



**Universidade de Brasília  
Departamento de Estatística**

**Interpretação de redes neurais**

**Davi Guerra Alves**

Projeto apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília  
2023**

**Davi Guerra Alves**

**Interpretação de redes neurais**

Orientador(a): Thais Carvalho Valadares Rodrigues

Projeto apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

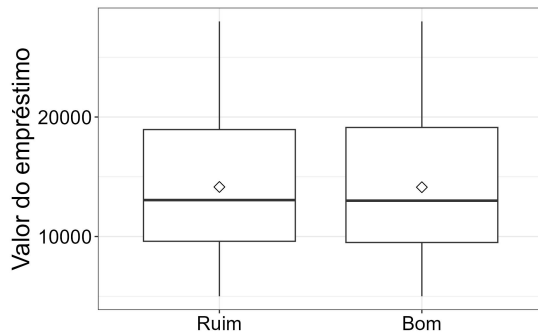
**Brasília  
2022**

## Sumário

<b>1 Resultados . . . . .</b>	<b>4</b>
<b>2 Conclusão . . . . .</b>	<b>14</b>
<b>3 Anexo . . . . .</b>	<b>15</b>

# 1 Resultados

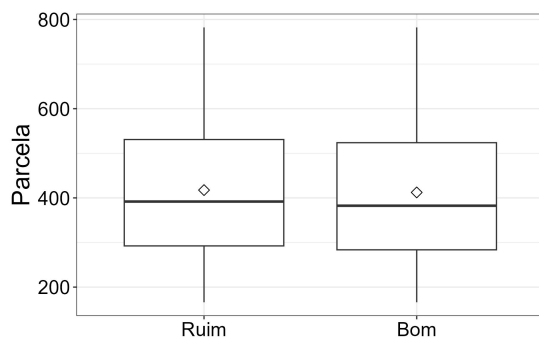
## Análise descritiva



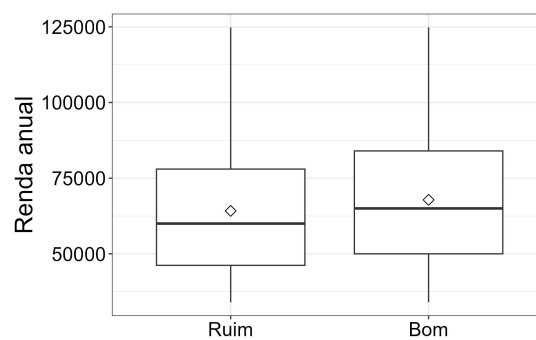
(a) Valor do empréstimo



(b) Tempo de emprego (em anos)



(c) Valor da parcela do empréstimo



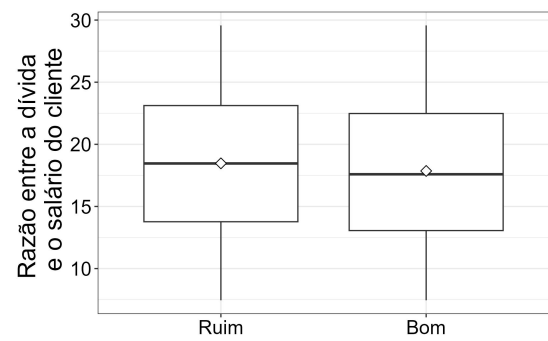
(d) Renda anual (em dólar)

