SMARTBUILD Optimizing Production: Predictive Analytics Report

Analysis of Production Log & Machine Settings

Presented By:
Mehmet Fatih Özdemir
Kasun Gayashan Pinto Ranwalage
Maulik Dilipbhai Chopda

11 December 2025

Executive Summary: The Bottom Line

THE FINANCIAL REALITY

We identified a massive cost disparity that drives our strategy.

- Cost to produce a faulty unit: **150 EUR** (Machine + Material)
- Cost to discard raw material: **10 EUR**
- **Opportunity:** We save **140 EUR** for every error predicted before production.

OUR SOLUTION

Developed a predictive models to catch defects early.

- **Key Driver:** "Ionization Class" is the #1 cause of critical failures.
- **Performance:** Weight predicted with >99% precision.
- **Impact:** Critical errors flagged before expensive machining begins.

Data Cleaning

DISCONNECTED DATA

Production Logs and Machine Settings were siloed. We successfully merged 9,000+ records into a single source of truth.

SENSOR GLITCH

Detected impossible outliers weighing 10^{21} kg. Applied Interquartile Range (IQR) filtering and Hard Caps (<1e9) to restore data integrity.

MISSING LABELS

The `Error_Type` field was often nan. We systematically imputed these missing values as "None" to create a clean baseline.

Q2 Error Type

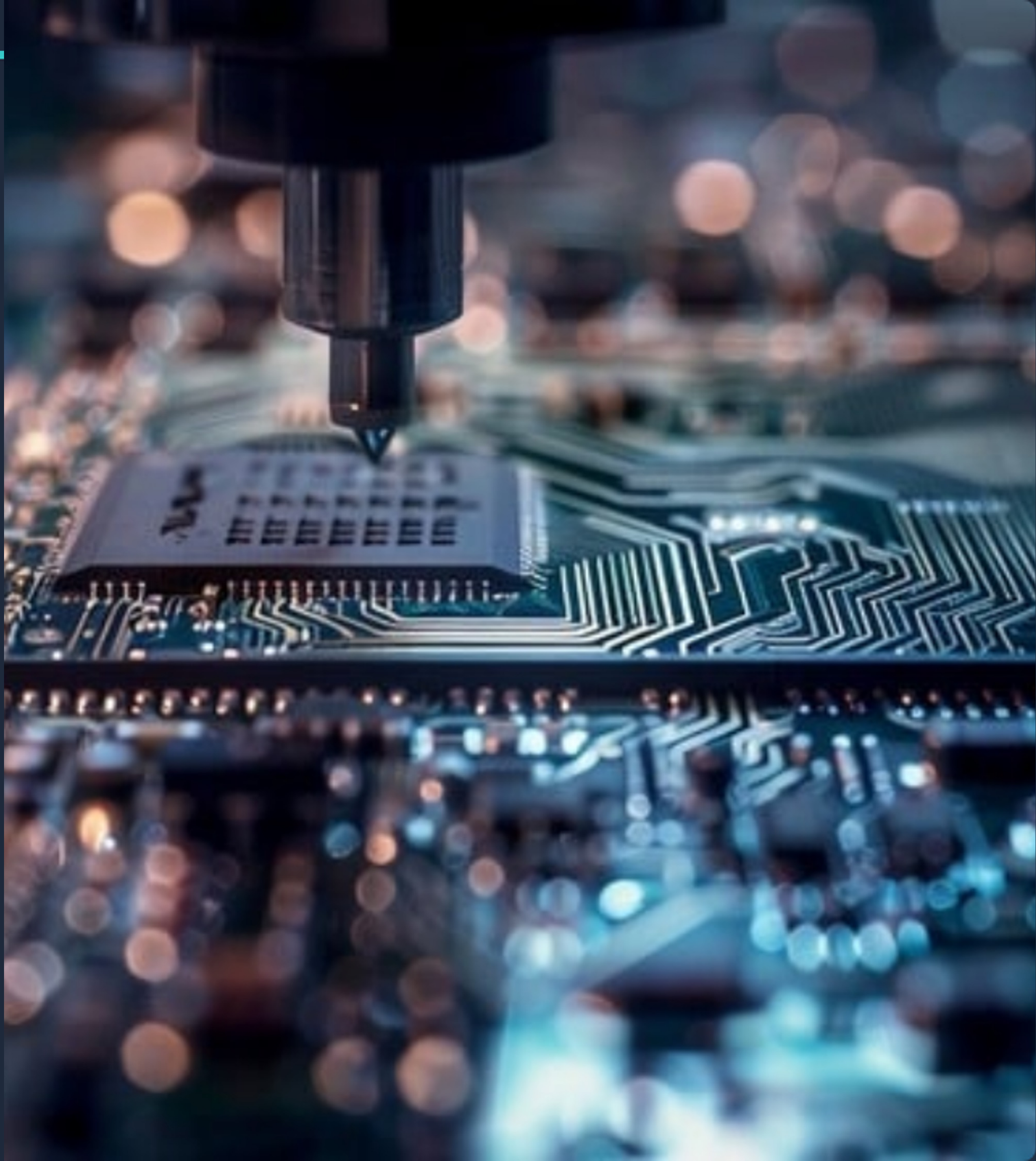
PREDICTING ERRORS BEFORE RUNNING MACHINE

