

Tema:

Corporate Credit Rating via Graph Neural Network

Introdução

A concessão de crédito é uma importante fonte de receita para bancos, envolvendo análise detalhada de empresas e cenários econômicos. Os bancos determinam taxas de juros com base em análises de balanços, fluxos de caixa e fatores macroeconômicos, resultando em ratings de crédito que refletem a capacidade de pagamento. Na era atual, a inteligência artificial surge como uma ferramenta promissora para aprimorar a construção de ratings de crédito, permitindo uma análise mais abrangente e precisa ao processar eficientemente grandes volumes de dados, incluindo informações não estruturadas como redes sociais, proporcionando decisões de crédito mais refinadas.

OBJETIVO: Tropicalização da GNN para combinar fatores financeiros e fatores qualitativos referentes as empresas a fim de garantir ratings financeiros assertivos.

Metodologia

Trabalho americano base

O trabalho base foi elaborado por alunos de pós-doutorado da Universidade do Texas, que desenvolveram um código robusto para a análise de score de crédito utilizando Grafos de Redes Neurais (GNN). Nele, são utilizados os conjuntos de dados relevantes das empresas listadas na bolsa americana, sendo eles:

Dados de rede: A solução utiliza uma rede de empresas conectadas usando a seção MD&A de arquivamentos SEC 10-K/Q. Assumimos que empresas com declarações prospectivas semelhantes têm maior probabilidade de estar conectadas em eventos de crédito.

Dados tabulares: Para as características dos nós do grafo, a solução utiliza as variáveis no modelo Altman Z-score e a categoria da indústria de cada empresa.

A classificação usando Redes Neurais em Grafos (GNNs) foi feita de forma binária, dividida entre investimento (níveis de classificação AAA, AA, A, BBB) e não investimento (níveis de classificação BB, B, CCC, CC, C, D).

Resultado

Foi percebido um aumento significativo entre a utilização de apenas dados tabulares e a utilização de GNN, saindo de um F1 Score de aproximadamente 0,7 para 0,8.

	Mean	SD	AutoGluon
F1 Score	0.790061	0.013024	0.719251

Figura 1: Resultado comparativo entre métodos

Integrantes: - Marcos Hora Gomes de Sá
- Pedro Mota Hoertel

Professor(a) Orientador(a): Profa. Dra. Anna Reali

Dados financeiros

Para a coleta de dados financeiros acerca das empresas de capital aberto brasileiras, foi utilizada a biblioteca do Yahoo Finance para python, sendo possível a extração dos dados de:

- Balanço patrimonial
- Fluxo de caixa
- Demonstração do resultado do exercício (DRE)

Com essas informações, calcula-se o Z-score para indicar a saúde financeira da empresa em questão.

Fatos relevantes

A divulgação de fatos relevantes pelas empresas para a Comissão de Valores Mobiliários (CVM) desempenha um papel crucial na transparência do mercado financeiro brasileiro. A CVM exige que as empresas listadas na Bolsa de Valores do Brasil (B3) forneçam informações precisas, oportunas e significativas que possam influenciar o mercado de capitais.

Partindo de uma série de arquivos PDF de fatos relevantes, foi necessário uma série de tratativas para resultar no dado final. Inicialmente, é necessário rodar um programa para coletar os textos dos PDFs, e depois a utilização do método Doc2Vec para geração dos embeddings

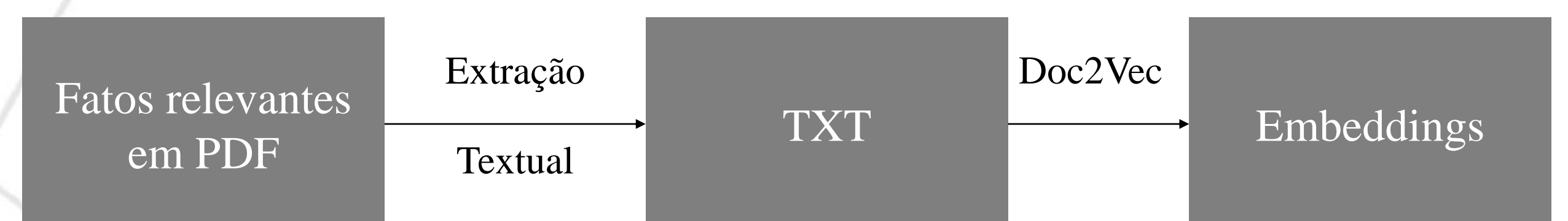


Diagrama 1: Tratativa dos fatos relevantes

Arquitetura

A seguir, podemos verificar a arquitetura completa do projeto de tropicalização da GNN para Rating de crédito.

