INTELLIHACK 5.0

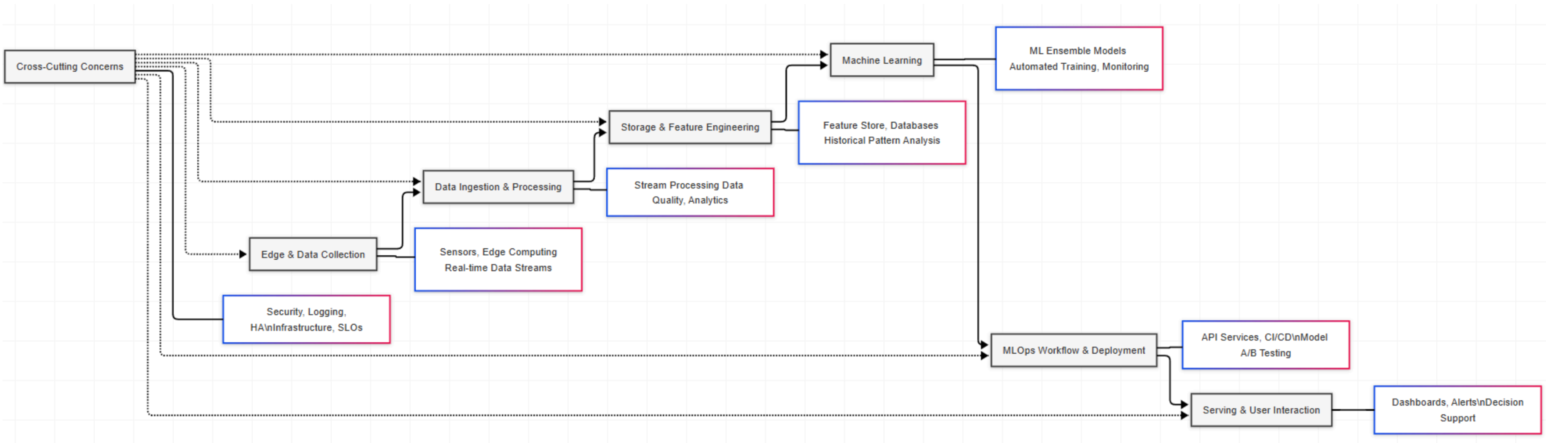
# Task 01

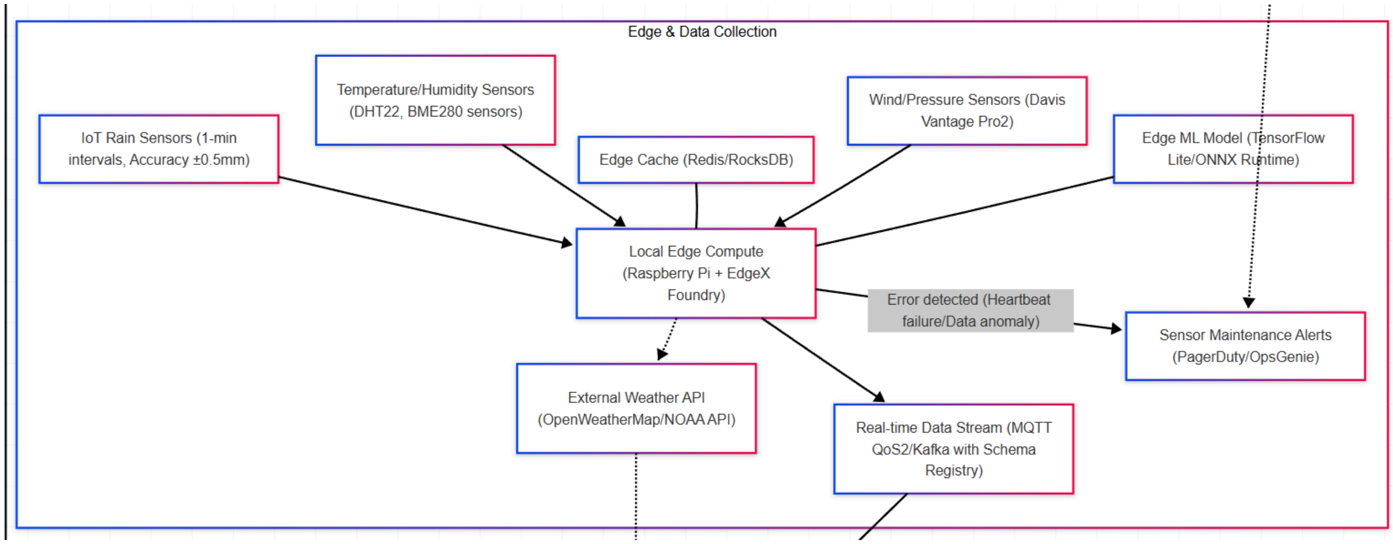
Team Cognic AI

# Part 02

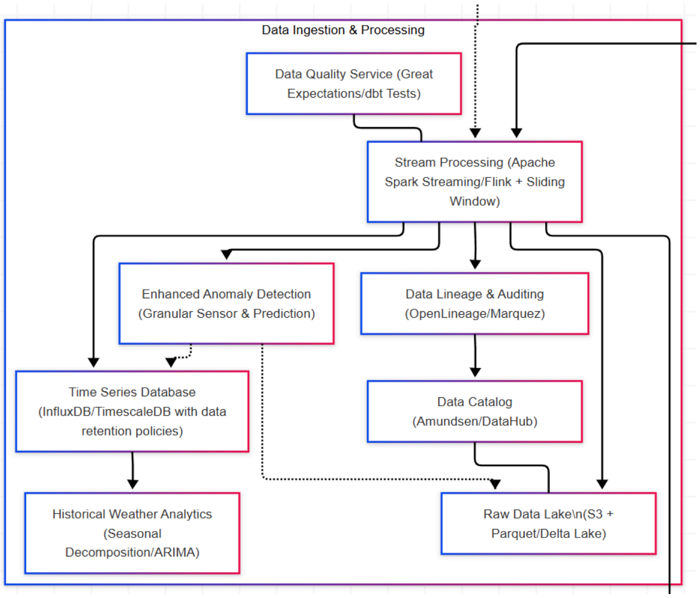
**Introduction**

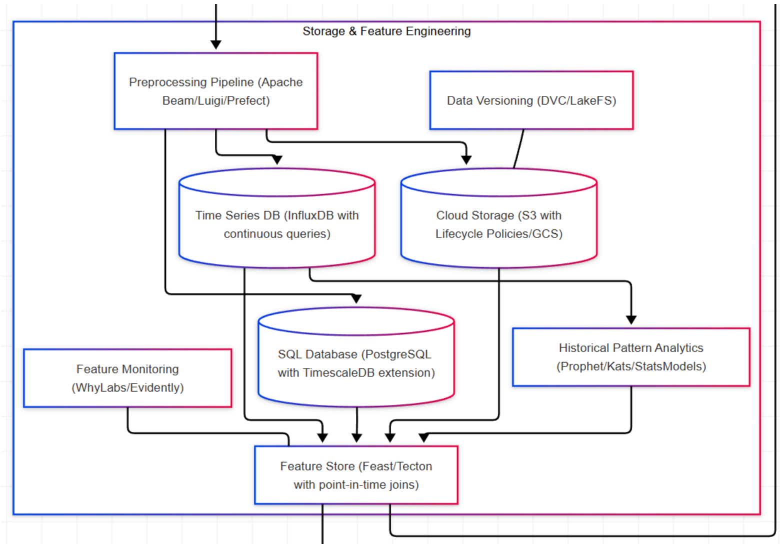
This report outlines a system architecture for real-time weather forecasting in smart agriculture. It integrates edge sensor data with robust processing, storage, and feature engineering pipelines, and leverages advanced machine learning with an automated MLOps workflow to deliver 21-day rain probability forecasts. Emphasizing scalability, fault tolerance, and real-time performance, the system features end-to-end monitoring, security measures, and feedback loops to ensure data quality and reliability, ultimately supporting informed decision-making for irrigation, planting, and harvesting. The following diagram provides an overview of the architecture, with subsequent sections detailing each component.



