UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Detecção de Fraudes no Setor de Saúde:

Métodos de Inteligência Artificial

Felipe de Mello Almeida

Trabalho de Conclusão de Curso
MBA em Inteligência Artificial e Big Data

ICMC USP
SÃO CARLOS

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Detecção de Fraudes no Setor de Saúde:

Métodos de Inteligência Artificial

USP - São Carlos

2024

Felipe de Mello Almeida

Detecção de Fraudes no Setor de Saúde:

Métodos de Inteligência Artificial

Trabalho de conclusão de curso apresentado ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial

Orientador: Prof. Dr. Marcelo Manzato

USP - São Carlos

2024

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi e
Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

Almeida, Felipe de Mello
A447d Detecção de Fraudes no Setor de Saúde: Métodos de Inteligência Artificial / Felipe de Mello Almeida; orientador Marcelo Garcia Manzato. -- São Carlos, 2024.
96 p.

Trabalho de conclusão de curso (MBA em Inteligência Artificial e Big Data) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2024.

1. Detecção de Fraudes. 2. Redes Neurais para Grafos. 3. Aprendizado de Máquina. 4. Aprendizado Profundo. 5. Medicare. I. Manzato, Marcelo Garcia , orient. II. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2: Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

A minha esposa pela compreensão, paciência e cumplicidade em todos os momentos de nossa jornada. E aos meus pais por tudo o que fizeram por mim.

AGRADECIMENTOS

Ao Prof. Dr. Marcelo Manzato, pela orientação, aprendizado, direcionamento e apoio neste projeto.

“Com tutores de IA, o potencial de cada criança, não importa quão pobre, pode ser realizado. O custo por criança seria insignificante, e essa criança poderia viver uma vida muito mais rica e produtiva”.

Russell, Stuart (2019)

RESUMO

ALMEIDA, F. M. **Detecção de Fraudes no Setor de Saúde: Métodos de Inteligência Artificial**. 2024. 69 f. Trabalho de conclusão de curso (MBA em Inteligência Artificial e Big Data) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2024.

A fraude na saúde tem um impacto significativo na economia do setor de saúde, aumentando os custos para os consumidores e diminuindo a qualidade do atendimento. Essas perdas são eventualmente repassadas aos consumidores na forma de prêmios de seguro mais altos, tornando a saúde suplementar menos acessível para muitos brasileiros. A detecção de fraudes tem sido um tópico de pesquisa ativo nas últimas décadas. Diversas técnicas têm sido exploradas para identificar e prevenir fraudes, incluindo métodos estatísticos e técnicas de detecção de anomalias. Recentemente, o foco tem se voltado para o uso de Inteligência Artificial (IA) para detecção de fraudes. A IA oferece várias vantagens sobre as técnicas tradicionais, incluindo a capacidade de processar grandes volumes de dados e a capacidade de identificar padrões complexos entre diferentes estruturas de dados. Estas estratégias de IA, quando implementadas corretamente, podem ajudar a reduzir significativamente as fraudes e desperdícios no setor de saúde, resultando em economia de custos e melhoria na qualidade do atendimento ao usuário final, sendo um tema mundial para os governos e empresas do setor. Apesar de ser um tema bem relevante, existem poucas pesquisas sobre a eficácia dos modelos de Inteligência Artificial na detecção de fraudes médicas. Apuraremos as fraudes cometidas pelos prestadores de serviços de saúde na falsificação de laudos de exames e em atendimentos falsos utilizando os dados públicos do sistema americano de saúde Medicare. Abordaremos os modelos de *Machine Learning* (*Logistic Regression, Random Forest, XGBoost*, *LightGBM* e *Multi-Layer Perceptron*) e os modelos *Graph Neural Networks* (GraphSAGE, GAT, HAN e HGT), pois são algoritmos amplamente utilizados na modelagem de detecção de fraudes em diversos setores. Desta forma, poderemos avaliar a eficácia dos modelos na detecção de fraudes no Medicare, e avaliar se essa eficácia poderá ser aprimorada com a adição de características derivadas de grafos, apresentando os resultados, análises e comparativos de performance em cada caso.

Palavras-chave: Detecção de Fraudes; Redes Neurais para Grafos; Aprendizado de Máquina; Medicare; Análise de Grafos; Aprendizado Profundo.

ABSTRACT

ALMEIDA, F. M. **Fraud Detection in Healthcare: Artificial Intelligence Methods** .2024. 69 f. Trabalho de conclusão de curso (MBA em Inteligência Artificial e Big Data) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2024.

Healthcare fraud has a significant impact on the economy of the healthcare sector, increasing costs for consumers and decreasing the quality of care. These losses are eventually passed on to consumers in the form of higher insurance premiums, making supplemental health less accessible to many Brazilians. Fraud detection has been an active research topic in recent decades. Various techniques have been explored to identify and prevent fraud, including statistical methods and anomaly detection techniques. Recently, the focus has shifted to the use of Artificial Intelligence (AI) for fraud detection. AI offers several advantages over traditional techniques, including the ability to process large volumes of data and the ability to identify complex patterns among different data structures. These AI strategies, when implemented correctly, can help significantly reduce fraud and waste in the healthcare sector, resulting in cost savings and improved quality of care for the end user, being a global theme for governments and companies in the sector. Despite being a very relevant topic, there is little research on the effectiveness of Artificial Intelligence models in detecting medical fraud. We will investigate frauds committed by healthcare service providers in the falsification of medical reports and in false services using public data from the American health system, Medicare. We will approach both Machine Learning models (Logistic Regression, Random Forest, XGBoost, LightGBM, and Multi-Layer Perceptron) and Graph Neural Networks models (GraphSAGE, GAT, HAN, and HGT), as they are algorithms widely used in fraud detection modeling across various sectors. Consequently, we will conduct a rigorous evaluation of the models' effectiveness in identifying fraudulent activities within the Medicare system. Furthermore, we will investigate whether the incorporation of graph-derived features can enhance this effectiveness. Our findings, analyses, and comparative performance metrics will be meticulously presented for each case under consideration.

Keywords: Fraud Detection; Graph Neural Networks; Machine Learning; Medicare; Graph Analysis; Deep Learning.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

AHIN	Attributed Heterogeneous Information Network
CMS	Centers for Medicare & Medicaid Services
CNN	Convolutional Neural Network
DP	Deep Learning
XGBoost	eXtreme Gradient Boosting
FBI	Federal Bureau of Investigation
FDA	Food and Drug Administration
GAT	Graph Attention Network
GCN	Graph Convolutional Network
GNN	Graph Neural Network
GraphSAGE	Graph Sample and Aggregate
HAN	Heterogeneous Graph Attention Network
HGT	Heterogeneous Graph Transformer
HGSampling	Heterogeneous Mini-Batch Graph Sampling
IESS	Instituto de Estudos de Saúde Suplementar
ICLR	International Conference on Learning Representations
IoT	Internet of Things
LGPD	Lei Geral de Proteção de Dados
ML	Machine Learning
MLP	Multi-Layer Perceptron
NLP	Natural Language Processing
OECD	Organization for Economic Co-operation and Development
PUF	Public Use Files
RL	Regressão Logística
RPA	Robotic Process Automation
RNN	Recurrent Neural Network
MCC	Mobile Cloud Computing

LISTA DE SÍMBOLOS

λ Lambda

σ Sigma

SUMÁRIO

1 INTRODUÇÃO

1.1 Contextualização

O mercado de Saúde Suplementar é um pilar fundamental para o bem-estar social, sendo utilizado como complementação e como alternativa ao Sistema Único de Saúde no Brasil. Antes de 1988, a assistência médica do Estado Brasileiro era destinada somente àqueles que contribuíam para a Previdência Social, deixando milhões de brasileiros sem acesso aos serviços básicos de saúde. Mesmo aos contribuintes, o atendimento era precário e limitado. Para preencher essa lacuna, surgiram diferentes iniciativas. Assim foi iniciado o setor de saúde suplementar, configurado por planos e seguros de saúde, que sustentam o atendimento organizado através de rede privada. A Constituição Federal de 1988 [1] que estabeleceu o direito fundamental à saúde, garantiu a criação do Sistema Único de Saúde (SUS) [2] na década de 90. Assim, os cuidados com a saúde tornaram-se mais amplos e inclusivos no Brasil, possibilitando que milhões de pessoas tivessem acesso aos serviços. Entretanto, o setor público não foi suficiente para atender à demanda completa e em constante grau de qualidade. Então, a Saúde Suplementar foi criada como pilar fundamental para o bem-estar social, sendo utilizada como complementação e, principalmente, como alternativa ao SUS. Com mais de duas décadas de estruturação regulada, o setor segue agregando valor ao bem-estar público.

Atualmente, o Setor possui 50,4 milhões de beneficiários [3] na rede (26% da população [4] brasileira) atendidos por 679 operadoras ativas. O setor demonstrou um leve crescimento na quantidade de beneficiários do período da pandemia até hoje, atingindo 50 milhões de beneficiários em sua rede, o que não ocorria desde 2014. Os gastos com saúde em 2019, somando os setores privado e público, foram de R$ 711,4 bilhões representando 9,6% do Produto Interno Bruto (PIB), podendo ultrapassar R$ 950 bilhões em 2024 [3]. A sinistralidade em 2022 foi de 89% na média, com alguns casos acima de 100%. Considerando a evolução do setor, observamos movimentos de ganho de escala, principalmente, com a verticalização, fusões e aquisições. Os planos de saúde figuram no TOP 3 bens/serviços mais desejados pela população brasileira segundo a Pesquisa Vox Populi (2021) em parceria com o Instituto de Estudos de Saúde Suplementar (IESS) [5].

O setor é concentrado em planos coletivos, com 82% dos beneficiários, principalmente através dos coletivos empresariais. E essas operadoras estão distribuídas em 5 modalidades, responsáveis por uma receita de R$ 234 bilhões em 2022 e de R$ 66 bilhões no 1º trimestre de 2023 [6].

É impreterível, porém, a atualização de práticas afim de garantir a qualidade e a sustentabilidade financeira do sistema de saúde. O setor de saúde está em momento de transformação constante e possui importantes desafios que merecem atenção. Existem muitos fatores que impulsionam os custos dos cuidados de saúde e dos seguros de saúde, incluindo fraudes, desperdícios e abusos no sistema, sendo um problema nacional e internacional [7].

A fraude na saúde tem um impacto significativo na economia do setor de saúde, aumentando os custos para os consumidores e diminuindo a qualidade do atendimento. Essas perdas são eventualmente repassadas aos consumidores na forma de prêmios de seguro mais altos, tornando a saúde suplementar menos acessível para muitos brasileiros [7].

Existem diversos estudos e análises nacionais e internacionais de estimativas de fraude e desperdícios no setor de Saúde Suplementar. O estudo do IESS de 2017 [9], considera uma estimativa em torno de R$ 28 bilhões que corresponde aos gastos das operadoras médico hospitalares no Brasil com contas hospitalares e exames consumidos indevidamente por fraudes e desperdícios com procedimentos desnecessários. A atualização das estimativas mostra que entre 12% e 18% das contas hospitalares apresentam itens indevidos e 25% a 40% dos exames laboratoriais não são necessários. Portanto, houve um gasto na saúde de aproximadamente R$ 15 bilhões com fraudes em contas hospitalares e R$ 12 bilhões em pedidos de exames laboratoriais não necessários.

Outro estudo, já a nível internacional, "*Tackling Wasteful Spending on Health*", elaborado pela *Organization for Economic Co-operation and Development* (OECD), ou em português, Organização para a Cooperação e Desenvolvimento Econômico [10], estima que as fraudes e desperdícios representa 6% das despesas totais do setor nos Estados Unidos da América (EUA). Tendo em vista a generalidade de fontes consultadas, o percentual de fraude ronda em cerca de 7% das receitas nos países desenvolvidos. O *Medicare* é um programa de saúde dos EUA estabelecido e financiado pelo governo federal americano, que oferece seguro saúde acessível para indivíduos com 65 anos ou mais e outros indivíduos selecionados com deficiências permanentes [40]. De acordo com o Relatório dos Administradores do *Medicare* de 2018 [41], o programa forneceu cobertura a 58,4 milhões de beneficiários e consumiu 710 bilhões de dólares em despesas totais. A Agência Federal de Investigação Americana, *Federal Bureau of Investigation* (FBI) estima que a fraude representa 3~10% de todas as faturas [43], e a *Coalition Against Insurance Fraud*, ou em português, a Coalizão Contra Fraude em Seguros, uma organização que reúne diversas entidades, incluindo companhias de seguros, consumidores, agências governamentais e órgãos legislativos, com o objetivo de combater a fraude em seguros, estima que a fraude custa a todas as linhas de seguros cerca de 80 bilhões

de dólares por ano [44]. Exemplos de fraude incluem cobrança por consultas que o paciente não compareceu, cobrança por serviços mais complexos do que os realizados ou cobrança por serviços não prestados. A prática abusiva é inconsistente com a prestação de serviços médicos necessários aos pacientes de acordo com padrões reconhecidos, por exemplo, cobrança por serviços médicos desnecessários ou uso indevido de códigos de serviços não prestados para ganho pessoal.

Apesar do tema de fraudes e desperdícios ser uma preocupação para o setor de Saúde, nos últimos anos os governos e empresas vem buscando soluções em tecnologia para mitigar o problema. Apesar de algumas operadoras de saúde possuírem infraestrutura e processos robustos que facilitam a captura e o armazenamento de dados, na maioria dos casos, não há uma gestão eficaz, tratamento, sistemas e análise desses dados para evitar o problema de fraudes e desperdícios [7].

1.2 Inovação e Proposta de Solução do Problema

Um dos grandes desafios do setor, é garantir a privacidade e a proteção dos dados de saúde dos pacientes, e a sua correta utilização. Os dados de saúde devem ser coletados, armazenados e tratados de forma segura e de acordo com as regulamentações e padrões de privacidade aplicáveis à Lei Geral de Proteção de Dados (LGPD) [11].

A detecção de fraudes na saúde tem sido um tópico de pesquisa ativo nas últimas décadas. Diversas técnicas têm sido exploradas para identificar e prevenir fraudes na saúde, incluindo métodos estatísticos e técnicas de detecção de anomalias [12]. No contexto da saúde, várias técnicas têm sido discutidas para a detecção de fraudes em seguros de saúde, destacando a importância das técnicas de mineração de dados para identificar padrões de fraude [13] [14].

Recentemente, o foco tem se voltado para o uso de Inteligência Artificial (IA) para detecção de fraudes. A IA oferece várias vantagens sobre as técnicas tradicionais, incluindo a capacidade de processar grandes volumes de dados e identificar padrões complexos [15] [16] [17].

As aplicações de Inteligência Artificial têm contribuído na verificação da veracidade de milhões de sinistros, além de identificar, analisar e corrigir problemas de codificação incorreta [18]. Uma das aplicações é possibilidade da criação de scores de fraudes possibilitando a identificação de perfis com maior ou menor potencial de fraude. O agrupamento por características dos beneficiários é uma outra possibilidade que permitiria realizar correlações com a quantidade de consultas por período, tipos de procedimentos, dentre outras análises

possíveis, com o objetivo de definir planos de ação para investigar eventuais desvios em relação ao padrão do grupo.

A IA também pode ser aplicada também na automação de detecção de fraudes por meio do uso de técnicas de *Machine Learning*. Nesse âmbito, os modelos de *Machine Learning* têm contribuído para a detecção de padrões de fraudes, utilizando modelos de regressão em uma base de dados de larga escala de despesas médicas para investigar as características de sinistros fraudulentos[19]. O algoritmo aplicado pelos pesquisadores mostrou que, em comparação com os padrões de sinistros normais, os fraudulentos ocorreram com maior frequência nos registros de despesas hospitalares. Aplicações desse tipo estão sendo testadas em vários países, principalmente em sistemas privados. Essas estratégias, quando implementadas corretamente, podem ajudar a reduzir significativamente as fraudes e desperdícios no setor de saúde, resultando em economia de custos e melhoria na qualidade do atendimento ao usuário final, sendo um tema mundial para os governos e empresas do setor.

Em fraudes no setor de saúde, é comum que um diagnóstico falso seja o resultado de uma colaboração entre os prestadores de serviços, médicos e pacientes. Esse falso diagnóstico é usado para obter ilegalmente benefícios do seguro de saúde ou reembolsos que são posteriormente distribuídos entre os fraudadores [3][7][9].

Nosso trabalho apresenta métodos de Inteligência Artificial para a detecção de fraudes na área da saúde. Apuraremos fraudes cometidas pelos prestadores de serviços de saúde na falsificação de laudos de exames e em atendimentos falsos utilizando os dados públicos do sistema americano de saúde Medicare [112]. Abordaremos os modelos de *Machine Learning* (*Logistic Regression, Random Forest, XGBoost, LightGBM* e *Multi-Layer Perceptron*) e os modelos *Graph Neural Networks* (GraphSAGE, GAT, HAN e HGT), pois são algoritmos amplamente utilizados na modelagem de detecção de fraudes em diversos setores [117] e [118]. Desta forma, poderemos avaliar a eficácia dos modelos na detecção de fraudes no Medicare, e se essa eficácia poderá ser aprimorada com a adição de características derivadas de grafos. Essa análise é importante para entendermos como novas técnicas podem contribuir para aprimorar as estratégias existentes de detecção de fraudes, apresentando os resultados e comparativos de performance e análises em cada caso.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Conceitos: Inteligência Artificial, Algoritmos e *Machine Learning*

A inteligência artificial (IA) é um campo de estudo desde a década de 1940 no século passado, quando vários pesquisadores passaram a se dedicar às chamadas máquinas inteligentes, sendo o artigo de McCulloch e Pitts sobre neurônios artificiais, de 1943, reconhecido como o primeiro trabalho nesse campo da Inteligência Artificial, ao mesclar conhecimentos de fisiologia básica e do funcionamento dos neurônios no cérebro humano, análise formal da lógica proposicional e a teoria da computação de Alan Turing [20] [21].

Do ponto de vista da Ciência da Computação, não há uma definição única e consensual acerca do que é IA, uma vez que as definições existentes dependem do objetivo que se pretende alcançar. Russel e Norvig (2010) [20] apontam quatro diferentes focos para a IA que foram propostos e desenvolvidos ao longo das décadas: pensar como humanos, pensar racionalmente, agir como humanos e agir racionalmente. Assim, a depender do foco que é selecionado, haverá diferentes caminhos possíveis para desenvolver e conceituar IA.

Apesar de um começo promissor nas décadas de 1950 e 1960, a IA logo caiu em descrédito ao não ser capaz de entregar os avanços prometidos [20]. Todavia, com o avanço computacional do final do século XX, tornou-se possível processar um volume maior de dados, muitas vezes em um intervalo de tempo da ordem de minutos ou segundos. Por sua vez, também houve um aumento da quantidade de dados disponíveis para utilização em modelos estatísticos, em um fenômeno conhecido como *Big Data* [22].

Essas condições criaram um cenário favorável para o desenvolvimento de uma nova geração de algoritmos e sistemas computacionais para implementar ferramentas de IA mais poderosas, mais baratas, mais eficazes e de fácil utilização. Assim, difundiu-se o uso de IA em cada vez mais campos do conhecimento e setores da economia, como educação, mercado financeiro, indústria automobilística, segurança pública, saúde e entretenimento [20] [23].

A inteligência artificial (IA) é a capacidade que uma máquina tem para reproduzir competências semelhantes às humanas, como a aprendizagem, planejamento e criatividade [25]. A IA utiliza alguns tipos específicos de algoritmos para aprender a reconhecer padrões e fazer previsões estatísticas de resultados. As ferramentas da IA podem ser construídas com base em diversos métodos [21]. Um dos métodos mais utilizados atualmente é o *Machine Learning* (ML), ou em português, aprendizado de máquina, que permite ao sistema

computacional aprender por si próprio e fazer previsões a partir de um de grupo finito de dados históricos, ou inputs, iniciais [24]. Usa-se um algoritmo, que é tem a capacidade de melhorar seu desempenho com o tempo, treinando-se por meio de métodos de análise de dados e modelagem analítica [21]. Há ainda o *Deep Learning* (DP), ou em português, aprendizado profundo, uma evolução do *Machine Learning* que "consiste em múltiplas camadas em cascata, modeladas a partir do sistema nervoso humano", em um arranjo conhecido como rede neural [21]. O *Natural Language Processing* (NLP), ou em português, Processamento de Linguagem Natural, também é um subcampo interdisciplinar da ciência da computação e da linguística, com foco em dar aos computadores a capacidade de entender e manipular a linguagem humana. O NLP combina linguística computacional, *Machine Learning* e modelos de *Deep Learning* para processar a linguagem humana em dados de texto ou voz e extrair seu significado completo [26] [27]. Um breve resumo pode ser encontrada na Figura 1:

Figura 1 – Principais Tipos de Técnicas Usadas em Sistemas de Inteligência Artificial

Fonte: Adaptado Osaki (2018) [26].

Na área da saúde, a IA foi inicialmente introduzida em 1976, quando um algoritmo de computador foi usado para determinar as razões da dor abdominal intensa [28]. Desde então, essa tecnologia tem sido aplicada em diversos tipos de implementações no setor de saúde. Isso está relacionado ao grande processo de digitalização que o setor tem passado nos últimos anos, que resulta em um grande volume de dados em toda sua cadeia de valor, tornando acessível o

chamado *Big Data*, desde a área de pesquisa em saúde, passando pela indústria de equipamentos médicos até os planos de saúde [29].

A maior disponibilidade de dados tem permitido ao setor de saúde a conversão dessas informações em insights significativos para a tomada de decisões e para a melhor utilização de recursos, o aprimoramento de serviços e procedimentos, e, mais recentemente, a criação e o uso de soluções analíticas, não apenas da análise de dados históricos (análise descritiva), mas também para prever resultados futuros (análise preditiva com uso de IA) [29].

2.2 Áreas da Saúde que a IA tem sido Utilizada

Os sistemas alimentados por IA na área da saúde podem realizar de forma autônoma ou semiautônoma uma ampla variedade de tarefas desde diagnósticos médicos, tratamentos, automonitoramentos e coaching de pacientes, análises de imagens médicas, detecção de interações medicamentosas, identificação de pacientes de alto risco, codificação de anotações médicas, detecção de irregularidades, dentre outros [30] [31]. No infográfico abaixo podemos ver algumas áreas que a IA têm sido mais frequentemente aplicada na área de Saúde, mas não é exaustiva, pois é uma área dinâmica e as pesquisas em IA na área de saúde estão em pleno desenvolvimento.

Figura 2 – A Inteligência Artificial no Setor de Saúde: Conceitos e Aplicações

Fonte: Adaptado IESS - Texto para Discussão nº 99 (2023) [39].

2.2.1 Área preditiva na saúde

Uma aplicação que tem sido bastante utilizada experimentalmente na saúde é a modelagem preditiva em hospitais, que avançou de como evoluirá o estado de um paciente, as taxas de internações, as taxas de mortalidade e também na área de cuidados paliativos. O uso de predições em saúde tem contribuído para o aperfeiçoamento de documentações, já que demandam dados em quantidade e qualidade, e no reconhecimento e detecção de doenças, para melhoria da movimentação de pacientes dentro da jornada do paciente e no gerenciamento da assistência a doentes crônicos [32]. Esse avanço tem sido possível em parte devido ao desenvolvimento concomitante da computação em nuvem, *Big Data* e IoT: *Internet of Things* , ou em português, Internet das Coisas. Essas tecnologias em conjunto permitem agregar grandes quantidades de dados de várias fontes, fornecer computação em larga escala e armazenamento de recursos, além de permitir a conexão e a troca dados por meio de dispositivos com sensores, softwares e outras tecnologias remotas. Sistemas vestíveis e inteligentes são exemplos comuns de dispositivos que produzem fluxos de dados constantes que podem ser aplicados para uma compreensão mais profunda da saúde e estilo de vida de um paciente [32].

