

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
NÚCLEO DE EDUCAÇÃO A DISTÂNCIA
Pós-graduação Lato Sensu em Ciência de Dados e Big Data



ANÁLISE DE HISTÓRICO DE JOGOS DE POKER PARA AUMENTO DE RENDIMENTO

RODRIGO XAVIER DE ALMEIDA LEÃO

2022



ANÁLISE DE HISTÓRICO DE JOGOS DE POKER PARA AUMENTO DE RENDIMENTO

Trabalho de Conclusão de Curso apresentado ao
Curso de Especialização em **Ciência de Dados e
Big Data** como requisito parcial à obtenção do
título de especialista..

1 INTRODUÇÃO

1.1 Contextualização

- O poker é um dos jogos de carta mais populares do mundo;
- Poker online permite jogar uma grande quantidade de mãos;
- É possível eliminar consideravelmente o fator sorte tornando-o um jogo de probabilidade e estatísticas;
- Grandes jogadores sabem como aplicar a probabilidade a seu favor fazendo com que no longo prazo seus ganhos superem as perdas.

1.2 O problema proposto

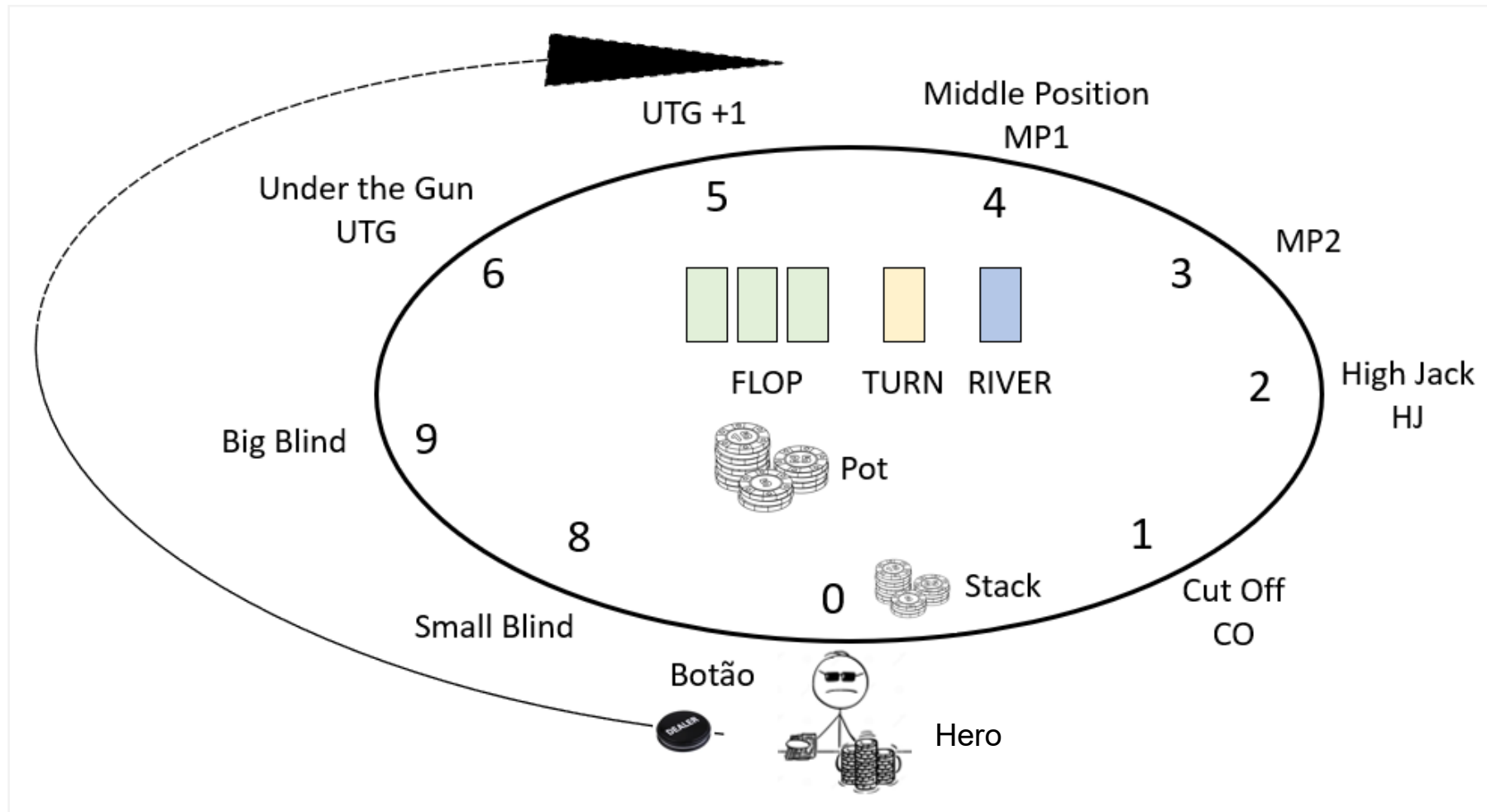
- Desenvolver um método capaz de fornecer uma análise global do estilo de jogo do cliente a partir de seu histórico de jogos online.
- Apontar falhas no estilo de jogo e sugerir alterações que resultem em maior lucratividade.



No ano de 2014, o serviço Global Poker Index reconheceu Daniel Negreanu como o maior jogador de poker da década.

1.3 Fundamentos do Poker

- Neste trabalho considera-se torneios do estilo Texas Hold'em de 9 jogadores como ilustrado abaixo.



1.3 Fundamentos do Poker

Jogadas do poker:

ROYAL STR. FLUSH	10 ♥	J ♥	Q ♥	K ♥	A ♥
STRAIGHT FLUSH	4 ♣	5 ♣	6 ♣	7 ♣	8 ♣
QUADRA	K ♠	K ♥	K ♣	K ♦	3 ♠
FULL HOUSE	10 ♥	10 ♠	10 ♦	A ♠	A ♣
FLUSH	10 ♠	K ♠	2 ♠	6 ♠	7 ♠
SEQUÊNCIA	7 ♣	8 ♠	9 ♦	10 ♠	J ♥
TRINCA	5 ♠	5 ♥	5 ♣	J ♦	A ♦
2 PARES	A ♠	A ♥	3 ♣	3 ♠	J ♣
1 PAR	Q ♦	Q ♥	2 ♥	8 ♠	9 ♣

Fatores que afetam o resultado de uma mão.

