

INTELIGÊNCIA ARTIFICIAL · UNIFOR · 12 DE NOVEMBRO DE 2025

Semantic and Structural Drift in Financial Knowledge Graphs: A Robustness Analysis of GNN-Based Fraud Detectors

Por Renner S. Menezes e Raimir Holanda Filho

Fonte: OpenAlex · Unifor · [Publicação original](#)

INOVAÇÃO	APLICABILIDADE	POTENCIAL ECONÔMICO	MATURIDADE
7/10	8/10	9/10	Média

Pesquisa da UNIFOR analisa a robustez de detectores de fraude baseados em Graph Neural Networks (GNNs) frente ao drift de dados em knowledge graphs financeiros. Estudo demonstra que arquiteturas mais simples como R-GCN são mais resilientes à degradação temporal em comparação com modelos complexos como HGT, estabelecendo benchmark crucial para modelos drift-aware no setor financeiro.

NÚCLEO DA ANÁLISE

Ideia de startup ou produto

Plataforma SaaS que monitora e corrige automaticamente o drift em detectores de fraude baseados em GNNs, combinando técnicas de detecção de mudanças no conhecimento com re-treinamento adaptativo, modelo de assinatura baseado no volume de transações monitorados.

Aplicações práticas

Sistemas de detecção de fraude adaptável para instituições financeiras, plataformas de monitoramento contínuo de qualidade de modelos, ferramentas de diagnóstico de drift para manutenção preditiva de sistemas de IA, frameworks de benchmarking para avaliação de robustez.

Potencial de mercado

Altíssimo potencial no setor financeiro global, onde perdas por fraude ultrapassam bilhões anualmente. Mercado de soluções anti-fraude em IA cresce 20-25% ao ano, com demanda crescente por sistemas mais robustos frente a mudanças comportamentais e de padrões.

FUNDAMENTO CIENTÍFICO

Problema abordado

Detectors de fraude baseados em GNNs sofrem significativa degradação de desempenho quando expostos a drift semântico e estrutural em knowledge graphs ao longo do tempo, representando risco crítico para aplicações financeiras em produção.

Metodologia

Treinamento de três modelos GNN (R-GCN, HGT e HAN) em dados históricos do IEEE-CIS Fraud Detection dataset, avaliação da degradação de desempenho em 50 janelas temporais consecutivas para comparar robustez de diferentes arquiteturas.

Principais descobertas

Todos os modelos experimentaram significativa e volátil queda de desempenho; HGT teve maior queda (F1-score de 0.747 para 0.455); R-GCN mostrou maior resistência ao drift com F1-score médio de 0.868; Existe trade-off entre desempenho de pico e estabilidade temporal nas arquiteturas GNN.

LEITURA PARA GESTÃO PÚBLICA

Políticas regulatórias exijam testes contínuos de robustez para sistemas de IA em instituições financeiras, padrões para monitoramento de drift em tempo real, incentivos para desenvolvimento de tecnologias anti-fraude adaptável, protegendo consumidores e sistema financeiro.

PARCERIA UNIVERSIDADE × EMPRESA

Parceria entre universidades (como UNIFOR), instituições financeiras e fintechs para desenvolvimento de benchmarks de robustez contra drift em detectors de fraude, criando ecossistema de pesquisa aplicada com benefícios mútuos de inovação e validação real.

OPORTUNIDADES DERIVADAS

4 direções estratégicas identificadas

STARTUP · IMPACTO ALTO · INTELIGÊNCIA ARTIFICIAL

Plataforma de Detecção Adaptativa de Fraude com Monitoramento de Drift

SaaS que monitora em tempo real o desempenho de modelos GNN para detecção de fraude, implementando correções automáticas quando drift é detectado.

PARCERIA · IMPACTO MÉDIO · SOFTWARE

Consortio de Pesquisa em Robustez de Modelos Financeiros

Parceria entre universidades, instituições financeiras e fintechs para desenvolvimento de benchmarks e soluções avançadas contra drift em sistemas de detecção de fraude.

Normatização de Testes de Robustez para IA Financeira

Políticas públicas que estabeleçam requisitos de monitoramento e teste de robustez para modelos de IA em sistemas financeiros críticos.

Framework de Avaliação de Drift para Detecção de Fraude

Solução corporativa que fornece ferramentas para avaliação e mitigação de drift em modelos existentes de detecção de fraudes, aumentando vida útil e precisão.

ABSTRACT ORIGINAL

The detection of financial fraud with Graph Neural Networks (GNNs) on Knowledge Graphs (KGs) is a state-of-the-art approach. However, the performance of these models in production is threatened by data drift, where the statistical properties of the data change over time. This study provides a foundational analysis of the robustness of different GNN architectures (convolutional and attention-based) when exposed to natural drift in a KG built from the IEEE-CIS Fraud Detection dataset (Kaggle). We trained three models, Relational Graph Convolutional Network (R-GCN), Heterogeneous Graph Transformer (HGT) and Heterogeneous Attention Network (HAN) -on a historical dataset and evaluated their performance degradation in 50 subsequent temporal windows. Our results quantify a significant and volatile performance drop over time for all models, with the F1-score of HGT falling from a baseline of 0.747 to a minimum of 0.455. In contrast, the simpler convolutional architecture of R-GCN proved to be more resistant to drift, achieving a high average F1-score of 0.868 across the monitored windows. We show that architectural choice directly affects the degradation profile, revealing a trade-off between peak performance and temporal stability. This provides strong evidence that drift analysis is a crucial step in the lifecycle of the GNN model for financial applications, establishing a vital benchmark for future work on drift-aware models.