2.2.2 Monitoramento remoto

A combinação de sensores de captura de dados médicos e inteligência artificial (IA) tem gerado uma ampla gama de interesses, pois os sensores convertem os parâmetros biomédicos em sinais facilmente mensuráveis. O monitoramento por dispositivos vestíveis cria oportunidades para a telemedicina e para o monitoramento contínuo, pois facilita a continuidade do cuidado e ainda é minimamente invasivo, além de permitir o acompanhamento de mais de um indicador ao mesmo tempo. No monitoramento remoto do paciente comumente se medem sinais vitais ou outros parâmetros fisiológicos, como movimento, que podem auxiliar em julgamentos clínicos ou no planejamento de tratamento de algumas condições de saúde [18].

Dispositivos vestíveis como *smartwatches* rastreiam continuamente os sinais vitais das pessoas e são exemplos de ferramentas utilizadas para capturar os dados no monitoramento remoto. Um sistema foi desenvolvido por Bekiri et al. (2020) para monitorar o estado de saúde dos indivíduos usando *smartwatches* conectados. Os *smartwatches* coletaram os sinais vitais dos pacientes e os enviam ao sistema de IA que possui um modelo de decisão. Os resultados do estado dos pacientes foram informados aos médicos. O modelo de *Machine Learning* alcançou

uma precisão de 90% em pacientes acometidos por doenças cardiovasculares, que são acompanhados por este monitoramento [18].

2.2.3 Detecção de Doenças

A inteligência artificial tem sido testada amplamente na área de detecção de doenças. A detecção por meio de processamento e análise de imagens médicas, por exemplo, usa o modelo de IA de redes neurais e vários algoritmos de processamento de imagens já são prontamente acessíveis, tais como: *Le-NET3*, *AlexNet4*, *VGGNet5*, *GoogLeNet* e *ResNet*. Por exemplo, o *Google* desenvolve uma grande pesquisa sobre detecção de doenças nos olhos usando fotografias externas [28]. Ainda no âmbito oftalmológico, um sistema de IA autônoma é usado nos Estados Unidos no atendimento primário para o exame de retina diabética. Esse sistema é capaz de diagnosticar retinopatia diabética e edema macular diabético sem supervisão humana [34]. Em vez de um paciente ser encaminhado a um oftalmologista para o exame de retina diabética, o sistema autônomo de IA, que inclui uma câmera robótica de fundo de olho, faz um diagnóstico em tempo real e emite um relatório de diagnóstico individualizado sem a necessidade de revisão por um ser humano. Somente se o resultado for anormal, o paciente é encaminhado para um oftalmologista. Esse sistema recebeu autorização do *Food and Drug Administration* (FDA), ou em português, Administração de Alimentos e Medicamentos, que é uma agência do governo dos Estados Unidos responsável pela proteção e promoção da saúde pública através do controle e supervisão da segurança alimentar, produtos de tabaco, suplementos dietéticos, medicamentos de prescrição e venda livre (medicamentos), vacinas, biofarmacêuticos, transfusões de sangue, dispositivos médicos, radiação eletromagnética, entre outros, em abril de 2018, após extensos testes clínicos, e desde então tem sido amplamente utilizado. Há estudos que estimam que em muitas áreas que usam imagens médicas para diagnósticos, o uso de métodos assistidos por IA aumentaria significativamente a eficiência do fluxo de trabalho, pois há o potencial de processar mais de 250 milhões de imagens por dia [31]. Outros experimentos apontam benefícios relacionados a custos ao se utilizar IA no diagnóstico. Um estudo clínico [28] comparou um método de diagnóstico assistido por IA e um método tradicional para distinguir entre pólipos colorretais (tumores não cancerosos) e tecidos normais em alguns países. Este estudo mostrou que a IA pode ajudar a reduzir o custo anual bruto nesse tipo de diagnóstico em até 18,9% no Japão, 10,9% nos Estados Unidos, 7,6% na Noruega e 6,9% na Inglaterra [28].

2.2.4 Saúde Odontológica

Na odontologia, diferentes tipos de modelos de IA [35] estão sendo utilizados. Um dos modelos mais desenvolvidos tem sido o de redes neurais, principalmente, para analisar imagens odontológicas. Em periodontologia, há exemplos de aplicação de redes neurais para detectar perda óssea periodontal. Além disso, redes neurais também têm sido utilizadas para a detecção de gengivite pertencente ao tratamento ortodôntico em andamento e aplicações em cirurgia oral para a detecção de diferentes tipos de anomalias, cistos e tumores [36] [37].

2.2.5 Área Administrativa

O gerenciamento administrativo no setor de saúde costuma ser complexo, mas a literatura tem indicado que o uso de IA tem um grande potencial de melhorar a eficiência nessa área. Nos EUA, por exemplo, estima-se que uma enfermeira gasta em média 25% de seu tempo em tarefas administrativas [38]. Um exemplo de tecnologia de IA aplicada para endereçar esse problema é a *Robotic Process Automation* (RPA), ou em português, a Automação Robótica de Processos, que consiste na implementação de softwares robôs para tornar os processos de rotina, como registros clínicos, históricos de pacientes e processamento de cadastro de fornecedores, mais eficientes. Estima-se que aproximadamente 60% das tarefas manuais do setor podem ser automatizadas com o auxílio de RPA. Além do RPA, aplicativos baseados em NLP, como *chatbots*, têm sido empregados em tarefas como preencher uma receita ou acompanhar um cronograma [38].

2.2.6 Detecção de Fraudes e Abusos

No âmbito de operadoras de planos e seguros de saúde, o *Machine Learning* é uma tecnologia de IA bastante associada ao gerenciamento de sinistros e pagamentos, pois pode ser usado para combinar dados de diferentes lugares. As aplicações de IA têm contribuído na verificação da veracidade de milhões de sinistros, além de identificar, analisar e corrigir problemas de codificação incorreta [33].

A IA tem sido aplicada na automação da detecção de fraudes por meio do uso de técnicas de *Machine Learning*. A fraude em seguros de saúde influencia não apenas as organizações que a sofrem, mas acaba por sobrecarregar os sistemas de saúde, pois gera aumentos nos custos e ameaçam a base econômica do setor, impactando os indivíduos beneficiários em termos de

aumento de prêmios dos seguros de saúde e prejudicando a qualidade da assistência [19]. Nesse âmbito, os modelos de *Machine Learning* têm contribuído para a detecção de padrões de fraudes. Li et al (2022) [19], por exemplo, utilizaram regressão logística binária em uma base de dados de larga escala de despesas médicas para investigar as características de sinistros fraudulentos. O algoritmo aplicado pelos pesquisadores mostrou que, em comparação com os padrões de sinistros normais, os fraudulentos ocorreram com maior frequência nos registros de despesas hospitalares. Aplicações desse tipo estão sendo testadas em vários países.

Um ranking consolidando as principais modalidades de fraudes de acordo com seu impacto e frequência nas operações dos planos de saúde foi elaborado com base no estudo, realizado pela consultoria Ernst & Young, em 2023, sobre Fraudes e Desperdícios na Saúde Suplementar no Brasil a pedido da IESS [7].

Tabela 1 – Impacto x Frequência de Fraudes e Desperdícios em Saúde Suplementar

Tabela de impacto e frequência de fraudes e Desperdícios

Agente	Descrição	Impacto	Frequência
Diversos	Desperdícios	Alto	Alta
Prestador	Adulteração de procedimento	Alto	Alta
Prestador	Superutilização	Alto	Alta
Beneficiário	Ocultação de condição pré-existente	Alto	Alta
Beneficiário	Fornecimento de dados de acesso a terceiros	Alto	Alta
Prestador	Reembolso sem desembolso	Alto	Alta
Prestador	Falsificação de laudos de exames	Alto	Alta
Prestador	Atendimento falso	Alto	Alta
Prestador	Fracionamento de recibo	Alto	Alta
Beneficiário	Reembolso duplicado	Alto	Alta
Outros	Boleto falso	Alto	Alta
Beneficiário	Empréstimo da carteirinha para não-beneficiário	Alto	Média
Prestador	Recebimento de propina	Alto	Média
Outros	Judicialização premeditada	Alto	Baixa
Prestador	Enquadramento forçado em critérios de cirurgias	Médio	Alta
Beneficiário	Falsificação de informações/documentos	Médio	Média
Prestador	Falso atendimento público	Médio	Baixa
Prestador	Empréstimo de matrícula	Médio	Baixa
Outros	Falso coletivo	Médio	Baixa
Outros	Estipulante fantasia	Médio	Baixa
Outros	Adulteração de cadastro de beneficiário	Baixo	Baixa
Outros	Prestador fantasma	Baixo	Baixa
Outros	Falso médico ou dentista	Baixo	Baixa

Impacto	Alto - Acima de 10 milhões	Médio - Entre 5 e 10 milhões	Baixo - Abaixo de 5 milhões
Frequência	Alta - Ocorre mensalmente	Média - Ocorre semestralmente	Baixa - Raramente ocorre

Fonte: IESS (2023)[7].

Umas das formas mais efetivas para melhorar o custo, a eficiência e a qualidade dos serviços de Saúde é reduzir a quantidade de fraudes, desperdícios e abusos do Setor. Auditar e investigar manualmente todos os dados de sinistros em busca de fraude é muito demorado, oneroso e ineficiente quando comparado às abordagens de *Machine Learning* e mineração de dados [47].

No Brasil, a LGPD trouxe um grande avanço em relação ao direto de proteção os dados pessoais dos cidadãos brasileiros, incluindo dados de saúde, estabelecendo regras para a coleta, armazenamento, uso e compartilhamento de dados pessoais, com o objetivo de garantir a privacidade e a segurança dos indivíduos. Mas em certos casos, a dificuldade de acesso dos pesquisadores aos dados de saúde, públicos e privados, poderá acarretar atrasos no ritmo das pesquisas, análises e no desenvolvimento de novos tratamentos [45].

Um caminho encontrado para avançar nas análises, estudos e pesquisas sobre o tema, foi buscar os dados de arquivos de uso público do setor. Nos EUA, o *Centers for Medicare & Medicaid Services* (CMS), ou em português, Centros de Serviços Medicare & *Medicaid*, é uma agência federal que fornece cobertura de saúde para mais de 160 milhões de pessoas através do *Medicare*, e do *Medicaid*, que é um programa de seguro de saúde infantil e do mercado de seguros de saúde. O CMS lançou o *Public Use Files* (PUF), ou em português, Arquivos de Uso Público, em 2014, consistindo em diversos conjuntos de dados que são disponibilizados ao público para análises e pesquisas. Esses arquivos são geralmente anonimizados para proteger a privacidade dos indivíduos representados nos dados [48]. Desde então, foram realizados vários estudos relacionados às anomalias no sistema americano *Medicare* e às detecções de fraudes na Saúde.

Em 2017, 86% dos médicos americanos que trabalharam em consultórios e mais de 96% dos hospitais americanos relatados adotaram os sistemas de registro eletrônico de saúde de acordo com a Lei Americana de Tecnologia da Informação em Saúde para Saúde Econômica e Clínica de 2009 e de acordo com o Plano Estratégico Federal Americano de TI em Saúde [48 ,49]. Esta explosão de dados relacionados à saúde incentivou a utilização da mineração de dados e da aprendizagem automática para detectar padrões e fazer previsões. O CSM juntou-se a este esforço orientado por dados, disponibilizando publicamente os conjuntos de dados do *Medicare* & *Medicaid*, afirmando que as diversas intenções de abusar dos programas federais de cuidados de saúde custaram aos contribuintes bilhões de dólares e arriscam o bem-estar dos beneficiários [46].

A detecção de fraudes usando dados do CSM ainda apresenta vários desafios. O problema é caracterizado pelos 4Vs do *Big Data*: volume, variedade, velocidade e veracidade

[50,51]. Os milhões de registros lançados pelo CSM a cada ano satisfazem tanto o alto volume quanto a velocidade. A variedade surge dos recursos de alta dimensão do tipo misto e da combinação de múltiplas fontes de dados. Esses conjuntos de dados também apresentam veracidade ou confiabilidade, pois são fornecidos por recursos governamentais confiáveis, com controles de qualidade transparentes e documentação detalhada [52,53]. O processamento de Big Data muitas vezes excede as capacidades dos sistemas tradicionais e exige arquiteturas especializadas ou sistemas distribuídos [54]. Outro desafio é que a classe de interesse positiva representa apenas 0,03% de todos os registros, criando uma distribuição gravemente desequilibrada. Aprender com essas distribuições pode ser muito difícil, e algoritmos de *Machine Learning* padrão normalmente superestimam a classe majoritária [55].

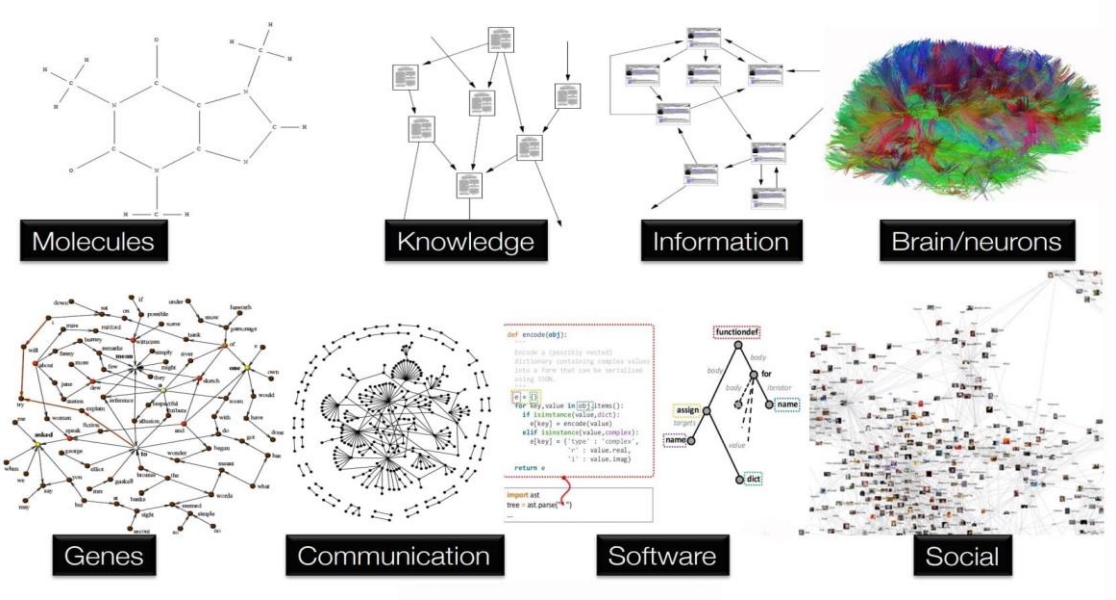
2.3 Panorama Geral - *Graph Neural Network* (GNN)

2.3.1 Conceito

O *Deep Learning* é uma importante área de pesquisa que desempenha um papel crítico no futuro da modelagem de *Big Data* de classes desequilibradas. Nos últimos 10 anos, os métodos de *Deep Learning* cresceram em popularidade à medida que melhoraram o que há de mais moderno em reconhecimento de fala, visão computacional e outros domínios [56]. Seu sucesso recente pode ser atribuído a uma maior disponibilidade de dados, melhorias em *hardware* e *software* [57,58,59,60,61] e vários avanços nos algoritmos que aceleraram o treinamento e melhoraram a generalização para novos dados [62].

Com o avanço recente dos algoritmos de *Deep Learning* , as *Graph Neural Networks* (GNNs), ou em português, Redes Neurais para Grafos, têm sido objeto de pesquisas significativas. Esses algoritmos são capazes de aprender a partir de conjuntos de dados estruturados em Grafos, utilizando a inteligência artificial. As GNNs constituem uma classe especial de algoritmos de *Machine Learning* que se concentram na análise de dados estruturados em grafos. Elas empregam o poder preditivo do *Deep Learning* em estruturas de dados complexas, que representam objetos e suas relações como pontos interconectados por linhas em um Grafo. Essa abordagem permite que as GNNs capturem e processem informações de maneira mais eficaz quando os dados são intrinsecamente Grafos, abrindo novas possibilidades para a análise de redes sociais, sistemas de recomendação, detecção de fraudes, dentre outros [63].

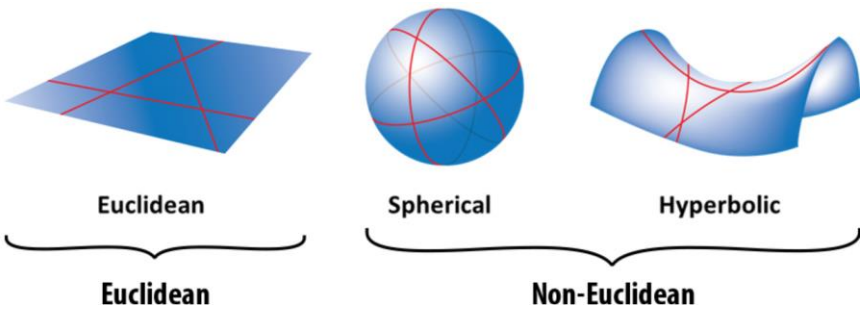
Figura 3 – Campos da Ciência e da Indústria que Podem ser Expressos como Grafos



Fonte: NVIDIA (2022) [63].

Os dados, em geral, podem ser categorizados em duas classes principais com base na geometria do espaço em que são representados: euclidiana e não-euclidiana. A geometria euclidiana, nomeada em homenagem ao matemático Euclides, é a geometria do espaço plano. Ela é caracterizada por cinco axiomas e cinco postulados, e qualquer geometria que satisfaça todos esses princípios é considerada euclidiana [65]. Dados como imagens, sons, vídeos e texto são exemplos de dados que podem ser representados em um espaço euclidiano. Esses dados podem ser representados por sequências não esparsas de valores em uma dimensão (vetores) ou duas dimensões (matrizes).

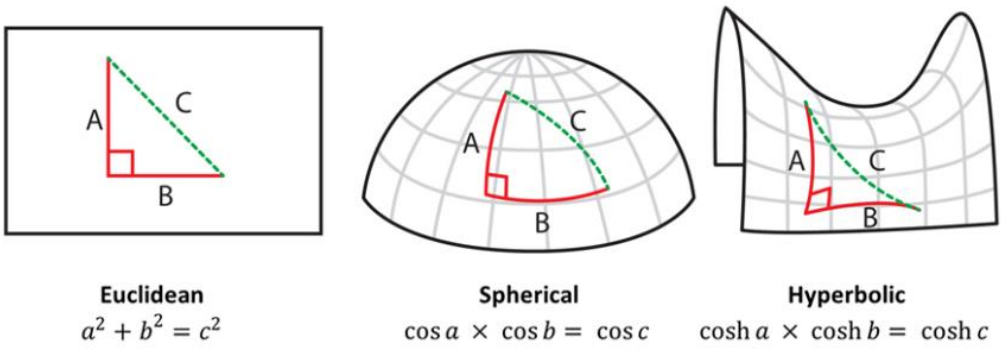
Figura 4 – Espaço Euclidiano vs Não-Euclidiano



Fonte: Boyer (2012) [65].

Por outro lado, a geometria não-euclidiana é a matemática moderna das superfícies curvas. Desenvolvida no século XIX, ela forçou os matemáticos a entenderem que as superfícies curvas possuem regras e propriedades geométricas completamente diferentes das superfícies planas [65]. Grafos, estruturas hierárquicas e objetos tridimensionais são exemplos de tipos de dados não euclidianos.

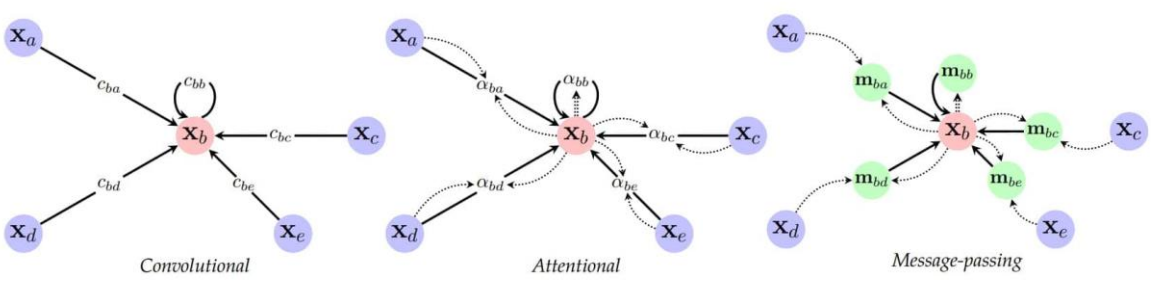
Figura 5 - Espaço Euclidiano vs Não-Euclidiano em Perspectiva Geométrica



Fonte: Boyer (2012) [65].

O *Deep Learning* se concentrou principalmente em imagens e textos, tipos de dados estruturados que podem ser descritos como sequências de palavras ou grades de pixels. Os Grafos, por outro lado, não são estruturados. Eles podem tomar qualquer forma ou tamanho e conter qualquer tipo de dados, incluindo imagens e textos. Usando um processo chamado "transmissão de mensagens", as GNNs organizam Grafos para que algoritmos de *Machine Learning* possam usá-los. A transmissão de mensagens incorpora informações em cada nó sobre seus vizinhos. Os modelos de IA empregam as informações incorporadas para encontrar padrões e fazer previsões [63].

Figura 6 – Exemplo de Fluxos de Dados em 3 Tipos de (GNNs)



Fonte: NVIDIA (2022) [63].

As *Graph Neural Networks* aplicam o poder do *Deep Learning* a grafos, permitindo a análise e a previsão de dados complexos e não euclidianos de maneiras que não seriam possíveis com técnicas de *Machine Learning* tradicionais. A compreensão das diferenças entre os dados em espaços euclidianos e não-euclidianos é fundamental para a escolha das técnicas de análise de dados mais apropriadas e eficazes. Esta compreensão pode levar a *insights* mais profundos e a melhores resultados preditivos em uma ampla gama de aplicações.

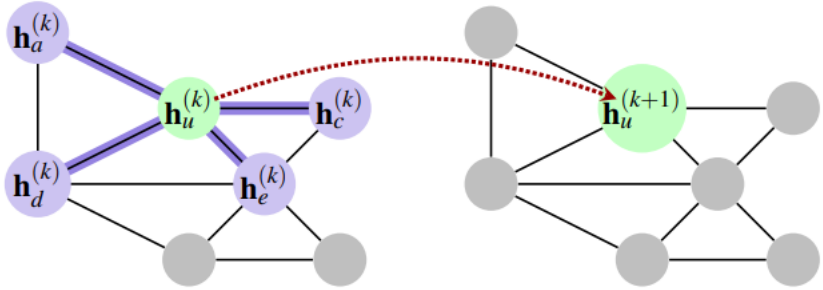
2.3.2 *Graph Neural Network* (GNN)

Os primeiros trabalhos que usaram o termo *Graph Neural Networks* (GNNs) e introduziram o aprendizado via redes neurais em grafos foram abordados por Gori, Monfardini e Scarselli (2005) e Scarselli et al (2009) [66] [67]. Uma maneira de definir modelos de GNN é utilizando a ideia de *message passing* entre os vértices. Estes vértices possuem representações vetoriais e essas representações são utilizadas para trocar mensagens entre os vértices e suas conexões. O modelo que recebe um grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ e representações vetoriais $\mathbf{h}_u^{(0)}$ iniciais para todo $u \in \mathcal{V}$ e aprende representações vetoriais para os vértices de tal forma a performar alguma específica aplicação como predição de links [68] ou classificação de nós [69]. Durante cada iteração em um modelo de GNN temos uma representação vetorial $\mathbf{h}_u^{(k)}$ para cada vértice e esta representação é atualizada de acordo com alguma regra de agregação sobre os vértices vizinhos. Neste caso o índice *k* é usado para denotar o número da camada escondida do modelo. De forma geral a regra de atualização das representações pode ser escrita da seguinte maneira [70]:

$$\mathbf{h}_u^{(k+1)} = UPDATE^{(k)}\left(\mathbf{h}_u^{(k)}, \mathbf{m}_{\mathcal{N}(u)}^{(k)}\right) \tag{3.1}$$

Onde $\mathbf{m}_{\mathcal{N}(u)}^{(k)}$ é a mensagem agregada da vizinhança do vértice *u* e *UPDATE* é uma função diferenciável arbitrária.

