Relatório de Análise de Anomalias – Projeto TrustShield

Data: 23 de julho de 2025

Autor: Douglas de Souza, com suporte da IA Gemini

Status: Análise Concluída

1. Resumo Executivo

Este relatório apresenta os resultados da análise exploratória e da modelagem de detecção de anomalias sobre o conjunto de dados de transações financeiras. Diante da ausência de exemplos de fraude rotulados, a estratégia foi pivotada com sucesso para uma abordagem de aprendizado não supervisionado utilizando o algoritmo **Isolation Forest**. O modelo analisou mais de 13 milhões de transações e identificou **13.306 (aproximadamente 0,1%)** como anômalas. A análise subsequente revelou um perfil claro e consistente para essas transações suspeitas, focado em três eixos principais: **valor da transação, horário de ocorrência e perfil de risco do cliente**. As descobertas aqui apresentadas fornecem uma base sólida para a implementação de regras de negócio imediatas e para o desenvolvimento de um sistema de monitoramento em tempo real.

2. Contexto e Metodologia

O objetivo inicial do projeto era treinar um modelo supervisionado para classificar transações como fraudulentas ou legítimas. Contudo, a análise revelou que o arquivo de rótulos (train_fraud_labels.json) não continha nenhum exemplo de fraude (label=1).

Para superar este desafio e ainda assim extrair valor dos dados, adotamos a técnica de **Detecção de Anomalias**. O algoritmo **Isolation Forest** foi escolhido por sua eficiência em grandes volumes de dados e sua capacidade de identificar observações atípicas sem a necessidade de rótulos. O modelo funciona isolando as transações que menos se parecem com a "norma", partindo do princípio de que anomalias são "poucas e diferentes".

3. Perfil das Anomalias Detectadas

A análise comparativa entre as transações normais e as anômalas (potenciais fraudes) revelou padrões distintos e estatisticamente significativos.

Característica 1: Valor da Transação (Amount)

As transações sinalizadas como anômalas possuem valores drasticamente superiores aos das transações normais.

- **Média do Valor Anômalo:** $ 418,73
- **Média do Valor Normal:** $ 42,60

Isso representa uma diferença de quase **10 vezes** na média. O gráfico de distribuição de valores ilustra que, enquanto as transações normais se concentram em valores baixos, as anomalias são responsáveis pela "cauda longa" da distribuição, representando os gastos mais extremos.

(Gráfico de Comparação da Distribuição de Valores)

Característica 2: Horário da Transação

O comportamento temporal das anomalias é um dos indicadores mais fortes de atividade suspeita.

- **Padrão Normal:** A maior parte das transações ocorre em horário comercial, com picos claros durante o dia.
- **Padrão Anômalo:** Apresenta picos de atividade em horários não convencionais, especificamente de madrugada (entre **2h e 5h**) e no final da noite (após as **22h**).

Essa inversão de padrão sugere que as atividades fraudulentas são realizadas em momentos de menor vigilância.

(Gráfico de Comparação da Distribuição das Horas)

Característica 3: Perfil de Risco do Cliente

O modelo também identificou um perfil de cliente específico associado às transações anômalas.

- **Score de Crédito:** Clientes envolvidos em anomalias tendem a ter um score de crédito **inferior** (média de 665 vs. 714).
- **Renda Anual:** Contra intuitivamente, esses mesmos clientes possuem uma renda anual **muito superior** (média de $123k vs. $46k).

Este perfil combinado pode indicar que contas de clientes de alta renda, porém com maior risco de crédito, são alvos preferenciais para atividades fraudulentas de alto valor.

4. Recomendações Acionáveis

Com base nos padrões identificados, recomendamos as seguintes ações:

1. **Implementação de Regras de Alerta (Curto Prazo):**
    a. Criar um sistema de alerta para o time de risco que sinalize transações que atendam a múltiplos critérios de anomalia. Exemplo de regra: SE (valor > $400 E horário entre 22h-06h) ENTÃO gerar alerta de alta prioridade.
    b. Considerar a implementação de um passo de verificação adicional (ex: 2FA) para transações que se encaixem no perfil de risco detectado.
2. **Aprovação do Modelo para Produção (Médio Prazo):**
    a. O modelo IsolationForest demonstrou sua eficácia em identificar padrões de risco relevantes. O próximo passo é seguir o plano de engenharia de software e implantar este modelo em um ambiente de produção para que ele possa pontuar transações em tempo real.
3. **Análise de Casos Específicos:**
    a. Realizar uma análise manual (deep dive) em uma amostra das 13.306 transações sinalizadas para confirmar a natureza da anomalia e, se possível, rotular alguns casos como fraude confirmada para futuras iterações do modelo.