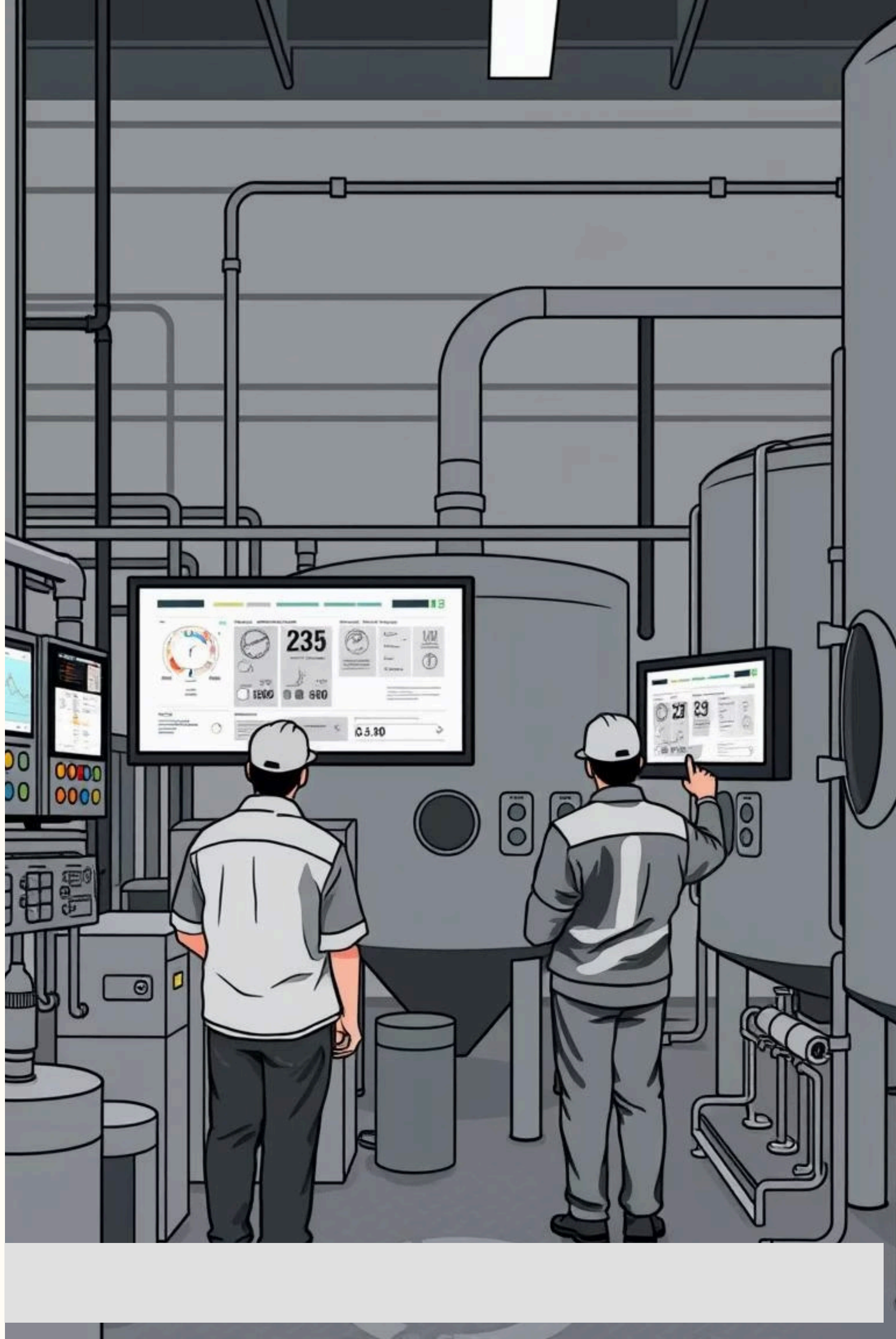
Predictive Quality Control for F&B Manufacturing

A Brewery Case Study

AMAN CHAUHAN(22BCE0476)

HONEYWELL CAMPUS CONNECT HACKATHON

An end-to-end predictive system that identifies potential quality anomalies in a brewery's manufacturing process in real-time, providing actionable Quality Alerts.



Project Overview

Objective

Build an end-to-end predictive system that identifies potential quality anomalies in a brewery's manufacturing process in real-time, providing actionable Quality Alerts.

