

Trabalho 1 de AMMCI

RA: 109882 - Gabriel Colli Pavan

RA: 107115 - Matheus Augusto Schiavon Parise

Maringá

2023

Introdução

Esse documento descreve um relatório sobre o desenvolvimento do primeiro trabalho da disciplina de aprendizado de máquina e modelagem do conhecimento incerto (AMMCI-9806/01) da Universidade Estadual de Maringá.

O trabalho consiste na modelagem de um problema com cerca de 3 a 5 variáveis, considerando de 3 a 5 decisões diferentes, considerando os conceitos apreendidos da teoria de decisões simples e de redes bayesianas.

Os bancos incorrem em prejuízos quando os clientes não pagam seus empréstimos em dia. Por causa disso, todos os anos, os bancos têm prejuízos de milhões, e isso também afeta em grande medida o crescimento econômico do país. Foi escolhida uma base de dados de previsão de empréstimo que contém 67 mil registros fictícios, esses registros são os dados dos supostos clientes que pediram empréstimo. O objetivo do trabalho é determinar se o banco deverá conceder empréstimo de acordo com os dados do cliente. A base de dados utilizada neste trabalho pode ser obtida em:

<https://www.kaggle.com/datasets/hemanthsai7/loandefault>

Explicação das variáveis, decisões e do modelo

Foram escolhidas quatro variáveis para decisão final, elas são *Loan Amount* (“Valor do empréstimo” em dólares), Valor já pago da dívida (esta variável foi criada ao dividir o valor pago da dívida (*Funded Amount*) pelo valor do empréstimo total (*Loan Amount*) pois o valor de cada dívida é diferente então estamos considerando a porcentagem da dívida paga) *Debit to Income* (“Débito que está para vir”, é a % da dívida total que será paga por mês) e por fim *Interest Rate* (“Taxa de Juros”, em % do valor total emprestado).

Os intervalos das categorias de cada uma das variáveis foram escolhidas com base nos dados dos clientes que são de fato inadimplentes. Para o bom funcionamento das funções estatísticas utilizadas, na maioria dos casos o intervalo é concentrado na categoria do meio.

Valor do empréstimo (em dólares)

- empréstimo ≤ 10.000
- $10.000 < \text{empréstimo} \leq 25.000$
- $25.000 < \text{empréstimo}$

Valor já pago da dívida (% da dívida paga)

- valor pago $< 50\%$
- $> 50\%$

Débito que está para vir (% da dívida total que será paga por mês)