Figura 1: Variáveis explicativas em relação à condição do empréstimo

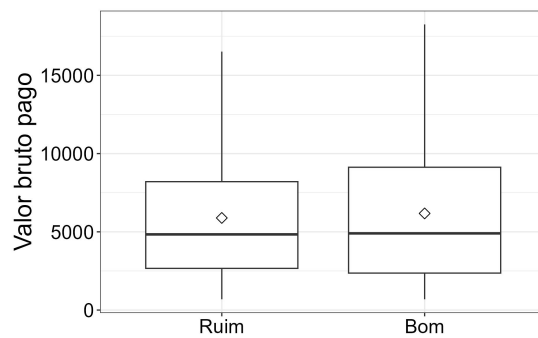
O comportamento da variável resposta nas Figuras 1(a) e 1(c) demonstrou semelhanças, onde, em ambos os casos, não foi evidenciada uma clara diferença entre o valor do empréstimo e o valor da parcela em relação às categorias da variável resposta. A Figura 1(b) também apresenta um comportamento semelhante entre as classes "Empréstimo ruim" e "Empréstimo bom", mas com um detalhe: a mediana do tempo de trabalho dos clientes rotulados como "Empréstimo ruim" foi inferior em comparação ao outro caso. Por fim, a Figura 1(d) indica que clientes com uma renda anual elevada tendem a ser categorizados como "Empréstimo bom".



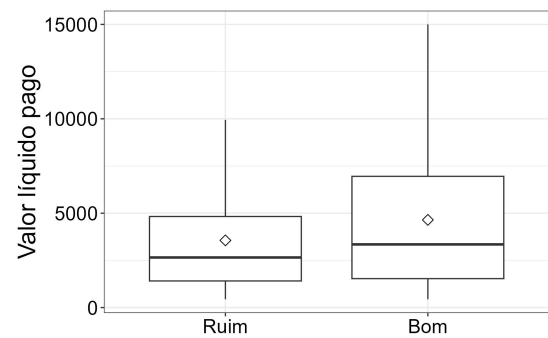
(a) Taxa de juros do empréstimo



(b) Razão entre a dívida e o salário do cliente



(c) Valor bruto do empréstimo pago



(d) Valor líquido do empréstimo pago

Figura 2: Variáveis explicativas em relação à condição do empréstimo

A Figura 2(a) evidencia uma relação significativa entre taxas de juros elevadas e empréstimos considerados ruins. A Figura 2(b) complementa a informação fornecida pela Figura 1(d), indicando que clientes com renda mais elevada tendem a cumprir adequadamente com seus pagamentos. As Figuras 2(c) e 2(d) seguem padrões semelhantes, sugerindo que clientes que quitaram a maior parte do empréstimo são frequentemente rotulados como bons pagadores.