- “**Hand_Power**” que também depende se as cartas são do mesmo naipe ou não “*Swited*”;
- “**BT_pos**” indica a posição em que o botão se encontra em relação ao jogador, quanto menor seu valor melhor para jogar;
- “**Pot_BB**” quantidade de fichas em jogo;
- “**Stack_BB**” quantidade de fichas que o jogador possui naquele momento.

1.4 Objetivos

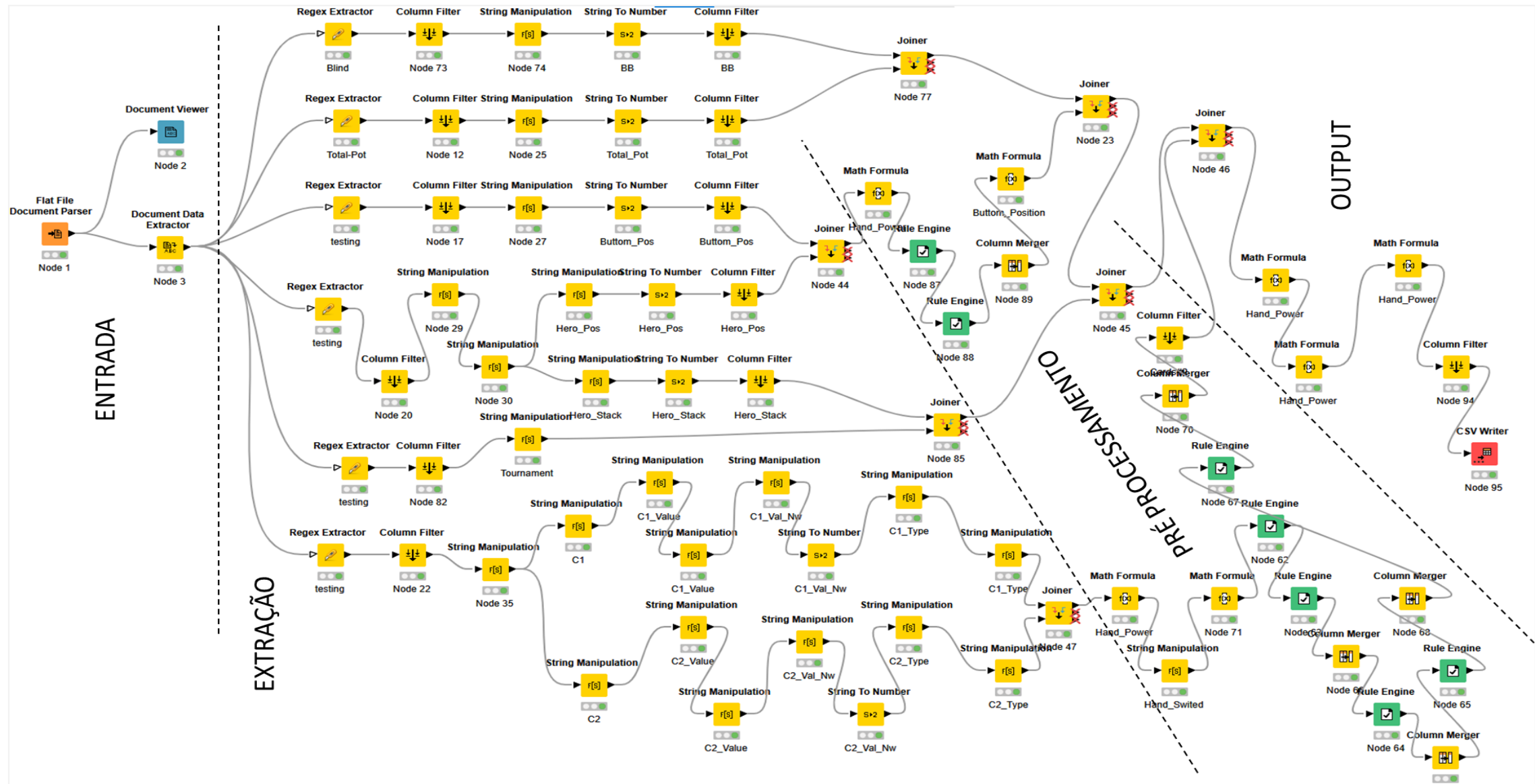
- Definir pesos e variáveis de relevância;
- Desenvolver um método de extração das variáveis a partir do histórico de jogos;
- Desenvolver uma forma de análise que permita relacionar as variáveis aos resultados obtidos pelo jogador;
- Aplicar métodos de aprendizado de máquina e avaliar a capacidade do modelo em tomar decisões consistentes com os fundamentos do Poker.

2 Coleta de dados

Histórico de jogos que pode ser obtido na plataforma de jogo utilizada.

```
PokerStars Hand #236216463916: Tournament #3413806416, $0.23+$0.02 USD Hold'em No Limit  
Table '3413806416 1' 3-max Seat #1 is the button  
Seat 1: zoriank (500 in chips)  
Seat 2: Asecos (500 in chips)  
Seat 3: rodxal (500 in chips)  
Asecos: posts small blind 10  
rodxal: posts big blind 20  
*** HOLE CARDS ***  
Dealt to rodxal [4h 9s]  
zoriank: folds  
Asecos: calls 10  
rodxal: checks  
*** FLOP *** [5h Td 5s]  
Asecos: checks  
rodxal: checks  
*** TURN *** [5h Td 5s] [Jc]  
Asecos: checks  
rodxal: bets 29  
Asecos: folds  
Uncalled bet (29) returned to rodxal  
rodxal collected 40 from pot  
rodxal: doesn't show hand  
*** SUMMARY ***  
Total pot 40 | Rake 0  
Board [5h Td 5s Jc]  
Seat 1: zoriank (button) folded before Flop (didn't bet)  
Seat 2: Asecos (small blind) folded on the Turn  
Seat 3: rodxal (big blind) collected (40)
```


