Observation Summary – Layoff Risk Classification Project

EDA Observations:

- The layoff ratio is heavily right-skewed, with most companies having ratios below 0.2 while a small subset shows extreme values.
- The dataset is highly imbalanced: 81% of companies fall into the high-risk class, requiring class-weighting during modeling.
- Industry differences are significant. Sectors such as crypto, SaaS/analytics, AI/ML, and cloud computing consistently show higher layoff risk rates.
- Layoff activity displays clear temporal peaks in late 2022 and early 2023, coinciding with macroeconomic tightening and sector-wide corrections.

Model Observations:

- The Decision Tree produced the strongest balance of interpretability and performance (Accuracy 0.931, F1 0.955, ROC-AUC 0.971).
- It produced zero false positives, indicating high precision in identifying stable companies.
- The two false negatives were companies whose last major layoff occurred over 1000 days ago, which the model reasonably down weighted due to time decay.
- The Random Forest achieved perfect recall but introduced four false positives, making it more suitable when avoiding missed risky cases is more important than precision.

Prediction Observations:

- Companies with layoff ratio ≥ ~0.07 were consistently classified as high-risk, aligning with patterns identified in the EDA.
- Predictions for unseen companies followed clear industry trends: security, blockchain, cloud, and ecommerce sectors frequently appeared high-risk.
- Low layoff ratio companies (~0.02–0.03) were correctly flagged as low-risk unless they belonged to historically volatile industries.

Overall Insight:

The project demonstrates that layoff risk can be effectively predicted using engineered layoff metrics and industry information. The model's behavior aligns closely with underlying data patterns, reinforcing the interpretability and real-world usefulness of the approach.