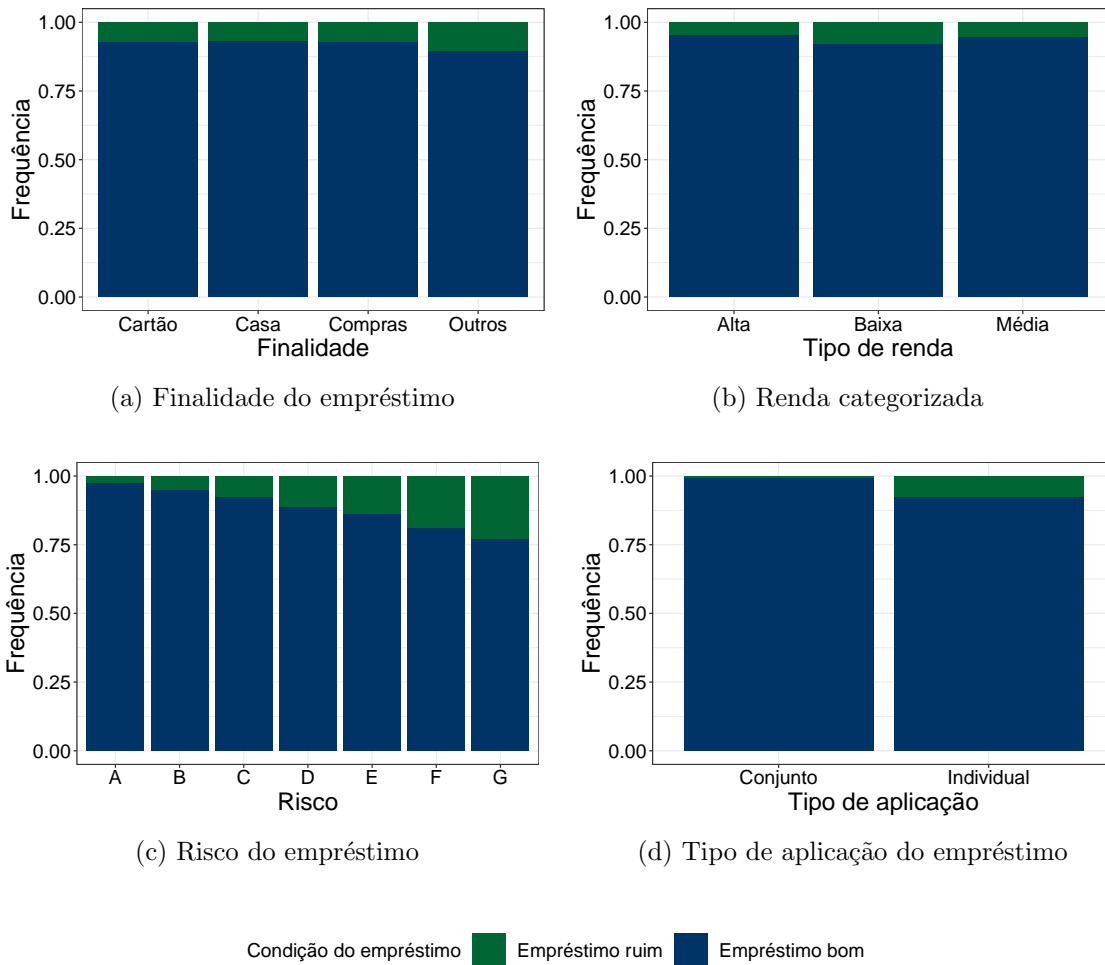


Figura 3: Variáveis explicativas em relação à condição do empréstimo

A Figura 3(a) ilustra que as categorias da variável "Finalidade" seguem a proporção natural da condição do empréstimo, conforme indicado na Tabela de Condição do Empréstimo. Na Figura 3(b), as categorias "Alta" e "Média" exibem proporções menores de empréstimos ruins em comparação com a categoria "Baixa", que apresenta uma proporção de quase 10% de empréstimos ruins. A Figura 3(c) revela um padrão de "cascata", indicando que à medida que o risco do empréstimo aumenta, a proporção de empréstimos ruins nas últimas categorias também aumenta, sendo a categoria G a mais afetada, com quase 25% de empréstimos classificados como ruins. Na Figura 3(d), a categoria "Empréstimo conjunto" não registrou observações de empréstimos ruins, concentrando a maioria desses empréstimos na categoria "Empréstimo individual".

