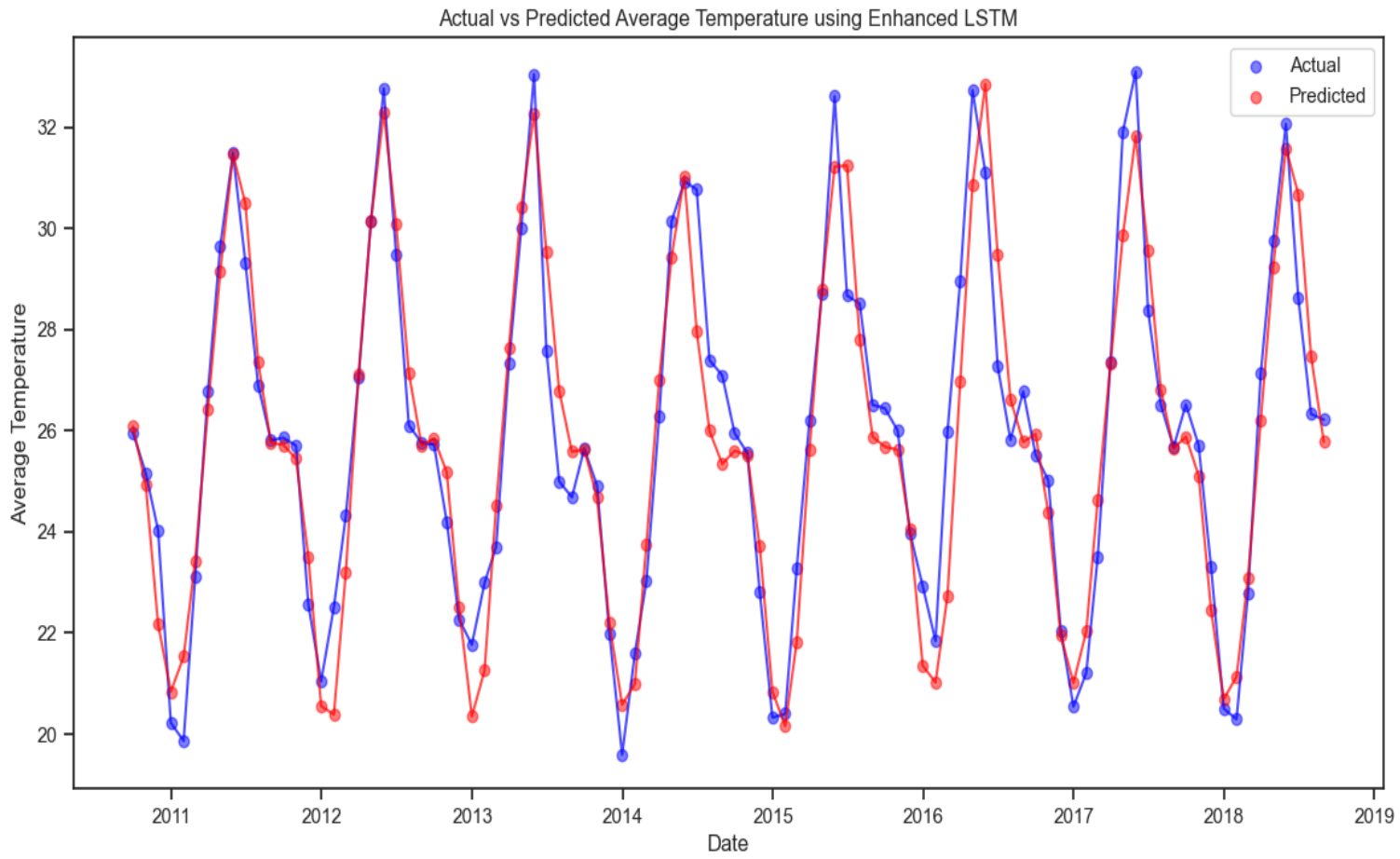
Best Machine learning Model-Temperature

- Mean Squared Error of 0.90 in predicting rainfall
- This shows average squared difference between predicted and actual rainfall values.
- A lower MSE indicates a better fit



