

Uncertainty-Calibrated AI Decision Intelligence from Supply Chain Data: A Late-Delivery Risk Triage Framework

Sampath Kumar Puvvada

Short Communication | Empirical research article

Abstract

Late-delivery classifiers are often evaluated as rankers, although operational triage requires probabilities that can be interpreted as risk. This short communication develops and tests an uncertainty-calibrated artificial intelligence (AI) decision-intelligence framework using 180,519 line-item records from the public DataCo supply-chain dataset. Uncertainty calibration is defined narrowly as empirical probability reliability rather than Bayesian decomposition of epistemic and aleatoric uncertainty. Twenty-six leakage-controlled order-time predictors were transformed into 120 model inputs. Order identifiers were separated across fit (115,231 rows), calibration (28,822), and test (36,466) partitions to prevent order-level contamination. Logistic regression, random forest, and XGBoost were evaluated uncalibrated and after sigmoid or isotonic calibration. Random forest with isotonic calibration achieved the strongest discrimination-calibration balance (ROC-AUC 0.7617; Brier score 0.1861; expected calibration error 0.0111), while XGBoost with isotonic calibration produced the lowest Brier score (0.1859) and log loss (0.5329). Isotonic calibration reduced Brier loss for all three models in paired order-cluster bootstraps; sigmoid calibration improved XGBoost but worsened logistic regression and did not materially improve random forest. In the prespecified random-forest/sigmoid triage, the high-risk tier contained 37.0% of test orders and showed an 85.3% observed late rate. The findings demonstrate probability reliability and risk-tier separation, not causal delay reduction or resilience gains. The contribution is a reproducible bridge from predictive supply-chain data to human-governed decision intelligence, with explicit calibration, triage, and claim boundaries.

Keywords: artificial intelligence; decision intelligence; supply-chain data; uncertainty calibration; probability calibration; late-delivery risk; risk triage; random forest; XGBoost