Microenterprise Density Prediction System: Robust Architecture, Quality Assurance, and Project Management Plan

Daniel Felipe Gómez Miranda
Dept. of Computer Engineering
Universidad Distrital Francisco José de Caldas
Email: {dfgomezm}@udistrital.edu.co

Julian David Cabrera Barragán
Dept. of Computer Engineering
Universidad Distrital Francisco José de Caldas
Email: {jdcabrerab}@udistrital.edu.co

Andrés Julián Vargas Medina
Dept. of Computer Engineering
Universidad Distrital Francisco José de Caldas
Email: {ajvargasm}@udistrital.edu.co

Geraldine Alejandra Vargas Moreno
Dept. of Computer Engineering
Universidad Distrital Francisco José de Caldas
Email: {geavargasm}@udistrital.edu.co

***Abstract*—This paper, the third in the Microenterprise Density Prediction System series, consolidates the evolution of the project by refining its architecture, ensuring quality through risk analysis, and formalizing a project management plan. The updated system integrates enterprise-grade robustness principles—modularity, fault-tolerance, and scalability—guided by ISO 9000, CMMI Level 3, and Six Sigma standards. Key risks related to data integrity, model drift, and security are assessed, with mitigation strategies and monitoring processes defined. The paper concludes with a structured management plan based on Agile-Scrum methodology, outlining roles, milestones, and continuous improvement strategies.**

***Index Terms*—System architecture, Fault tolerance, Quality assurance, Risk management, Project management, Forecasting, Scalability.**

I. Introduction

The Microenterprise Density Prediction System aims to forecast regional microbusiness density using socioeconomic and temporal data. Following the conceptual (Workshop 1) and technical (Workshop 2) developments, this third stage focuses on transforming the system into a production-ready architecture aligned with quality and project management standards.

The objectives of this paper are:

- Refine the system architecture for modularity, fault-tolerance, and scalability.
- Conduct a comprehensive quality and risk analysis.
- Establish a structured project management plan with defined roles and deliverables.
- Integrate lessons learned from Workshops 1 and 2 into a mature, operational design.

II. Refined System Architecture

A. Evolution from Previous Workshops

The foundational six-layer design proposed in Workshop 2 has evolved into an eight-layer architecture that addresses reliability, scalability, and monitoring deficiencies identified in earlier versions [1], [2]. The refined architecture now incorporates redundancy, continuous monitoring, and automatic recovery mechanisms, ensuring system stability under dynamic socioeconomic conditions.

B. Robust Design Principles

The architecture follows three key enterprise-grade principles:

- **Modularity:** Each component operates independently, adhering to ISO/IEC 25010 maintainability standards. Modules expose API contracts and support parallel development.
- **Fault Tolerance:** Redundant databases, service replication, and cached fallbacks guarantee operational continuity. These mechanisms align with CMMI reliability practices.
- **Scalability:** Horizontal scaling via microservice orchestration (Kubernetes) and vertical scaling via GPU-accelerated model training enable elastic resource management.

C. Enhanced Architectural Layers

The updated architecture consists of eight integrated layers:

1) **Data Ingestion and Validation:** Connects to multiple data sources (Kaggle, economic datasets) and enforces schema validation.
2) **Data Processing and Cleaning:** Handles missing data using robust interpolation and anomaly detection.
3) **Feature Engineering:** Implements feature versioning and lag variable generation.
4) **Predictive Modeling:** Combines ARIMA, XGBoost, and LSTM models, stored in a version-controlled model registry.

5) **Ensemble and Calibration:** Aggregates model predictions, adjusts bias, and estimates confidence intervals.
6) **Evaluation and Monitoring:** Measures performance (RMSE, MAE, IoU) and tracks drift using Prometheus and Grafana dashboards.
7) **Deployment and API:** Exposes RESTful endpoints with authentication, rate limiting, and caching layers.
8) **Feedback and Retraining:** Supports automated retraining, A/B testing, and CI/CD integration.

