Quarterly Layoff Prediction for Publicly-Traded Companies

An Early Warning System Using Financial Indicators and Machine Learning

PROBLEM

Layoffs represent significant disruptions to employees, investors, and the broader economy. Existing approaches rely on retrospective analysis rather than proactive prediction. This project develops a machine learning model to classify publicly-traded companies as High Risk vs Low Risk for layoffs at the quarterly level, enabling stakeholders to take preventive action.

DATA

Sources: (1) Kaggle layoffs dataset (3,642 layoff events), (2) SEC EDGAR API (financial statements for 160 companies), (3) FRED API (8 macroeconomic indicators)
Final Dataset: 2,736 company-quarter observations from Q1 2020 to Q1 2024
Companies: 152 publicly-traded companies with ≥80% SEC data coverage
Features Engineered: 55 features including liquidity ratios (current ratio), profitability ratios (ROE, ROA), leverage ratios (debt-to-equity), YoY growth rates, QoQ changes, financial distress indicators, and economic interaction terms
Target: Binary classification - Layoff Event (1) vs No Layoff (0), with 10.8% positive class rate

MODEL

Algorithms Evaluated: Decision Tree, Random Forest, XGBoost, LightGBM
Best Model: XGBoost with hyperparameter tuning and threshold optimization
Optimization Strategy: GridSearchCV with F2-Score (β=2) to emphasize recall over precision
Final Parameters: max_depth=3, learning_rate=0.01, n_estimators=200, scale_pos_weight=10, subsample=0.6, reg_lambda=1
Threshold: 0.243 (optimized for maximum F2-Score)
Class Imbalance Handling: Scale_pos_weight, threshold adjustment, forward-fill imputation for temporal stability

RESULTS

Metric	Value	Interpretation
Recall	95%	Caught 96 out of 101 actual layoffs
Precision	18%	96 correct out of 520 predictions
F2-Score	0.52	Strong performance for imbalanced data
Accuracy	27%	Trade-off for maximizing recall
AUC-ROC	0.63	Good discrimination ability

Key Finding: The model successfully identifies 95% of layoffs with only 5 false negatives, making it highly effective as an early warning system. The high false positive rate (424 false alarms) is an acceptable trade-off for a system where missing a layoff is more costly than a false alarm.

LIMITATIONS

- **Data lag:** Predictions use financial data from the previous quarter (3-month lag)
- **Coverage:** Limited to publicly-traded companies with consistent SEC filings (152 companies)
- **Precision:** High false positive rate (81.5%) may cause alert fatigue
- **Temporal scope:** Model trained on 2020-2024 data; may not generalize to different economic conditions
- **Missing factors:** Does not incorporate news sentiment, social media signals, or management commentary

NEXT STEPS

Short-term: (1) Implement industry-specific thresholds to improve precision, (2) Add confidence intervals for risk scores, (3) Create quarterly monitoring dashboard
Medium-term: (1) Incorporate real-time news sentiment from GDELT/News API, (2) Add hiring freeze signals from job board data (LinkedIn, Indeed), (3) Include earnings call transcript sentiment analysis
Long-term: (1) Extend to private companies using alternative data (Crunchbase, PitchBook), (2) Build company-specific models for major tech firms, (3) Deploy as production API for investor use