MATÉRIA PARA LEIGOS

Para leigos: Como detectar fraudes financeiras mesmo quando o comportamento criminoso muda

O cenário atual

O combate à fraude financeira é uma batalha constante entre instituições financeiras e criminosos. Com o avanço da tecnologia, abordagens baseadas em inteligência artificial, especialmente redes neurais em grafos (GNNs), têm se tornado eficazes na identificação de padrões suspeitos. Essas redes funcionam como detetives digitais que conectam informações diferentes para encontrar relações que escapariam a análises tradicionais. No entanto, como os criminosos também se adaptam e mudam suas estratégias, os modelos de detecção precisam ser robustos para acompanhar essas mudanças.

O que os pesquisadores fizeram

Pesquisadores da Universidade de Fortaleza (UNIFOR) no Brasil analisaram como diferentes arquiteturas de redes neurais em grafos se comportam quando expostas a mudanças naturais nos dados ao longo do tempo. Eles treinaram três modelos distintos em um conjunto de dados histórico de detecção de fraudes do IEEE-CIS (disponível no Kaggle) e depois avaliaram como o desempenho desses modelos decaía em 50 períodos subsequentes. Os modelos testados foram: Rede de Convolução em Grafos Relacional (R-GCN), Transformer em Grafo Heterogêneo (HGT) e Rede de Atenção Heterogênea (HAN).

Como funciona na prática

Imagine um mapa complexo que conecta pessoas, empresas, transações e outros elementos do sistema financeiro. Esse mapa é o que os pesquisadores chamam de "grafo de conhecimento". As redes neurais em grafos funcionam como um navegador especializado que aprende a navegar por esse mapa, identificando padrões suspeitos. Por exemplo, podem detectar quando uma transação parece desproporcional para um determinado perfil de cliente ou quando uma rede de pessoas e empresas está se comunicando de forma anômala. O problema é que, ao longo do tempo, o mapa muda: novas relações surgem, antigas se tornam irrelevantes e os padrões de fraude evoluem. A pergunta dos pesquisadores era: como esses navegadores reagem quando o mapa se modifica continuamente?

Resultados e evidência

Os resultados revelaram uma tensão interessante entre desempenho máximo e estabilidade ao longo do tempo. O modelo mais complexo (HGT) teve desempenho inicial mais alto (F1-score de 0.747), mas mostrou uma queda significativa e volátil ao longo dos 50 períodos monitorados, chegando a um mínimo de 0.455 em alguns momentos. Por outro lado, o modelo mais simples, a R-GCN, não teve um desempenho inicial tão alto quanto o HGT, mas provou ser muito mais resistente às mudanças ao longo do tempo, mantendo uma média de F1-score de 0.868 nos períodos monitorados. Isso demonstra um claro trade-off entre o desempenho máximo em condições ideais e a capacidade de se manter estável quando o ambiente muda.

Implicações práticas

Para as instituições financeiras, essa pesquisa sugere que ao escolher modelos de detecção de fraude, não basta apenas considerar qual tem o melhor desempenho em testes estáticos. É crucial avaliar como esses modelos se comportam ao longo do tempo quando os padrões de fraude naturalmente evoluem. Em alguns casos, um modelo um pouco menos preciso inicialmente pode ser mais útil a longo prazo devido à sua maior resistência a mudanças. Além disso, a análise de como os modelos decaem pode fornecer um alerta precoce sobre quando é necessário re-treiná-los ou ajustá-los, ajudando a manter a segurança do sistema financeiro mesmo em cenários dinâmicos.

Limitações e próximos passos

Os pesquisadores mencionam que sua análise representa um primeiro passo fundamental na compreensão da robustez desses modelos em ambientes dinâmicos. Trabalhos futuros podem focar no desenvolvimento de modelos "aware of drift" – ou seja, modelos que possam detectar quando está ocorrendo uma mudança significativa nos dados e se adaptar automaticamente. Além disso, outros tipos de arquiteturas e abordagens de detecção de mudanças ainda precisam ser explorados. A pesquisa também sugere que o momento do re-treinamento é uma questão crucial que pode otimizar a relação entre custo computacional e desempenho do sistema ao longo do tempo.

QUEM SÃO OS PESQUISADORES

Quem são os pesquisadores

Os autores deste estudo são Renner S. Menezes e Raimir Holanda Filho, pesquisadores da Universidade de Fortaleza (UNIFOR), no Brasil. O paper não detalha informações específicas sobre suas formações, titulações ou trajetórias profissionais, limitando-se a identificar seus nomes e instituição de afiliação. A pesquisa foi realizada em colaboração e publicada através da plataforma OpenAlex com vinculação à UNIFOR.

PALAVRAS-CHAVE

Robustness (evolution) · Convolutional neural network · Graph · Concept drift · Data modeling · Knowledge graph · Detector · Baseline (sea)