Figura 7 – Representação Esquemática da Atualização da Representação entre Camadas



Fonte: Lima de Carvalho (2022) [71]

A Figura 7 representa visualmente como funciona o método. O vértice *u* cuja representação vetorial na camada *k* é $\mathbf{h}_u^{(k)}$ recebe uma mensagem dos seus vizinhos *a*,*c*,*d* e *e* e com a sua própria representação é feita uma operação para gerar a nova representação vetorial $\mathbf{h}_u^{(k+1)}$.

No caso específico do modelo GNN de Scarselli et al (2009) [67], pode-se definir a operação de *message passing* da seguinte maneira simplificada:

$$\mathbf{h}_u^{(k+1)} = \sigma\left(\mathbf{w}_{self}^{(k+1)}\mathbf{h}_u^{(k)} + \mathbf{w}_{neigh}^{(k+1)} \sum_{v\in\mathcal{N}(u)} \mathbf{h}_v^{(k)} + \mathbf{b}^{(k+1)}\right) \tag{3.2}$$

Em que $\mathbf{w}_{self}^{(k+1)}, \mathbf{w}_{neigh}^{(k+1)} \in \mathbb{R}^{d^{(k+1)}\times d^{(k)}}$ são matrizes de parâmetros treináveis e $\mathbf{b}^{(k+1)} \in \mathbb{R}^{d^{(k+1)}}, \sigma(.)$ é uma função de ativação que atua elemento a elemento. Os inteiros $d^{(k)}$ e $d^{(k+1)}$ são utilizados para denotarem a dimensão das representações vetoriais para os vértices das camadas *k* e *k +1*. Portanto, entre as camadas é possível alterar a dimensão das representações vetoriais para os nós dependendo da escolha específica das dimensões dos parâmetros treináveis.

Note que no modelo de atualização das representações descrito na equação (3.2) a mensagem agregada da vizinhança do vértice *u* pode ser escrita como:

$$\mathbf{m}_{\mathcal{N}(u)}^{(k)} = \sum_{v\in\mathcal{N}(u)} \mathbf{h}_v^{(k)} \tag{3.3}$$

Portanto, se tivermos um vértice que tem muitos vizinhos a mensagem agregada descrita na equação (3.3) pode resultar em vetores cujos elementos possuem valor alto. Este fato pode

resultar em dificuldade no processo de otimização dos parâmetros treináveis e também prejudica a performance de maneira geral.

Uma característica interessante dos modelos de GNN em geral é que os parâmetros treináveis utilizados em cada camada são compartilhados. Ou seja, para atualizar as representação de todos os nós em uma camada específica *k* teremos os mesmos parâmetros $\mathbf{w}_{self}^{(k+1)}, \mathbf{w}_{neigh}^{(k+1)}$ e $\mathbf{b}^{(k)}$, $\forall\ v \in \mathcal{V}$. Sendo assim, é possível escrever a equação (3.2) na forma matricial para a atualização conjunta de todas as representações vetoriais para os vértices, omitindo o termo $\mathbf{b}^{(k+1)}$ por simplicidade:

$$\mathbf{H}^{(k+1)} = \sigma\Big(\mathbf{H}^{(k)}\mathbf{W}_{self}^{(k+1)} + \mathbf{A}\mathbf{H}^{(k)}\mathbf{W}_{neigh}^{(k+1)}\ \Big) \tag{3.4}$$

A equação (3.4) realiza a atualização das representações vetoriais para os vértices na camada *k* para a camada *k +1* simultaneamente.

2.3.3 *Graph Convolutional Network* (GCN)

O modelo de *Graph Convolutional Network* (GCN), ou em português, Rede Neural Convolucional, foi introduzido por Kipf e Welling (2017) [69]. Este modelo é similar ao de Scarselli et al (2009) [67] com a diferença de utilizar uma normalização simétrica da operação de agregação e utilizar apenas uma matriz de parâmetros treináveis em cada camada. Pode-se definir a regra de atualização da seguinte maneira:

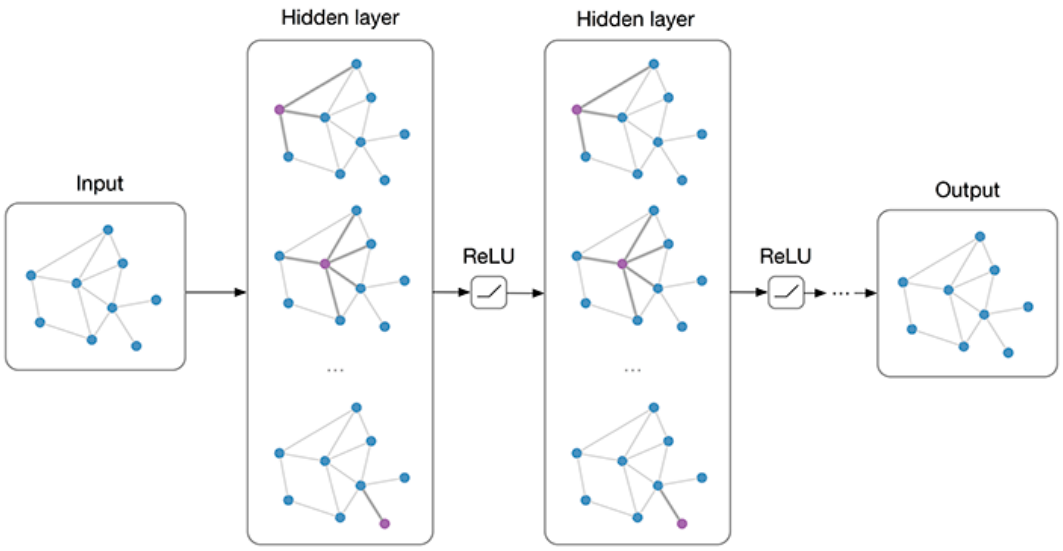
$$\mathbf{h}_u^{(k+1)} = \sigma\Big(\mathbf{W}^{(k+1)}\ \textstyle\sum_{v\in\mathcal{N}(u)\cup\{u\}} \mathbf{h}_v^{(k)} / \sqrt{|\mathcal{N}(u)||\mathcal{N}(v)|} + \mathbf{b}^{(k+1)}\Big) \tag{3.5}$$

Em comparação com a equação 3.2, além de introduzir a normalização, a camada GCN possui menos parâmetros treináveis, sendo apenas a matriz $\mathbf{w}^{(k+1)} \in \mathbb{R}^{d^{(k+1)}\times d^{(k)}}$ e $\mathbf{b}^{(k+1)} \in \mathbb{R}^{d^{(k+1)}}$. Isso acontece devido ao fato de que a mensagem agregada dos vizinhos é incorporada com a mensagem do próprio vértice, basta observar que a soma das representações vetoriais acontece para $v \in \mathcal{N}(u) \cup \{u\}$. O termo de normalização simétrica $1/\sqrt{|\mathcal{N}(u)||\mathcal{N}(v)|}$ introduzido ajuda a reduzir instabilidades numéricas ou dificuldades no processo de otimização [70]. A equação (3.5) também pode ser escrita na forma matricial já que os parâmetros da mesma camada são compartilhados, então é possível fazer a atualização das representações para todos os vértices de maneira conjunta:

$$\mathbf{H}^{(k+1)} = \boldsymbol{\sigma}\left(\widetilde{\mathbf{D}}^{-1/2}\,\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{-1/2}\mathbf{H}^{(k)}\mathbf{W}^{(k+1)}\right) \tag{3.6}$$

Em que $\widetilde{\mathbf{A}} = \mathbf{A} + \mathbb{I}_N$ é a matriz de adjacências do grafo em questão com *auto-loop* e $\widetilde{\mathbf{D}}$ é a matriz diagonal de graus dos vértices em que os elementos da diagonal são dados por $\widehat{D}_{i,i} = \sum_j \tilde{A}_{i,j}$. Na equação (3.6) foi omitido o termo $\mathbf{b}^{(k+1)}$ por simplicidade.

Figura 8 – Esquema de uma GCN de Múltiplas Camadas



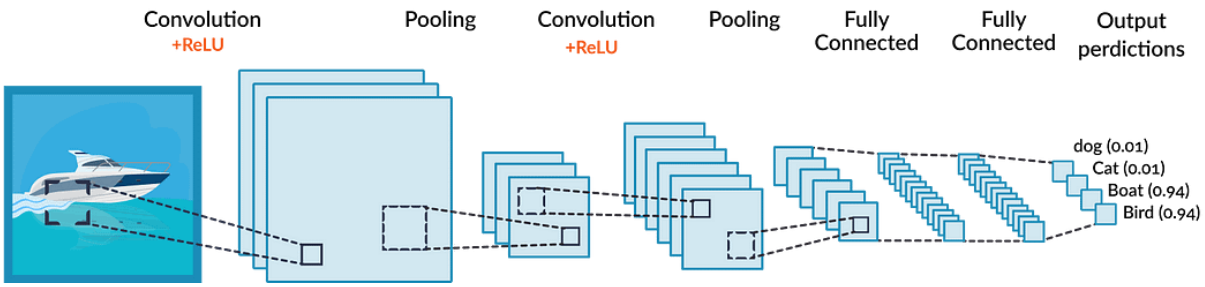
Fonte: Kipf (2016) [72]

Uma GCN é composta por uma série de camadas convolucionais e de *pooling*. As camadas convolucionais são responsáveis pela aplicação de uma série de filtros, ou kernels, às entradas. Estes filtros são capazes de extrair características importantes de cada parte da entrada, através da aplicação de operações de convolução. As camadas de *pooling*, por outro lado, realizam operações de redução nos dados extraídos pelas camadas convolucionais. Estas operações de redução têm como objetivo reter apenas as informações mais relevantes, reduzindo assim a dimensionalidade dos dados. Após a passagem pelas camadas convolucionais e de *pooling*, os dados são encaminhados para uma camada totalmente conectada. Esta camada realiza a tarefa supervisionada, que pode ser, por exemplo, uma tarefa de classificação.

O processo de aprendizado em uma GCN envolve a otimização dos pesos dos filtros utilizados nas camadas convolucionais. Esta otimização é geralmente realizada através de um

algoritmo de otimização, que ajusta os pesos dos filtros de forma a minimizar uma função de perda.

Figura 9 – Camadas da *Graph Convolutional Networks*

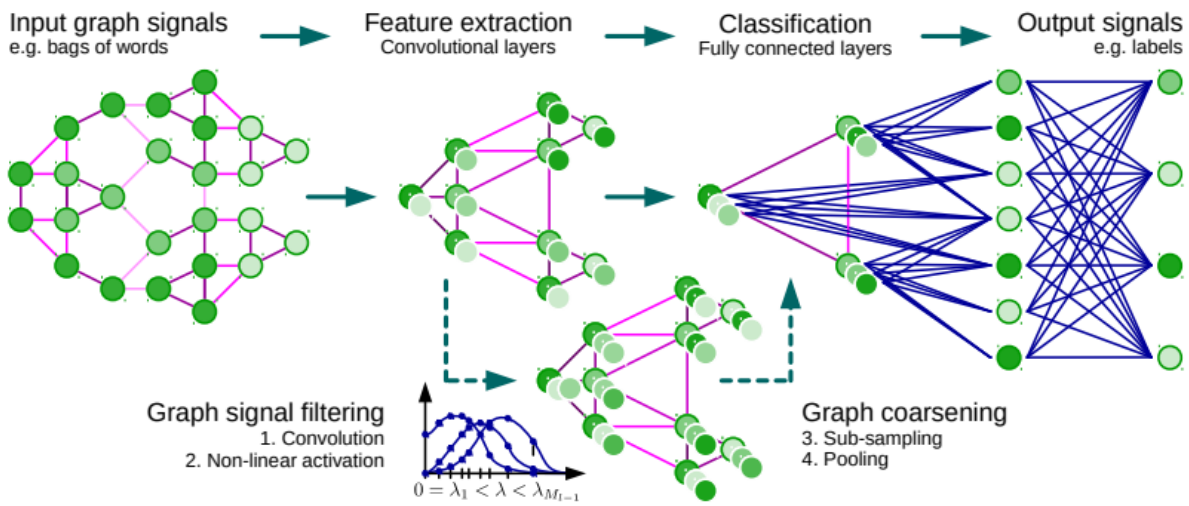


Fonte: Fonte: Kipf (2016) [72]

As *Graph Convolutional Networks* (GCNs)**,** ou em português, as Redes Neurais Convolucionais de Grafos, trabalham não apenas com os nós (ou pontos) de um gráfico, mas também com os sinais que são passados entre esses nós. E esses sinais são como mensagens sendo transmitidas de um nó para outro. As GCNs realizam as chamadas 'convoluções' nessas mensagens. O objetivo dessas convoluções é aplicar filtros para destacar as conexões mais importantes para a representação de um nó específico. A camada de *pooling* é uma técnica utilizada para reduzir a dimensionalidade dos dados, a ideia é simplificar as informações da camada anterior, preservando as características mais relevantes. Geralmente, a camada de *pooling* é inserida em intervalos regulares entre as camadas convolucionais. Ao reduzir o tamanho dos *feature maps*, ou em português, mapas de características, e, portanto, o número de parâmetros da rede, a camada de *pooling* acelera o tempo de cálculo e reduz o risco de ajustes, extraindo apenas as informações mais relevantes para um problema específico [73].

A Figura 10 mostra um esquema de como as convoluções são aplicadas em gráficos. O processo começa com um gráfico e sua matriz de adjacência (uma tabela que mostra quais nós estão conectados entre si). As camadas convolucionais aplicam filtros para extrair características importantes. Depois que as convoluções são aplicadas, uma função de ativação não linear, como a ReLU, é aplicada aos resultados. Em seguida, são aplicadas camadas de redução, incluindo subamostragem e *pooling*, para simplificar e extrair as características mais importantes de partes do gráfico. Finalmente, as saídas são enviadas para uma camada totalmente conectada, que é a última etapa do processo [73].

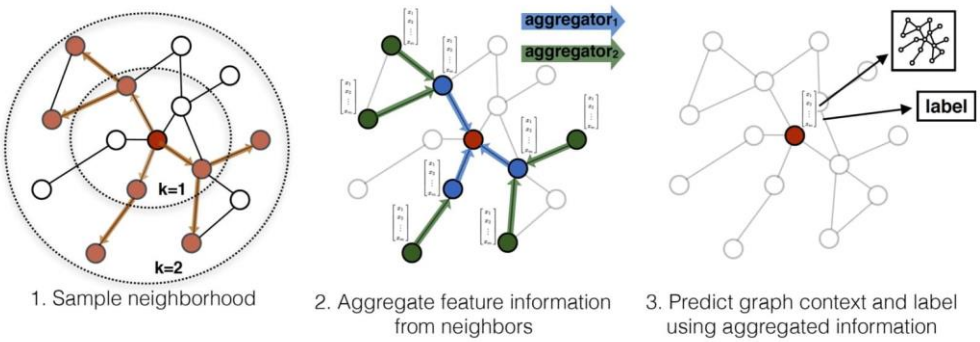
Figura 10 – Arquitetura de uma CCN em grafos.



Fonte: Defferrard, Bresson e Vandergheynst (2016) [73]

O trabalho da GCN inspirou Leskovec e dois colegas da Universidade de Stanford, Hamilton e Ying, a criarem o *Graph Sample and Aggregate* (GraphSAGE) [74]. Um framework indutivo geral que aprende representações de baixa dimensão para nós em grafos. Ele faz isso agregando características dos nós vizinhos. Em vez de treinar *embeddings* individuais para cada nó, o GraphSAGE aprende uma função que gera *embeddings* ao amostrar e agregar características de uma vizinhança local de um nó. Ele colocou sua criação à prova no verão de 2017 na Pinterest, uma empresa americana que opera uma plataforma de compartilhamento de imagens que permite aos usuários criar e gerenciar, em painéis pessoais temáticos, coleções de imagens como eventos, interesses, hobbies e muito mais onde atuou como cientista-chefe [63].

Figura 11 – Ilustração Visual da Abordagem de Amostragem e Agregação do GraphSAGE



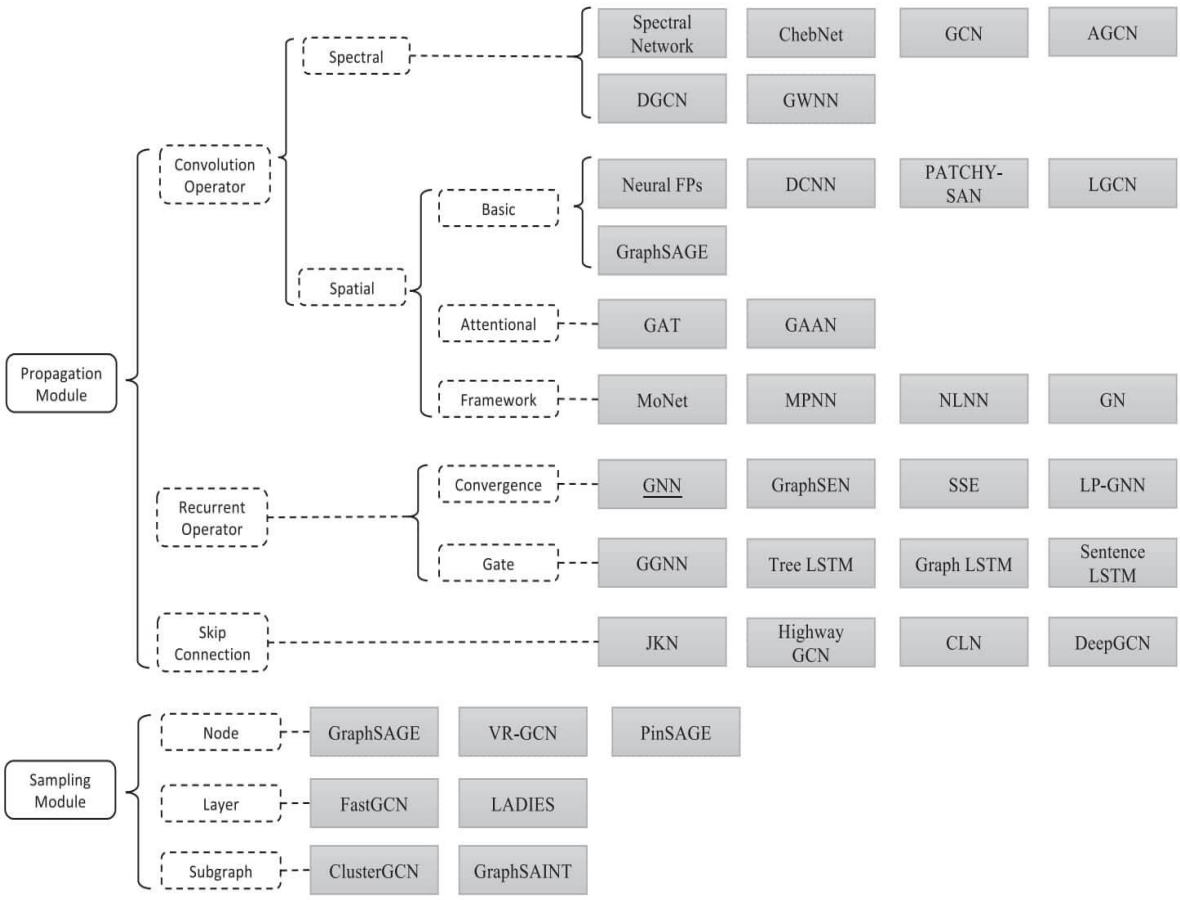
Fonte:Hamilton, Ying e Leskovec (2017) [74]

Em 2018, Leskovec et al (2018) implementou um sistema de recomendação, conhecido como *PinSage* [75], que contava com 3 bilhões de nós e 18 bilhões de arestas que superou outros modelos de IA na época. O método utiliza a exploração em um determinado nó e se move aleatoriamente para os nós vizinhos, repetindo esse processo por um número fixo de passos. Isso permite que o algoritmo explore a estrutura local do grafo e capture as relações entre os nós. O Pinterest aplica-o hoje em mais de 100 casos de uso em toda a empresa [63].

As *Graph Neural Networks* (GNNs) evoluíram significativamente ao longo dos anos, com várias variantes e modelos sendo desenvolvidos para lidarem com diferentes tipos de problemas e conjuntos de dados, tais como: *Graph Convolutional Network* (GCN), *Graph Sample and Aggregate* (GraphSAGE), *Graph Attention Network* (GAT), *Heterogeneous Graph Attention Network* (HAN), e *Heterogeneous Graph Transformer* (HGT), como poderemos observa na Figura 12 [63].

.

Figura 12 – Visão geral das GNNs e suas Variantes.



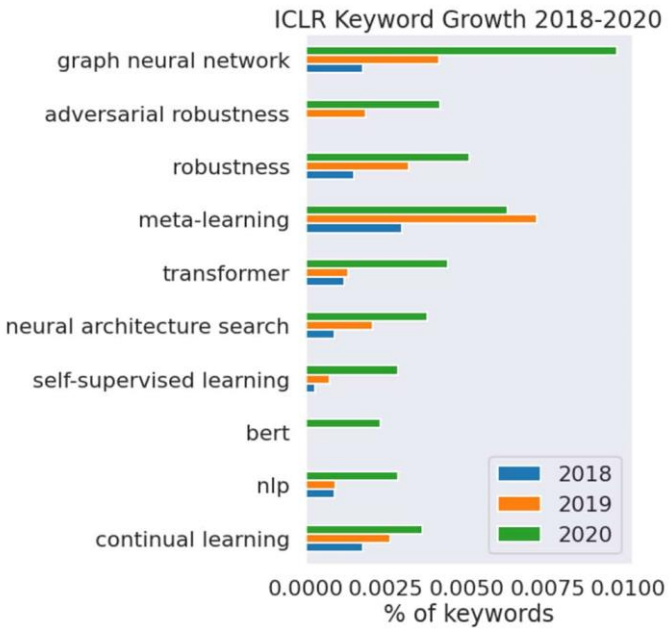
Fonte: NVIDIA (2022) [63].

3 DESENVOLVIMENTO

3.1 Panorama do Mercado de Detecção de Fraudes na Área da Saúde

A *International Conference on Learning Representations* (ICLR), ou em português, Conferência Internacional sobre Representações de Aprendizado, é a principal reunião de profissionais dedicados ao avanço do ramo da inteligência artificial, geralmente referido como *Deep Learning* [63]. A ICLR é mundialmente conhecida por apresentar e publicar pesquisas de ponta em todos os aspectos do *Deep Learning* usados nos campos da inteligência artificial, estatística e ciência de dados, bem como em áreas de aplicação importantes como visão computacional, biologia computacional, reconhecimento de fala, compreensão de texto, jogos e robótica. As GNNs estão "sendo motivo de entusiasmo devido à sua flexibilidade para modelar relações complexas, algo que as redes neurais tradicionais não podem fazer", disse Jure Leskovec, Professor Associado de Stanford, em uma palestra no fórum de Computação na ICRL de 2021, onde apresentou o gráfico, Figura 13, referente aos artigos de IA que as mencionam:

Figura 13 – Principais Palavras-Chave nas conferências da ICLR (2018 a 2020)



Fonte: NVIDIA (2022) [63].

A análise de grafos ou análise de rede é um dos métodos mais utilizados para capturar fraudes no *Medicare*. Um estudo utilizou metodologias de *Knowledge Discovery in DataBases*

(KDD), ou em português, Descoberta de Conhecimento em Banco de Dados, para detectar prestadores de serviços de saúde fraudulentos utilizando dados de seguro médico [76]. Nesse estudo, recursos de rede, como medidas de centralidade da rede entre prestadores de serviço de saúde, contribuíram para a previsão de fraude no *Medicare*. Em outro estudo, após a criação de um grafo composto por nós de prestadores de serviços de saúde a partir de dados de planos de saúde, a detecção de fraudes foi realizada utilizando as semelhanças em procedimentos médicos e prescrições de medicamentos entre prestadores de serviços de saúde fraudulentos e não fraudulentos [77]. Características baseadas em grafos, como tamanho dos agrupamentos, densidade dos agrupamentos e a quantidade média em dólares, foram utilizada a partir de um grafo heterogêneo consistindo de prestadores de serviços médicos, pacientes e médicos como nós [78].