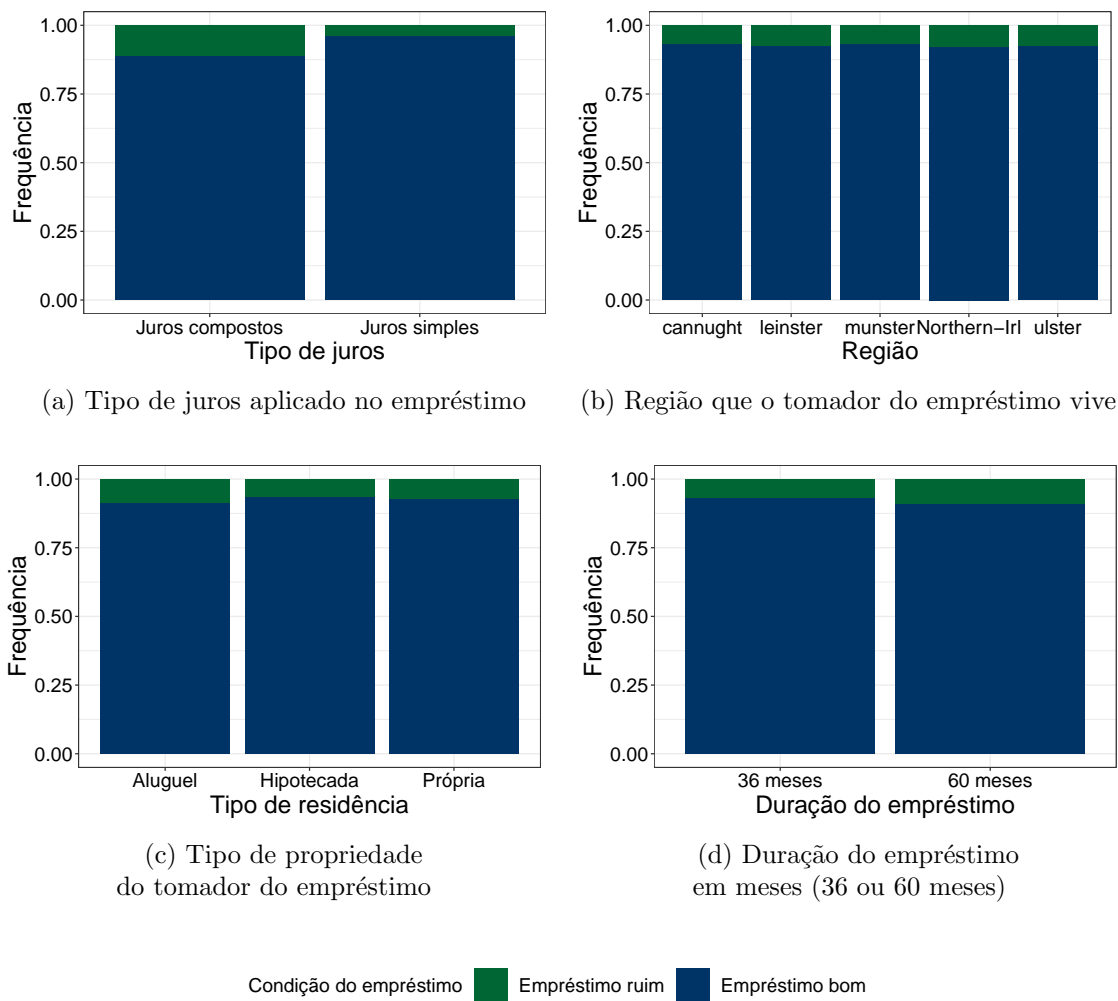


Figura 4: Variáveis explicativas em relação à condição do empréstimo

Na Figura 4(a), o gráfico evidencia que empréstimos obtidos sob juros compostos possuem uma proporção mais elevada de rotulações ruins em comparação com empréstimos sob juros simples. As Figuras 4(b) e 4(c) destacam uma proporção natural refletida pela distribuição das categorias da variável resposta, conforme apresentado na Tabela xxxx. Já a Figura 4(d) revela uma proporção mais significativa de empréstimos ruins quando estes tendem a demorar mais para serem pagos.

Covariáveis	Coefficiente de correlação
Tempo de trabalho	-0.02
Renda anual	-0.03
Valor do empréstimo	0.00
Taxa de juros	0.18
DTI	0.01
Valor bruto pago	-0.04
Valor líquido pago	-0.10
Parcela	-0.01

Tabela 1: Valores do coeficiente de Pearson entre as covariáveis e a variável resposta

Covariáveis	Coefficiente de contingência
Tipo de residência	0.04
Tipo de aplicação	0.01
Finalidade	0.03
Tipo de juros	0.14
Risco	0.15
Região	0.01

Tabela 2: Valores do coeficiente de contingência entre as covariáveis e a variável resposta

## Modelagem da regressão logística

Falar do modelo utilizado, a normalização dos dados, os resultados métricas de avaliação e interpretação dos coeficientes



Covariáveis	Coeficientes	Erro padrão
Valor líquido pago	-4.733	0.034
Valor bruto pago	3.321	0.028
Tipo de aplicação	-1.848	0.106
Taxa de juros	1.412	1.412
Valor do empréstimo	-1.406	0.039
Risco	-1.049	0.014
Tipo de juros	-0.462	-0.462
Prazo	-0.203	0.028
Renda anual	-0.195	0.010
DTI	-0.151	0.011
Categoria renda	-0.111	0.015
Duração do empréstimo	-0.061	0.009
Região	0.038	0.003
Duração do empréstimo em dias	0.033	0.009
Tipo de residência	0.019	0.003
Finalidade	0.019	0.002
Parcela	0.005	0.000