3 Processamento / Tratamento de Dados



Banco de Dados

```
#DataSet -> Dataframe
df=pd.read_csv('PokerData.csv')
print('Dimensões do DataSet --> ', df.shape)
df.head()
```

✕ Dimensões do DataSet --> (1812, 6)

	BT_Pos	Tournamnet_Number	Swited	Stack_BB	Pot_BB	Hand_Power
0	5	3400213716	1	75.0	6.50	60
1	6	3400213716	1	75.0	21.00	224
2	7	3400213716	1	65.5	80.60	88
3	8	3400213716	1	58.5	48.15	21
4	0	3400213716	1	28.3	101.43	22

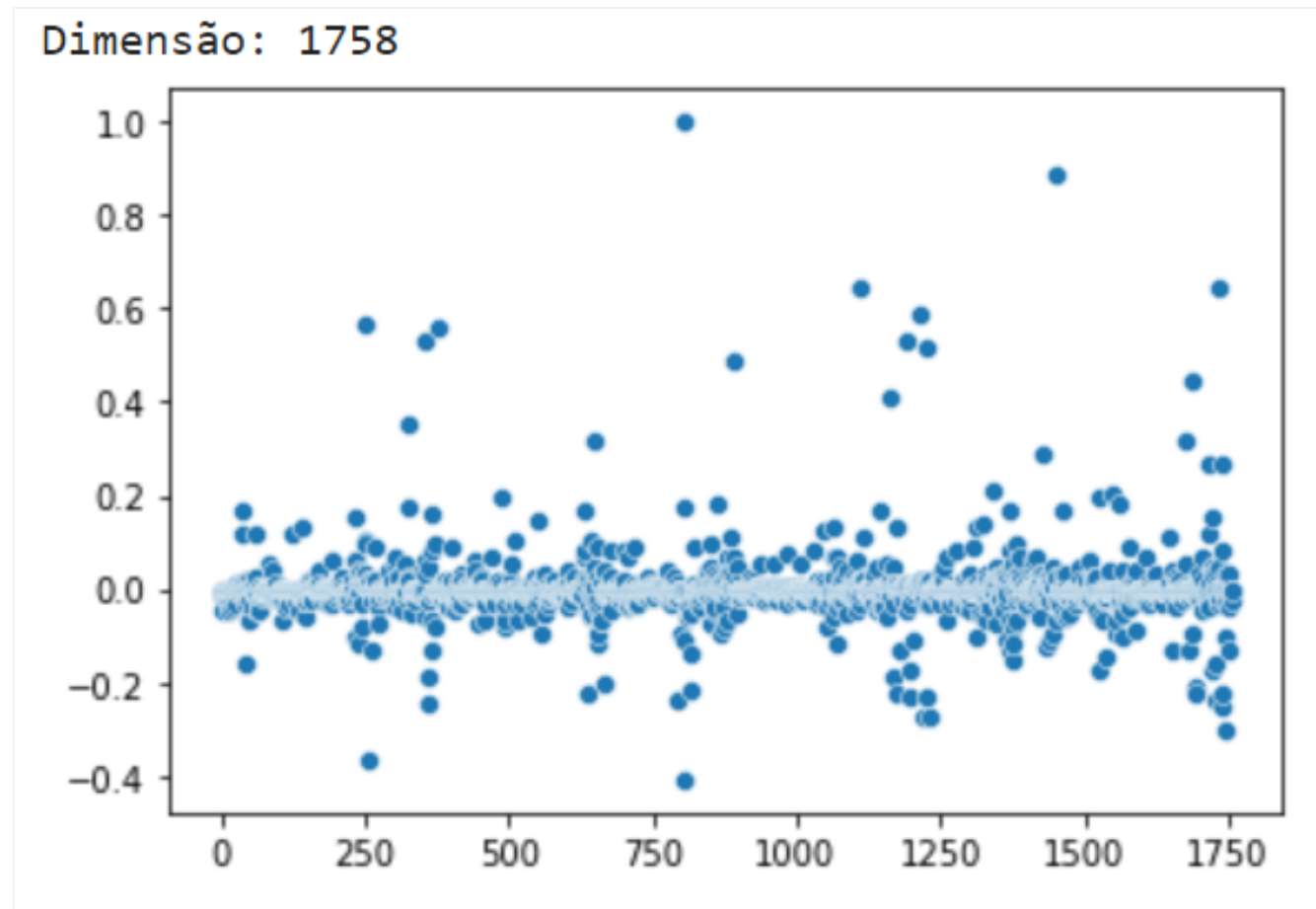
Hand_Power = Pc1 x Pc2 x Cat

Categorias:

- 5. Cartas iguais (par de mão);
- 4. Ás e outra carta qualquer;
- 3. Duas figuras;
- 2. Figura e outra carta que não seja o Ás;
- 1. Dois números diferentes.

