Weather Forecasting

Part 2 -

System Diagram

1. IoT Sensors & Data Collection:

- IoT-based weather sensors continuously collect temperature, humidity, pressure, and precipitation data.
- These sensors transmit data to the cloud every minute via an API.

2. Data Ingestion & Preprocessing Layer:

- A **Message Queue (e.g., Kafka, RabbitMQ)** ensures reliable and scalable real-time data ingestion.
- A **Data Validation Module** filters out corrupted or missing sensor readings.
- A **Preprocessing Pipeline** normalizes and aggregates data for daily-based predictions.

3. Storage & Database Management:

- **Time-Series Database (e.g., InfluxDB, TimescaleDB)** stores raw and processed weather data.
- **Data Lake (e.g., AWS S3, Google Cloud Storage)** stores historical data for model training and retraining.

4. Machine Learning Pipeline:

- A **Trained LSTM Model** predicts the probability of rainfall for the next 21 days based on historical and real-time data.
- A **Model Monitoring System** detects drifts and schedules automatic retraining when necessary.
- **Batch Processing & Scheduling (e.g., Airflow, Prefect)** ensures model inference runs daily.

5. API Layer & Prediction Service:

- A **FastAPI or Flask-based Prediction API** serves real-time predictions to clients.
- An **Error Handling & Fallback Mechanism** ensures system robustness when sensors malfunction.

6. User Interface & Dashboard:

- A **Web & Mobile Dashboard (e.g., Streamlit, React, Grafana)** visualizes real-time and future rain probabilities.
- **Alert System (e.g., SMS, Email Notifications)** notifies stakeholders if extreme weather conditions are predicted.

Component Descriptions

1. **IoT Sensors & API**: Collect real-time weather data and send it to the cloud.
2. **Message Queue**: Ensures smooth data ingestion and prevents data loss.
3. **Data Validation & Preprocessing**: Cleans, normalizes, and aggregates incoming data.
4. **Storage & Database**: Manages historical and real-time data efficiently.
5. **ML Model & Pipeline**: Runs predictions and retrains when necessary.
6. **Prediction API**: Serves predictions to applications and users.
7. **Dashboard & Alerts**: Provides a user-friendly interface and notifications for stakeholders.