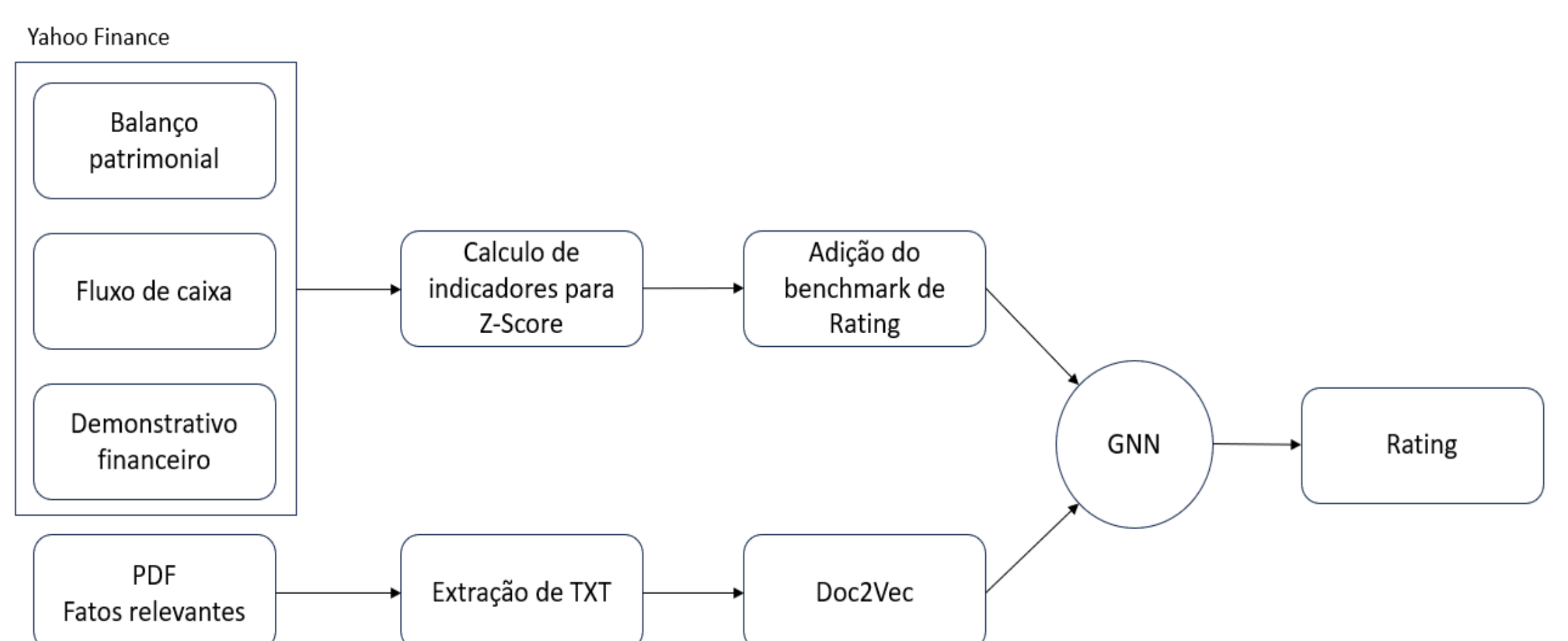


Diagrama 21: Arquitetura completa da predição

Conclusão

Com a utilização da GNN é possível notar uma significativa melhora no F1 Score, devido a criação do vínculo entre empresas em cenários parecidos. Para a tropicalização são encontradas uma série de desafios, como falta de dados padronizados, e relatórios excassos, porém, com as adapções sugeridas, é possível chegar em uma melhora significativa da predição.