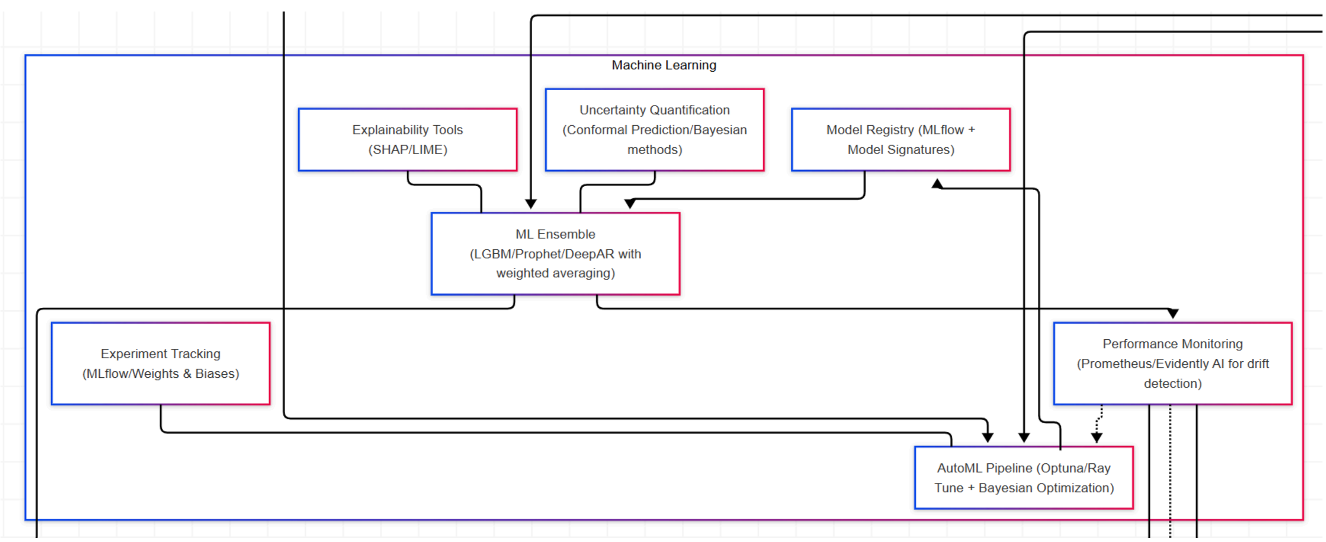

The **Edge & Data Collection** layer gathers real-time weather data from IoT sensors, including rain, temperature, humidity, wind, and pressure sensors. These sensors transmit data to a local edge compute unit, which performs initial validation, caching, and lightweight ML processing before streaming the data via MQTT/Kafka. If anomalies or sensor failures are detected, an external weather API serves as a fallback, and alerts are triggered for maintenance.