Recentemente, com o desenvolvimento de algoritmos de *machine learning*, pesquisas significativas para a detecção de fraudes têm sido conduzidas com *Graph Neural Networks* (GNNs). Diversos modelos têm demonstrado um alto nível de desempenho de previsão de fraudes em um determinado conjunto de dados estruturados por grafos [79], [80], [81].

Em fraudes no setor de saúde, é comum que um diagnóstico falso seja o resultado de uma colaboração entre pacientes, médicos, prestadores de serviços e hospitais. Esse diagnóstico é usado para obter ilegalmente benefícios de seguro médico ou reembolsos que são posteriormente distribuídos entre os fraudadores. No entanto, apesar desses estudos recentes sobre a detecção de fraudes usando GNNs, poucos estudos tentaram detectar fraudes no *Medicare* por meio de colaboração entre pacientes, médicos, prestadores de serviços e hospitais com GNN.

Em estudo recente, o algoritmo GraphSAGE demonstrou um desempenho melhor na detecção de fraudes no *Medicare* do que os métodos tradicionais de Machine Learning [18]. Um outro estudo relacionado à detecção de fraudes no *Medicare*, utiliza algoritmos convencionais de *Machine Learning*, tais como: Regressão Logística (RL), *eXtreme Gradient Boosting* (XGBoost) e *Multi-Layer Perceptron* (MLP), que adicionaram recursos gerados a partir de Grafos de uma rede de dados entre pacientes, médicos, prestadores de serviços e hospitais [82]. Em seguida, examinamos se o desempenho da detecção de fraude do *Medicare* melhora introduzindo o modelo GNN e adicionando recursos gerados a partir de Grafos a algoritmos convencionais de aprendizado de máquina. Este estudo aborda o desempenho de modelos de *Machine Learning*, algoritmos GGNs, e modelos de *Machine Learning* com recursos adicionais gerados a partir de Grafos para a detecção de fraudes no *Medicare*. As vantagens e desvantagens de cada método serão apresentados.

3.2. Trabalhos Relacionados

3.2.1 Detecção de Fraudes Utilizando Análise de Grafos

Muitos estudos tentaram melhorar o desempenho da tarefa de classificação, considerando a relação entre os objetos incluídos no conjunto de dados. Por exemplo, em redes sociais, existem casos em que benefícios econômicos são obtidos influenciando as decisões de compra de outras pessoas por meio da criação de falsas avaliações ou contas através de uma colaboração coletiva. Para refletir essa relação de colaboração, eles definem revisores, avaliações e lojas como nós e usam um modelo de Grafo para prever avaliações fraudulentas [83]. Outro estudo detectou contas de usuários falsas com base em um Grafo de relacionamento de usuários. Especificamente, para novas contas criadas dentro de 7 dias e com menos de 50 solicitações de amizade, a probabilidade de uma conta falsa é calculada com base na resposta à solicitação de amizade e no resultado de aceitar ou rejeitar a solicitação [84]. Além disso, algoritmos *Random Walk* foram propostos para detectar contas de usuários falsas através de caminhadas aleatórias de uma série de nós, com base na suposição de que contas de usuários falsas precisam de mais saltos em média do que contas padrão no Grafo da rede social [85].

A análise de Grafos também é amplamente utilizada na previsão de risco, uma tarefa de classificação, na área financeira. Por exemplo, o desempenho da classificação de crédito pode ser melhorado extraindo a medida de centralidade da rede de correlação para empresas que tomaram dinheiro emprestado de instituições financeiras e usando-a como uma variável de regressão logística [86]. O desempenho do modelo de classificação de crédito corporativo usando métodos estatísticos e de *Machine Learning* pode ser melhorado usando adicionalmente análise de Grafos considerando a relação de transação entre empresas. Neste caso, uma pontuação ponderada do cliente empregando análise de Grafos e as informações de centralidade do Grafo são usadas [87]. Além do campo de risco de crédito, existem estudos analisando o risco sistêmico do mercado de derivativos de commodities, representando a correlação entre os preços futuros ao longo do tempo em uma estrutura de Grafos [88].

3.2.2 Detecção de Fraude Usando GNN

Recentemente, estudos sobre GNNs, que utilizam modelos que podem aprender informações de recursos sobre nós vizinhos a partir de um conjunto de dados estruturado em Grafos, surgiram em vários campos, como classificação de crédito, detecção de fraude e

detecção de anomalias. Por exemplo, um modelo de classificação de crédito foi desenvolvido aprendendo a relação entre instituições aplicando GNN a uma rede de garantia de empréstimo composta por instituições financeiras em unidades de Grafo [89]. Em tarefas de recomendação, a GNN pode ser empregada. Assim, tanto os usuários quanto os itens que são recomendados são representados como nós em um gráfico. A GNN então aprende a relação entre esses nós. Em outras palavras, ela descobre como os usuários e os itens recomendados estão interligados dentro dessa estrutura de gráfico. Isso ajuda a realizar as recomendações mais precisas e personalizadas. [90].

Em particular, vários estudos sobre detecção de fraude usando a GNN apareceram nas redes sociais e nos campos financeiros. Primeiro, para detectar visualizações falsas ou contas falsas em redes sociais, um estudo mediu a possibilidade de visualizações falsas em unidades de sub-Grafos após agrupar todos os dados do Grafo com um GCN e uma rede *Deep Learning* [91]. Usando o GCN, eles detectam avaliações fraudulentas no sistema de revisão de aplicativos online, considerando os textos, comportamentos e relacionamentos dos revisores [92]. Segundo, em finanças, existem vários tipos de fraude, como fraudes de seguros, contas fraudulentas, fraudes em transações bancárias, fraudes em *Bitcoin*, dentre outras. Um estudo apresentou um método de aprendizado de rede chamado *node2vec* para detectar um grupo organizado de fraudes de seguros que cometia reivindicações fraudulentas para o seguro de frete de devolução da empresa Alibaba [93]. Para detectar as contas maliciosas do sistema de pagamento do Alibaba, Alipay, que continuamente enviavam spam e causavam danos financeiros, um estudo propôs uma GNN heterogênea que poderia considerar o comportamento entre contas [94]. Um estudo propôs um GAT usando um Grafo heterogêneo que poderia considerar interações de nível de transação para detectar transações fraudulentas em plataformas de comércio eletrônico [95]. Os criminosos usam o anonimato da criptomoeda para prejudicar o sistema financeiro, foi elaborado um estudo que define transações de Bitcoin e fluxos de pagamento como nós e arestas para desenvolver um modelo GCN para detectar transações fraudulentas relacionadas às criptomoedas [96].

A partir desses estudos anteriores, podemos observar que a GNN é utilizada para detecção de fraude em vários campos. Desta forma, as GNNs se tornaram um método potencial para tarefas de detecção de fraude, detectando nós fraudulentos agregando informações vizinhas para considerar diversas relações [94].

3.2.3 Literatura Anterior sobre Algoritmos GNN

Diversos algoritmos de GNNs dependem de como eles aprendem a estrutura do Grafo. Primeiro, temos a *Graph Convolutional Network* (GCN), que é um dos modelos mais básicos de GNNs. O GCN é semelhante a uma rede neural convolucional, que é um tipo de rede neural comumente utilizada para analisar imagens [69]. O GCN propaga informações dos nós vizinhos usando camadas convolucionais de Grafos. No entanto, o GCN precisa de uma estrutura de Grafo completa para funcionar, o que pode resultar em alto consumo de memória quando o Grafo é muito grande [97]. Em segundo lugar, temos o *Graph Sample and Aggregate* **(**GraphSAGE)**,** que é um algoritmo de *Machine Learning* que foi projetado para trabalhar com grandes volumes de Grafos. Ele é uma versão generalizada do *Graph Convolutional* (GCN). O GraphSAGE é um modelo de aprendizado indutivo, que pode aprender a partir de exemplos e generalizar para situações não vistas antes. Isso é útil quando você tem um Grafo extenso e quer realizar previsões sobre os nós que não estavam presentes quando o modelo foi treinado. Ele emprega uma técnica denominada 'amostragem de nós vizinhos'. Em vez de coletar todas as informações de todos os nós vizinhos (o que pode ser demais para gráficos grandes), ele seleciona um número fixo de nós vizinhos para levar em consideração. Isso o ajuda a tornar o algoritmo mais eficiente. O GraphSAGE agrega, ou combina, recursos dos nós vizinhos para criar uma representação do nó de interesse. Isso permite que o modelo capture informações importantes sobre a vizinhança de um nó [74]. Terceiro, temos *o Graph Attention Network* (GAT), que é uma arquitetura de rede neural que opera com dados estruturados em Grafos [98]. Ele utiliza um mecanismo chamado *self-attention* para calcular a importância relativa dos nós vizinhos ao calcular a representação oculta de cada nó. Desta forma, ele permite que cada nó 'preste atenção' aos nós que são mais relevantes. Isso permite que o modelo capture informações importantes sobre a vizinhança de um nó. Ao permitir que cada nó atribua diferentes pesos a diferentes nós em uma vizinhança, o GAT fornece um modelo mais flexível e potencialmente mais robusto. Esta abordagem baseada em atenção oferece um nível de compreensão muito significativo, pois os coeficientes de atenção podem ser vistos como indicativos de importância para cada vizinhança [99].

Os algoritmos GNNs geralmente aprendem a partir de um único tipo de nó e aresta. No entanto, isso pode ser um problema quando lidamos com gráficos heterogêneos, que têm vários tipos de nós e arestas. Por isso, foram propostos algoritmos GNNs específicos para gráficos heterogêneos [100]. Um desses algoritmos é o *Heterogeneous Graph Attention Network* (HAN)**.** O HAN é um tipo especial de rede neural que aprende a importância de diferentes tipos

de nós. Ele faz isso utilizando o que é conhecido como ‘atenção em nível de nó’ e ‘atenção em nível semântico’, baseado em *meta-paths*. Os *meta-paths* são como sequências de tipos de relações que conectam os nós em um gráfico heterogêneo. A atenção em nível de nó ajuda a HAN a aprender a importância entre um nó e seus vizinhos que estão conectados pelos *meta-paths*. Enquanto isso, a atenção em nível semântico ajuda a HAN a aprender a importância de diferentes *meta-paths*. Com a importância aprendida tanto da atenção em nível de nó quanto da atenção em nível semântico, a HAN pode considerar totalmente a importância do nó e dos *meta-paths*. Assim, o modelo proposto pode criar uma representação de um nó, reunindo características de nós vizinhos baseados em *meta-paths* de maneira hierárquica. Isso permite que o HAN entenda a importância relativa de diferentes tipos de nós e relações em um gráfico heterogêneo [101]. Já o algoritmo HGT (Heterogeneous Graph Transformer) é projetado para estruturas de gráficos heterogêneos com dois ou mais tipos de nós e arestas. O HGT reúne informações dos nós de origem para obter uma representação contextualizada para o nó de destino, construindo um HGT utilizando parâmetros que dependem do tipo de nó e aresta. O HGT pode ser dividido em três componentes: *heterogeneous mutual attention, heterogeneous message-passing, e target-specific aggregation*, que modelam a heterogeneidade. Além disso, o algoritmo *Heterogeneous Mini-Batch Graph Sampling* (HGSampling) é usado para um aprendizado eficiente e escalável de dados de gráfico em escala da web [102]. Esses algoritmos ajudam a entender e representar a estrutura complexa de gráficos heterogêneos, que são gráficos com diferentes tipos de nós e conexões. Eles fazem isso prestando atenção a diferentes partes do gráfico e reunindo informações de maneira inteligente.

3.2.4 Desempenho da Detecção de Fraudes na Área da Saúde

Várias pesquisas têm sido realizadas na área de detecção de fraudes no *Medicare* utilizando abordagens de *Machine Learning* e *Graph Neural Networks* (GNNs). Cada abordagem tem a sua própria vantagem e desvantagem. Os modelos de *Machine Learning*, por exemplo, são conhecidos pelo seu rápido tempo de treinamento e por serem relativamente pequenos e simples. Por outro lado, os modelos GNNs têm a vantagem de incorporar diretamente as relações entre os objetos. Isso é realizado através da representação dos objetos como nós e das relações entre os objetos, arestas, que conectam esses nós. Um outro método, é agregar as informações de Grafos em modelos de *Machine Learning* já existentes, que podem oferecer a vantagem de uma aplicação mais simples dessas informações com uma excelente

compatibilidades entre os modelos. A Tabela 2 resume vários métodos para a detecção de fraudes no *Medicare*.

Tabela 2 – Resumo das Diferentes Categorias de Detecção de Fraudes no *Medicare*

Categoria	Artigo	Metodologia	Resultados	Prós	Contras
Aprendizado de Máquina	Kumar et al. (2010) [104]	**SVM**	Precisão melhor (taxa de acerto) sobre as abordagens existentes, que é precisa o suficiente para resultar em economias de $15-25 milhões para uma seguradora típica.	* Treinamento curto * Modelo Relativamente simples e pequeno	* Métodos de Aprendizado profundo tem melhores performance
	Shin et al. (2012) [105]	**DT**	O modelo proposto foi em grande parte consistente com as técnicas de detecção manuais atualmente usadas para identificar potenciais fraudadores e a detecção automática de fraudes é flexivel e simples de atualização.		
	Aral et al. (2012) [106]	**Abordagem de mineração de dados baseada em distância**	Taxa de verdadeiros positivos de 77,4% e taxa de falsos positivos de 6% para as prescrições médicas fraudulentas		
	Ortega et al. (2006) [107]	**MLP**	Uma taxa de detecção de aproximadamente 75% dos casos realmente fraudulentos por mês, tornando a detecção 6,6 meses mais cedo do que sem a utilização do sistema		
	Mayaki et al. (2022) [110]	**LSTM Autoencoder**	A AUC (PRC) do método proposto e de alguns métodos de última geração é de 0,745		
	Johnson et al. (2019) [108]	**ROS-RUS (Random OverSampling - Random UnderSampling) + NN (Redes Neurais)**	O ROS e o ROS-RUS apresentam desempenho significativamente melhor do que os métodos de base e de nível de algoritmo com pontuações AUC médias de 0,8505 e 0,8509, enquanto o ROS-RUS maximiza a eficiência com uma aceleração de 4x no tempo de treinamento		
Rede Neural de Grafo	Yoo et al. (2022) [33]	**RL/GNN**	O modelo GraphSAGE usando dados estruturados em grafo superou o modelo de base em todas as métricas de desempenho	* Pode melhorar o desempenho considerando a relação de conexão	* Problema de oversmoothing ocorre ao empilhar várias camadas de redes neurais de grafo * Custo computacional é muito alto
	Lu et al. (2023) [109]	**GNN**	Este modelo baseado em atenção pode melhorar os resultados interpretáveis para fornecer mais insights sobre a tarefa de detecção de fraude em saúde		* Dificuldade em configurar o conjunto de dados estruturados em grafos
Aprendizado de Máquina com Informação de Grafo	Liu et al. (2016) [78]	**Graph-Theoretical Feature**	As ferramentas e algoritmos foram capazes de melhorar a produtividade do usuário e permitir que os usuários produzam resultados que antes eram difíceis ou demorados de produzir	* A informação do grafo pode ser aplicada mais facilmente do que GNN	* Dificuldade em configurar o conjunto de dados estruturados em grafos
	Branting et al. (2016) [77]	**Graph-Theoretical Feature**	Uma combinação de 11 features atingiu um f-score de 0,919 e uma área ROC de 0,960 na previsão de exclusão.		
	Herland et al. (2018) [82]	**ML + Graph-Theoretical Feature**	Nossos resultados mostram que o conjunto de dados combinado com o modelo de Regressão Logística (LR) rendeu o score de 0,816.	* Relações adicionais podem ser consideradas em modelos de ML existentes	

Fonte: Adaptado YOO et al (2023) [64].

Os modelos de *Machine Learning* estão se tornando cada vez mais populares para classificar conjuntos de dados tabulares, como na detecção de fraudes no *Medicare*, graças à sua capacidade de fazer previsões com precisão. Além disso, os algoritmos *Convolutional Neural Networks* (CNNs)**,** ou em português, Redes Neurais Convolucionais, são frequentemente utilizados para analisar dados de imagem [103]. Por exemplo, um estudo anterior usou o algoritmo *Support Vector Machine*, ou em português, Máquina de Vetores de Suporte, para detectar anomalias em processos de sinistros de seguros [104]. Em outro estudo, o modelo *Decision Tree*, ou em português, Árvore de Decisão, foi utilizado para detectar fraudes no Medicare, identificando padrões anormais em sinistros de seguros e dados de pacientes [105]. Outro estudo usou métricas de risco e correlações baseadas em distância para detectar fraudes de prescrição do Medicare [106]. Um modelo *Multi-Layer Perceptron* (MLP), ou em português, Perceptron Multicamadas, foi usado para desenvolver um modelo de detecção de fraude para sinistros de seguros no Chile [107]. Além disso, muitos estudos demonstraram desempenhos significativos das redes neurais na detecção de fraudes médicas [108] [110].

Os modelos GNNs têm sido utilizados para analisar as informações estruturais sobre a relação entre prestadores de serviços médicos, hospitais, médicos e pacientes. Em um estudo anterior, um gráfico heterogêneo foi criado, onde os prestadores e os beneficiários do *Medicare* foram definidos como nós. O algoritmo GraphSAGE foi então aplicado para detectar fraudes no *Medicare* [33]. Em outro estudo, o modelo *Attributed Heterogeneous Information Network* (AHIN), ou em português, Rede de Informação Heterogênea Atribuída, foi utilizado. Este modelo foi usado para construir as relações comportamentais entre as várias visitas dos pacientes [109]. Há estudos adicionais que deram um passo além do modelo tradicional de *Machine Learning* ao adicionarem recursos baseados em teoria de grafos para detectar as fraudes no sistema *Medicare* [76] [77][82].

3.2.5 Resultados e Contribuições de Nosso Estudo

Durante o desenvolvimento de modelos de detecção de fraude a partir da revisão da literatura, identificamos alguns pontos que poderiam ser apreciados. Primeiro, os modelos de *Machine Learning*, apresentam rápido tempo de treinamento e são relativamente pequenos e simples. No entanto, eles tem a dificuldade em capturar as relações complexas entre as diferentes estruturas de dados, o que limita a sua eficácia na detecção de fraudes. Segundo, os modelos *Graph Neural Networks* (GNNs) e as suas variantes têm a vantagem de incorporar diretamente as relações entre os objetos, representando os objetos como nós (ou vértices) e as

relações entre os objetos como arestas (ou linhas) que conectam esses nós. No entanto, eles são mais complexos e demorados para treinar, necessitando de esforço computacional significativo, especialmente quando abordamos Grafos grandes. Terceiro, o método que agrega as informações de Grafos em modelos de *Machine Learning* já existentes oferece a vantagem de uma aplicação mais fácil dessas informações com uma excelente compatibilidade. No entanto, essa abordagem pode não ser capaz de capturar todas as nuances das relações no gráfico em certos casos. Quarto, como há poucas pesquisas sobre a eficácia dos modelos de Inteligência Artificial na detecção de fraudes médicas. Desta forma, abordaremos os modelos de *Machine Learning* **(**Logistic Regression, Random Forest, XGBoost, LightGBM e Multi-Layer Perceptron), modelos *Graph Neural Networks* (GraphSAGE, GAT, HAN e HGT), e os mesmos modelos *Machine Learning* com os recursos de Grafos, apresentando os resultados e comparativos de performance e análises em cada caso.

4 MATERIAIS E MÉTODOS

4.1 Conjuntos de Dados

4.1.1 Descrição do Conjunto de Dados Inicial

Realizamos os experimentos utilizando o conjunto de dados de código aberto do serviço de saúde Medicare fornecidos pelo repositório Kaggle [112]. O conjunto de dados é composto basicamente por 4 estruturas de dados:

- **Dados de Internação (1):** Informações dos beneficiários (pacientes) hospitalizados:

Tabela 3 – Informações dos Beneficiários Internados

Variável	Descrição
ID_SOLICITACAO	Identificador exclusivo para cada solicitação de reembolso.
ID_PRESTADOR	Identificador exclusivo para cada prestador de serviços médicos.
ID_BENEFICIARIO	Identificador exclusivo para cada beneficiário (paciente).
DT_ADMISSAO	Data em que o paciente foi admitido no hospital.
DT_ALTA	Data em que o paciente recebeu alta do hospital.
CD_DIAGNOSTICO	Códigos que representam o diagnóstico do paciente.
CD_PROCEDIMENTO	Códigos que representam os procedimentos médicos realizados durante a internação.
GDR	Grupo de Diagnóstico Relacionado: Classificação que agrupa os pacientes com condições semelhantes e ao uso de recursos.
TEMPO_HOSPITALIZADO	Duração da internação do paciente, medida em dias.
CUSTO_TOTAL	Custo total da internação.
DT_INICIO_SOLICITACAO	A data em que a solicitação de reembolso foi iniciada.
DT_FINAL_SOLICITACAO	A data em que a solicitação de reembolso foi encerrada.
VLR_REEMBOLSO_SEGURO	O valor do dinheiro reembolsado em uma reivindicação de seguro.
ID_MEDICO_CHEFE	Identificador exclusivo do médico que supervisionou o atendimento ao paciente.
ID_MEDICO_ASSISTENTE	Identificador exclusivo do médico que atendeu o paciente.

Fonte: Elaborado pelo Autor.

- **Dados de Ambulatório (2):** Informações dos beneficiários (pacientes) que receberam atendimento médico sem a necessidade de internação:

Tabela 4 – Informações dos Beneficiários no Ambulatório

Variável	Descrição
ID_SOLICITACAO	Identificador exclusivo para cada solicitação de reembolso.
ID_PRESTADOR	Identificador exclusivo para cada prestador de serviços médicos.
ID_BENEFICIARIO	Identificador exclusivo para cada beneficiário (paciente).
DT_ADMISSAO	Data em que o paciente foi admitido no hospital.
DT_ALTA	Data em que o paciente recebeu alta do hospital.
CD_DIAGNOSTICO	Códigos que representam o diagnóstico do paciente.
CD_PROCEDIMENTO	Códigos que representam os procedimentos médicos realizados durante o período no ambulatório.
GDR	Grupo de Diagnóstico Relacionado: Classificação que agrupa os pacientes com condições semelhantes e ao uso de recursos.
TEMPO_HOSPITALIZADO	Duração da internação do paciente, medida em dias.
CUSTO_TOTAL	Custo total da internação.
DT_INICIO_SOLICITACAO	A data em que a solicitação de reembolso foi iniciada.
DT_FINAL_SOLICITACAO	A data em que a solicitação de reembolso foi encerrada.
VLR_REEMBOLSO_SEGURO	O valor do dinheiro reembolsado em uma reivindicação de seguro.
ID_MEDICO_CHEFE	Identificador exclusivo do médico que supervisou o atendimento ao paciente.
ID_MEDICO_ASSISTENTE	Identificador exclusiva do médico que atendeu o paciente.

Fonte: Elaborado pelo Autor.

- **Dados de Beneficiários (3):** Informações dos beneficiários (pacientes) dos serviços médicos.

Tabela 5 – Informações Básicas dos Beneficiários

Variável	Descrição
ID_BENEFICIARIO	Identificador exclusivo para cada beneficiário (paciente).
DT_NASC_BENEF	Data de nascimento do beneficiário.
DT_FALEC_BENEF	Data de falecimento do beneficiário, caso ocorra.
SEXO_BENEF	Sexo do beneficiário.
ESTADO_BENEF	Localização geográfica do beneficiário.
REGIÃO_BENEF	Localização geográfica do beneficiário.
EST_SAUDE_BENEF	Informações sobre o estado de saúde atual do beneficiário.
HIST_MED_BENEF	Registro do histórico médico do beneficiário.
MEDICAM_BENEF	Registro dos medicamentos que o beneficiário está utilizando.
VIS_HOSP_BENEF	Frequência das visitas ao hospital ou consultas médicas.
PROC_RLZ_BENEF	Registro dos procedimentos médicos que o beneficiário passou.