4 Análise e Exploração dos Dados

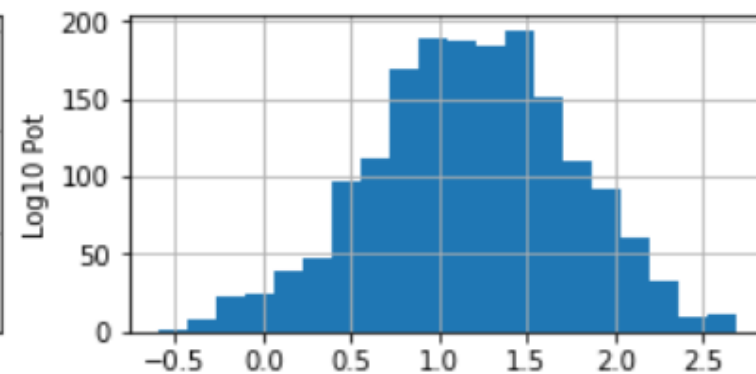
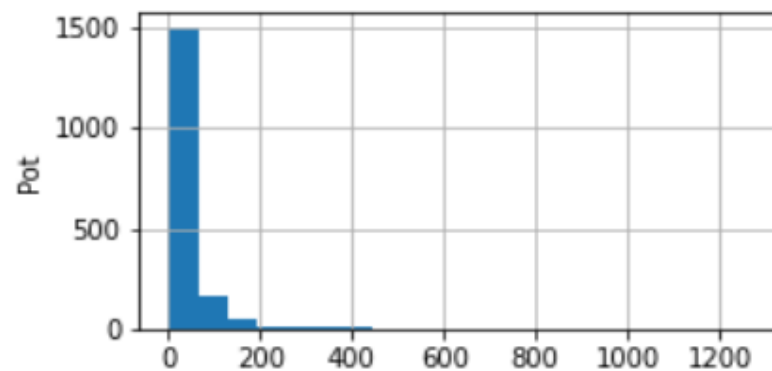
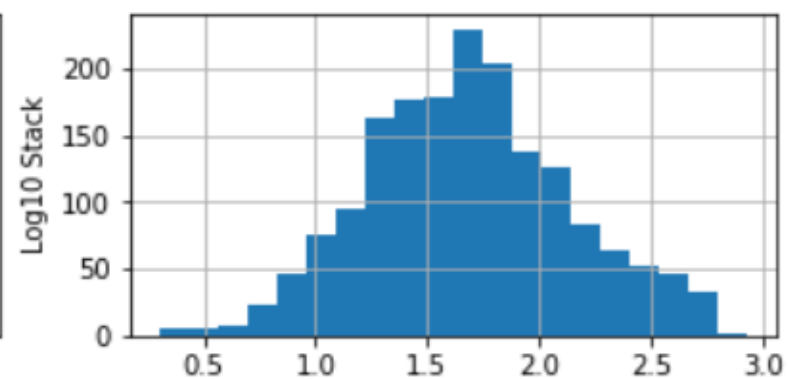
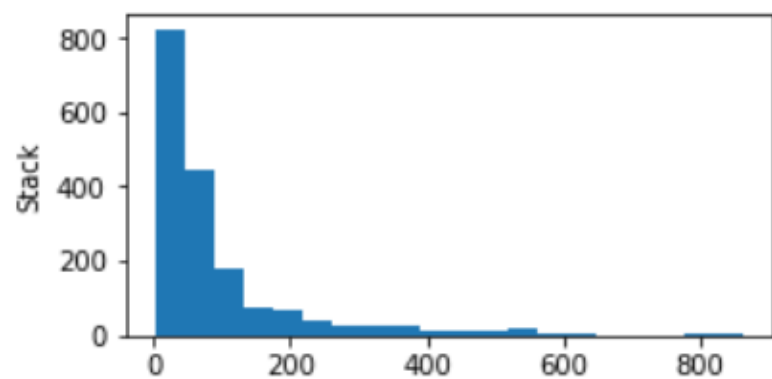
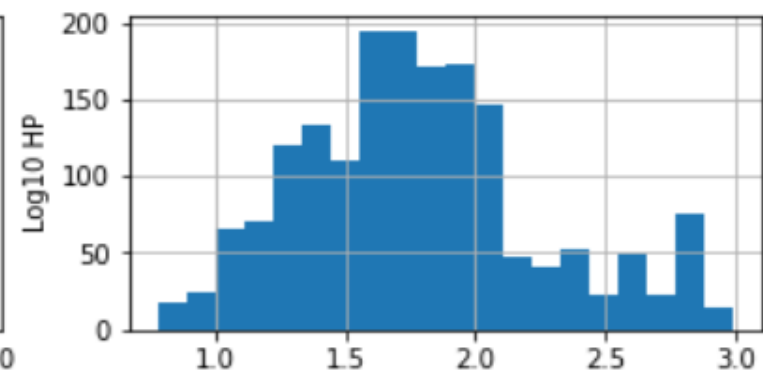
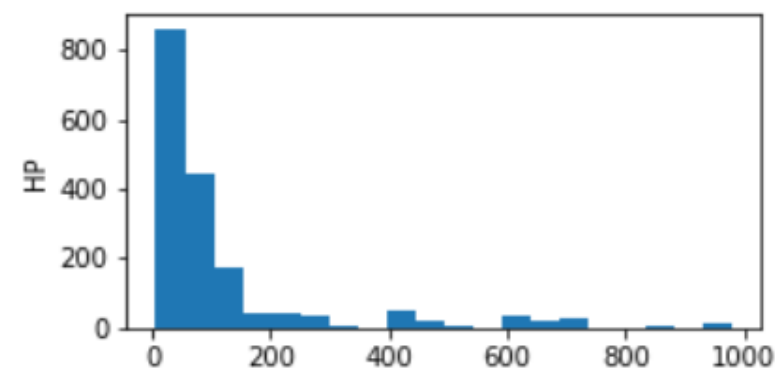
- O parâmetro *Stack* é transformado em ganho por mãos “**Gain**” causando uma redução do banco de dados de 2.98%.



