

Project Proposal Draft — Smart Agriculture: 24-Hour Irrigation Demand Forecasting

1) Brief Problem Statement

In large agricultural fields, we aim to accurately and timely forecast the irrigation needs of each zone over the next 24 hours in order to reduce water waste and increase crop yield. The system produces zone-level predictions using soil moisture sensors collected every 30 minutes, local weather data, and historical irrigation records. Our targets are to reduce water usage by approximately 20%, increase yield by approximately 5%, and classify irrigation needs with 90% accuracy. Predictions must be delivered to the decision support system within a maximum latency of 5 minutes. Key constraints include geospatial data privacy, sensor reliability, and regulatory compliance.

2) Scope (Brief)

In Scope:

- Data collection, cleaning, and integration from existing sensors and data sources (soil moisture, weather stations, irrigation logs, field boundaries, optional NDVI).
- Development of a 24-hour irrigation demand forecasting model, real-time delivery via API (\leq 5-minute latency target), and generation of actionable recommendations for farm managers.

Out of Scope:

- Procurement of new sensors or field installations.
- Physical installation of irrigation equipment; integration specifications will be provided, while field implementation remains the customer's responsibility.
- Acquisition of legal permits, insurance, and licenses.

3) Key Risks and Mitigation Strategies

- **Data privacy:** Data anonymization, zone-level or grid-level location abstraction, encryption, and strict access control.
- **Missing or faulty sensor data:** Data validation, imputation techniques, sensor health monitoring, and backup data sources (weather APIs, satellite data).
- **Incorrect labeling:** Human-in-the-loop validation and label cleaning processes.
- **Model drift and seasonality:** Continuous performance monitoring and scheduled retraining triggers.
- **Latency:** Lightweight models, edge inference/caching, optimized pipelines, and scalable APIs.

- **Compliance:** Legal review, data processing agreements, and acquisition of required approvals.

4) Methodology (Summary)

- **Data:** Soil moisture (30-minute intervals), weather station data, historical irrigation records, field boundaries, optional NDVI.
- **Features:** Time-series summaries, 24-hour meteorological trends, soil/topographic characteristics, and historical irrigation patterns.
- **Models:** XGBoost/LightGBM with time-series approaches for rapid prototyping; lightweight ensembles and model optimization for production if required.
- **Validation:** Zone-based cross-validation, latency measurements, and pilot simulations to analyze water savings and yield impact.

5) Deliverables

- Brief Problem Statement (1 paragraph).
- Scope and Out-of-Scope definitions (two item groups).
- Risk analysis and mitigation documentation.
- Technical proposal draft: methodology, timeline, and budget range.
- Pilot PoC: 4–8 week implementation report.
- Final package: model code, data pipeline scripts, API documentation, and maintenance/user guide.

6) Proposed Timeline (Accelerated)

- **Week 0–1:** Data access setup, sample data acquisition, clarification of privacy requirements.
- **Week 2–4:** Data preparation and rapid prototype models.
- **Week 5–8:** Model maturation and integration API development.
- **Week 9–12:** Pilot validation, performance reporting, and final recommendations.

7) Resources (Proposed)

- Project Architect / Proposal Lead: you
- Data Engineer: 1
- Data Scientist: 1–2
- Backend Engineer: 1

- Field Coordinator: provided by the customer

8) Rough Budget Range

- **PoC (8–12 weeks): \$25k–\$60k** (human resources + infrastructure).
Production costs will be evaluated separately based on scope expansion.

9) Recommended First Step

Obtain pilot field data access and clarify privacy/permission requirements from the customer (Week 0–1). Once these are secured, we can validate data connections and begin rapid prototyping.