Goal: Predicting error types to identify root causes and stop defects.

Approach A (Decision Tree Classifier): Visualizes the decision tree for clear logic.

Approach B (XGBoost): Confirmed that the inability to classify specific error is a limit of the data, not the model.

Approach C (Random Forest Classifier): Delivered the best balance of high detection accuracy and clear root cause analysis.



Q2 Model Performance

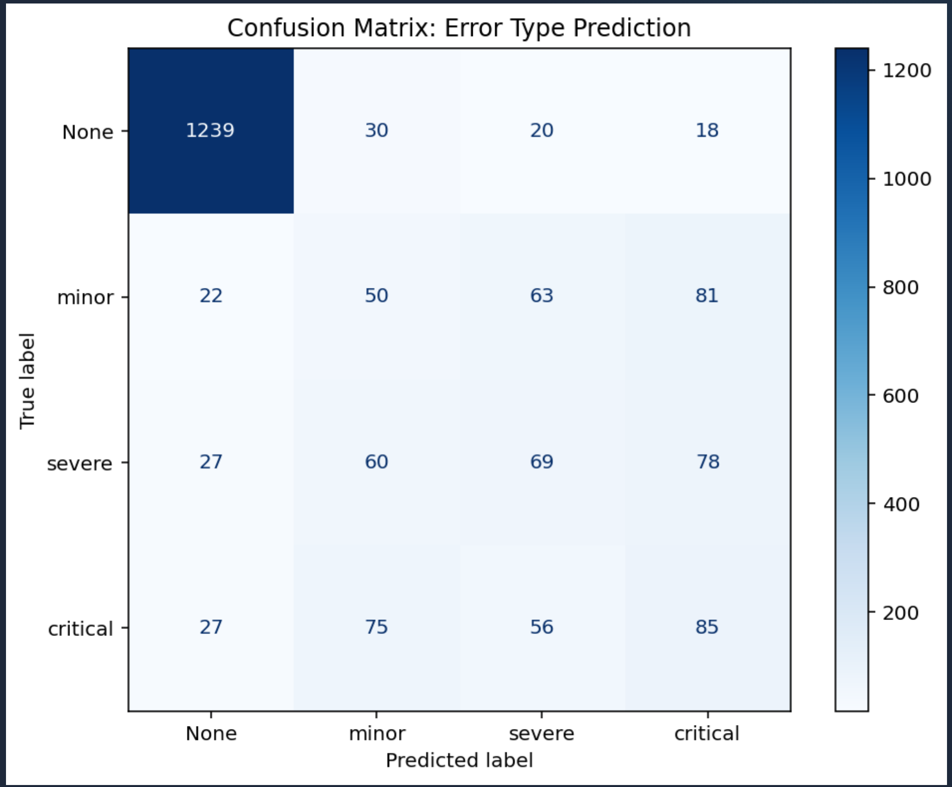
Random Forest Classifier

The model are highly effective at predicting if a failure will occur, even if distinguishing the specific type of failure remains complex..

- **Error Detection: 93% Accuracy**
  We successfully predict if a product will fail.

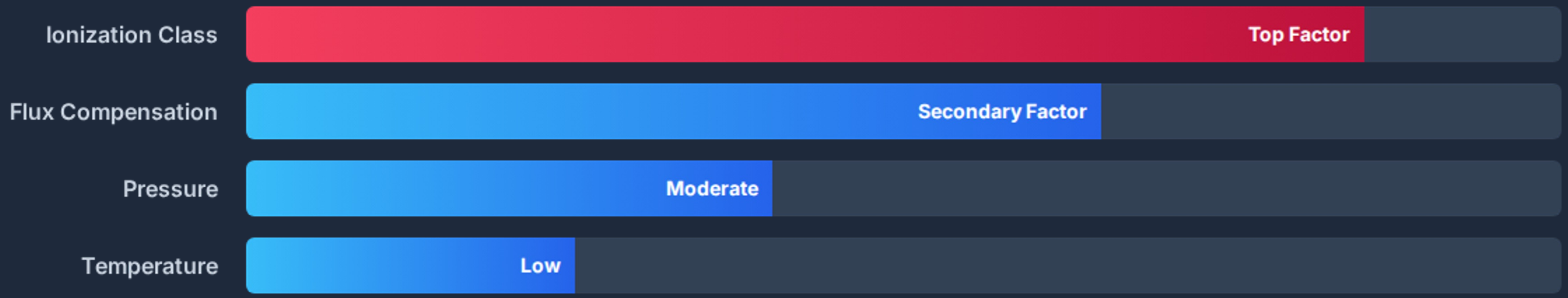
- **Error Classification: ~30% Accuracy**
  We struggle to predict the exact failure type.