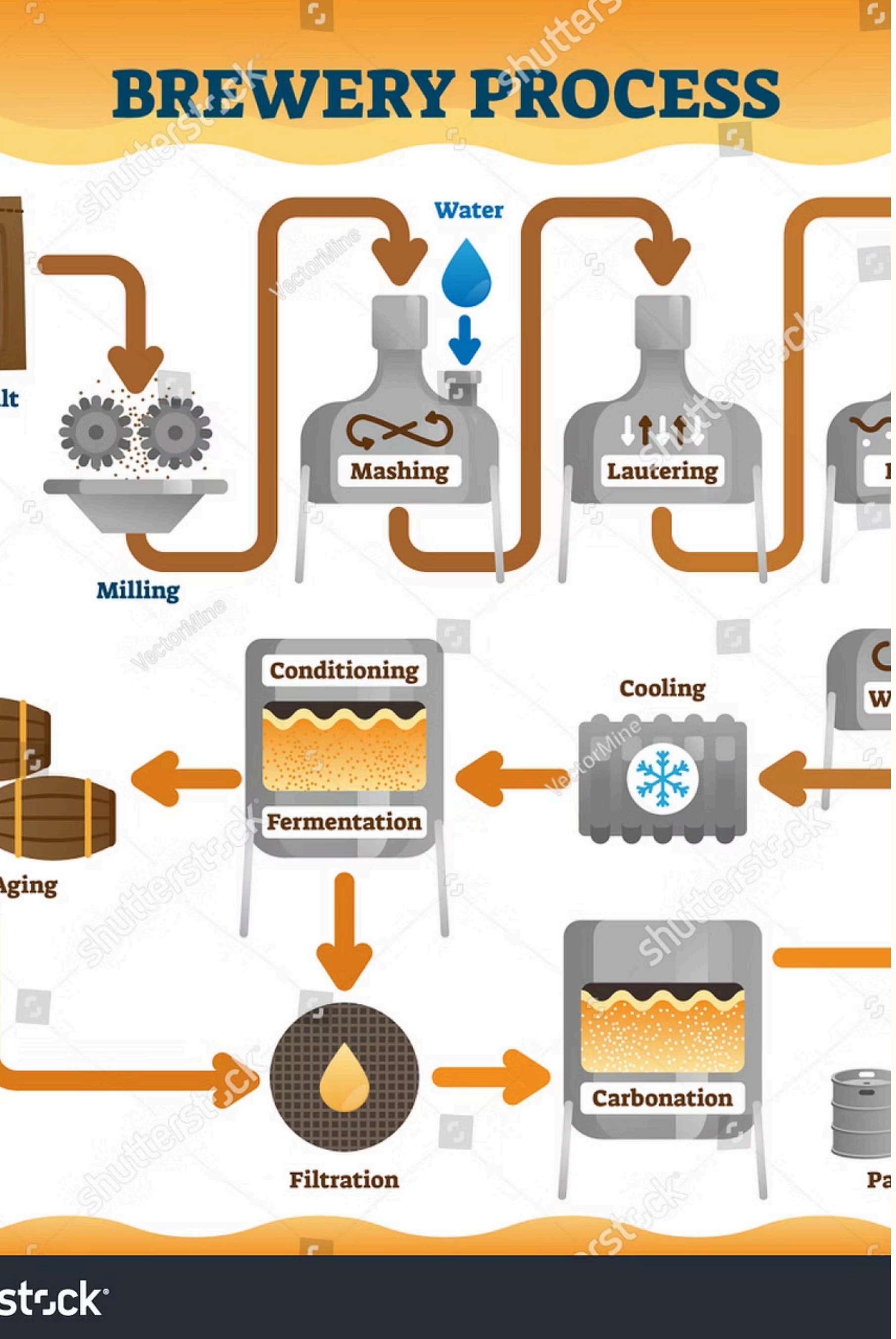
Dataset

Selected the Brewery Operations and Market Analysis Dataset for its rich process parameters including fermentation variables, product attributes, and key performance indicators.

Solution

Machine learning models analyze in-process data to classify batches as "Normal" or "Anomaly" via an interactive Streamlit dashboard with SHAP explanations.