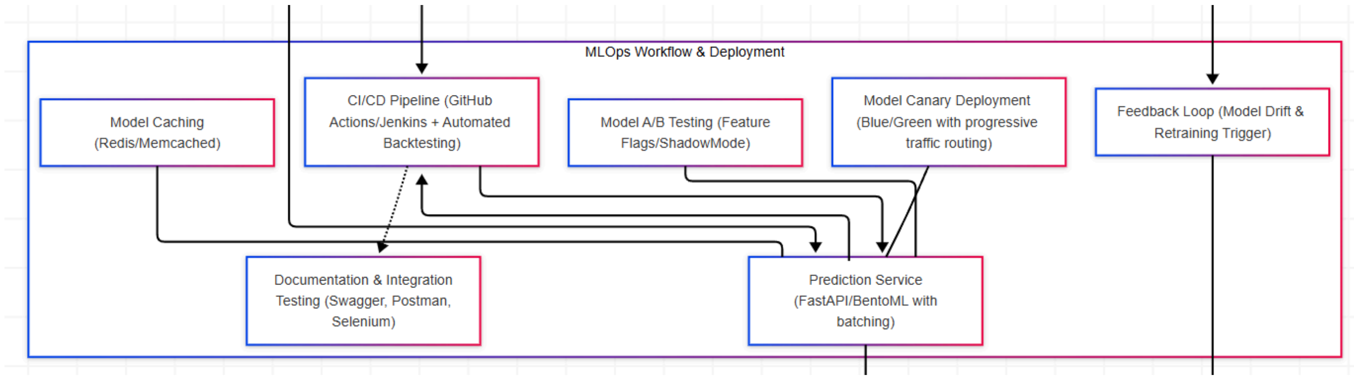




The **Data Ingestion & Processing** layer ensures efficient handling of incoming sensor data through stream processing frameworks like Apache Spark or Flink. The data undergoes cleaning, anomaly detection, and transformation before being stored in a raw data lake and a time-series database. Quality checks, metadata cataloging, and historical analytics further enhance data integrity and accessibility for downstream tasks.

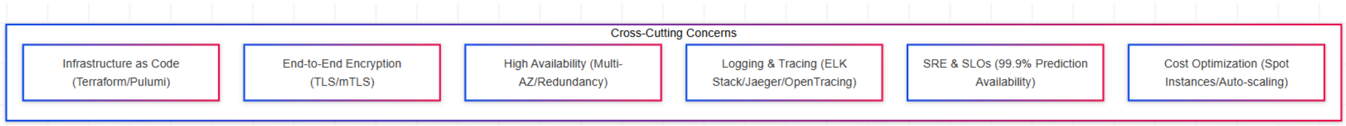
The **Storage & Feature Engineering** layer organizes processed data into structured storage solutions, including SQL databases, cloud storage, and time-series databases. Historical patterns and weather analytics are derived to generate meaningful features. A feature store centralizes and manages these engineered features, ensuring consistency across model training and inference. Versioning and monitoring mechanisms maintain data integrity and feature reliability.



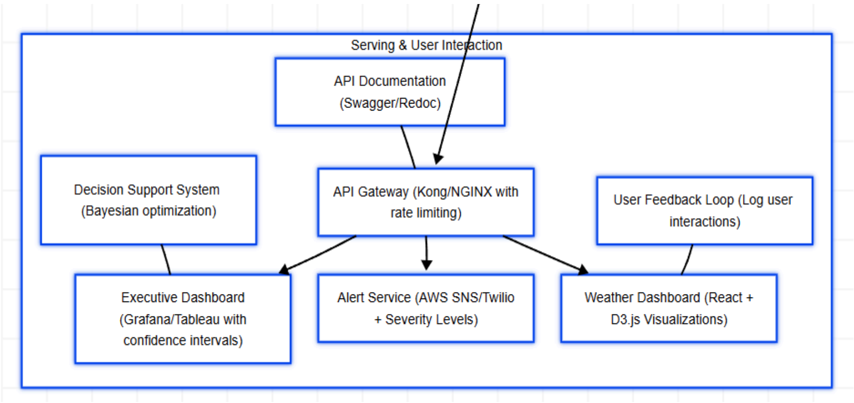
The **Machine Learning** layer utilizes an ensemble of models, including statistical and deep learning techniques, to generate accurate 21-day rainfall forecasts. Automated hyperparameter tuning and model versioning ensure optimal performance. Model explainability, uncertainty quantification, and performance monitoring help maintain reliability by detecting data drift and improving predictions over time.




The **MLOps Workflow & Deployment** ensures seamless model deployment, monitoring, and continuous integration. Predictions are served via an optimized API with caching and load balancing, while CI/CD pipelines automate model updates and backtesting. Canary deployments, A/B testing, and feature flagging enable controlled rollouts, ensuring stability and performance




**Cross-Cutting Concerns** address security, reliability, and efficiency across the system. End-to-end encryption ensures data protection, while logging, tracing, and SRE practices maintain system observability and uptime. Infrastructure as code and cost optimization strategies enhance scalability and resource management.




**Serving & User Interaction** ensures seamless access to predictions through an API gateway, delivering forecasts access to predictions through an API gateway, delivering forecasts via a weather dashboard, alert services, and executive insights. User feedback loops and decision support systems enhance usability, while API documentation ensures accessibility for developers.

This system enables accurate, real-time weather forecasting by integrating edge data collection, machine learning, and automated deployment, providing reliable decision support for smart agriculture. It ensures scalability, fault tolerance, and continuous improvement through a robust MLOps pipeline.