Tabela 3: Estimativa dos coeficientes do modelo logístico e o erro padrão associado

	Precisão	Recall	F1-Score	Tamanho da amostra
0	0.928081	0.995603	0.960657	163990
1	0.536334	0.0618419	0.110897	13486
Média macro	0.732208	0.528723	0.535777	177476
Média ponderada	0.898313	0.924649	0.896086	177476
Acurácia			0.924649	

Tabela 4: Report do modelo logístico

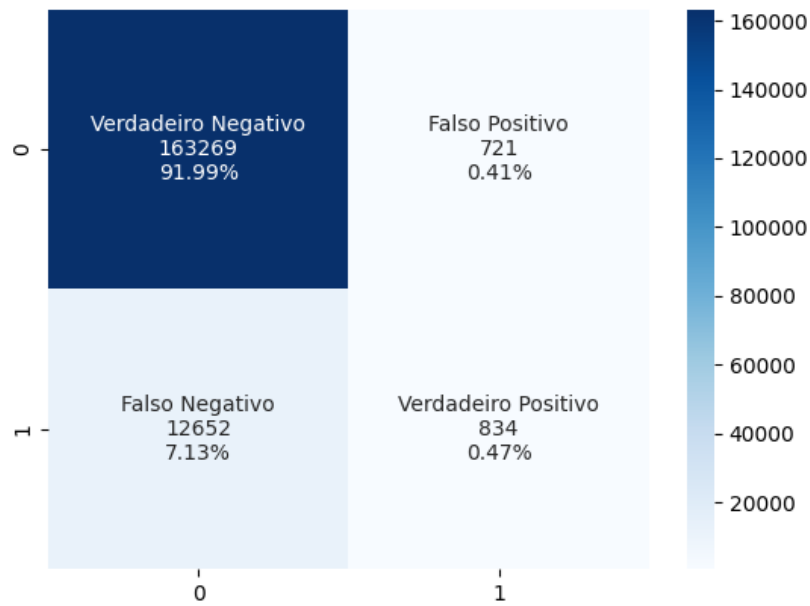


Figura 5: Matrix de confusão do modelo logístico

## Modelagem da rede neural

Falar sobre Arquitetura inicial, suas variações junto com os resultados avaliativos e por fim falar qual modelo foi utilizado e porque

## Interpretação de rede neural

- mostrar gráfico da média dos shap vs regressao logistica

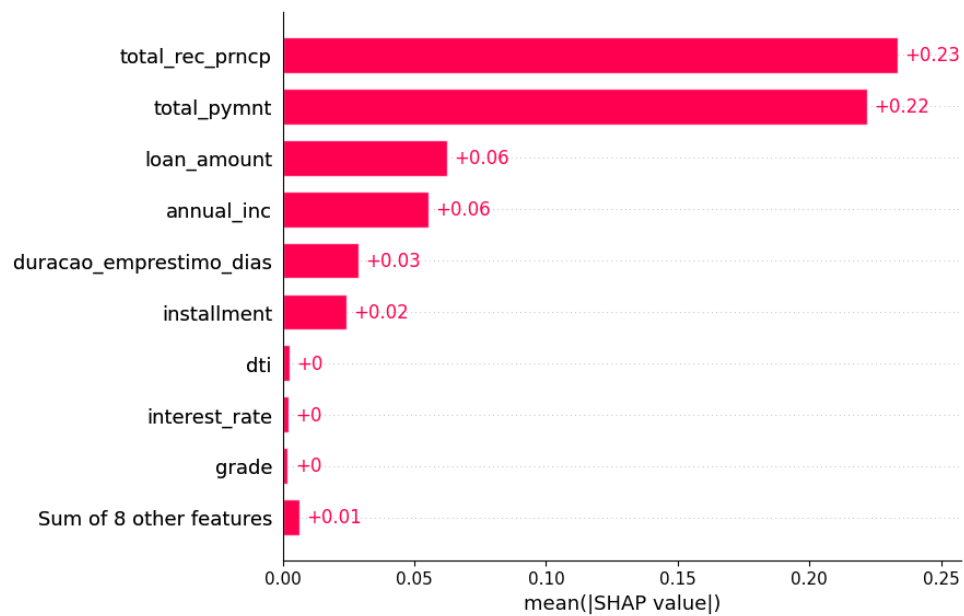


Figura 6: Média absoluta dos valores de shap

- mostrar o grafico de dependencia entre valor da variável com resultado do modelo variando ela