Primary Dataset: **Brewery Operations Dataset (Kaggle)**



Brewing Process Understanding

Process Flow

Brewhouse (Mashing & Boiling) → Fermentation → Packaging

Key raw materials: Malted Barley, Hops, Yeast, Water

Control Parameters

- Brewhouse: Efficiency, Bitterness (IBU)
- Fermentation: Temperature, pH Level, Gravity, Time
- Packaging: Loss During Bottling/Kegging

Our analysis confirmed that process variations are complex; no single parameter guarantees quality. Instead, it's the complex interaction between multiple variables that determines the final quality.

Data Processing & Preparation

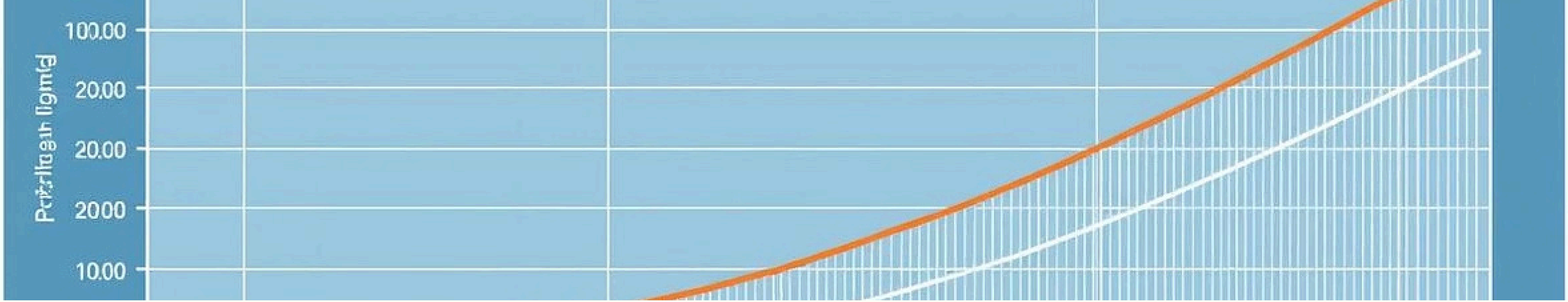
Data Quality Analysis

Statistical checks confirmed zero missing values and zero duplicate rows in the dataset. This robust initial quality ensured a strong foundation for subsequent modeling, preventing common data-driven issues.

Outlier Detection

Used Interquartile Range (IQR) method to identify outliers in key process parameters such as temperature fluctuations during fermentation and variations in ingredient quantities. Instead of removing this valuable data, we marked these deviations as new features. This approach allowed the machine learning model to learn from unusual but potentially informative patterns, distinguishing between 'normal' operational boundaries and actual anomalies.

To create an actionable Quality Alert, we converted the continuous Quality_Score into a binary target: Normal (0) for 'OK' batches and Anomaly (1) for 'Warning' or 'Reject' batches. This critical transformation aligned the model's output directly with operational decision-making, providing clear signals for intervention.



Machine Learning Approach

Baseline Models

Established performance benchmarks using Logistic Regression, Isolation Forest, PCA, and Statistical Process Control (SPC) charts.

Model Selection

Conducted a comparative analysis of Random Forest, LightGBM, and XGBoost. Each was trained on the same balanced dataset to ensure fair comparison.