Fonte: Elaborado pelo Autor.

- **Dados do Prestador de Serviços Médicos (4):** Informações dos Prestadores de Serviços médicos, incluindo o indicativo de potenciais fraudes:

Tabela 6 – Informações dos Prestadores de Serviços Médicos

Variável	Descrição
ID_PRESTADOR	Identificador exclusivo para cada prestador de serviços médicos do Medicare.
ESPEC_PREST	A especialidade médica do prestador.
ESTADO_PREST	A localização geográfica do prestador de serviços médicos.
REGIÃO_PREST	A localização geográfica do prestador de serviços médicos.
PT_FRAUDE_PREST	Indicação de se o prestador de serviços médicos é um potencial fraudador.

Fonte: Elaborado pelo Autor.

As solicitações de serviços médicos para os beneficiários (pacientes) são os registros que detalham os serviços de saúde que um paciente requisitou ou recebeu. Esses registros são fundamentais para o processo de faturamento e administração do sistema de saúde Medicare [52]. Eles incluem informações, tais como:

- **Tipo de Serviço**: Descreve o procedimento ou tratamento que o beneficiário (paciente) recebeu, tais como consultas médicas, exames de diagnóstico, cirurgias, dentre outros.
- **Data do Serviço**: A data em que o serviço foi prestado ao paciente.
- **Custo do Serviço**: O valor que foi cobrado pelo serviço médico.

Essas solicitações são documentadas e utilizadas para verificar se os serviços foram devidamente prestados e se os custos cobrados estão corretos. No contexto de detecção de fraude, a análise dessas solicitações ajuda a identificar os padrões anormais que podem indicar práticas fraudulentas, como cobranças por serviços não realizados ou superfaturamento [52].

4.1.2 Elaboração do Conjunto de Dados Final

As informações para desenvolver os modelos de detecção de fraude foram elaboradas a partir de algumas etapas de enriquecimento e de cruzamento das 4 estruturas de dados apresentadas. Nesses conjuntos de dados utilizamos as 3 colunas-chave: **ID_PRESTADOR**, **ID_BENEFICIARIO** e **ID_SOLICITACAO**.

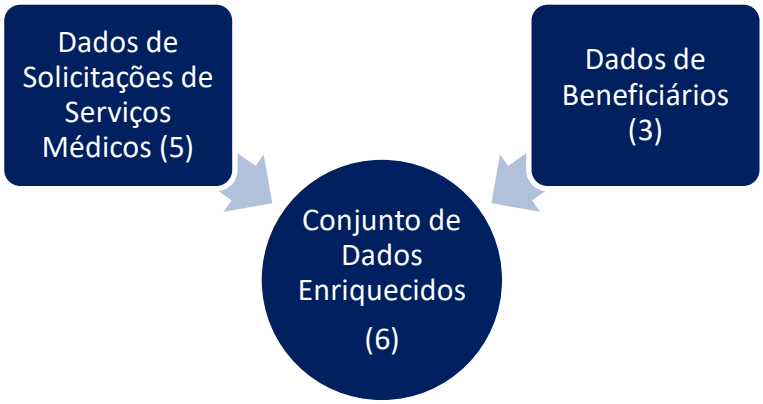
Na primeira etapa, combinamos os **Dados de Internação (1)** e os **Dados de Ambulatório (2)**, pois compartilham os mesmos atributos dos beneficiários (pacientes) relacionados aos **Dados de Solicitações de Serviços Médicos (5)**.

Figura 14 – Etapa 1: Enriquecimento dos Dados de Solicitações de Serviços Médicos

Fonte: Elaborado pelo Autor.

Na segunda etapa, enriquecemos os **Dados de Beneficiários (3)**, com o conjunto de **Dados de Solicitações de Serviços Médicos (5)** utilizando a coluna **ID_BENEFICIARIO** como chave, para reunir as informações dos beneficiários (pacientes).

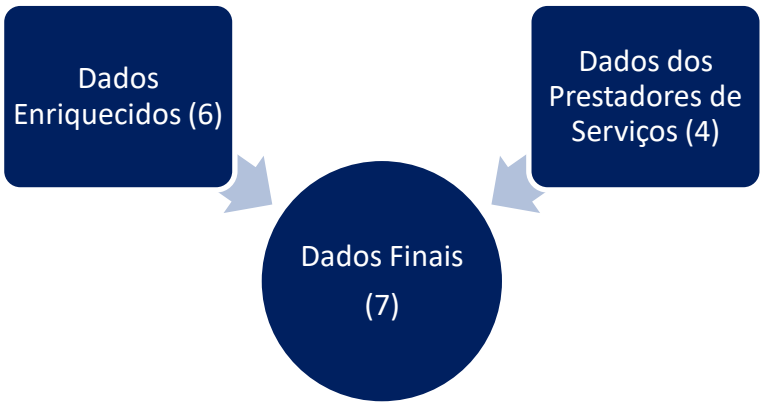
Figura 15 – Etapa 2: Enriquecimento do Conjunto de Dados do Beneficiário



Fonte: Elaborado pelo Autor.

Na última etapa, integramos **os Dados do Prestador de Serviços Médicos (4)** ao **Conjunto de Dados Enriquecidos (6)** utilizando a coluna **ID_PRESTADOR** como chave. O **Conjunto de Dados Final (7)**, que engloba as informações completas sobre as solicitações de serviços médicos do Medicare referentes aos beneficiários (pacientes), aos prestadores de serviços e aos médicos, é obtido a partir do enriquecimento e do cruzamento das quatro estruturas de dados, **Dados de Internação (1)**, **Dados de Ambulatório (2)**, **Dados de Beneficiários (3), Dados do Prestador de Serviços Médicos (4)**, em três etapas distintas.

Figura 16 – Etapa Final: Enriquecimento com os Dados dos Prestadores de Serviços Médicos



Fonte: Elaborado pelo Autor.

Tabela 7 – Informações do Conjunto de Dados Final

Variável	Descrição
ID_SOLICITACAO	Identificador único para cada solicitação de reembolso.
ID_BENEFICIARIO	Identificador único para cada beneficiário (paciente).
DT_NASC_BENEF	Data de nascimento do beneficiário.
DT_FALEC_BENEF	Data de falecimento do beneficiário, caso ocorra.
SEXO_BENEF	Sexo do beneficiário.
ESTADO_BENEF	Localização geográfica do beneficiário.
REGIÃO_BENEF	Localização geográfica do beneficiário.
EST_SAUDE_BENEF	Informações sobre o estado de saúde atual do beneficiário.
HIST_MED_BENEF	Registro do histórico médico do beneficiário.
MEDICAM_BENEF	Registro dos medicamentos que o beneficiário está utilizando.
VIS_HOSP_BENEF	Frequência das visitas ao hospital ou consultas médicas.
PROC_RLZ_BENEF	Registro dos procedimentos que o beneficiário passou.
ID_PRESTADOR	Identificador exclusivo para cada prestador de serviços médicos.
ESPEC_PREST	A especialidade médica do prestador.
ESTADO_PREST	O estado geográfico do prestador de serviços médicos.
REGIÃO_PREST	A região geográfica do prestador de serviços médicos.
PT_FRAUDE_PREST	Indicação de se o prestador de serviços médicos é um potencial fraudador.
DT_ADMISSAO	Data em que o paciente foi admitido no hospital.
DT_ALTA	Data em que o paciente recebeu alta do hospital.
CD_DIAGNOSTICO	Códigos que representam o diagnóstico do paciente.
CD_PROCEDIMENTO	Códigos que representam os procedimentos médicos realizados durante a internação.
GDR	Grupo de Diagnóstico Relacionado: Classificação que agrupa os pacientes com condições semelhantes e ao uso de recursos.
TEMPO_HOSPITALIZADO	Duração da internação do paciente, medida em dias.
CUSTO_TOTAL	Custo total da internação.
DT_INICIO_SOLICITACAO	A data em que a solicitação de reembolso foi iniciada.
DT_FINAL_SOLICITACAO	A data em que a solicitação de reembolso foi encerrada.
VLR_REEMBOLSO_SEGURO	O valor reembolsado em uma reivindicação de seguro.
ID_MEDICO_CHEFE	Identificador exclusivo do médico que supervisionou o atendimento ao paciente.
ID_MEDICO_ASSISTENTE	Identificador exclusivo do médico que atendeu o paciente.

Fonte: Elaborado pelo Autor.

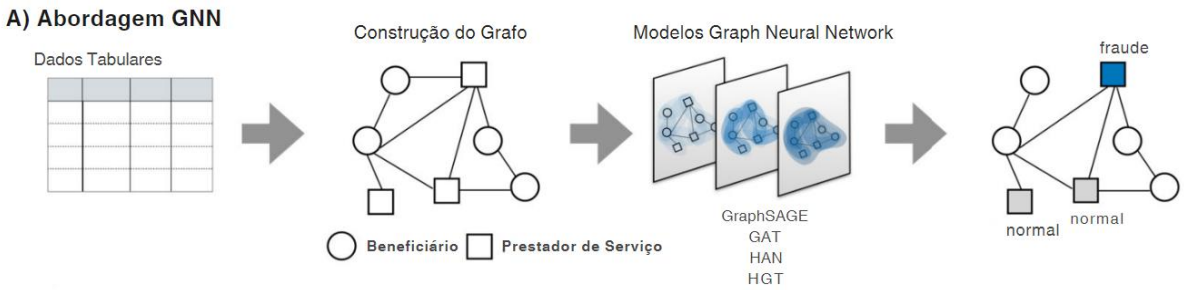
O conjunto de dados rotulados como “Dados finais” referem-se às entradas utilizadas para o desenvolvimento dos modelos de detecção de fraudes neste TCC. Este conjunto compreende aproximadamente 1 milhão de registros, e dentre eles, aproximadamente 50% dos registros são identificados como fraudes no sistema de saúde Medicare, indicando pouco desequilíbrio entre as classes.

4.2 Desenvolvimento dos Modelos de Detecção de Fraude

Para desenvolver os modelos de detecção de fraudes utilizando os dados do *Medicare*, consideramos duas abordagens diferentes. A primeira foi desenvolver os modelos *Graph Neural Networks* (GraphSAGE, GAT, HAN e HGT), e a segunda foi desenvolver os modelos de *Machine Learning* (*Logistic Regression*, *Random Forest*, XGBoost, LightGBM e *Multi-Layer Perceptron*) com características de grafos.

A Figura 17 mostra um esquemático da primeira abordagem. Na abordagem *GNN*, transformamos os dados enriquecidos em dados estruturados em grafos. Em seguida, desenvolvemos os modelos *Graph Neural Networks* para detectar as fraudes por meio de uma tarefa de classificação de nós (Seção 4.3).

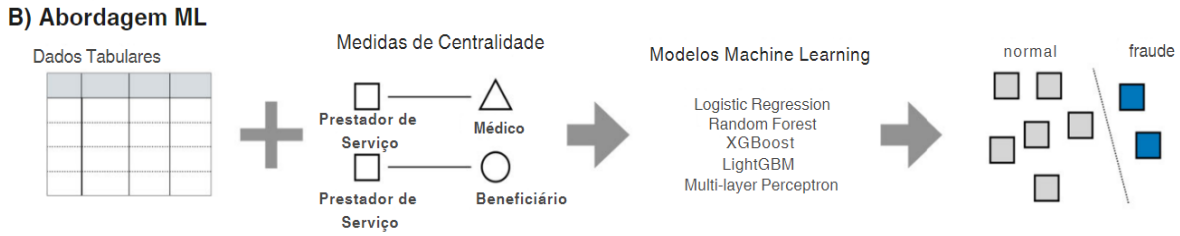
Figura 17 – Desenvolvimento do Modelo de Fraude : Abordagem *Graph Neural Networks*



Fonte: Adaptado YOO et al (2023) [64].

A Figura 18 mostra um esquemático da segunda abordagem. Na abordagem de *Machine Learning*, extraímos dois tipos de relações entre Prestadores-Médicos e Prestadores-Beneficiários do conjunto de dados enriquecidos e, em seguida, criamos dois grafos considerando essas relações, como mostrado na Figura 18, para introduzir informações de centralidade de grafos como características de entrada para os algoritmos de *Machine Learning* (Seção 4.4).

Figura 18 – Desenvolvimento do Modelo de Fraude : Abordagem *Machine Learning*



Fonte: Adaptado YOO et al (2023) [64].

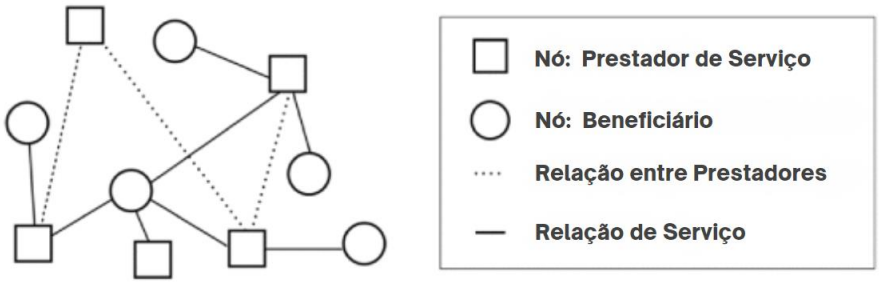
4.3 Modelos de Detecção de Fraude Baseados *Graph Neural Networks* (GNNs)

4.3.1 Construção dos Grafos

Quando os dados enriquecidos são fornecidos, eles devem ser convertidos em dados estruturados em grafos para o treinamento das *Graph Neural Networks*. No entanto, não há um formato padronizado para a representação dos grafos em um determinado conjunto de dados específico. Portanto, transformar um conjunto de dados em grafos é uma espécie de atividade de modelagem, e a melhor representação é aquela que absorve o algoritmo de interesse [77].

A Figura 19 mostra o conjunto de dados estruturados em grafos neste estudo. Estabelecemos os prestadores de serviços e os beneficiários como nós e as conexões entre prestadores, beneficiários e médicos como arestas. As relações entre os prestadores de serviços e os beneficiários foram conectadas através das arestas “Relação de Serviço”. As relações entre os prestadores de serviços foram conectadas através das arestas “Relação entre Prestadores”, extraídas da projeção do grafo entre prestadores. Neste estudo, os modelos de classificação de nós foram construídos utilizando este grafo para identificar os prestadores de serviços médicos fraudulentos por meio de modelos GNN.

Figura 19 – Estrutura dos Dados em Grafos



Fonte: Elaborado pelo Autor.

4.3.2 Modelos *Graph Neural Network* (GNNs)

As *Graph Neural Networks* (GNNs) têm se destacado como uma ferramenta robusta na detecção de fraudes, capazes de revelarem possíveis nós fraudulentos por meio da agregação de informações de vizinhança através de diferentes relações e combinações. De maneira geral, uma GNN aprende através de um processo denominado "passagem de mensagens", que consiste na "agregação" de informações dos nós vizinhos e na "atualização" dessas informações para a camada subsequente. Suponha que $\boldsymbol{H}_{\boldsymbol{u}^{(l)}}$ seja a representação do nó *u* na *l-ésima* camada da GNN. $\boldsymbol{H}_{\boldsymbol{u}^{(l+1)}}$ na próxima camada (*l*+1) é atualizado com base na camada (*l*) da seguinte forma:

$$\boldsymbol{H}_{\mathrm{u}}^{(\mathrm{l}+1)} = \boldsymbol{Up}^{(\mathrm{l})}\left(\boldsymbol{H}_{u}^{(l)}, \boldsymbol{Agg}^{(\mathrm{l})}\left(\left\{\mathrm{H}_{\mathrm{v}}^{(\mathrm{l})}, \forall_{\mathrm{v}} \in \boldsymbol{N}(u)\right\}\right)\right) \tag{4.1}$$

onde $\boldsymbol{N}(u)$ denota os nós vizinhos do nó *u*, $\boldsymbol{Agg}$ é um operador que coleta informações, através de uma rede neural, sobre nós vizinhos utilizando operações de agregação como média, soma e máximo. E $\boldsymbol{Up}$ é um operador que transforma as mensagens agregadas na próxima camada [113].

Os modelos GNN considerados neste TCC variaram de acordo com a estrutura de atualização na equação (4.1). O GCN coleta e combina informações dos nós próximos usando uma técnica chamada convolução. Depois, ele ajusta essas informações para que fiquem padronizadas e atualiza cada nó com esses dados novos e organizados. O GraphSAGE propõe um método de amostragem que define um tamanho fixo para os nós vizinhos e emprega diversas técnicas — como média, soma, máximo e *Recurrent Neural Network* (RNN), ou em português, Rede Neural Recorrente — para agregar as informações dos nós selecionados [74]. O GAT usa

um sistema de atenção para identificar quais nós são mais importantes, atribuindo-lhes pesos específicos [99]. O HAN melhora a forma como os nós são representados ao considerar tanto as características individuais de cada nó quanto as relações entre eles, seguindo caminhos pré-definidos [101]. O HGT, por sua vez, ajusta os pesos de cada nó e de suas conexões de forma diferenciada, garantindo que a informação seja passada de maneira eficaz. Além disso, o HGT organiza as informações recebidas de cada nó de forma a beneficiar os nós que recebem essas informações [102].

Neste trabalho, um conjunto de dados enriquecido é utilizado para a construção de um conjunto de dados estruturados em grafos para considerar a relação entre prestadores de serviços, médicos e beneficiários. Além disso, desenvolvemos quatro modelos GNN - GraphSAGE, GAT, HAN e HGT - que são capazes de aprender com dados específicos e aplicar esse conhecimento para classificar novos nós de maneira eficaz [74][99][101][102], conforme Figura 17.

As GNNs sofrem instabilidades e "sobreajustes" (*overfitting*) em tarefas de classificação de nós à medida que a profundidade do modelo aumenta [114]. O "sobreajuste" ocorre quando utilizamos um modelo superparametrizado para ajustar uma distribuição com dados de treinamento limitados. O modelo que treinamos se ajusta bem aos dados de treinamento e validação, mas generaliza mal para os dados de teste [115]. O *Dropout* é uma técnica frequentemente utilizada para mitigar o sobreajuste, pois atenua o fenômeno de coadaptação entre os neurônios [119]. Desta forma, incorporamos o *Dropout* em todos os nossos modelos como uma estratégia eficaz para combater o "sobreajuste".

4.4 Modelos de Detecção de Fraude Baseados em Machine Learning

4.4.1 Medidas de Centralidade de Grafos

Uma vez que conhecemos a estrutura de um grafo, podemos gerar uma variedade de medidas valiosas para quantificar a centralidade, o nível de interações e a similaridade relacionada a outras qualidades da estrutura do grafo. A centralidade, é uma informação teórica de grafos, define o nível de importância de um nó dentro de uma rede [116].

Neste estudo, nosso objetivo é extrair as características de centralidade nodal de dois grafos: o grafo que obtém informações de arestas de prestadores e médicos e o grafo que obtém informações de arestas de prestadores e beneficiários. Calculamos a centralidade de grau, centralidade de autovetor, centralidade de proximidade e *PageRank* para a medida de centralidade. Em seguida, adicionamos essas medidas de centralidade como características de entrada dos modelos de aprendizado de máquina convencionais.

A centralidade de grau é a medida mais direta que conta o número de arestas conectadas a um nó. A centralidade de grau C_d é definida da seguinte forma:

$$C_d(v_i) = d_i \tag{4.2}$$

onde v_i é o *i*-ésimo nó, d_i é o grau do *i*-ésimo nó.

A centralidade de autovetor representa a centralidade de um nó refletindo a importância dos nós vizinhos por meio de uma matriz de adjacência. A centralidade de autovetor é proporcional à soma das centralidades dos nós vizinhos; assim, a definição da centralidade de autovetor C_e do nó v_i é a seguinte:

$$C_e(v_i) = \frac{1}{\lambda}\sum_{j=1}^{n} A_{i,j} C_e(v_j) \tag{4.3}$$

onde $A_{i,j}$ é a matriz de adjacência, e λ é um autovalor da matriz de adjacência. Note que o maior autovalor é selecionado usando o teorema de Perron–Frobenius, mesmo que uma matriz possa ter vários autovalores [116].

A centralidade de proximidade refere-se ao conceito de que quanto mais próximo está do nó central, mais rápido pode alcançar os outros nós. A centralidade de proximidade se torna maior à medida que os outros nós se aproximam porque é o inverso das distâncias médias mais curtas para os outros nós. A centralidade de proximidade C_c é definida da seguinte forma:

$$C_c(v_i) = \frac{1}{\overline{I_{vi}}} \tag{4.4}$$

onde $\overline{I_{vi}}$ é o comprimento médio do caminho mais curto do nó vi para outros nós.

O *PageRank* é uma medida de centralidade que melhora as limitações da centralidade de Katz; quando um nó se torna central em uma rede, ele passa sua centralidade para todos os seus nós vizinhos. Ao calcular a influência de cada nó usando o *PageRank*, o número de arestas

de saída é usado como um fator de normalização para evitar a propagação de uma importância excessivamente alta. O PageRank C_p é definido da seguinte forma:

$$C_p(v_i) = \alpha \sum_{j=1}^{n} A_{j,i} \left(\frac{C_p(v_j)}{d_j^{\text{out}}} \right) + \beta \tag{4.5}$$

onde α é uma constante, β é o termo de viés que evita o valor zero de centralidade, e d_j^{out} é o número de arestas de saída (grau de saída).

4.4.2 Modelos de *Machine Learning*

Em nosso estudo, buscamos avaliar a eficácia dos algoritmos convencionais de *Machine Learning* na detecção de fraudes no Medicare, e se essa eficácia poderia ser aprimorada com a adição de características derivadas de grafos. Essa análise é importante para entender como as técnicas avançadas de análise de grafos podem contribuir para aprimorar as estratégias existentes de detecção de fraudes. Em contraste com as *Graph Neural Networks* (GNNs), que aprendem diretamente a partir de conjuntos de dados estruturados como grafos, os modelos de *Machine Learning* que utilizamos se beneficiam do conjuntos de dados enriquecidos que elaboramos e das características de centralidade de grafos como variáveis independentes adicionais. Nesse contexto, calculamos medidas de centralidade de grafos, incluindo centralidade de grau, centralidade de autovetor, centralidade de proximidade e *PageRank*.

Para a tarefa de classificação voltada para a detecção de fraudes no Medicare, aplicamos diversos modelos de Machine Learning, incluindo *Logistic Regression*, *Random Forest*, *XGBoost*, *LightGBM* e *Multi-Layer Perceptron*, conforme Figura 18. Esses algoritmos são amplamente utilizados na modelagem de detecção de fraudes financeiras [117] e [118]. Comparamos o desempenho de cada um desses modelos para entender melhor as suas capacidades e limitações no contexto específico da detecção de fraudes no Medicare.