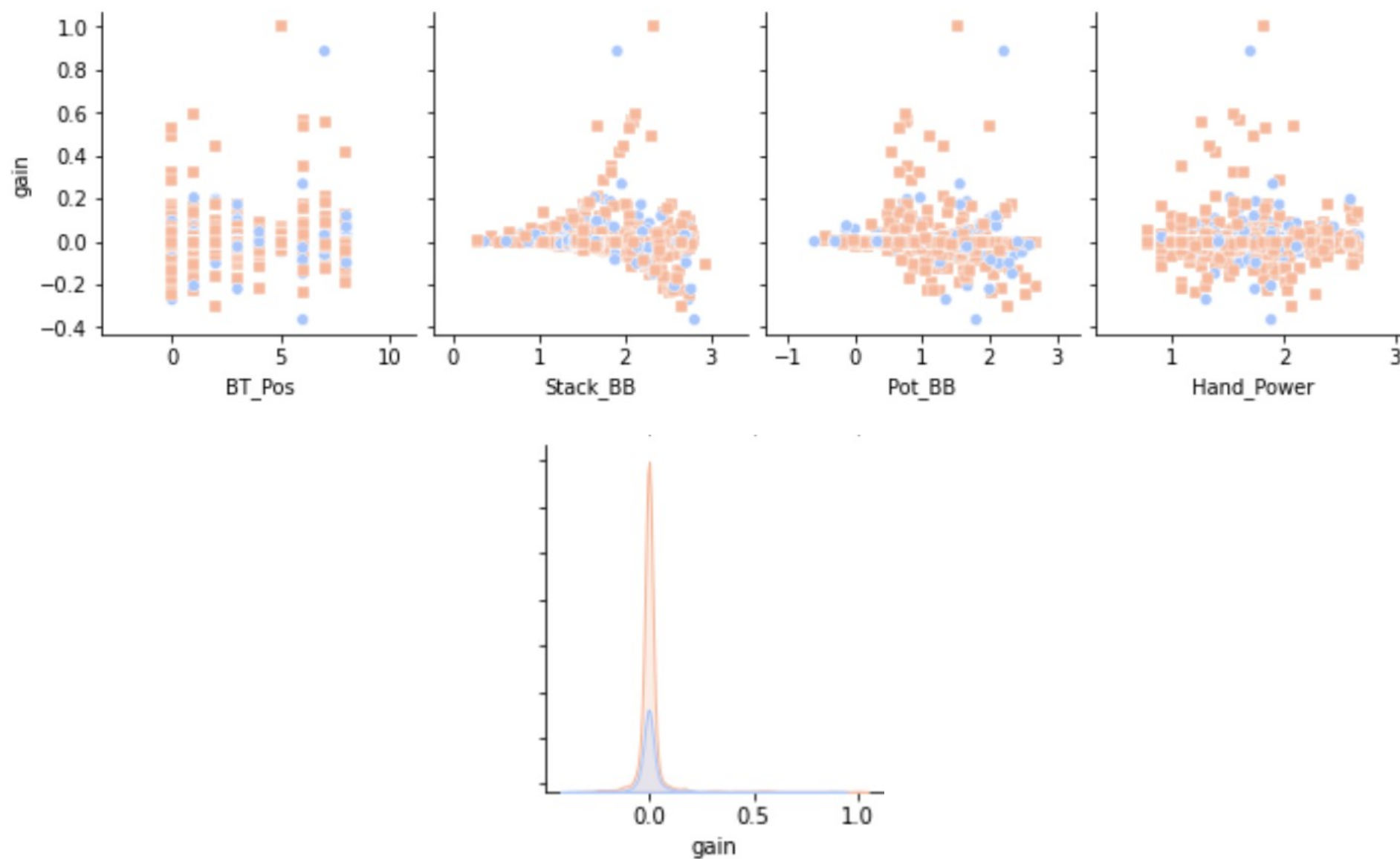
Corte de Outliers

- Mão envolvidas em pots **maiores que 500 BB** que normalmente representam os instantes finais do jogo em que os *blinds* já estão muito altos ou situações extraordinárias de início de jogo.
- Mãos em que o jogador inicia com cartas que possuam “**Hand_Power > 605**” -> **maior que JJ** (par de valetes) que normalmente envolverão muitas fichas.
- O corte de *outliers* reduz o banco de dados para 1656 mãos, representando uma **perda de 5.15 %**.

Harmonização



Harmonização



5 Criação de Modelos de Machine Learning

Regressão Linear Simples

```
#Features - Target
y=dflog['gain'].to_numpy()
X=dflog.iloc[:,5]

from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
y=scaler.fit_transform(y.reshape(-1,1))

print(y)
X.head()
```

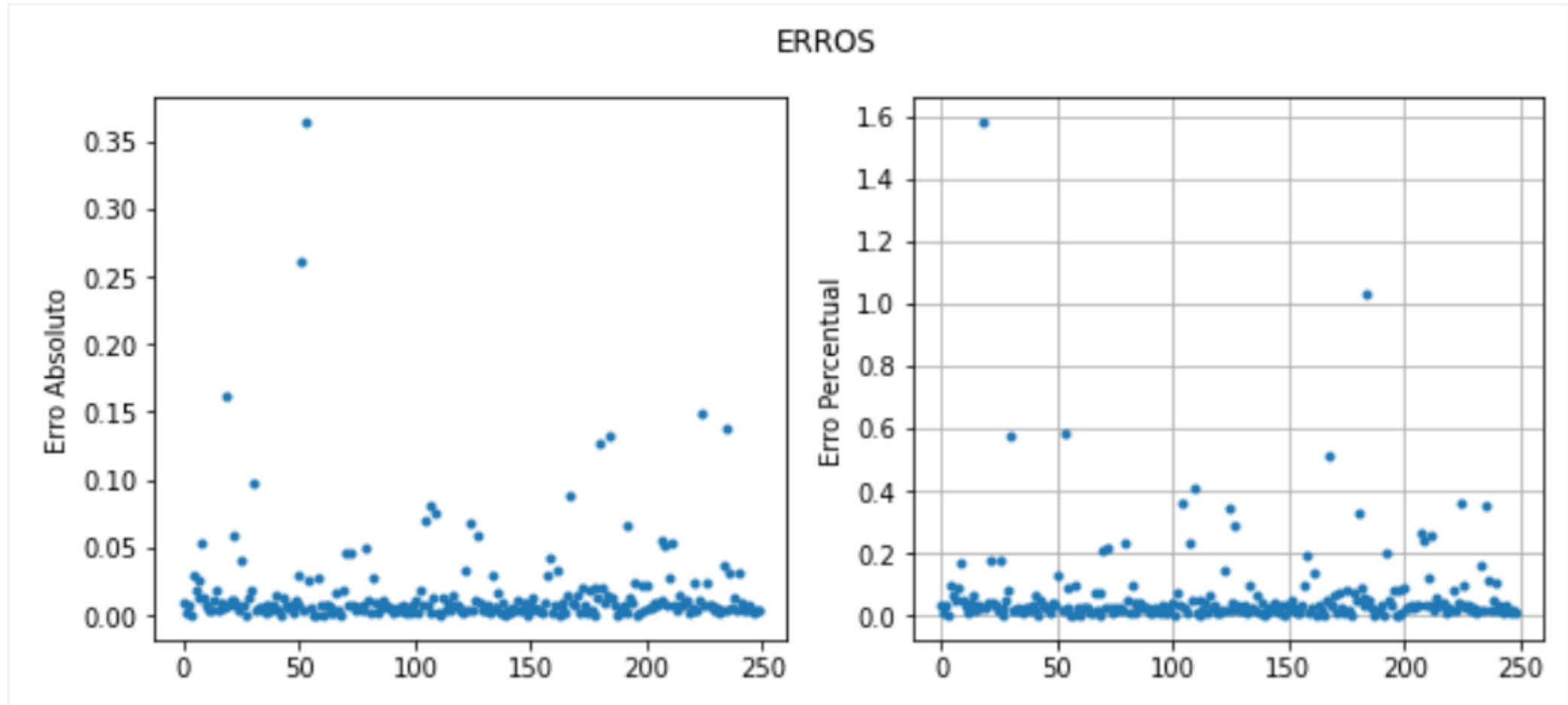
```
[[0.2670118 ]
 [0.25624215]
 [0.25907627]
 ...
 [0.26630327]
 [0.24929856]
 [0.26630327]]
```