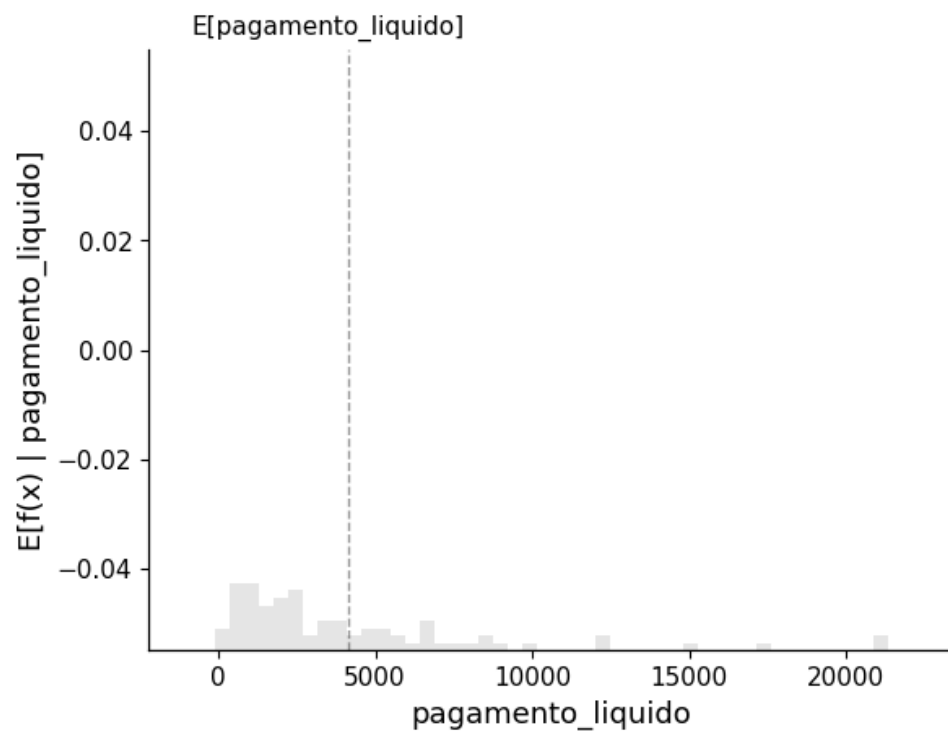
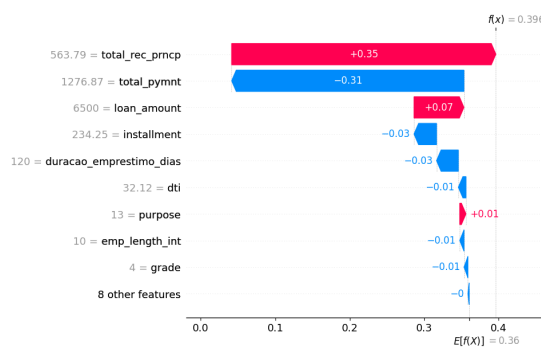
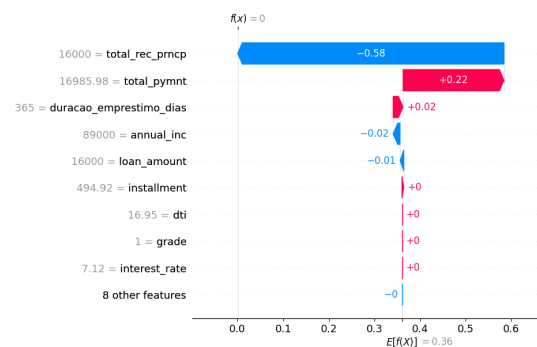