5 RESULTADOS

5.1 Estruturação

5.1.1 Medidas de Desempenho para os Modelos

Para avaliar a eficácia dos modelos *Graph Neural Networks* (GraphSAGE, GAT, HAN e HGT), e dos modelos de *Machine Learning* (*Logistic Regression*, *Random Forest*, XGBoost, LightGBM e *Multi-Layer Perceptron*) com características de grafos para a detecção de fraudes no Medicare, utilizaremos quatro métricas de desempenho:

- **Precisão**: É a proporção de *True Positive* (TP), ou em português, Verdadeiros Positivos, sobre o Total de Previsões Positivas: (TP) + *False Positive* (FP), ou em português, Falsos Positivos. A precisão é uma medida da qualidade das previsões positivas do modelo. A fórmula para a precisão é:

$$\text{Precisão} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5.1}$$

- **Revocação**: Também conhecida como sensibilidade, é a proporção de *True Positive* (TP) sobre o Total de casos Positivos Reais: (TP) + *False Negative* (FN), ou em português, Falsos Negativos. A Revocação é uma medida da capacidade do modelo de encontrar todos os exemplos positivos. A fórmula para a Revocação é:

$$\text{Revocação} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5.2}$$

- **F1 Score**: É a média harmônica entre a Precisão e a Revocação. Quanto mais próxima de 1, melhor é o desempenho do modelo. A fórmula para o F1 Score é:

$$F1\ Score = \frac{2}{\frac{1}{Precisão} + \frac{1}{Revocação}} \tag{5.3}$$

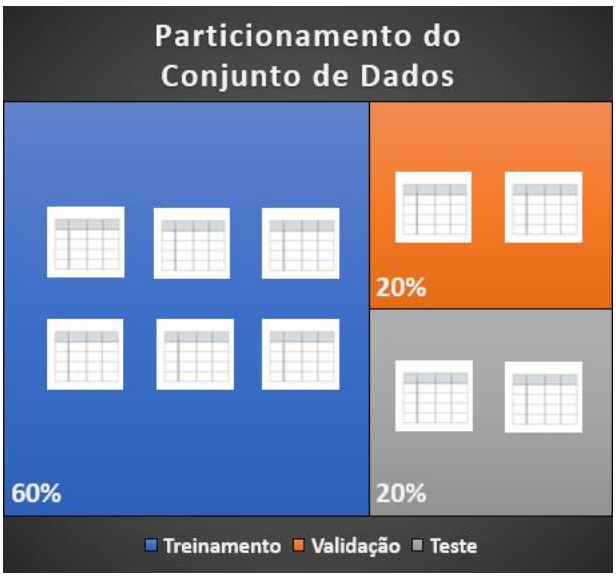
- **AUROC**: Esta métrica é a área sob a curva do gráfico de taxa de verdadeiros positivos (sensibilidade) versus taxa de falsos positivos (1 - especificidade). Quanto maior o valor de AUROC e mais próximo de 1, melhor a precisão do modelo.

Um excelente modelo é caracterizado pelo seu desempenho através dos valores elevados de Precisão e Revocação [120]. Desta forma, por meio do F1 Score, que abrange a média harmônica entre a Precisão e a Revocação, selecionamos o modelo que apresentou a melhor pontuação deste estudo.

5.1.2 Conjuntos de Treinamento, Validação e Teste

Em consonância com a metodologia estabelecida em [33], a construção de um modelo robusto para detecção de fraudes no Medicare requer um particionamento estratégico dos dados em três conjuntos distintos:

Figura 20 – Particionamento do Conjunto de Dados



Fonte: Elaborado pelo Autor.

- **Conjunto de Treinamento:** Composto por aproximadamente 60% dos dados, este conjunto é fundamental para o aprendizado do modelo. Através de técnicas de otimização, os parâmetros do modelo são ajustados iterativamente, buscando minimizar a função de perda e capturar os padrões subjacentes aos dados, incluindo relações legítimas e potenciais indicadores de fraude.

- **Conjunto de Validação:** Representando cerca de 20% dos dados, este conjunto atua como um "termômetro" durante o treinamento. O desempenho do modelo é avaliado periodicamente neste conjunto, permitindo a seleção criteriosa de hiperparâmetros e a adoção de estratégias para mitigar o *overfitting*, fenômeno em que o modelo se torna excessivamente especializado nos dados de treinamento, comprometendo sua capacidade de generalização.

- **Conjunto de Teste:** O restante dos dados, aproximadamente 20% das informações, constitui o conjunto de teste, reservado exclusivamente para a avaliação final do modelo. Uma vez que o modelo não teve acesso a esses dados durante o treinamento e a validação, o desempenho neste conjunto oferece uma estimativa realista de sua capacidade de generalização e, portanto, de sua efetividade ou não na detecção de fraudes em novos casos. A distribuição dos dados foi parametrizada com foco nos prestadores de serviços, pois eles desempenham um papel central na rede de relacionamentos do sistema Medicare.

A Tabela 8 detalha a estrutura do grafo para cada um desses conjuntos, especificando o número de nós (Prestadores e Beneficiários) e arestas (as Relações de Serviços e as Relações entre os Prestadores). O *setup* da distribuição da quantidade do conjunto de dados foi direcionado ao nó dos prestadores de serviços conforme abaixo:

Tabela 8 - Estrutura Básica do Grafo de Cada Conjunto de Dados

Descrição	Qtd. do conjunto de Dados		
	Treinamento	**Validação**	**Teste**
Número de Nós: Prestador de Serviços	5.843	1.948	1.948
Número de Nós: Beneficiários	200.225	107.226	96.512
Número de Arestas: Relação de Serviços	594.519	220.578	186.970
Número de Arestas: Relação entre Prestadores	6.275	616	555

Fonte: Elaborado pelo Autor.

5.2 Desempenho

5.2.1 Desempenho dos Modelos GNN

Elaboramos os seguintes modelos de GNN, para detectar as fraudes no Medicare: GraphSAGE, GAT, HAN e HGT, utilizando (60%) do conjunto de dados para treinamento e (20%) para validação. O desempenho de cada modelo foi comparado e avaliado utilizando (20%) do conjunto de dados para teste.

A necessidade de empregar técnicas de amostragem em GNNs decorre da natureza inerentemente complexa do aprendizado em grafos, onde a escalabilidade computacional pode ser comprometida ao lidar com *datasets* de grande porte. O processo de amostragem viabiliza um treinamento mais eficiente ao selecionar, de forma estratégica, subconjuntos relevantes do

grafo para cada iteração de atualização dos parâmetros do modelo, mitigando o custo computacional associado ao processamento da estrutura completa do grafo. Desta forma, utilizamos três métodos de amostragem distintos:

- **Sem Mini-Lote**: O modelo é treinado usando todo o conjunto de dados de uma só vez.

- **Mini-Lote (Vizinhança)**: O conjunto de dados é dividido em pequenos lotes, ou "mini lotes", e os resultados em cada mini lote são selecionados com base em sua proximidade ou similaridade.

- **Mini-Lote (Heterogêneo)**: O conjunto de dados é dividido em mini lotes, mas a seleção dos resultados é feita de forma não uniforme, ou seja, algumas amostras têm mais probabilidades de serem escolhidas do que outras, dependendo de suas características.

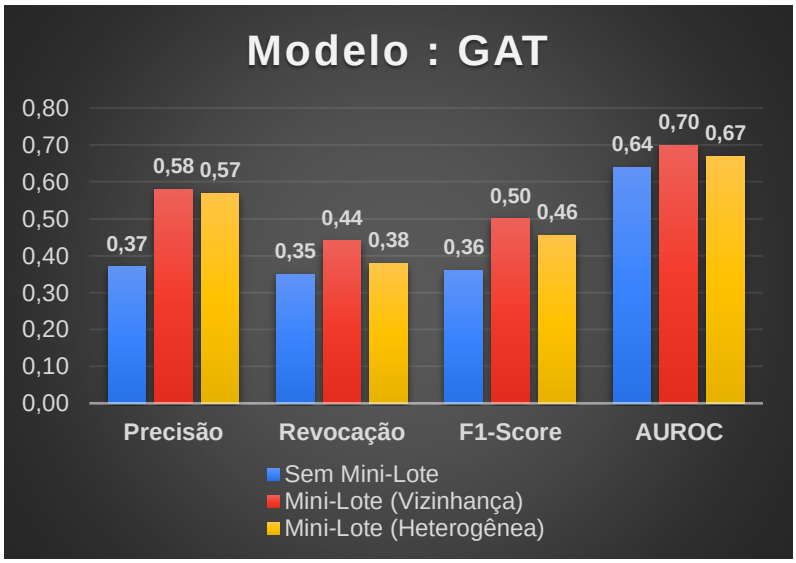
A Tabela 9 apresenta o desempenho dos modelos GraphSAGE, GAT, HAN e HGT para a detecção de fraude no Medicare. Cada um dos modelos GNN foi treinado, aproximadamente, quinze vezes. O valor médio das métricas de desempenho foi calculado com base no conjunto de dados de teste. O modelo que apresentou o melhor F1 Score foi selecionado, conforme ilustrado na Tabela 9. O modelo GNN com o melhor desempenho foi o HAN, obtido a partir da amostragem Mini-Lote (Heterogênea). Este modelo apresentou os seguintes valores para Precisão, Revocação, F1-Score e AUROC: 0,53; 0,54; 0,53 e 0,74, respectivamente.

Tabela 9 - Desempenho dos modelos GNN

Modelo	Amostragem	Precisão	Revocação	F1-Score	AUROC
GAT	Sem Mini-Lote	0,37	0,35	0,36	0,64
	Mini-Lote (Vizinhança)	0,58	0,44	0,50	0,70
	Mini-Lote (Heterogênea)	0,57	0,38	0,46	0,67
GraphSAGE	Sem Mini-Lote	0,44	0,48	0,46	0,71
	Mini-Lote (Vizinhança)	0,53	0,49	0,51	0,71
	Mini-Lote (Heterogênea)	0,62	0,39	0,48	0,66
HAN	Sem Mini-Lote	0,51	0,52	0,51	0,72
	Mini-Lote (Vizinhança)	0,62	0,36	0,46	0,56
	Mini-Lote (Heterogênea)	0,53	**0,54**	**0,53**	**0,74**
HGT	Sem Mini-Lote	0,50	0,48	0,49	0,71
	Mini-Lote (Vizinhança)	0,66	0,40	0,50	0,69
	Mini-Lote (Heterogênea)	**0,70**	0,34	0,46	0,66

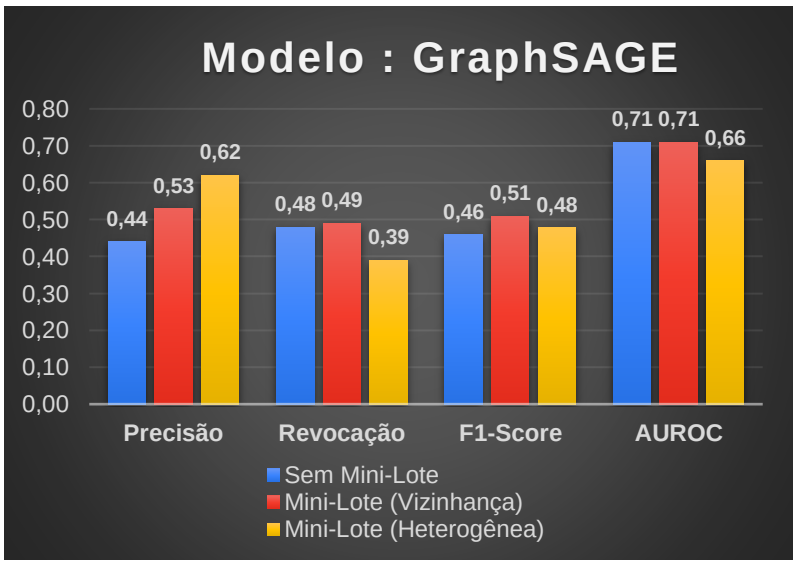
Fonte: Elaborado pelo Autor.

Figura 21 - Desempenho do modelo GAT



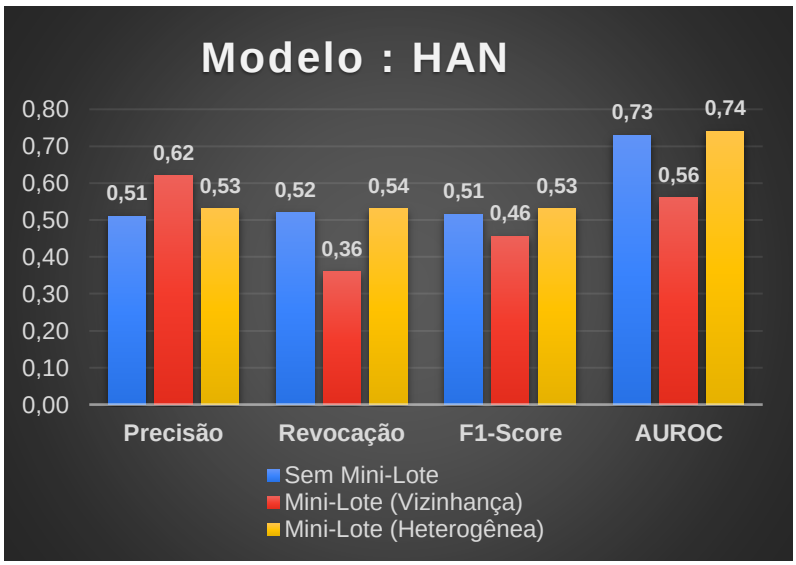
Fonte: Elaborado pelo Autor.

Figura 22 - Desempenho do modelo GraphSAGE



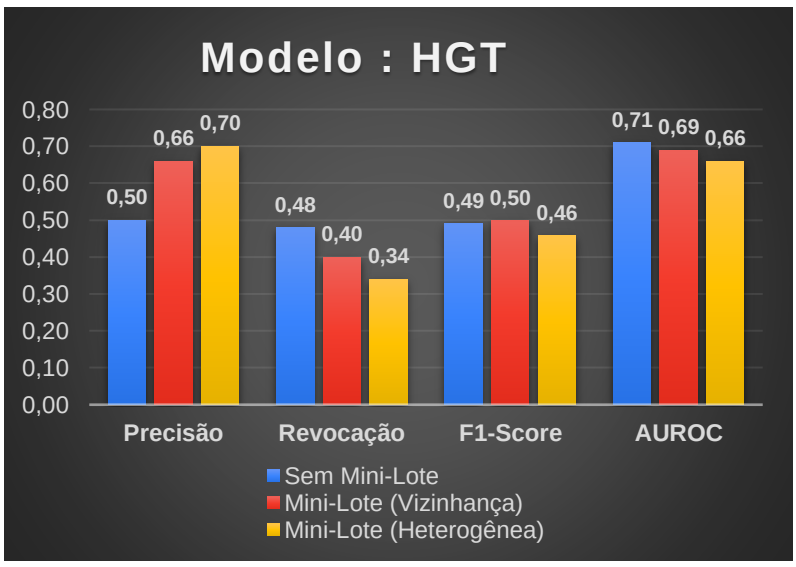
Fonte: Elaborado pelo Autor.

Figura 23 - Desempenho do modelo HAN



Fonte: Elaborado pelo Autor.

Figura 24 - Desempenho do modelos HGT



Fonte: Elaborado pelo Autor.

5.2.2 Desempenho dos Modelos de *Machine Learning*

Elaboramos os seguintes modelos de *Machine Learning*, para detectar as fraudes no Medicare: *Logistic Regression*, *Random Forest*, XGBoost, LightGBM e *Multi-Layer Perceptron*, utilizando (60%) do conjunto de dados para treinamento e (20%) para validação. O desempenho de cada modelo foi comparado e avaliado utilizando (20%) do conjunto de dados para teste.

Para avaliarmos o desempenho dos modelos de *Machine Learning*, utilizamos como referência os dados com e sem recursos de centralidade de grafos como *features* de entrada e avaliamos a performance dessas *features*. As características de centralidade nodal no grafo que obtém informações de arestas entre Prestadores de Serviços e Médicos e no grafo que obtém informações de arestas entre Prestadores de Serviços e Beneficiários foram extraídas, onde calculamos a centralidade de grau, centralidade de autovetor, centralidade de proximidade e PageRank para a medida de centralidade. Avaliamos o desempenho dos modelos da seguinte forma:

- **Referência Padrão – Sem Recursos de Centralidade:** Este é o modelo mais básico, que não leva em consideração as relações entre os diferentes atores envolvidos no Medicare (Prestadores de Serviços, Médicos e Beneficiários). Ele serve como ponto de referência para avaliar se a inclusão de informações sobre a rede de relacionamentos melhora a capacidade de detecção de fraudes.

- **Referências com Medidas de Centralidade (Prestadores-Médicos) :** Este modelo incorpora as métricas que quantificam a importância de cada Prestador de Serviço dentro da rede de relacionamentos com os Médicos. A ideia é que Prestadores com conexões mais "centrais" ou influentes nessa rede podem ter maior probabilidade de envolvimento em fraudes.

- **Referências com Medidas de Centralidade (Prestadores-Beneficiários):** Similar ao anterior, mas aqui as medidas de centralidade são calculadas a partir da rede que conecta Prestadores de Serviços e Beneficiários. A lógica é a mesma: Prestadores mais "centrais" ou influentes nessa rede podem ser mais propensos a fraudes.

A Tabela 10 apresenta o desempenho dos modelos *Logistic Regression*, *Random Forest*, XGBoost, LightGBM e *Multi-Layer Perceptron* para a detecção de fraude no Medicare. O

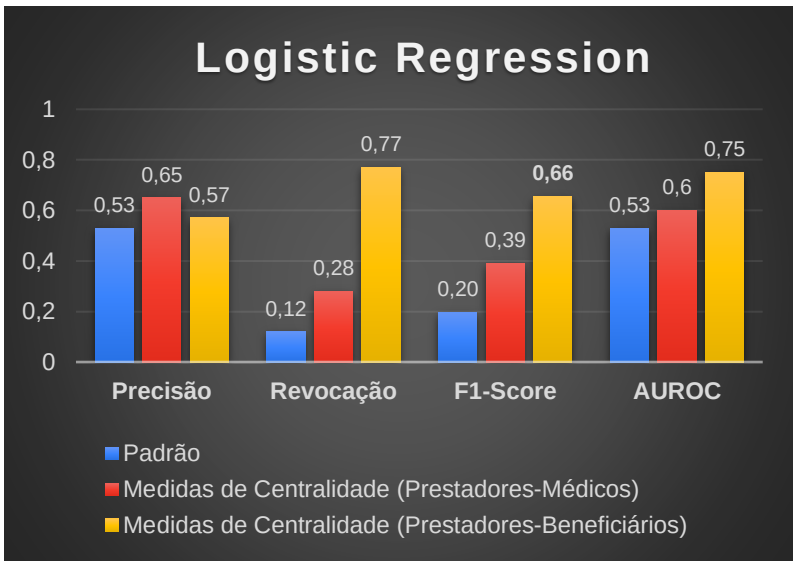
modelo que apresentou o melhor F1 Score foi selecionado, conforme ilustrado na Tabela 10. O modelo de *Machine Learing* com o melhor desempenho foi o *Logistic Regression*, obtido a partir da feature de entrada com recurso de centralidade (Prestadores-Beneficiários). Este modelo apresentou os seguintes valores para Precisão, Revocação, F1-Score e AUROC: 0,57; 0,77; 0,66 e 0,75, respectivamente.

Tabela 10 - Desempenho dos modelos de aprendizado de máquina com e sem recursos de centralidade do gráfico

Model	Referência	Precisão	Revocação	F1-Score	AUROC
Logistic Regression	Padrão	0,53	0,12	0,20	0,53
	Medidas de Centralidade (Prestadores-Médicos)	**0,65**	0,28	0,39	0,6
	Medidas de Centralidade (Prestadores-Beneficiários)	0,57	0,77	**0,66**	**0,75**
Random Forest	Padrão	0,46	0,25	0,32	0,56
	Medidas de Centralidade (Prestadores-Médicos)	0,52	0,35	0,42	0,6
	Medidas de Centralidade (Prestadores-Beneficiários)	0,43	0,76	0,55	0,64
XGBoost	Padrão	0,49	0,33	0,39	0,59
	Medidas de Centralidade (Prestadores-Médicos)	0,45	0,38	0,41	0,59
	Medidas de Centralidade (Prestadores-Beneficiários)	0,46	0,83	0,59	0,6
LightGBM	Padrão	0,51	0,34	0,41	0,6
	Medidas de Centralidade (Prestadores-Médicos)	0,52	0,31	0,39	0,59
	Medidas de Centralidade (Prestadores-Beneficiários)	0,48	**0,84**	0,61	0,69
Multi-Layer Perceptron	Padrão	0,43	0,2	0,27	0,54
	Medidas de Centralidade (Prestadores-Médicos)	0,31	0,29	0,30	0,48
	Medidas de Centralidade (Prestadores-Beneficiários)	0,48	0,66	0,56	**0,75**

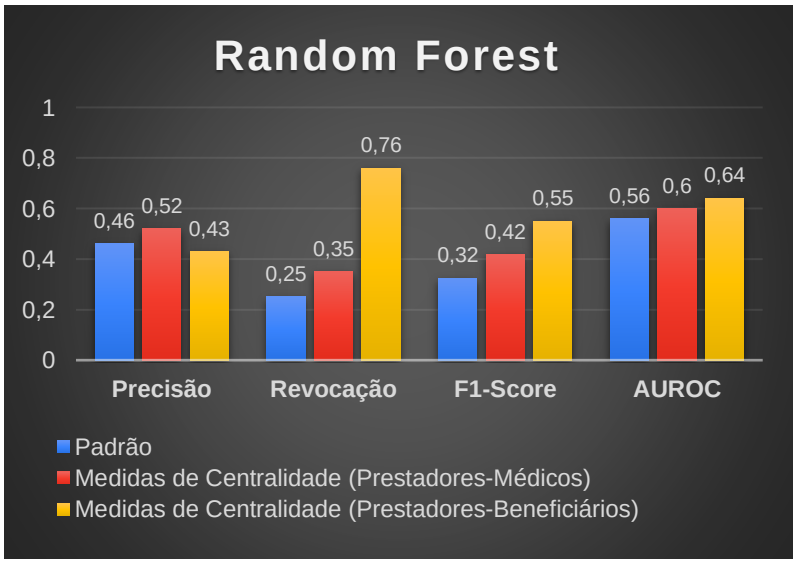
Fonte: Elaborado pelo Autor.

Figura 25 - Desempenho do modelo *Logistic Regression*



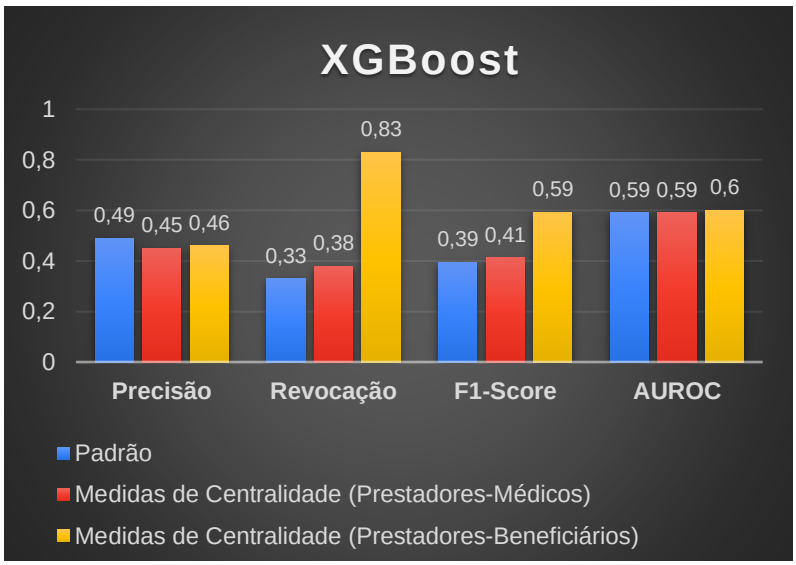
Fonte: Elaborado pelo Autor.

Figura 26 - Desempenho do modelo *Random Forest*



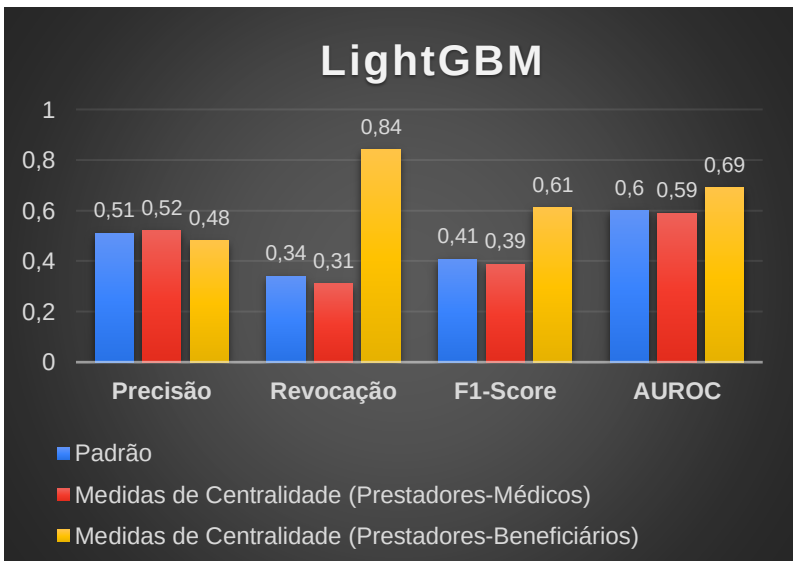
Fonte: Elaborado pelo Autor.