	BT_Pos	Swited	Stack_BB	Pot_BB	Hand_Power
0	5	1	1.88	0.81	1.78
1	6	1	1.88	1.32	2.35

Coefs * 10⁴

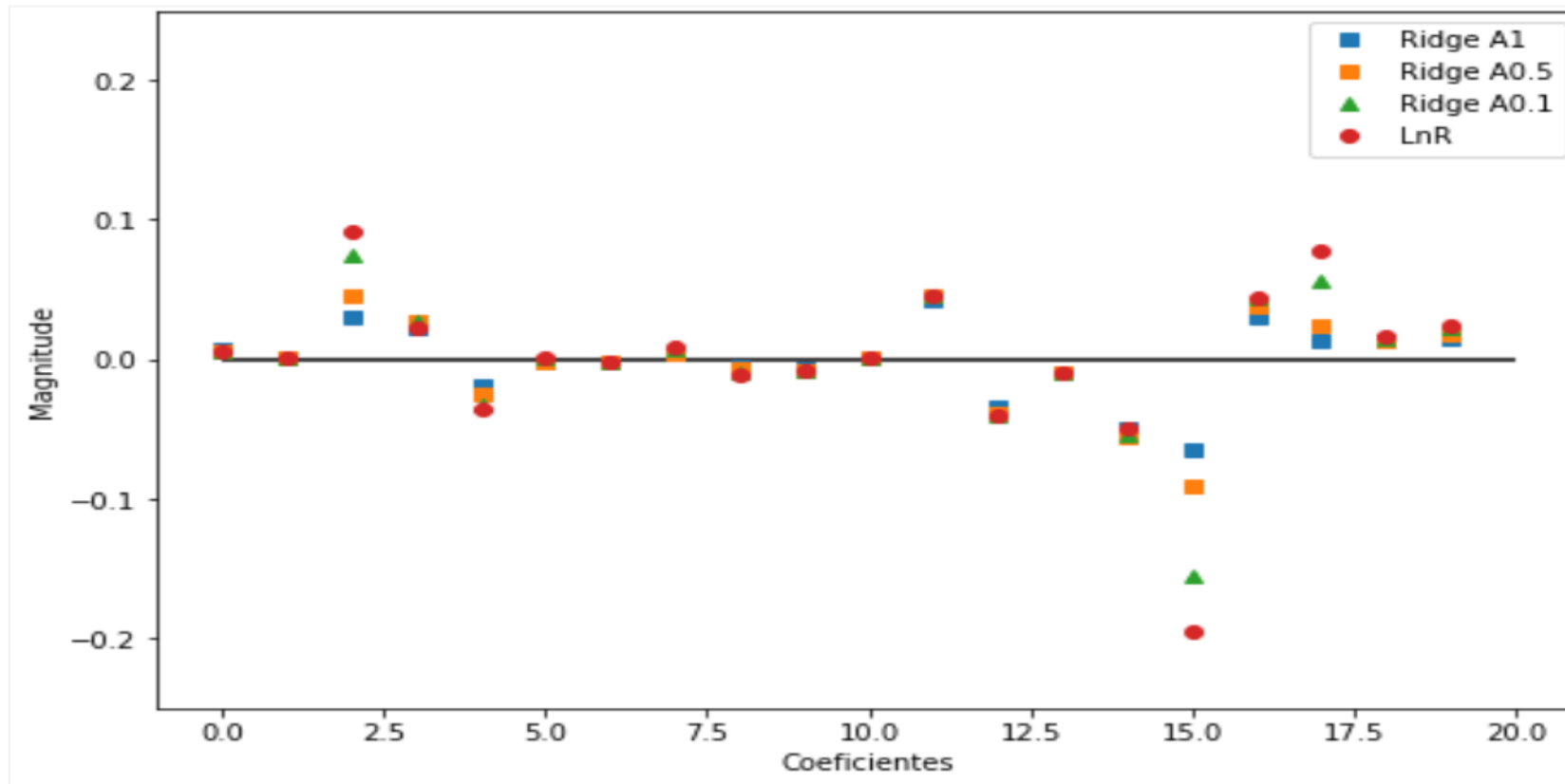
	0
BT_Pos	2.5
Swited	0.6
Stack_BB	-23.2
Pot_BB	-13.4
Hand_Power	1.3