Confusion Matrix (Sample)



Root Cause Analysis: What Drives Defects?

Using Feature Importance extraction, we identified that errors are not random—they are tied to specific machine configuration settings.



Insight: Mechanical faults are likely occurring specifically during **Ionization Class B** and **Flux Compensation III** settings.

Strategic Next Steps

1

Fix Sensors

Address the specific sensor logging infinite weights. Recalibrate to ensure data integrity.

2

Deploy Q2 Model

Install the Error Classifier at the start of the line to automatically reject bad configurations.

3

Engineering Review

Investigate Ionization Class B and Flux Compensation III for underlying mechanical faults.

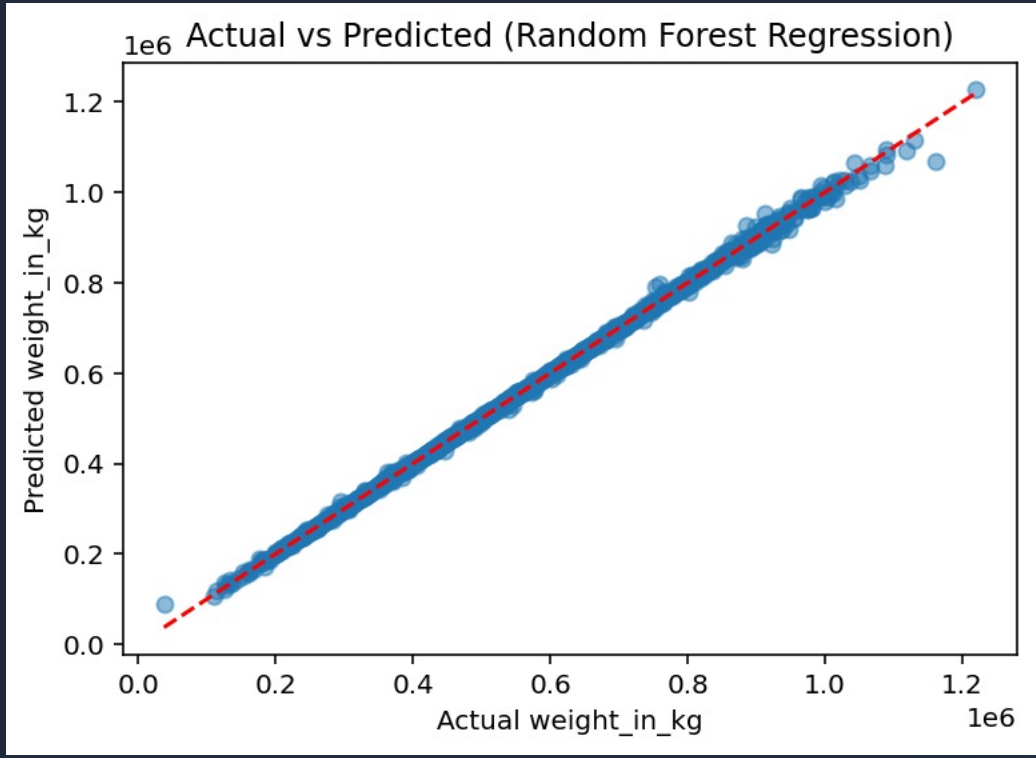
Q1: Weight Prediction

PREDICTING MATERIAL NEEDS

Goal: Forecast `weight_in_kg` based on machine settings to automate inventory planning.

- **Approach A (Linear):** Captured basic trends but missed complexity.
- **Approach B (Random Forest):** Good handling of non-linear physics.
- **Approach C (XGBoost):** The winner. Optimized for maximum precision.

Result: R^2 Score of **0.9993**. We can now forecast exact material usage.



Q & A

Thank you for your attention.

Ready for Discussion

SmartBuild Production Optimization Team