

Fraudulent Insurance Claim Detection

Objective:

Detect fraudulent insurance claims based on incident, customer, and claim details.

Models:

- Logistic Regression (RFECV)
- Random Forest

Key Findings:

- Logistic Regression outperformed with 85.3% accuracy
- Recall (Fraud): 86.5% (vs. 14.9% with RF)
- ROC AUC: 0.84

Business Implications:

- High-recall model ensures fewer fraudulent claims are missed
- Easy-to-deploy model can be integrated with claim processing pipeline

Next Steps:

- Use logistic regression for real-time checks

- Enrich dataset with behavioral and claim timing data