Final Evaluation

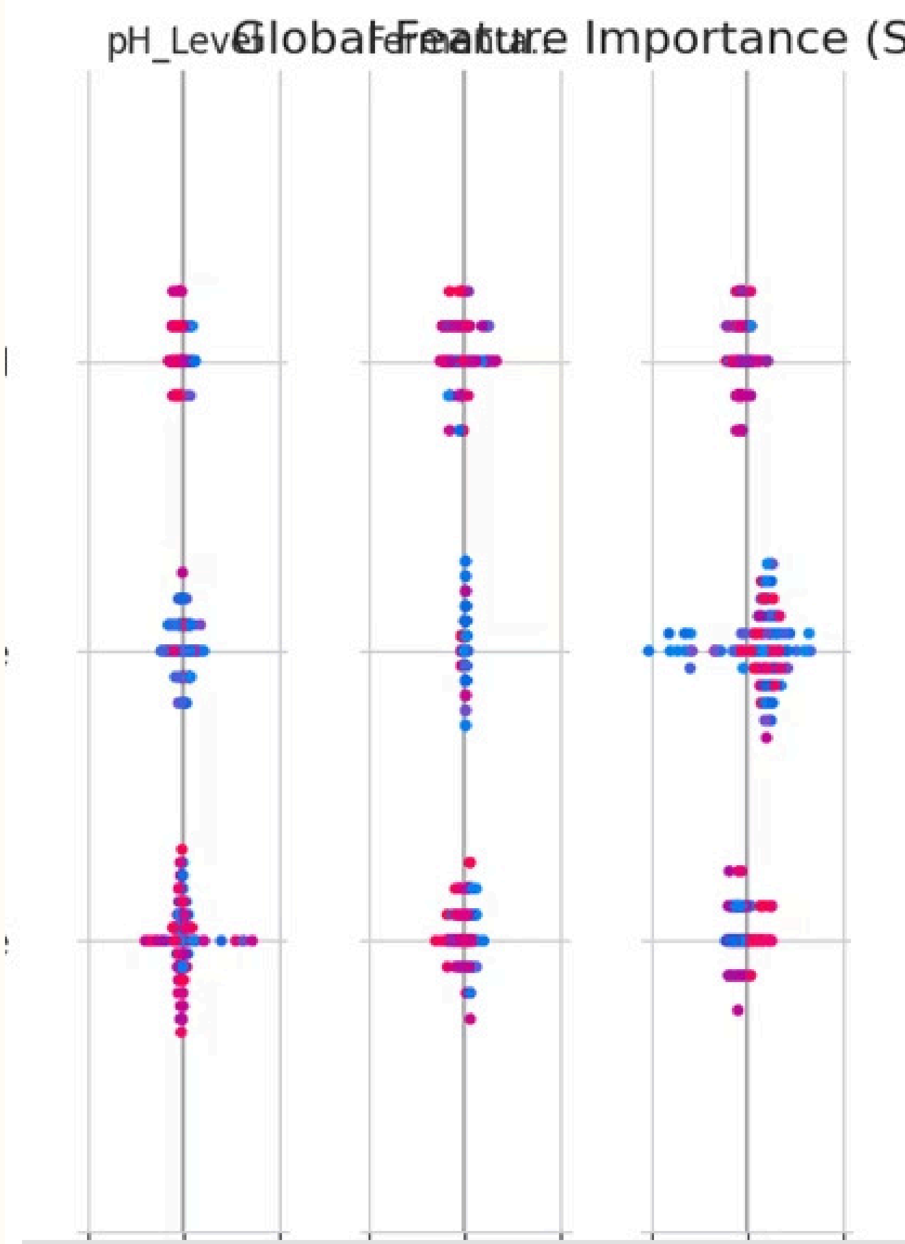
The winning Random Forest model was rigorously evaluated on the unseen test set, analyzing successes and failures to provide actionable engineering judgement.

Explainability & Root-Cause Analysis

To transform our model from a "black box" into a practical tool, we used SHAP (SHapley Additive exPlanations) to understand the reasoning behind predictions.

Analysis revealed that Brewhouse_Efficiency and Loss_During_Fermentation are the most significant factors in predicting anomalies.

> **Example Root-Cause Rule:** IF 'Fermentation_Time' is high AND 'Temperature' is low, THEN LIKELY ROOT CAUSE: Stalled or sluggish fermentation. SUGGESTED ACTION: Check yeast health, pitch rate, and glycol chilling system.