Deep Learning

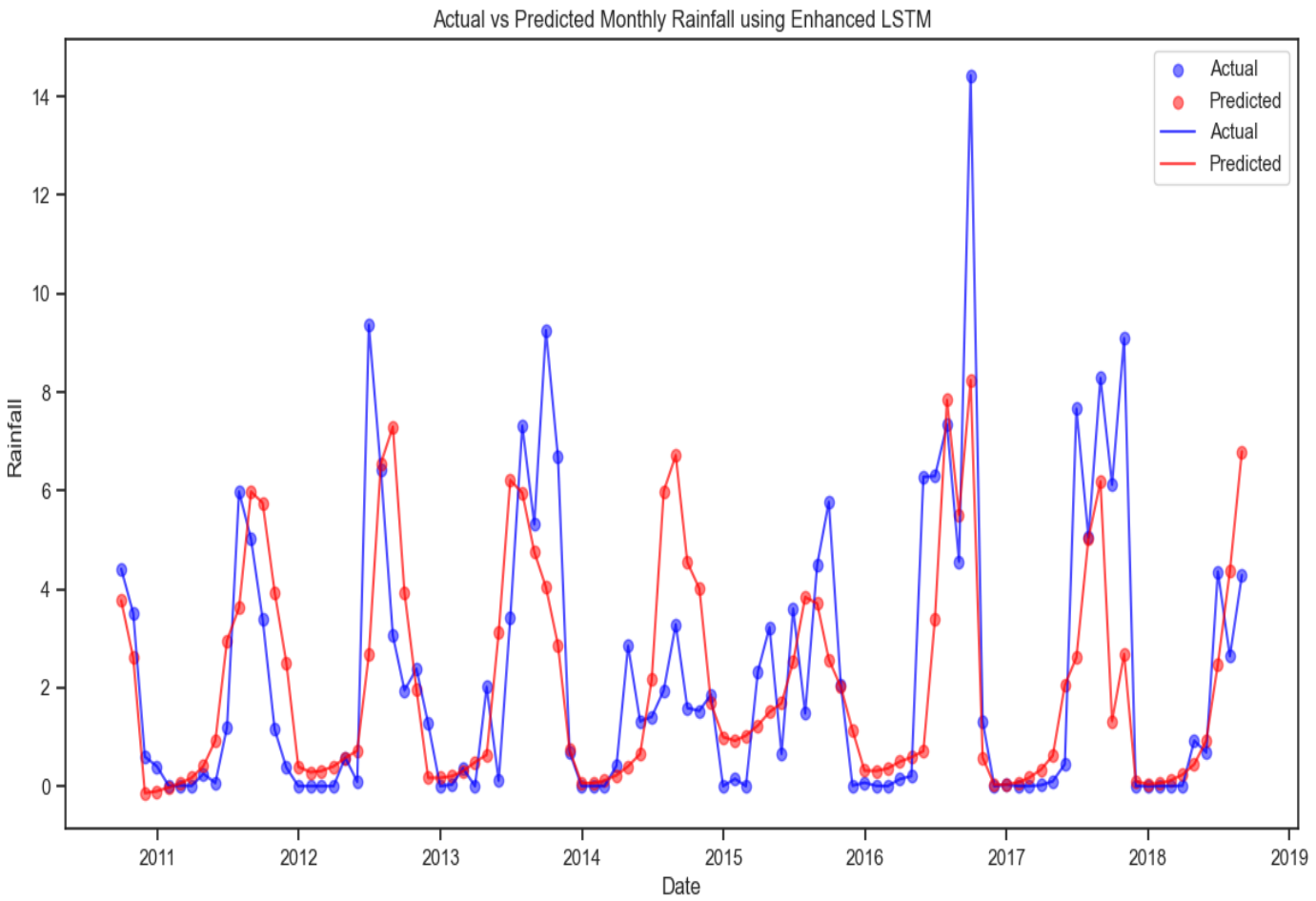
Best Deep Learning Model for Temperature



- **Model Architecture**: Enhanced LSTM with three layers (100 units each) and dropout of 0.2 after each LSTM layer, followed by a dense output layer.
- **Training**: Optimized with Adam optimizer (learning rate=0.001), trained for 400 epochs, batch size of 16, and 20% validation split.
- Metrics
  - MSE: 1.17
  - MAE: 0.83
  - RMSE: 1.08
  - R2: 0.90

Deep Learning

Best Deep Learning Model for Rainfall



- Model Architecture: Enhanced LSTM with three layers (100 units each), dropout of 0.2 after each LSTM layer, and a final dense layer.
- Training: Optimized with Adam optimizer (learning rate=0.001), trained for 200 epochs, batch size of 16, and 20% validation split.
- Metrics
  - MSE: 4.47
  - MAE: 1.37
  - RMSE: 2.11
  - R2: 0.46

Key findings

Rainfall by Month

- Peak Rainfall: June, July, August (>100 mm).
- Dry Periods: Jan, Feb, Mar (minimal rainfall).

Temperature Trends

- Warmer Months: Peak temperatures (up to 35°C) in May, June, July.
- Cooler Months: Lower temperatures (<25°C) in Jan, Feb, Nov, Dec

Recommendations

Crop Selection: Plant heat-tolerant crops in warmer months; cooler-temperature crops in cooler months.

Irrigation: Increase frequency during warmer months; efficient systems for water management.

Planting: Sow before rains start (Apr) for rain-fed crops.

Water Conservation: Harvest rainwater during peak rainfall.

Deep Learning model for rain

- Enhanced LSTM with three layers (100 units each), dropout of 0.2 after each LSTM layer, and a final dense layer."
- Optimized with Adam optimizer (learning rate=0.001), trained for 200 epochs, batch size of 16, and 20% validation split

Deep Learning model for Temperature

- LSTM model-Three layers (100 units each) with 0.2 dropout after each LSTM layer.
- Optimized with Adam (lr=0.001), trained 400 epochs, batch size 16, 20% validation split

Limitation of the study

Single Collection Point: Data from one location in Hyderabad may not generalize well across the region due to local weather variations and microclimates.

Temporal Variations: The model may not capture all temporal changes affecting weather patterns, like climate shifts and extreme events.

Data Quality: Accuracy is influenced by data completeness and potential biases, impacting the reliability of predictions.

Thank you