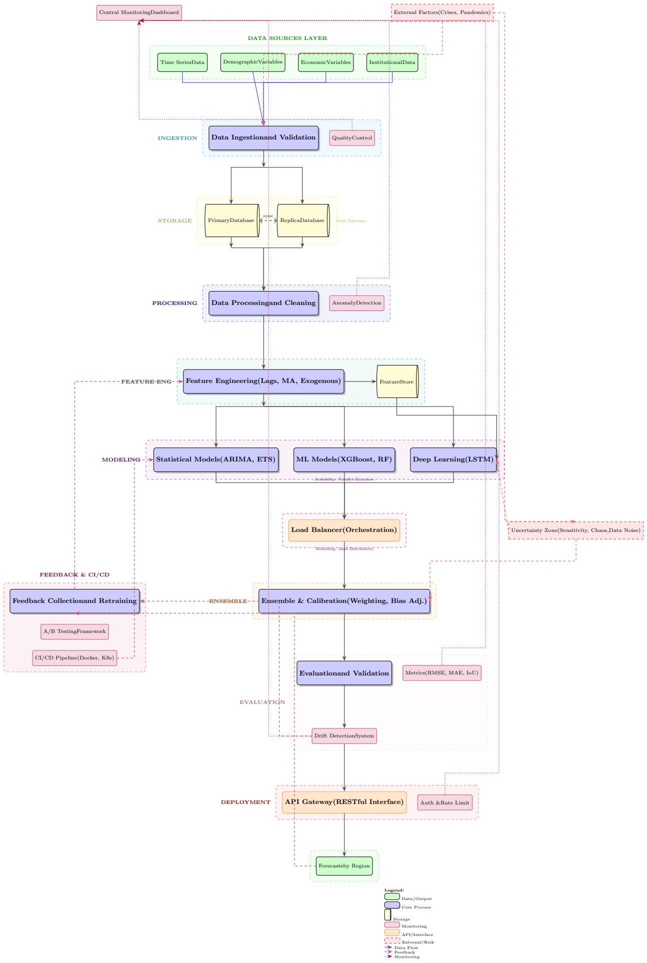

Fig. 1. Refined eight-layer system architecture highlighting redundancy, scalability, and continuous feedback.

D. Quality Standards Integration

The system architecture complies with internationally recognized quality frameworks:

- **ISO 9000:** Ensures process consistency and continuous improvement.
- **CMMI Level 3:** Promotes standardization, measurement, and documentation of development processes.
- **Six Sigma:** Focuses on statistical control and reduction of operational variability.

III. Quality and Risk Analysis

Table I summarizes key system risks, mitigation strategies, and monitoring responses, derived from the Quality and Risk Analysis document [3], [4].

TABLE I
Primary Risks and Mitigation Strategies

Risk	Mitigation Strategy	Monitoring
Data Integrity Failure	Use robust interpolation, anomaly detection, and data contracts to validate input sources.	Dashboard alerts, backup sources.
Model Drift (Performance Decay)	Apply adaptive retraining and ensemble stabilization. Schedule monthly updates even if drift undetected.	MLflow metrics, drift alerts.
Security and Ethics	Implement encrypted authentication and bias audits to ensure fairness.	Intrusion logs, fairness validation.

Each risk category has defined response procedures:

- **Detection:** Automated alerts through Prometheus and Grafana.
- **Response:** Immediate rollback or retraining triggers.
- **Recovery:** Backup data sources or cached models replace unavailable components.

IV. Project Management Plan

A. Roles and Responsibilities

Team roles were defined to ensure clear accountability and coverage across technical and managerial tasks [5]:

- **Geraldine Alejandra Vargas Moreno — Data Analyst:** Conducts EDA, feature engineering, and model validation.
- **Andrés Julián Vargas Medina — Project Manager:** Coordinates scheduling, communication, and milestone tracking.
- **Julian David Cabrera Barragán — Lead Developer:** Implements hybrid models, integration layers, and deployment pipelines.
- **Daniel Felipe Gómez Miranda — Tester & DevOps Assistant:** Executes performance tests, drift simulations, and deployment validation.

B. Methodology and Tools

A hybrid **Agile–Scrum** approach governs development. Weekly sprints focus on incremental goals (model retraining, testing, deployment), supported by a **Trello Kanban board** for task visibility.

Core tools include:

- **Collaboration:** GitHub, Google Drive
- **Visualization:** Streamlit, Power BI
- **Version Control:** Git, MLflow
- **Management:** Trello, Google Calendar

C. Key Milestones and Timeline

Table II outlines the one-month delivery plan.

TABLE II
KEY PROJECT MILESTONES

Milestone	Description	Deadline
M1: Data Validation	Finalize dataset and update features.	Week 1
M2: Model Optimization	Retrain hybrid models with Optuna.	Week 2
M3: Testing Phase	Evaluate metrics (RMSE, MAE, IoU).	Week 3
M4: Deployment	Dockerize model and finalize report.	Week 4

V. Incremental Improvements and Lessons Learned

Across Workshops 1–3, the system evolved significantly:

- Transitioned from conceptual modeling to enterprise-grade architecture.
- Added continuous monitoring and drift detection for resilience.
- Implemented Agile management to ensure transparency and iterative progress.
- Integrated quality frameworks (ISO 9000, CMMI, Six Sigma) for process maturity.

These refinements resulted in a robust, adaptive system with measurable quality assurance and streamlined collaboration.

VI. Conclusion

This workshop formalizes the Microenterprise Density Prediction System as a reliable, scalable, and ethically governed predictive platform. By uniting robust architectural design, quality assurance, and structured project management, the system achieves both technical resilience and organizational efficiency. Future work will focus on cloud deployment optimization, user interface enhancement, and real-time predictive analytics.

References

References

[1] C. E. Shannon, “A Mathematical Theory of Communication,” *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[2] Kaggle, “GoDaddy Microbusiness Density Forecasting Competition,” 2023. [Online]. Available: https://www.kaggle.com/competitions/godaddy-microbusiness-density-forecasting

[3] T. Akiba *et al.*, “Optuna: A Next-generation Hyperparameter Optimization Framework,” *KDD*, 2019.

[4] S. Taylor and B. Letham, “Forecasting at Scale,” *PeerJ Preprints*, 2017.

[5] Project Management Plan, Internal Document, Universidad Distrital, 2025.