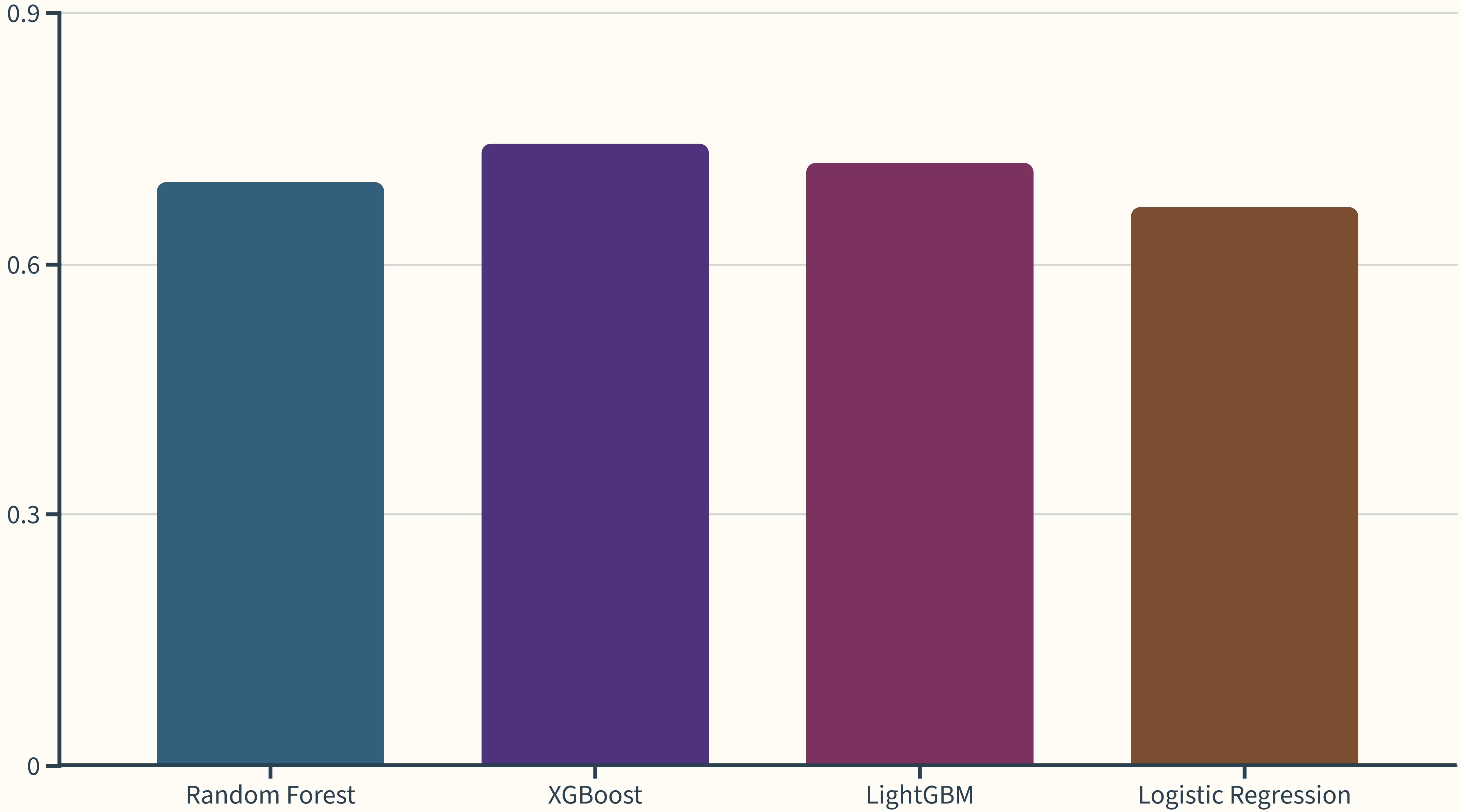
Real-Time Process Dashboard

We developed an interactive dashboard using Streamlit that allows users to:

- Select a specific batch from production history
- Choose which trained model to use for analysis
- View the model's live Normal/Anomaly prediction and confidence score
- Understand predictions through SHAP plots showing top contributing features
- Compare batch parameters against process averages using Radar Charts