Figura 27 - Desempenho do modelo *XGBoost*



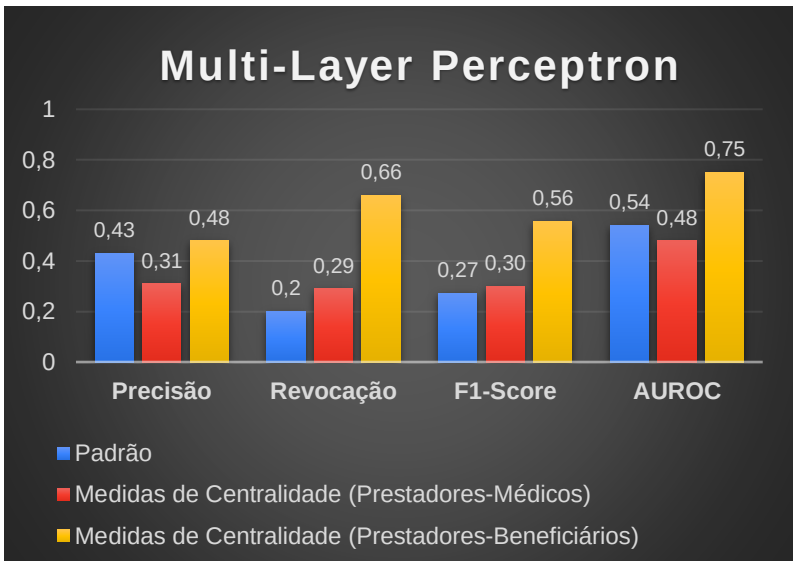
Fonte: Elaborado pelo Autor.

Figura 28 - Desempenho do modelo *LightGBM*



Fonte: Elaborado pelo Autor.

Figura 29 - Desempenho do modelo *Multi-Layer Perceptron*



Fonte: Elaborado pelo Autor.

A inclusão de medidas de centralidade de grafos como features de entrada melhorou significativamente o desempenho dos modelos de *Machine Learning*, principalmente em termos de Revocação e F1-Score.

As medidas de centralidade obtidas da relação entre os Prestadores de Serviços e os Beneficiários tiveram um impacto maior no desempenho do que as medidas da relação entre os Prestadores de Serviços e os Médicos, indicando que a relação (Prestadores-Beneficiários) é muito importante para detectar as fraudes.

5.2.3 Desempenho dos Modelos GNN x *Machine Learning*

Ao analisarmos o desempenho dos modelos *Graph Neural Networks* (GraphSAGE, GAT, HAN e HGT), e dos modelos de *Machine Learning* (*Logistic Regression*, *Random Forest*, XGBoost, LightGBM e *Multi-Layer Perceptron*) com características de grafos para a detecção de fraudes no Medicare, obtemos os seguintes resultados:

O modelo HAN apresentou o melhor F1-Score entre os modelos baseados em GNN a partir da amostragem Mini-Lote (Heterogênea), como ilustrado na Figura 22.

O modelo *Logistic Regression* obteve o melhor desempenho em termos de F1-Score entre os modelos de *Machine Learning* ao incluirmos as *features* de centralidade de grafos da rede (Prestadores-Beneficiários), como ilustrado na Figura 24.

Em suma, os modelos de *Machine Learning* que incluíram as de medidas de centralidade de grafos da rede (Prestadores-Beneficiários) como features de entrada demonstraram um desempenho superior aos modelos GNN em relação à Revocação e ao F1-Score.

Especificamente, o melhor modelo de *Machine Learning*(*Logistic Regression*) demonstra uma performance superior à **24,52%** no F1-Score, quando comparado em relação ao melhor modelo GNN (HAN).

Tabela 11 - Diferença de Desempenhos dos Modelos ML x GNN

Modelo	Referência/Amostragem	Precisão	Revocação	F1-Score	AUROC
Modelos Machine Learning	Padrão	0,43-0,53	0,12-0,34	0,2-0,41	0,53-0,6
	Medidas de Centralidade (Prestadores-Médicos)	0,31-0,65	0,28-0,38	0,3-0,42	0,48-0,6
	Medidas de Centralidade (Prestadores-Beneficiários)	0,43-0,57	0,66-**0,84**	0,55-**0,66**	0,6-**0,75**
Modelos GNNs	Sem Mini-Lote	0,37-051	0,35-0,52	0,36-0,51	0,64-0,73
	Mini-Lote (Vizinhança)	0,53-0,66	0,36-0,49	0,46-0,51	0,56-0,71
	Mini-Lote (Heterogênea)	0,53-**0,7**	0,34-0,54	0,46-0,53	0,66-0,74
HAN	Mini-Lote (Heterogênea)	0,53	0,54	0,53	0,74
Logistic Regression	Medidas de Centralidade (Prestadores-Beneficiários)	0,57	0,77	**0,66**	**0,75**

Fonte: Elaborado pelo Autor.

6 CONCLUSÕES

A fraude no setor de saúde representa um desafio global que impacta significativamente o custo e a qualidade do atendimento aos usuários. A detecção precoce e precisa dessas atividades fraudulentas é fundamental para mitigar os seus impactos negativos. Embora o tema seja uma preocupação crescente, governos e empresas têm buscado soluções tecnológicas para enfrentar esse problema. A Inteligência Artificial (IA) surge como uma ferramenta promissora nesse contexto, oferecendo vantagens como a capacidade de processar grandes volumes de dados e identificar padrões complexos.

Este trabalho concentrou-se na apuração de fraudes cometidas por prestadores de serviços de saúde em relação à falsificação de laudos de exames e em falsos atendimentos utilizando os dados públicos do sistema americano de saúde Medicare. O processo de coleta e preparação destes dados envolveu a combinação e o enriquecimento de diferentes conjuntos de dados, como informações sobre internações, atendimentos ambulatoriais, beneficiários e prestadores de serviços médicos.

O potencial da Inteligência Artificial foi especificamente explorado através de duas abordagens principais: Modelos de *Machine Learning* (como *Logistic Regression*, *Random Forest*, *XGBoost*, *LightGBM*, and *Multi-Layer Perceptron*) e Modelos de *Graph Neural Networks* (como GraphSAGE, GAT, HAN e HGT). Adicionalmente, foi investigada a inclusão de características derivadas de grafos no aprimoramento da eficácia dos Modelos de *Machine Learning*.

O estudo analisou fraudes no setor de saúde modelando as interações entre prestadores de serviços, médicos e beneficiários em um grafo. O modelo HAN, que utiliza redes neurais para analisar grafos heterogêneos, obteve o melhor desempenho na identificação de fraudes dentre os modelos GNNs testados. No entanto, os modelos de *Machine Learning* tradicionais, enriquecidos com métricas de centralidade da rede entre prestadores de serviços e beneficiários, superaram os modelos GNNs em termos de Revocação e F1-Score, com destaque para a *Logistic Regression*. Os resultados indicaram que a integração de técnicas de aprendizado de máquina com métricas de centralidade dos nós em redes de relacionamentos pode ser uma estratégia eficaz para a detecção de fraudes no setor de saúde.

O presente trabalho, apesar de sua contribuição inicial, abre diversas oportunidades para investigações futuras. É importante aprimorar a capacidade de generalização dos modelos, abordando o desequilíbrio de classes frequentemente observado em conjuntos de dados de

fraude. Análises mais detalhadas sobre o impacto das propriedades dos grafos nos modelos de *Machine Learning*, bem como a investigação de outras variantes de GNNs, podem proporcionar novas perspectivas. A aplicação dessas metodologias em outros contextos de fraudes no setor de saúde para avaliar o seu potencial de aplicabilidade se faz necessária.

Os modelos de detecção de fraude desenvolvidos poderiam ser implementados e testados em um ambiente simulado de *Mobile Cloud Computing* (MCC), ou em português, Computação em Nuvem Móvel. As métricas como latência, largura de banda e *throughput* poderiam ser monitoradas para avaliar o impacto do MCC no desempenho dos modelos, especialmente em cenários com grande volume de dados ou em situações que poderiam exigir processamento em tempo real.

Por fim, é importante considerar a integração da inteligência artificial na tomada de decisões de governos e empresas para o desenvolvimento de soluções inovadoras e eficazes para combater este problema complexo e persistente.

REFERÊNCIAS

[1] BRASIL. **Constituição da República Federativa do Brasil de 1988.** Promulgada em 5 de outubro de 1988. Diário Oficial da União, Brasília, DF, 5 out. 1988. Art.196. Disponível em: <https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 13 Fev. 2024.

[2] BRASIL. **LEI Nº 8.080, DE 19 DE SETEMBRO DE 1990**. 1990. Disponível em: <https://www.planalto.gov.br/ccivil_03/leis/l8080.htm>. Acesso em: 10 Fev. 2024.

[3] AGÊNCIA NACIONAL DE SAÚDE SUPLEMENTAR, 2024. **Setor fecha 2023 com 51 milhões de beneficiários em planos de assistência médica**. Disponível em: <https://www.gov.br/ans/pt-br/assuntos/noticias/numeros-do-setor/setor-fecha-2023-com-51-milhoes-de-beneficiarios-em-planos-de-assistencia-medica>. Acesso em: 10 Fev. 2024.

[4] INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2023. **População brasileira**. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/populacao.html >. Acesso em: 10 Fev. 2024.

[5] INSTITUTO DE ESTUDOS DE SAÚDE SUPLEMENTAR, 2021. **Pesquisa de Satisfação de Beneficiários de Planos de Saúde.** Disponível em: < https://iess.org.br/biblioteca/pesquisa-iess/pesquisa-iess/pesquisa-iess-2021>. Acesso em: 13 Fev. 2024.

[6] AGÊNCIA NACIONAL DE SAÚDE SUPLEMENTAR, 2023. **Painel Contábil**. Disponível em: < https://www.ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor>. Acesso em: 13 Fev. 2024.

[7] IESS, 2023. **Fraudes e Desperdícios em Saúde Suplementar**. Disponível em: <https://iess.org.br/biblioteca/tds-e-estudos/estudos-especiais-externos/fraudes-e-desperdicios-em-saude-suplementar>. Acesso em: 13 Fev. 2024.

[9] INSTITUTO DE ESTUDOS DE SAÚDE SUPLEMENTAR,2018. **Fraudes e Desperdícios em Saúde Suplementar: estimativa 2017.** Disponível em: < https://iess.org.br/publicacao/blog/r-28-bilhoes-em-fraudes-e-desperdicios1.> Acesso em: 13 Fev. 2024.

[10] OECD, 2017. **Tackling Wasteful Spending on Health, OECD Publishing**. Disponível em: <http://dx.doi.org/10.1787/9789264266414-en. 13 Fev. 2024.

[11] BRASIL. **LEI Nº 13.709, DE 14 DE AGOSTO DE 2018**. Lei Geral de Proteção de Dados Pessoais. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm.> Acesso em: 13 Fev. 2024.

[12] CHANDOLA, V.; BANERJEE, A.; KUMAR, V. Anomaly detection: A survey**. ACM computing surveys,** v. 41, n. 3, p. 1-58, 2009. Disponível em: <https://dl.acm.org/doi/10.1145/1541880.1541882>. Acesso em: 13 Fev. 2024.

[13] Kushal, K., Kurup, G., Urolagin, S. (2021). Data Mining Techniques for Fraud Detection—Credit Card Frauds. In: Gao, XZ., Kumar, R., Srivastava, S., Soni, B.P. (eds) **Applications of Artificial Intelligence in Engineering. Algorithms for Intelligent Systems.** Springer, Singapore. Disponível em:< https://doi.org/10.1007/978-981-33-4604-8_10>. Acesso em: 13 Fev 2024.

[14] LI, J.; HUANG, K. Y.; JIN, J.; SHI, J. A survey on statistical methods for health care fraud detection**. Health care management science**, v. 13, n. 3, p. 273-287, 2010. Disponível em:<https://www.researchgate.net/publication/23290716_A_survey_on_statistical_methods_for_health_care_fraud_detection>. Acesso em: 13 Fev 2024.

[15] WANG, Y.; WEI, J.; WANG, B.; LI, T. A survey on data mining for healthcare from the past to the future**. Comput. Math. Methods Med., 2019.** Disponível em:< https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8630112/ >. Acesso em: 13 Fev. 2024.

[16] LU, J.; LIN, K.; CHEN, R.; LIN, M.; CHEN, X.; LU, P. (2023**).** Fraud Detection in Health Insurance Using an Attributed Heterogeneous Information Network with a Hierarchical Attention Mechanism. **Journal of Medical Systems**, v. 47, n. 2. Disponível em:< https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-023-02152-0>. Acesso em: 13 Fev. 2024.

[17] MEJIA, N. (2019). **AI for Healthcare Fraud Detection - Current Applications.** Emerj, v. 6, n. 1. Disponível em:<https://emerj.com/ai-sector-overviews/ai-for-health-insurance-fraud-detection-current-applications/>. Acesso em: 13 Fev 2024.

[18] SHAIK, THANVEER, et al. Remote patient monitoring using artificial intelligence: Current state, applications, and challenges. **WIREs Data Mining and Knowledge Discovery. 2023**, Vol. 13, 2. Disponível em: <https://arxiv.org/ftp/arxiv/papers/2301/2301.10009.pdf>. Acesso em: 13 Abr. 2024.

[19] LI, JIE, et al. A Study of Health Insurance Fraud in China and Recommendations for Fraud Detection and Prevention. **Journal of Organizational and End User Computing**. 2022, Vol. 34, 4 Disponível em: <https://www.igi-global.com/article/a-study-of-health-insurance-fraud-in-china-and-recommendations-for-fraud-detection-and-prevention/301271>. Acesso em: 13 Fev. 2024.

[20] RUSSEL, STUART J.; NORVIG, PETER. **Artificial Intelligence: a modern approach.** 3. ed. Upper Saddler River: Pearson Education, 2010.

[21] LAGE, FERNANDA DE CARVALHO**. Manual de inteligência artificial no direito brasileiro.** Salvador: Juspodivm, 2021. Disponível em:
< https://bdjur.stj.jus.br/jspui/handle/2011/154607>. Acesso em: 13 Fev. 2024.

[22] TEFFÉ, CHIARA SPADACCINI DE. AFFONSO, FILIPE JOSÉ MEDON. A utilização de inteligência artificial em decisões empresariais: notas introdutórias acerca da responsabilidade civil dos administradores. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (coordenadoras). **Inteligência artificial e Direito: ética, regulação e responsabilidade. 2. Ed. São Paulo: Thomson Reuters Brasil, 2020**. p. 463-499. Disponível em:< https://bdjur.stj.jus.br/jspui/bitstream/2011/136945/inteligencia_artificial_direito_frazao_2.ed.pdf>. Acesso em: 13 Fev. 2024.

[23] O'NEIL, CATHY. **Weapons of math destruction: how big data increases inequality and threatens democracy**. Nova Iorque: Crown Publishers, 2016. Edição digital. Disponível em:< https://www.jstor.org/stable/26732553 >. Acesso em: 13 Fev. 2024.

[24] SAMUEL, ARTHUR. L. Some Studies in *Machine Learning* Using the Game of Checkers. **IBM Journal of Research and Development.** [S.L.], v. 3, n. 3, jul. 1959. Disponível em: < https://ieeexplore.ieee.org/document/5392560 >. Acesso em: 13 Fev. 2024.

[25] SAHNI, NIKHIL, et al. THE POTENTIAL IMPACT OF ARTIFICIAL INTELLIGENCE ON HEALTHCARE SPENDING. **NBER WORKING PAPER SERIES. January 2023**. Disponível em:< https://www.nber.org/papers/w30857>. Acesso em: 03 Mar. 2024.

[26] OSAK, MILTON. Inteligência artificial, prática médica e a relação médico-paciente. **Revista de Administração em Saúde. Setembro 2018**, Vol. 8, 72. Disponível em:< https://cqh.org.br/ojs-2.4.8/index.php/ras/article/view/134/164>. Acesso em: 03 Mar. 2024.

[27] O'CONNOR, JOSEPH; SEYMOUR, JOHN. **Introdução à Programação Neurolingüística. Como Entender e Influenciar as Pessoas.** São Paulo: Summus Editorial, 1995

[28] KUMAR, KAMLESH, et al. Artificial Intelligence and *Machine Learning* Based Intervention in Medical Infrastructure: A Review and Future Trends**. Healthcare. 2023**, Vol. 11, 207. Disponível em:< https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9859198/>. Acesso em: 03 Mar. 2024.

[29] HAQUE, SAZU AND AKTER, JAHAN. Big Data Analytics & Artificial Intelligence In Management Of Healthcare: Impacts & Current State**. Management of Sustainable Development Journal. 2022**, Vol. 14, 1. Disponível em:< https://msdjournal.org/article/big-data-analytics-artificial-intelligence-in-management-of-healthcare-impacts-current-state/>. Acesso em: 03 Mar. 2024.

[30] LEE, PETER, BUBECK, SEBASTIEN AND PETRO, JOSEPH. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. **The new england journal o f medicine. 388, 2023**, pp. 1233-1239. Disponível em:
< https://www.nejm.org/doi/full/10.1056/NEJMsr2214184>. Acesso em: 03 Mar. 2024.

[31] CHEW, HAN SHI JOCELYN and ACHANANUPARP, PALAKORN. Perceptions and Needs of Artificial Intelligence in Health Care to Increase Adoption: Scoping Review. **Journal of Medical Internet Research. 2022**, Vol. 24, 1. Disponível em:< https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8800095/>. Acesso em: 03 Mar. 2024.

[32] YANG, CHRISTOPHER C. Explainable Artificial Intelligence for Predictive Modeling in Healthcare. **Journal of Healthcare Informatics Research. February 11, 2022**, 6, pp. 228-239. Disponível em:< https://link.springer.com/article/10.1007/s41666-022-00114-1>. Acesso em: 03 Mar. 2024.

[33] Y. YOO, D. SHIN, D. HAN, S. KYEONG, and J. SHIN, ''Medicare fraud detection using graph neural networks,'' in Proc. **Int. Conf. Electr., Comput. Energy Technol. (ICECET),**

Jul. 2022, pp. 1–5, Disponível em: <https://ieeexplore.ieee.org/document/9872963>. Acesso em: 03 Mar.2024

[34] ABRÀMOFF, MICHAEL, et al. A reimbursement framework for artificial intelligence in healthcare. npj **Digital Medicine. 2022**, Vol. 5, 72. Disponível em:< https://www.nature.com/articles/s41746-022-00621-w>. Acesso em: 03 Mar. 2024.

[35] MAHDI, SYED, et al. How does artificial intelligence impact digital healthcare initiatives? A review of AI applications in dental healthcare. **International Journal of Information Management Data Insights. 3, 2023**. Disponível em:< https://www.sciencedirect.com/science/article/pii/S2667096822000878 >. Acesso em: 03 Mar. 2024.

[36] LEE, J. H., KIM, D. H. and JEONG, S. N. **Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network**. Oral Diseases. 26, 2020, Vol. 1, pp. 152-158. Disponível em:< https://onlinelibrary.wiley.com/doi/10.1111/odi.13223 >. Acesso em: 03 Mar. 2024.

[37] YANG, H., et al. Deep Learning for Automated Detection of Cyst and Tumors of the Jaw in Panoramic Radiographs. **Journal of Clinical Medicine. 9, 2020**, Vol. 6, pp. 1-14. Disponível em:< https://www.mdpi.com/2077-0383/9/6/1839>. Acesso em: 03 Mar. 2024.

[38] KARIMIAN, GOLNAR, PETELOS , ELENA and EVERS, SILVIA M. A. A. The ethical issues of the application of artificial intelligence in healthcare: a systematic scoping review. **AI and Ethics. 2022**, 2, pp. 539-551. Disponível em: <https://link.springer.com/article/10.1007/s43681-021-00131-7>. Acesso em: 03 Mar. 2024.

[39] IESS - Texto para Discussão nº 99 – 2023 | **A INTELIGÊNCIA ARTIFICIAL NO SETOR DE SAÚDE: CONCEITOS E APLICAÇÕES**. Disponível em: < https://iess.org.br/sites/default/files/2023-12/TD%20-%2099.pdf>. Acesso em: 17 Mar. 2024.

[40] U.S. GOVERNMENT, U.S. CENTERS FOR MEDICARE & MEDICAID SERVICES. **The Official U.S. Government Site for Medicare**. Disponível em: <https://www.medicare.gov/>. Acesso em: 17 Mar. 2024.

[41] CENTERS FOR MEDICARE & MEDICAID SERVICES. **Trustees report & trust funds**. Disponível em: < https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/ReportsTrustFunds/index.html>. Acesso em: 17 Mar. 2024.

[42] CENTERS FOR MEDICARE & MEDICAID SERVICES. **Medicare enrollment dashboard.** Disponível em: < https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Dashboard/Medicare-Enrollment/Enrollment%20Dashboard.html>. Acesso em: 17 Mar. 2024.

[43] MORRIS L. Combating fraud in health care: an essential component of any cost containment strategy**. Health Aff. 2009**;28:1351–6. Disponível em: < https://doi.org/10.1377/hlthaff.28.5.1351>. Acesso em: 17 Mar. 2024.

[44] **COALITION AGAINST INSURANCE FRAUD: BY THE NUMBERS: FRAUD STATISTICS**. Disponível em: < https://www.insurancefraud.org/statistics.htm>. Acesso em: 17 Mar. 2024.

[45] ROSA, HELENA et al. Bancos de Dados de Saúde e Pesquisa: Prós e Contras da LGPD. In: DALLARI, Analluza; MONACO, Gustavo. **LGPD na Saúde. São Paulo (SP): Editora Revista dos Tribunais, 2021**. Disponível em: <https://www.jusbrasil.com.br/doutrina/lgpd-na-saude/1250396534>. Acesso em: 17 Mar. 2024.

[46] MEDICARE FRAUD & ABUSE: PREVENTION, DETECTION, AND REPORTING. **Centers for Medicare & Medicaid Services. 2017**. Disponível em: < https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/Fraud_and_Abuse.pdf>. Acesso em: 17 Mar. 2024.

[47] LI J, HUANG K-Y, JIN J, SHI J. A survey on statistical methods for health care fraud detection. **Health Care Manag Sci. 2008**;11:275–87. https://doi.org/10.1007/s10729-007-9045-4. >. Acesso em: 17 Mar. 2024.

[48] THE OFFICE OF THE NATIONAL COORDINATOR FOR HEALTH INFORMATION TECHNOLOGY**: OFFICE-BASED PHYSICIAN ELECTRONIC HEALTH RECORD ADOPTION.** Disponível em: < https://dashboard.healthit.gov/quickstats/quickstats.php>. Acesso em: 17 Mar. 2024.

[49] THE OFFICE OF THE NATIONAL COORDINATOR FOR HEALTH INFORMATION TECHNOLOGY**: ADOPTION OF ELECTRONIC HEALTH RECORD SYSTEMS AMONG U.S**. Non-Federal Acute Care Hospitals: 2008–2015. Disponível em: < https://dashboard.healthit.gov/evaluations/data-briefs/non-federal-acute-care-hospital-ehr-adoption-2008-2015.php>. Acesso em: 17 Mar. 2024.

[50] DUMBILL E. **What is Big Data? : an introduction to the Big Data landscape.** Disponível em: < http://radar.oreilly.com/2012/01/what-is-big-data.html>. Acesso em: 17 Mar. 2024.