Figura 7: Relação entre a variável X com o resultado modelo quando a mesma varia

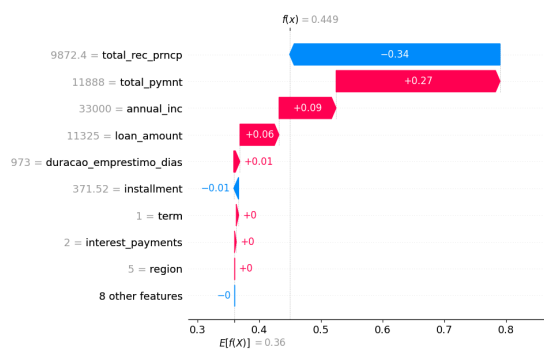
- mostrar 2 gráficos de shap específicos de 2 observações (para mau pagador e para bom pagador)



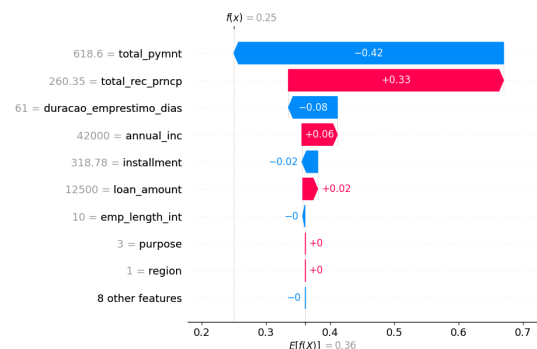
(a) Valor e classificação do emprestimo



(b) WSIR residuals.



(c) WSIR predictions in the covariate space.



(d) WSIR residuals

Figura 8: aloalo

- mostrar o gráfico com todos as amostras de shap(shap.plots.beeswarm)

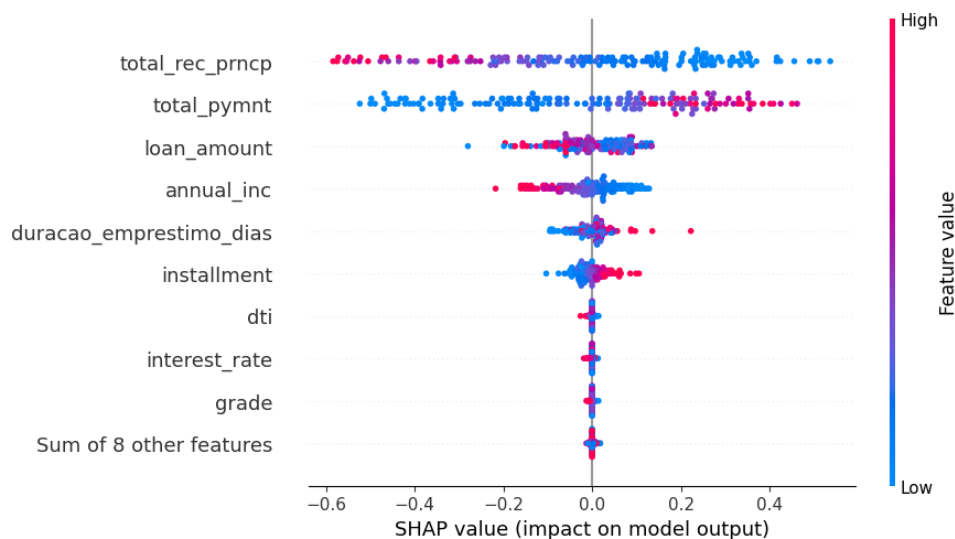


Figura 9: Valores de shap para as 80 observações utilizadas

- mostrar o gráfico de força pra apenas uma observaçã

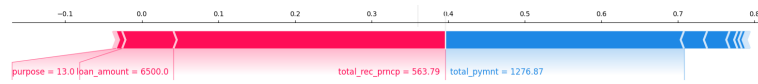


Figura 10: Gráfico de força em uma observação

- mostrar o gráfico de força para todas as observações

## Benchmark entre redes neurais e regressão logística

## 2 Conclusão

### **3 Anexo**