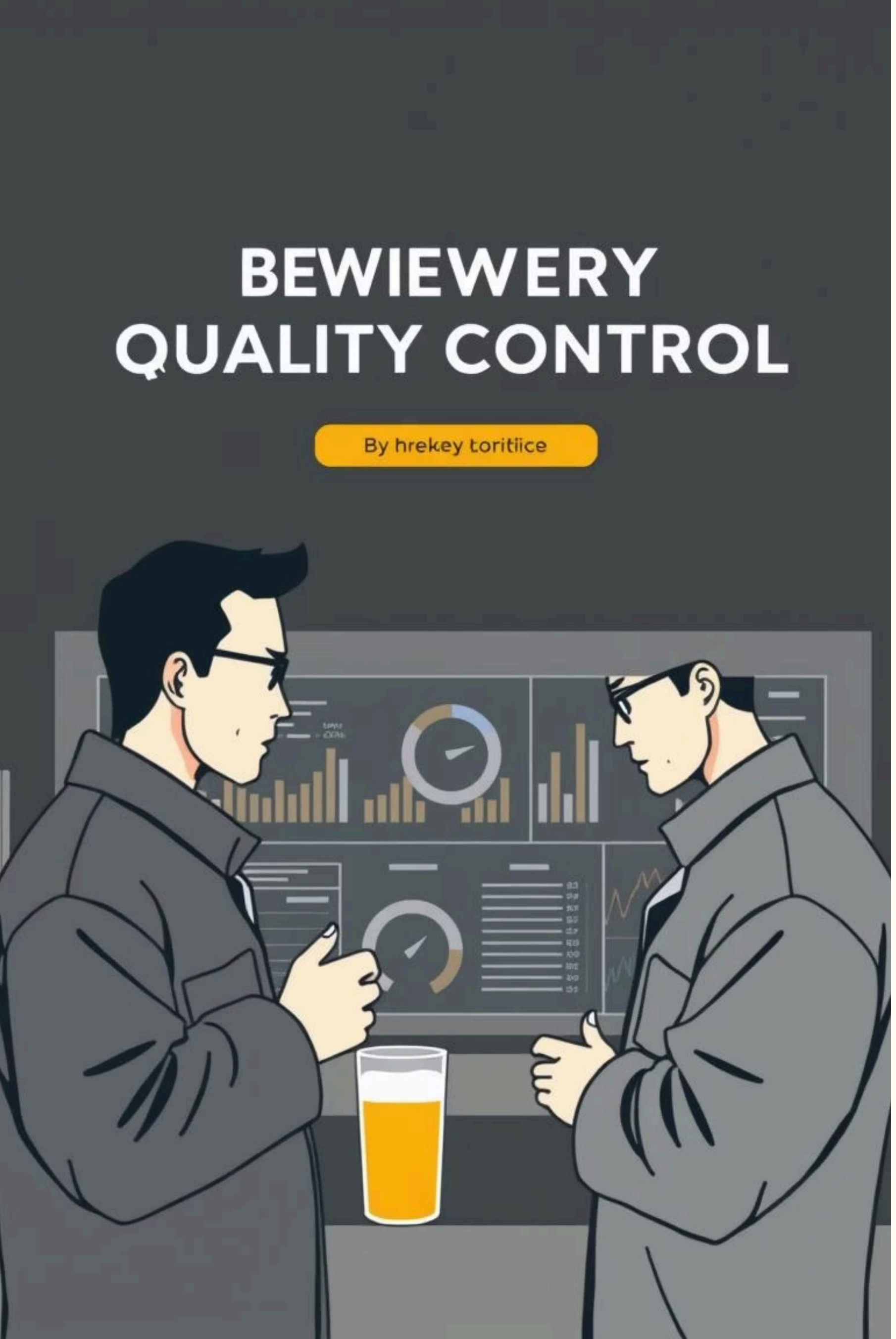
Dashboard demo: **https://fbmanufacturing.streamlit.app/**

Model Performance



The Random Forest model was the clear winner with an F1-Score of 0.95 on the test set, significantly outperforming baseline methods and other advanced models.

This demonstrates the value of using sophisticated machine learning approaches for complex manufacturing quality prediction tasks.



Project Viability & Implementation

Feasibility

The solution is highly feasible, built entirely with open-source Python libraries (Scikit-learn, SHAP, Streamlit) on a public dataset, and can be run on standard hardware.

Challenges & Risks

Primary risk in real-world deployment is the quality of input sensor data. Main challenge during development was the model's initial struggle with a 3-class problem.

Solutions

Overcame modeling challenge by strategically simplifying to a more robust binary classification, which significantly improved performance and practical value.

Resources & Next Steps

Project Links

- Dataset: **Brewery Operations Dataset**
- Demo dashboard: **https://fbmanufacturing.streamlit.app/**
- GitHub repo: **HoneyWell_F-BManufacturing**
- Video demo: **Watch on Google Drive**
- Google Collab: **Google Collab Link**

Future Enhancements

- Integrate with real-time brewery sensors
- Expand model to predict specific quality parameters
- Develop automated corrective action recommendations
- Apply transfer learning to other F&B manufacturing processes

This project demonstrates how machine learning can transform quality control in F&B manufacturing from reactive to predictive, saving time and resources while improving product consistency.