Erro - Regressão Linear Simples



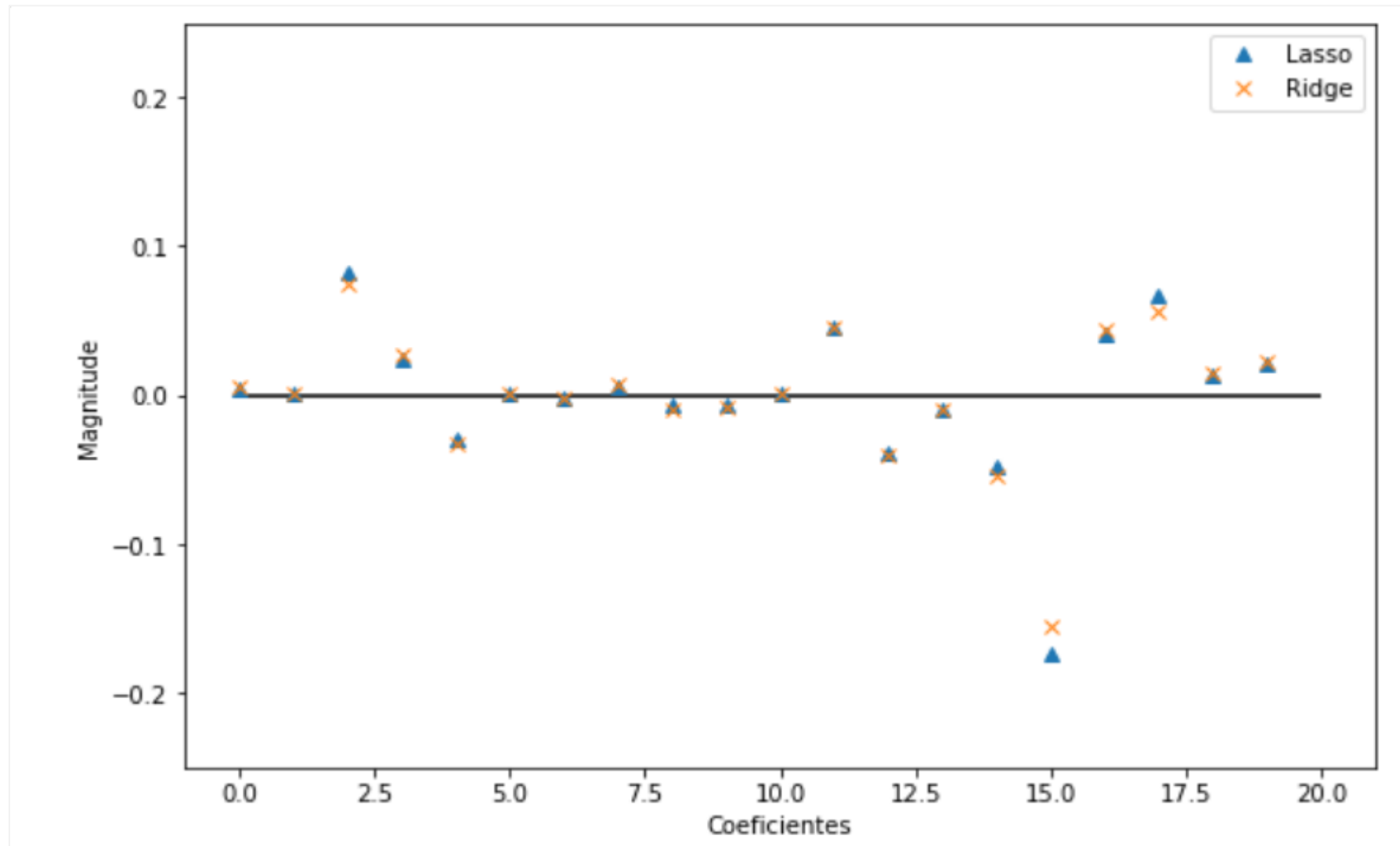
Regressão Linear com Atributos Polinomiais - *20 parâmetros

```
▶ from sklearn.preprocessing import PolynomialFeatures  
  
Xpoly=PolynomialFeatures(degree=2, include_bias=False).fit_transform(X)  
  
( 'Xpoly SHAPE --> ', Xpoly.shape)  
  
[ ] ( 'Xpoly SHAPE --> ', (1746, 20))
```



Regressão Linear com Atributos Polinomiais - *20 parâmetros

RIDGE x LASSO



Árvore de Decisão

```
SHAPE -> (1656, 5)
```

	BT_Pos	HP	SWT	POT	LDW
0	5	C	ST	A	D
1	6	D	ST	C	L
2	7	C	ST	D	L
3	8	A	ST	D	L
4	0	A	ST	D	D

#ENCODING

```
dfENC=dfEncode(dfLDW)  
dfENC.shape
```

```
[0, 1, 2, 3, 4, 5, 6, 7, 8]  
['A', 'B', 'C', 'D']  
['NS', 'ST']  
['A', 'B', 'C', 'D']  
['D', 'L', 'W']
```

```
ENCODED DF
```