[51] AHMED SE. Perspectives on Big Data analysis: methodologies and applications. **Providence: American Mathematical Society**; 2014.

[52] CENTERS FOR MEDICARE & MEDICAID SERVICES. **Medicare fee-for-service provider utilization & payment data physician and other supplier public use file: a methodological overview.** Disponível em: < https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/medicare-provider-charge-data/physician-and-other-supplier.html.> Acesso em: 17 Mar. 2024.

[53] OFFICE OF INSPECTOR GENERAL. **LEIE downloadable databases**. Disponível em: < https://oig.hhs.gov/exclusions/exclusions_list.asp. Acesso em: 17 Mar. 2024.

[54] LEEVY JL, KHOSHGOFTAAR TM, BAUDER RA, SELIYA N. A survey on addressing high-class imbalance in big data. **J Big Data. 2018**;5(1):42. Disponível em: < https://doi.org/10.1186/s40537-018-0151-6>. Acesso em: 01 Abr. 2024.

[55] VAN HULSE J, KHOSHGOFTAAR TM, NAPOLITANO A. Experimental perspectives on learning from imbalanced data. In: Proceedings of the **24th international conference on machine learning. ICML '07.** ACM, New York, NY, USA. 2007. pp. 935–42. Disponível em: < https://doi.org/10.1145/1273496.1273614 Acesso em: 01 Abr. 2024.

[56] LECUN Y, BENGIO Y, HINTON G. Deep learning. **Nature. 2015;521:436**. Disponível em: < https://www.nature.com/articles/nature14539>. Acesso em: 01 Abr. 2024.

[57] ABADI M, et al. **TensorFlow: large-scale *Machine Learning* on heterogeneous systems,2015**. Disponível em: < https://arxiv.org/abs/1603.04467>. Acesso em: 01 Abr. 2024.

[58] THEANO DEVELOPMENT TEAM. **''Theano: a Python framework for fast computation of mathematical expressions''**, **2016**. Disponível em: <arxiv:abs/1605.02688> Acesso em: 01 Abr. 2024.

[59] CHOLLET F, et al. **Keras, 2015**. Disponível em: < https://github.com/fchollet/keras>. Acesso em: 01 Abr. 2024.

[60] PASZKE A, GROSS S, CHINTALA S, CHANAN G, YANG E, DEVITO Z, LIN Z, DESMAISON A, ANTIGA L, LERER A. **Automatic differentiation in pytorch**. In: NIPS-W. 2017. Disponível em: < https://gwern.net/doc/www/openreview.net/54b149cefe1fbbe841975209b4840fa04086a701.pdf> Acesso em: 01 Abr. 2024.

[61] CHETLUR S, WOOLLEY C, VANDERMERSCH P, COHEN J, TRAN J, CATANZARO B, SHELHAMER E. **cudnn: efficient primitives for deep learning, 2014**. Disponível em: < https://arxiv.org/abs/1410.0759> Acesso em: 01 Abr. 2024.

[62] KRIZHEVSKY A, SUTSKEVER I, HINTON GE. Imagenet classification with deep convolutional neural networks. **Neural Inf Process Syst. 2012**. Disponível em: < https://doi.org/10.1145/3065386> Acesso em: 01 Abr. 2024.

[63] NVIDIA,2022. Disponível em: < **https://blog.nvidia.com.br/2022/10/31/o-que-sao-graph-neural-networks/** >. Acesso em: 01 Abr. 2024.

[64] Y. YOO et al.: Medicare Fraud Detection Using Graph Analysis. **IEEE Access, Jan,2023**. v.11, p. 88278 – 88294. Disponível em: < https://ieeexplore.ieee.org/abstract/document/10223204>. Acesso em: 01 Abr. 2024.

[65] BOYER, Carl B. **História da Matemática**. São Paulo: Blucher, 2012.

[66] GORI, M.; MONFARDINI, G.; SCARSELLI, F. A new model for learning in graph domains. In: Proceedings. **2005 IEEE International Joint Conference on Neural Networks**, 2005. [S.l.: s.n.], 2005. v. 2, p. 729–734 vol. 2. Disponível em: < https://ieeexplore.ieee.org/document/1555942>. Acesso em: 14 Abr. 2024.

[67] SCARSELLI, F.; GORI, M.; TSOI, A. C.; HAGENBUCHNER, M.; MONFARDINI, G. The graph neural network model. **IEEE Transactions on Neural Networks**, v. 20, n. 1, p. 61–80, 2009. Disponível em: < https://ieeexplore.ieee.org/document/4700287>. Acesso em: 14 Abr. 2024.

[68] ZHANG, M.; CHEN, Y. Link prediction based on graph neural networks. **In: Advances in Neural Information Processing Systems 31**. [S.l.]: Curran Associates, Inc., 2018. p. 5165–5175. Disponível em: <https://dl.acm.org/doi/10.5555/3327345.3327423>. Acesso em: 14 Abr. 2024.

[69] KIPF, T. N.; WELLING, M. Semi-supervised classification with graph convolutional networks. **5th International Conference on Learning Representations**, **Set 2017**. Disponível em: <https://arxiv.org/abs/1609.02907>. Acesso em: 14 Abr. 2024.

[70] HAMILTON, W. L. Graph representation learning. **Synthesis Lectures on Artificial Intelligence and Machine Learning, Morgan and Claypoo**l, v. 14, n. 3, p. 1–159, 2020. Disponível em: <https://www.cs.mcgill.ca/~wlh/grl_book/files/GRL_Book.pdf>. Acesso em: 14 Abr. 2024.

[71] LIMA DE CARVALHO, GUILHERME MICHEL**.** ***Graph Neural Networks*** **e suas aplicações aos sistemas complexos**. 2022. 90 p. Dissertação (Mestrado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022. Disponível em:
<https://repositorio.ufscar.br/bitstream/handle/ufscar/16236/Guilherme_Michel_Lima_de_Carvalho_revisada.pdf?sequence=1>. Acesso em: 14 Abr. 2024.

[72] KIPF, THOMAS.2016. **GRAPH CONVOLUTIONAL NETWORKS**. Disponível em: <https://tkipf.github.io/graph-convolutional-networks/ >. Acesso em: 14 Abr. 2024.

[73] DEFFERRARD, MICHAEL; BRESSON, XAVIER; VANDERGHEYNST, PIERRE. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. **Advances in Neural Information Processing Systems**, v. 29, 2016. Disponível em: < https://arxiv.org/abs/1606.09375>. Acesso em: 14 Abr. 2024.

[74] HAMILTON, W. L.; YING, R.; LESKOVEC, J. Inductive Representation Learning on Large Graphs. In: **ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 2017**. Disponível em: < https://arxiv.org/abs/1706.02216>. Acesso em: 14 Abr. 2024.

[75] YING, R.; HE, R.; CHEN, K.; EKSOMBATCHAI, P.; HAMILTON, W. L.; LESKOVEC, J. Graph Convolutional Neural Networks for Web-Scale Recommender Systems. In: KDD ’18: **The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining**, August 19–23, 2018, London, United Kingdom. Disponível em: <https://arxiv.org/abs/1806.01973>. Acesso em: 14 Abr. 2024.

[76] V. CHANDOLA, S. R. SUKUMAR, AND J. C. SCHRYVER, ‘‘Knowledge discovery from massive healthcare claims data,’’ in Proc. **19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining**, New York, NY, USA, Aug. 2013, pp. 1312–1320, . Disponível em: < https://dl.acm.org/doi/10.1145/2487575.2488205>. Acesso em: 14 Abr. 2024.

[77] L. K. BRANTING, F. REEDER, J. GOLD, and T. CHAMPNEY, ‘‘Graph analytics for healthcare fraud risk estimation,’’ in Proc. **IEEE/ACM Int. Conf. Adv. Social Netw. Anal.**

Mining (ASONAM), Aug. 2016, pp. 845–851, . Disponível em: < doi: 10.1109/ASONAM.2016.7752336>.
Acesso em: 14 Abr. 2024.

[78] J. LIU, E. BIER, A. WILSON, J. A. GUERRA-GOMEZ, T. HONDA, K. SRICHARAN, L. GILPIN, and D. DAVIES, ‘‘Graph analysis for detecting fraud, waste, and abuse in health-care data,’’ **AI Mag**., vol. 37, no. 2, pp. 33–46, Jun. 2016. Disponível em: < https://dl.acm.org/doi/10.1609/aimag.v37i2.2630>. Acesso em: 14 Abr. 2024.

[79] Z. XIE, R. ZHU, J. LIU, G. ZHOU, J. X. HUANG, and X. CUI, ‘‘GFCNet: Utilizing graph feature collection networks for coronavirus knowledge graph embeddings,’’ **Inf. Sci**., vol. 608, pp. 1557–1571, Aug. 2022. Disponível em: < https://www.sciencedirect.com/science/article/pii/S0020025522007204?via%3Dihub >.
Acesso em: 14 Abr. 2024.

[80] Y. DING, Z. ZHANG, X. ZHAO, D. HONG, W. LI, W. CAI, and Y. ZHAN, ‘‘AF2GNN: Graph convolution with adaptive filters and aggregator fusion for hyperspectral image classification,’’ **Inf. Sci**., vol. 602, pp. 201–219, Jul. 2022. Disponível em: < https://dl.acm.org/doi/abs/10.1016/j.ins.2022.04.006>. Acesso em: 14 Abr. 2024.

[81] S. FU, W. LIU, D. TAO, Y. ZHOU, and L. NIE, ‘‘HesGCN: Hessian graph convolutional networks for semi-supervised classification,’’ **Inf. Sci**., vol. 514,pp. 484–498, Apr. 2020. Disponível em: < https://dl.acm.org/doi/10.1016/j.ins.2019.05.019>.
Acesso em: 14 Abr. 2024.

[82] M. HERLAND, T. M. KHOSHGOFTAAR, and R. A. BAUDER, ‘‘Big data fraud detection using multiple medicare data sources,’’ **J. Big Data**, vol. 5, no. 1, p. 29, Dec. 2018. Disponível em: < https://journalofbigdata.springeropen.com/articles/10.1186/s40537-018-0138-3>. Acesso em: 14 Abr. 2024.

[83] G. WANG, S. XIE, B. LIU, and P. S. YU, ‘‘Review graph based online store review spammer detection,’’ in Proc. **IEEE 11th Int. Conf. Data Mining**, Dec. 2011, pp. 1242–1247. Disponível em: < https://dl.acm.org/doi/10.1109/ICDM.2011.124>.Acesso em: 14 Abr. 2024.

[84] A. BREUER, R. EILAT, and U. WEINSBERG, ‘‘Friend or faux: Graphbased early detection of fake accounts on social networks,’’ in Proc. **Web Conf**., New York, NY, USA, Apr. 2020, pp. 1287–1297. Disponível em: < https://dl.acm.org/doi/10.1145/3366423.3380204>. Acesso em: 14 Abr. 2024.

[85] J. JIA, B. WANG, and N. Z. GONG, ‘‘Random walk based fake account detection in online social networks,’’ in Proc. 47th Annu. **IEEE/IFIP Int. Conf. Dependable Syst. Netw**. (DSN), Jun. 2017, pp. 273–284. Disponível em: < https://ieeexplore.ieee.org/document/8023129>. Acesso em: 14 Abr. 2024.

[86] P. GIUDICI, B. HADJI-MISHEVA, and A. SPELTA, ‘‘Network based credit risk models,’’ **Qual. Eng**., vol. 32, no. 2, pp. 199–211, Apr. 2020. Disponível em: < https://www.tandfonline.com/doi/full/10.1080/08982112.2019.1655159>. Acesso em: 14 Abr. 2024.

[87] M. YILDIRIM, F. Y. OKAY, and S. ÖZDEMIR, ‘‘Big data analytics for default prediction using graph theory,’’ **Expert Syst. Appl**., vol. 176, Aug. 2021, Art. no. 114840. Disponível em: < https://dl.acm.org/doi/10.1016/j.eswa.2021.114840>. Acesso em: 14 Abr. 2024.

[88] D. LAUTIER and F. RAYNAUD, ‘‘Systemic risk in energy derivative markets: A graph-theory analysis,’’ **Energy J.**, vol. 33, no. 3, pp. 215–239, Jul. 2012. Disponível em: <http://www.jstor.org/stable/23268099>. Acesso em: 14 Abr. 2024.

[89] B. FENG, H. XU, W. XUE, and B. XUE, ‘‘**Every corporation owns its structure: Corporate credit ratings via graph neural networks,**’’ 2020. Disponível em: < https://arxiv.org/abs/2012.01933>. Acesso em: 14 Abr. 2024.

[90] S. WU, W. ZHANG, F. SUN, and B. CUI, ‘‘Graph neural networks in recommender systems: A survey,’’ **ACM Comput. Surv**., vol. 55, no.‘5, pp. 1–37, 2022. Disponível em: < https://dl.acm.org/doi/10.1145/3535101>. Acesso em: 14 Abr. 2024.

[91] C. CAO, S. LI, S. YU, and Z. CHEN, ‘‘**Fake reviewer group detection in online review systems**,’’ 2021. Disponível em: <https://arxiv.org/abs/2112.06403>. Acesso em: 14 Abr. 2024.

[92] J. WANG, R. WEN, C. WU, Y. HUANG, and J. XIONG, ‘‘FdGars: Fraudster detection via graph convolutional networks in online app review system,’’ in Proc. **Companion World Wide Web Conf**., New York, NY, USA, May 2019, pp. 310–316. Disponível em: < https://dl.acm.org/doi/10.1145/3308560.3316586>. Acesso em: 14 Abr. 2024.

[93] C. LIANG, Z. LIU, B. LIU, J. ZHOU, X. LI, S. YANG, and Y. QI, ‘‘Uncovering insurance fraud conspiracy with network learning,’’ in Proc. 42nd **Int. ACM SIGIR Conf. Res. Develop. Inf. Retr**., New York, NY, USA, Jul. 2019, pp. 1181–1184. . Disponível em: <https://dl.acm.org/doi/10.1145/3331184.3331372>. Acesso em: 14 Abr. 2024.

[94] Z. LIU, C. CHEN, X. YANG, J. ZHOU, X. LI, and L. SONG, ‘‘Heterogeneous graph neural networks for malicious account detection,’’ in Proc. 27th **ACM Int. Conf. Inf. Knowl. Manage**., New York, NY, USA, Oct. 2018, pp. 2077–2085. Disponível em: < https://dl.acm.org/doi/10.1145/3269206.3272010>. Acesso em: 14 Abr. 2024.

[95] C. LIU, L. SUN, X. AO, J. FENG, Q. HE, and H. YANG, ‘‘Intentionaware heterogeneous graph attention networks for fraud transactions detection,’’ in Proc. 27th **ACM SIGKDD** Conf. Knowl. Discovery Data Mining, New York, NY, USA, Aug. 2021, pp. 3280–3288. . Disponível em: < https://dl.acm.org/doi/10.1145/3447548.3467142>. Acesso em: 14 Abr. 2024.

[96] M. WEBER, G. DOMENICONI, J. CHEN, D. KARL I. WEIDELE, C. BELLEI, T. ROBINSON, and C. E. LEISERSON, ‘‘**Anti-money laundering in Bitcoin: Experimenting with graph convolutional networks for financial forensics**,’’ 2019.Disponível em: <https://arxiv.org/abs/1908.02591>. Acesso em: 14 Abr. 2024.

[97] W.-L. CHIANG, X. LIU, S. SI, Y. LI, S. BENGIO, and C.-J. HSIEH, ‘‘ClusterGCN: An efficient algorithm for training deep and large graph convolutional networks,’’ in Proc. 25th **ACM SIGKDD** Int. Conf. Knowl. Discovery Data Mining, Anchorage, AK, USA, Jul. 2019, pp. 257–266. Disponível em: < http://dx.doi.org/10.1145/3292500.3330925>. Acesso em: 14 Abr. 2024.

[98] A. VASWANI, N. SHAZEER, N. JONES, A. N. GOMEZ, L. KAISER, and I. POLOSUKHIN, ‘‘Attention is all you need,’’ in Proc. **Conference on Neural Information Processing Systems**, Long Beach, CA, USA, Dez. 2017. Disponível em: < https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf >. Acesso em: 14 Abr. 2024.

[99] P. VELIČKOVIĆ, G. CUCURULL, A. CASANOVA, A. ROMERO, P. LIOSS, and Y. BENGIO, **‘‘Graph attention networks**,’’ 2017. Disponível em: < https://arxiv.org/abs/1710.10903>. Acesso em: 14 Abr. 2024.

[100] J. ZHAO, X. WANG, C. SHI, B. HU, G. SONG, and Y. YE, ‘‘Heterogeneous graph structure learning for graph neural networks,’’ in Proc. **Innov. Appl. Artif. Intell. Conf**., May 2021, pp. 4697–4705. Disponível em: < https://ojs.aaai.org/index.php/AAAI/article/view/16600>. Acesso em: 14 Abr. 2024.

[101] S. MA, J.-W. LIU, X. ZUO, and W.-M. LI, ‘‘Heterogeneous graph gated attention network,’’ in Proc. **Int. Joint Conf. Neural Netw**., New York, NY, USA, Jul. 2021, pp. 1–6. Disponível em: < https://ieeexplore.ieee.org/document/9533711>. Acesso em: 14 Abr. 2024.

[102] Z. HU, Y. DONG, K. WANG, and Y. SUN, ‘‘Heterogeneous graph transformer,’’ in Proc. **Web Conf**., New York, NY, USA, Apr. 2020, pp. 2704–2710. Disponível em: < https://dl.acm.org/doi/10.1145/3366423.3380027 >. Acesso em: 14 Abr. 2024.

[103] F. YASMIN, MD. M. HASSAN, M. HASAN, S. ZAMAN, C. KAUSHAL, W. EL-SHAFAI, and N. F. SOLIMAN, ‘‘PoxNet22: A fine-tuned model for the classification of monkeypox disease using transfer learning,’’ **IEEE Access**, vol. 11, pp. 24053–24076, 2023. Disponível em: < https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10063857 >. Acesso em: 14 Abr. 2024.

[104] M. KUMAR, R. GHANI, and Z.-S. MEI, ‘‘Data mining to predict and prevent errors in health insurance claims processing,’’ in Proc. 16th **ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining**, Washington, DC, USA, Jul. 2010, pp. 65–74. Disponível em: < https://dl.acm.org/doi/abs/10.1145/1835804.1835816 >. Acesso em: 14 Abr. 2024.

[105] H. SHIN, H. PARK, J. LEE, and W. C. JHEE, ‘‘**A scoring model to detect abusive billing patterns in health insurance claims**,’’ Expert Syst. Appl., vol. 39, no. 8, pp. 7441–7450, Jun. 2012. Disponível em: < https://dl.acm.org/doi/10.1016/j.eswa.2012.01.105 >. Acesso em: 14 Abr. 2024.

[106] K. D. ARAL, H. A. GÜVENIR, İ. SABUNCUOĞLU, and A. R. AKAR, ‘‘A prescription fraud detection model,’’ **Comput. Methods Programs Biomed**., vol. 106, no. 1, pp. 37–46, Apr. 2012. Disponível em: < https://dl.acm.org/doi/10.1016/j.cmpb.2011.09.003>. Acesso em: 03 Mai. 2024.

[107] P. A. ORTEGA, C. J. FIGUEROA, and G. A. RUZ, ‘‘A medical claim fraud/abuse detection system based on data mining: A case study in Chile,’’ in Proc. **Int. Conf. Data Mining**, Las Vegas, NV, USA, 2006, pp. 1–12. Disponível em: < https://www.researchgate.net/publication/220704891_A_Medical_Claim_FraudAbuse_Detection_System_based_on_Data_Mining_A_Case_Study_in_Chile>. Acesso em: 05 Mai. 2024.

[108] J. M. JOHNSON and T. M. KHOSHGOFTAAR, ‘‘Medicare fraud detection using neural networks,’’ **J. Big Data**, vol. 6, no. 1, p. 63, Dec. 2019. Disponível em: < https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0225-0 >. Acesso em: 18 Mai. 2024.

[109] J. LU, K. LIN, R. CHEN, M. LIN, X. CHEN, and P. LU, ‘‘Health insurance fraud detection by using an attributed heterogeneous information network with a hierarchical attention mechanism,’’ **BMC Med. Informat. Decis. Making**, vol. 23, no. 1, p. 62, Apr. 2023. Disponível em: < https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-023-02152-0 >. Acesso em: 06 Jun. 2024.

[110] M. ZOUBEIROU, A. MAYAKI, and M. RIVEILL, “**Multiple inputs neural networks for medicare fraud detection**,” 2022. Disponível em: <https://arxiv.org/abs/2203.05842>. Acesso em: 10 Jun. 2024.

[111] RUSSELL, STUART. **Human Compatible: Artificial Intelligence and the Problem of Control**. 1. ed. [S.l.]: Viking, 2019.

[112] R. A. GUPTA, “**Kaggle Healthcare Provider Fraud detection Datasets**,” 2018, . Disponível em:< https://www.kaggle.com/datasets/rohitrox/healthcare-provider-fraud-detection-analysis>. Acesso em: 17 Jun. 2024.

[113] HAMILTON, W.L. The Graph Neural Network Model. In: Graph Representation Learning,2020. **Synthesis Lectures on Artificial Intelligence and Machine Learning**. Springer, Cham. Disponível em:< https://doi.org/10.1007/978-3-031-01588-5_5>. Acesso em: 21 Jun. 2024.

[114] K. ZHOU, Y. DONG, K. WANG, W. SUN LEE, B. HOOI, H. XU, et al., "**Understanding and resolving performance degradation in graph convolutional networks**". Disponível em:< https://dl.acm.org/doi/abs/10.1145/3459637.3482488>. Acesso em: 21 Jun. 2024.

[115] Y. RONG, W. HUANG, T. XU and J. HUANG, "DropEdge: Towards deep graph convolutional networks on node classification", **Proc**. **ICLR,** pp. 1-13, 2020. Disponível em:< https://arxiv.org/abs/1907.10903>. Acesso em: 21 Jun. 2024.

[116] M. NEWMAN, **Networks: An Introduction**, New York, NY, USA:Oxford Univ. Press, 2010. Disponível em:< https://books.google.com.br/books?hl=en&lr=&id=on0QBwAAQBAJ&oi=fnd&pg=PA3&dq=+Networks:+An+Introduction&ots=e4qJHYBuY9&sig=axN3zRFopDeOF_beXpy__fgwXWE&redir_esc=y#v=onepage&q=Networks%3A%20An%20Introduction&f=false>. Acesso em: 22 Jun. 2024.

[117] M. SEERA, C. P. LIM, A. KUMAR, L. DHAMOTHARAN and K. H. TAN, "An intelligent payment card fraud detection system", **Ann. Oper. Res**., vol. 10, pp. 1-23, Jan. 2021. Disponível em:< https://link.springer.com/article/10.1007/s10479-021-04149-2>. Acesso em: 22 Jun. 2024.

[118] V. N. DORNADULA AND S. GEETHA, "Credit card fraud detection using machine learning algorithms", **Proc. Comput. Sci.**, vol. 165, pp. 631-641, Jan. 2019. Disponível em:< https://www.sciencedirect.com/science/article/pii/S187705092030065X>. Acesso em: 23 Jun. 2024.

[119] N. SRIVASTAVA, G. HINTON, A. KRIZHEVSKY, I. SUTSKEVER and R. SALAKHUTDINOV, "Dropout: A simple way to prevent neural networks from overfitting", **J. Mach. Learn. Res**., vol. 15, no. 1, pp. 1929-1958, 2014. Disponível em:< https://www.jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf?utm_content=buffer79b43&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer>. Acesso em: 19 Jun. 2024.

[120] M. BUCKLAND AND F. GEY, ‘‘The relationship between recall and precision,’’ **J. Amer. Soc. Inf. Sci**., vol. 45, no. 1, pp. 12–19 2014. Disponível em:< https://escholarship.org/uc/item/0g80268t>. Acesso em: 01 Set. 2024.