

Redes Bayesianas aplicadas a previsão de *Ratings* de Crédito

Lucas Hattori Costa

PCS5708 - Técnicas de Raciocínio Probabilístico em IA

lucashattori@usp.br

24 de junho de 2021

Programa

- 1 Introdução
 - Motivação
 - Objetivo
- 2 Metodologia
 - Dataset
 - Modelos
- 3 Resultados
- 4 Conclusões e trabalhos futuros

Ratings de crédito

- Um rating de crédito é uma avaliação da solvência de uma instituição enquanto mutuária, muitas vezes com relação a um determinado título de dívida e outras obrigações financeiras.
- Calculados periodicamente e de forma extensiva por três principais agências: Fitch, S&P e a Moody's

Instituição	Rating (Moody's)	Rating (Fitch)
Banco ABC	Ba3	BB+
Banco Alfa	Ba2	BB-
Banco Itaú	Ba3	B

Tabela: Exemplos de Ratings para bancos brasileiros

Ratings de crédito

- Muitos *players* no mercado financeiro que estão interessados em estimar risco de crédito: investidores institucionais, outros emissores de crédito, agências governamentais e outras empresas.
- O cálculo desses ratings das agências é custoso, demorado e com conhecimento muito concentrado¹ nas próprias agências.
- Extensa literatura de técnicas estatísticas, ou de data mining ou de IA para tentar estimar ratings futuros.

¹Atos legislativos como Dodd-Frank (2010) vieram para tentar quebrar isso.

Objetivo

Desenvolver um algoritmo com técnica de raciocínio probabilístico que seja capaz de estimar o rating atribuído a uma dada instituição em um período de 1 ano.

Dataset

Descrição

- Proveniente do Kaggle
- 2029 observações de grandes empresas americanas listadas na NYSE ou NASDAQ:
 - 1 Setor da empresa
 - 2 Rating na escala global
 - 3 Data da avaliação
 - 4 25 Indicadores financeiros

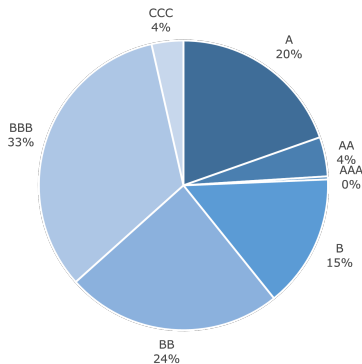


Figura: Distribuição dos Ratings no Dataset

Feature Selection

- Dos 25 indicadores financeiros, escolheu-se apenas 4 para representar as seguintes características [Hajek and Michalak, 2013]
 - 1 Liquidez: capacidade da empresa de pagar dívidas de curto prazo
 - 2 Rentabilidade: o percentual da receita da empresa que é lucro real
 - 3 Alavancagem: razão entre a dívida e os ativos da empresa
 - 4 Caixa da empresa: o quanto a empresa tem em caixa
- Variável categórica: Setor da empresa
- Dado macroeconômico: Retornos anuais do S&P 500

Modelo Inicial

- Baseado somente no rating atual da instituição
- Treinado com abordagem frequentista
- Tipo de modelo muito estudado por ser simples de computar e fácil de aplicar [Hadad et al., 2007]

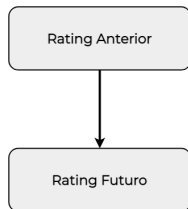


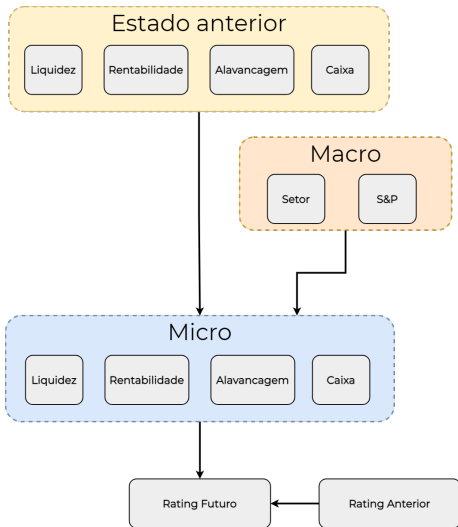
Figura: Modelo Inicial

Corporate rating transition matrix based on the cohort method
%, 2001–05

	Number of companies at period end	AAA	AA	A	BBB	BB	B	CCC	D	NR
AAA	1	100	0	0	0	0	0	0	0	0
AA	2	0	50	50	0	0	0	0	0	0
A	21	0	14.29	61.90	4.76	0	0	4.76	14.29	0
BBB	9	0	0	44.44	44.44	0	11.11	0	0	0
BB	3	0	0	0	66.67	0	0	0	0	33.33
B	3	0	33.33	33.33	0	0	33.33	0	0	0
CCC	1	0	0	100	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0
NR	0	0	0	0	0	0	0	0	0	0
Total	40									

Figura: Exemplo de matriz de transição

Modelo Bayesiano

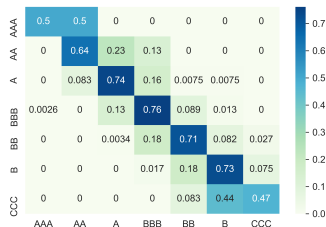


Discretização das variáveis:

- S&P: 1 se maior que média histórica, 0 caso não
- Ind. Fin.: [0,1,2] de acordo com média histórica do setor

Matrizes de transição

Gera-se as matrizes de transição para ambos modelos para efeito de comparação.



(a) Modelo Inicial



(b) Modelo Bayesiano

Figura: Matrizes de transição

Resultados

- Dataset de validação: 40 observações balanceadas entre as classes
- AUC ROC é *scale-invariant*. Como são parecidas, indica que ambos os modelos são precisos para estimar qual é o próximo estado/rating.
- MAE leva em consideração o *scale*, nesse caso, as probabilidades estimadas. Ou seja, o modelo Bayesiano estima o estado correto com maior probabilidade.

Model	MAE	AUC ROC
Inicial	0.1438	90.35%
Bayesiano	0.0968	92.92%

Tabela: Métricas de comparação

Conclusões e trabalhos futuros

- Como exposto, o modelo Bayesiano possui um AUC ROC levemente maior apenas que o modelo inicial, porém um MAE em quase 5 p.p. Para situações onde a probabilidade associada ao estado predito faça diferença, o modelo Bayesiano se mostra bem melhor.
- Futuras melhorias podem incluir:
 - Adicionar dados de mercado das próprias empresas (preço, retorno, volatilidade, etc).
 - Melhoria na redução da dimensionalidade de indicadores financeiros.
 - Fazer uma versão para dados de mercado brasileiro.

Referências



Muliaman D Hadad and Wimboh Santoso and Bagus Santoso and Dwityapoetra Besar and Ita Rulina (2007)

Rating migration matrices: empirical evidence in Indonesia
IFC Bulleting 31, 260 – 276.



Petr Hajek; Vladimir Olej; Ondrej Prochazka

Predicting Corporate Credit Ratings Using Content Analysis of Annual Reports – A Naïve Bayesian Network Approach
FinanceCom



Petr Hajek; Krzysztof Michalak

Feature selection in corporate credit rating prediction
Knowledge-Based Systems 51, 72 – 84

Obrigado!