	BT_Pos	HP	SWT	POT	LDW
0	5	2	1	0	0
1	6	3	1	2	1
2	7	2	1	3	1
3	8	0	1	3	1
4	0	0	1	3	0

```
(1656, 5)
```

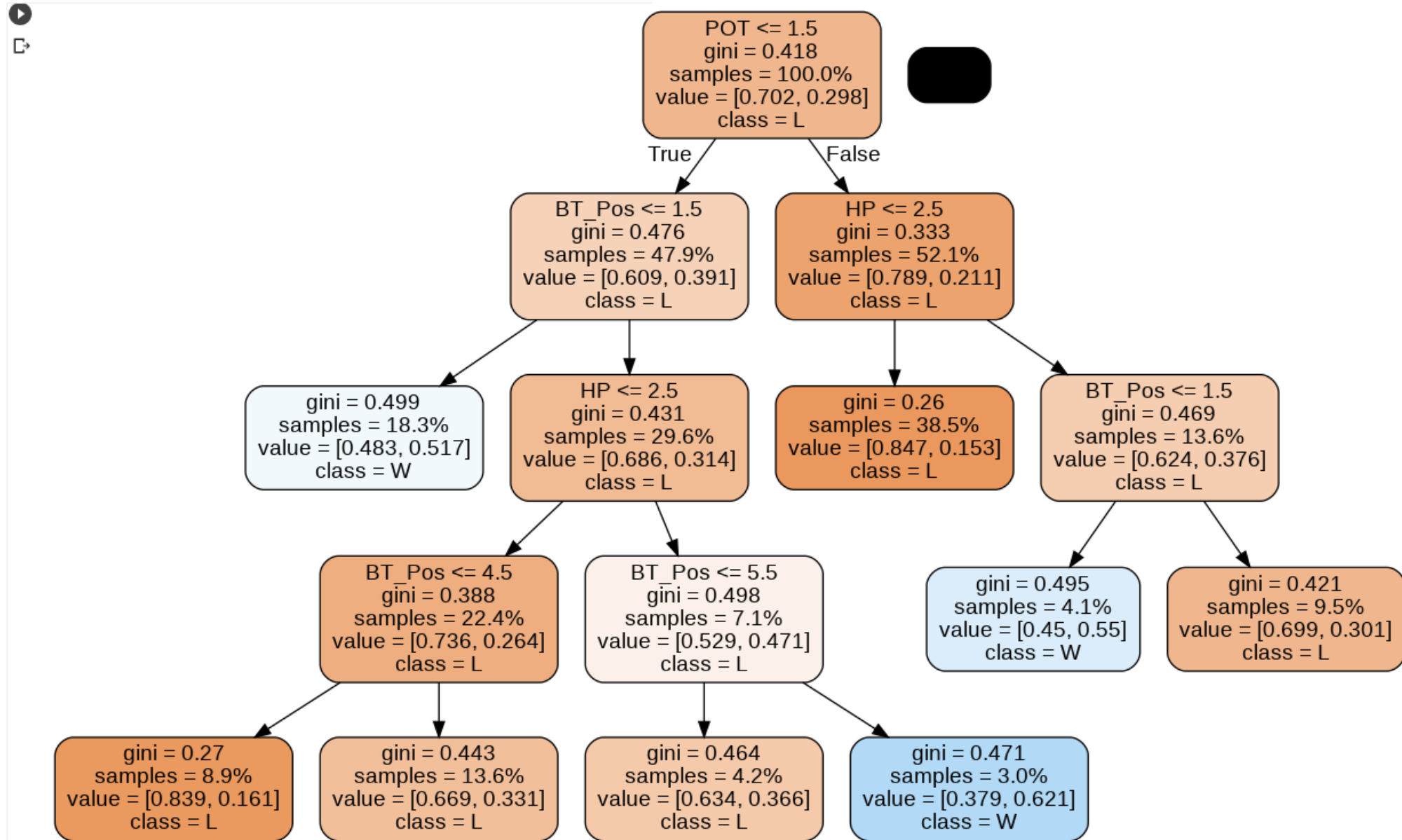
#REMOVENDO DRAWS

```
dfENC=dfENC[dfENC['LDW'] != 0]  
dfENC.shape
```

```
(1635, 5)
```

Árvore de Decisão

PARAMETROS -> {'criterion': 'gini', 'min_impurity_decrease': 0.002, 'min_weight_fraction_leaf': 0.017, 'splitter': 'best'}



Árvore de Decisão

	L	W
L	361	109
W	104	80

Acurácia do Treino: 72.0 %

Acurácia do Teste: 67.4 %

	precision	recall	f1-score	support
1	0.78	0.77	0.77	470
2	0.42	0.43	0.43	184
accuracy			0.67	654
macro avg	0.60	0.60	0.60	654
weighted avg	0.68	0.67	0.68	654

Naive-Bayes



Erros de Classificação: 23.9 %



Acurácia do Teste: 76.1 %

	precision	recall	f1-score	support
L	0.76	0.97	0.86	239
W	0.73	0.18	0.29	88
accuracy			0.76	327
macro avg	0.75	0.58	0.57	327
weighted avg	0.75	0.76	0.70	327
	L	W		
L	233	6		
W	72	16		

6 Interpretação dos Resultados

- O **ganho** do jogador praticamente nulo, indicando que seu padrão de jogo não tem refletido em lucro efetivo;
- Cliente joga preferencialmente com mãos '**Swited**' o que é vantajoso, porém tende a jogar com mão de naipes diferentes quando está com uma quantidade maior de fichas;
- **As maiores perdas** ocorrem quando o jogador se envolve em mão com muitas fichas ou quando seu Stack é muito superior ao Pot;
- **Resultados positivos** ocorrem preferencialmente em mãos cuja razão Pot/Stack seja inferior a 1.5 e em posições menores que 2;
- **Em posições acima de 2**, a força da mão deve ser superior a 2.5 que corresponde a **par de oitos "88"**.

6 Interpretação dos Resultados

- O modelo foi capaz de **compreender os princípios básicos** que levam à vitória: relação positiva para cartas naipadas e força das mão e negativa para Stack e Pot;
- **O melhor modelo** para regressão polinomial foi o RIDGE considerando alpha de 0.1 e 20 parâmetros;
- O modelo tem uma **boa capacidade de prever mão perdedora**, porém precisa evoluir para avaliar melhor as mãos vencedora;
- Melhores previsões caso o modelo utilize como banco de dados o **histórico de um jogador profissional** lucrativo;
- **Outros parâmetros** podem ser obtidos e considerados no modelo de forma a contribuir na tomada de decisão em cada etapa da mão;
- Evoluir para um **modelo autônomo**;

7 Apresentação dos Resultados





NING CANVAS

Designed for: TCC PUC Minas

Designed by: *Rodrigo Xavier*

Date: 14/08/2022

Iteration: 1

DECISÃO 	PÚBLICO ALVO 	FONTE DE DADOS 	ONDE CONSEGUIR 
Os modelos ajudam a tomar decisões durante as mãos de um jogo de poker com base da análise de ações vitoriosas do jogador.	Jogadores de Poker que queiram entender melhor as estatísticas do jogo e analisar suas ações na mesa de forma a torná-las mais lucrativas	Histórico de mãos obtido em Plataforma de jogos de poker.	Poker Stars 888 Poker Full Tilt Poker Party Poker

8 Links

<https://youtu.be/aArSSwsSamM>

https://github.com/rodrigoxal/TCC_PUCMinas

TABELA DADOS **JOGADOR** MODELO
CADA GANHOS APOSTA MÃOS **JOGO**
RELAÇÃO HISTÓRICO POKER LINEAR
MESA CARTA REGRESSÃO CARTAS STACK
FICHAS JOGAR ANÁLISE PARÂMETROS
ETAPA POSIÇÃO BOTÃO

Contato:

Rodrigo Xavier de Almeida Leão
rodrigoxal@gmail.com