FRAUDULENT CLAIM DETECTION

A MACHINE LEARNING APPROACH TO IDENTIFYING INSURANCE FRAUD

Submitted by:-

1. **Shreya Singh**
2. **Ayush Saraswat**
3. **Ashish Shrivastava**

Problem Statement

- Insurance fraud causes significant financial loss.
- Goal: Predict whether a claim is fraudulent (Y) or not (N).
- Challenge: Detect rare fraudulent events accurately.

Methodology Overview

- **Data Preparation:** Cleaning, encoding, handling missing values.
- **Model Selection:** Random Forest Classifier
- **Evaluation Metrics:** AUC-ROC, Accuracy, Precision, Recall, F1-Score
- **Visualization Tools:** Confusion Matrix, ROC Curve, Feature Importance Plot

Class Imbalance Problem

- Majority class: Non-fraudulent ('N’)
- Minority class: Fraudulent ('Y’)
- Significant imbalance affects model performance on 'Y' class

Model Performance Metrics

- **AUC-ROC Score:** 0.772
- **Accuracy:** 0.75
- **Precision (Y):** 0.40
- **Recall (Y):** 0.03
- **F1-Score (Y):** 0.05

Confusion Matrix Insights

- High number of True Negatives
- Very low True Positives
- Model struggles with identifying fraudulent claims

Feature Importance

- Top 10 features identified by Random Forest
- Useful for feature selection and model optimization
- Can be used to understand key drivers of fraud

Visualizations

- Confusion Matrix Heatmap
- ROC Curve
- Feature Importance Bar Plot

Key Insights

- Model performs well on majority class
- Poor recall and F1-score on minority class
- Current model not suitable for automated fraud flagging

Actionable Outcomes

- **Class Imbalance Handling:** Use SMOTE or class weighting
- **Algorithm Testing:** Try XGBoost or LightGBM
- **Threshold Tuning:** Lower probability threshold
- **Feature Engineering:** Use top features for new variable creation
- **Manual Review Strategy:** Flag high-risk claims for further investigation

Conclusion

- Good baseline model with potential for improvement
- Requires enhanced strategies to effectively detect fraudulent claims
- Future work: Address imbalance, optimize recall, improve business value

THANK YOU