- débito $\leq 15\%$
- $15\% < \text{débito} \leq 30\%$

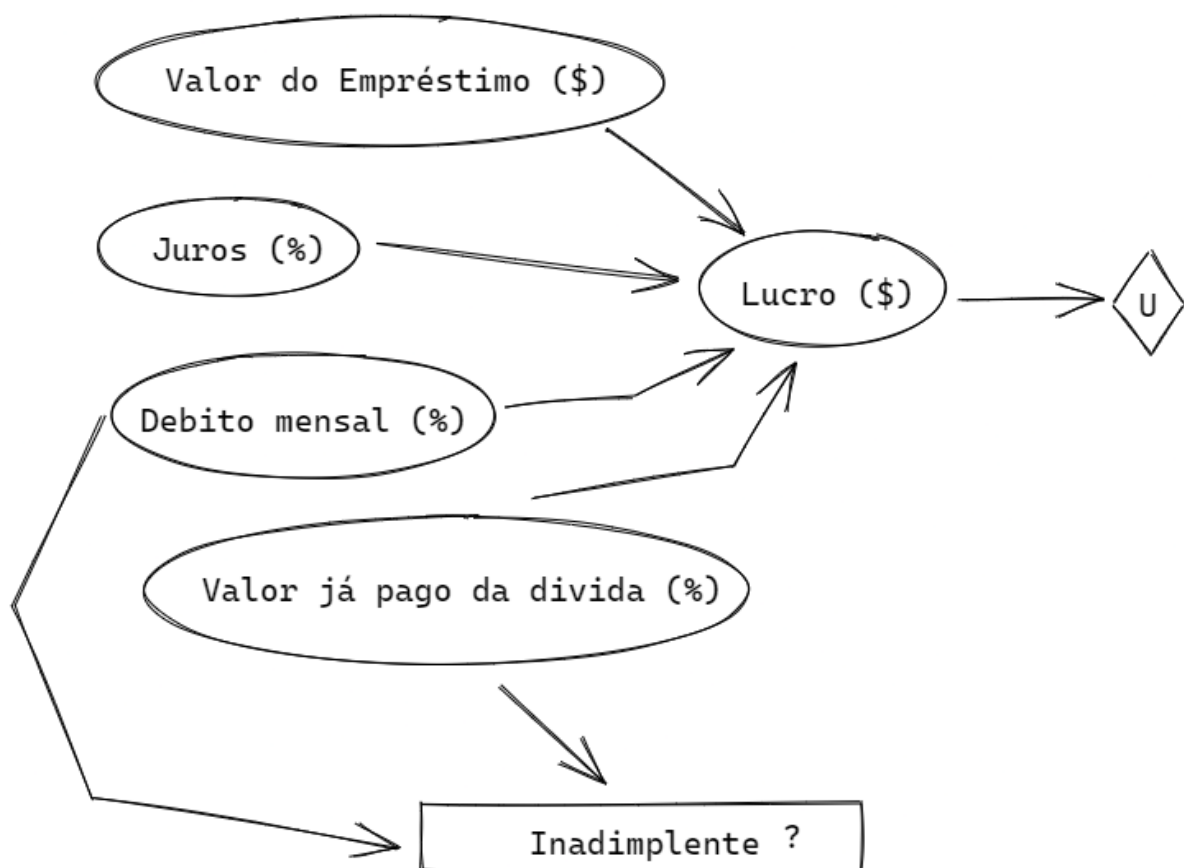
- $\text{débito} > 30\%$

Taxa de Juros (% do valor total emprestado que é cobrado mensalmente)

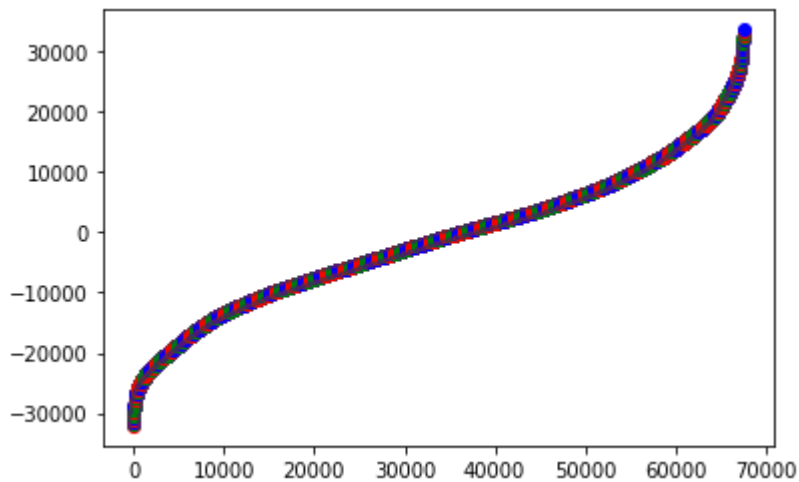
- $\text{Juros} \leq 8\%$
- $8\% < \text{Juros} \leq 15\%$
- $\text{Juros} > 15\%$

O programa decide a porcentagem de chance dele ser ou não inadimplente e também sua utilidade. A utilidade é considerada baixa caso não tenha pago ou tenha pago muito pouco o valor investido, média utilidade se pagou uma porção próxima ao valor emprestado e alta se pagou maior que o valor emprestado.

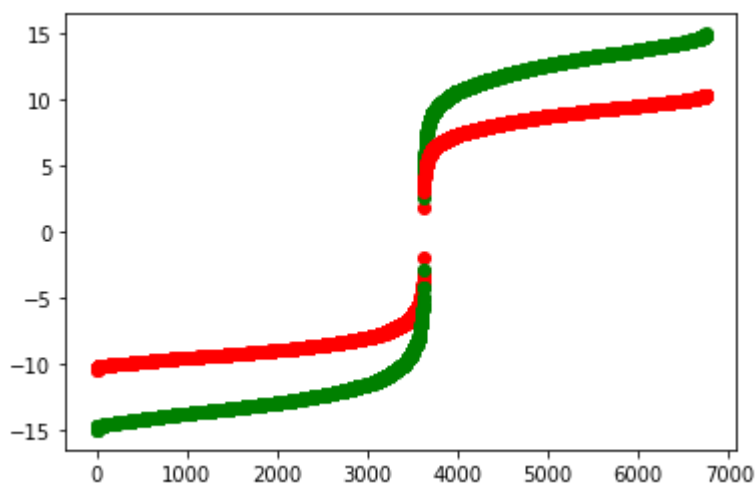
Modelo empréstimo bancário



Por fim, a utilidade é determinada pelo valor financeiro faturado pelo banco em relação a cada cliente. Na imagem a seguir a curva da utilidade do dinheiro



Em seguida, a linha vermelha do gráfico a seguir representa o logaritmo natural da quantia e em verde logaritmo na base dois da quantia.



Explicação de como o código implementado funciona

O código está comentado com sua explicação

no powerpoint colocar as coisas principais

Qual é a decisão que o código toma entre as diferentes decisões

O código decide se determinado cliente será ou não elegível para empréstimo, dessa forma isso irá minimizar os prejuízos bancários.

Bibliografia utilizada

- Slides da aula disponibilizados no classroom
- Vídeo aula disponibilizada no classroom
- Documentação da biblioteca pandas: <https://pandas.pydata.org/docs/>
- Documentação da biblioteca matplotlib:
https://matplotlib.org/3.5.3/api/_as_gen/matplotlib.pyplot.html