

# PREDIÇÃO DE LINKS

---

*Tópicos Avançados em Inteligencia Artificial*  
*Prof. Dr. Alneu de Andrade Lopes*

*Vinicius Alvarenga Lovato 7696455*

# TÓPICOS

---

- Introdução
- Tecnologias Utilizadas
- Descrição dos Dados
- Técnicas e Preditores
- Resultados e Conclusão

# INTRODUÇÃO

---

- Diversos problemas podem ser representados por grafo
  - Aeroportos, Logística, Redes Sociais, entre outros.
- Predição de Link
  - Utilizar estado atual da rede para “prever o futuro”
  - Utilizado para medir interações entre proteínas ou até mesmo detectar redes de terrorismo

# TECNOLOGIAS UTILIZADAS

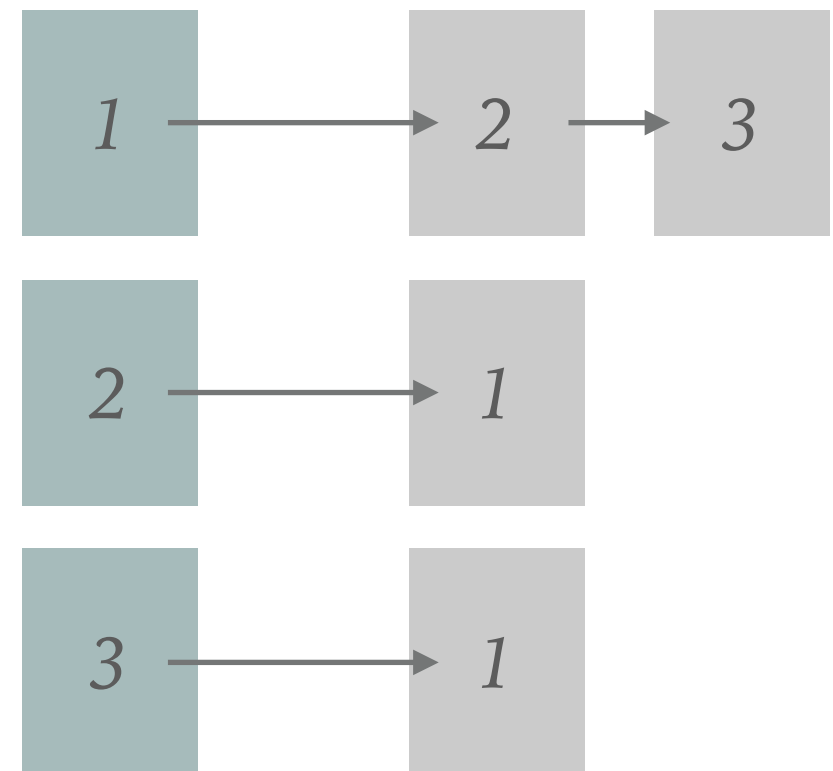
---

- Javascript
  - Flot
  - Papa parser

# DESCRIÇÃO DOS DADOS

---

- Conjunto de dados contendo conexões do facebook
- Cada linha uma conexão
- Amizade no facebook é mútua
- Representação escolhida: lista de adjacências

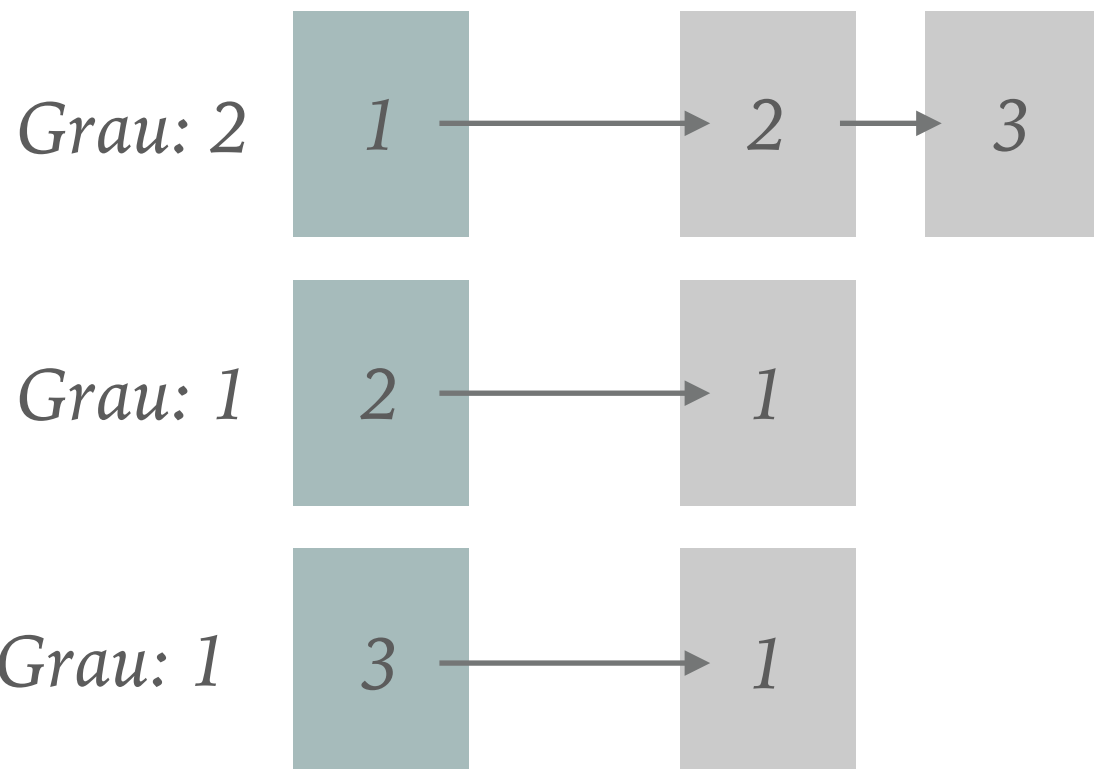


# TÉCNICAS UTILIZADAS E PREDITORES

---

## PROPRIEDADE DA REDE

- Grau do vértice a calculado a partir do tamanho da lista de adjacência



- Média de Grau

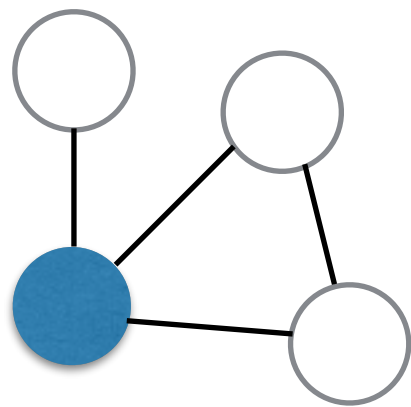
$$Md = \frac{(2+1+1)}{3} = \frac{4}{3}$$

# TÉCNICAS UTILIZADAS E PREDITORES

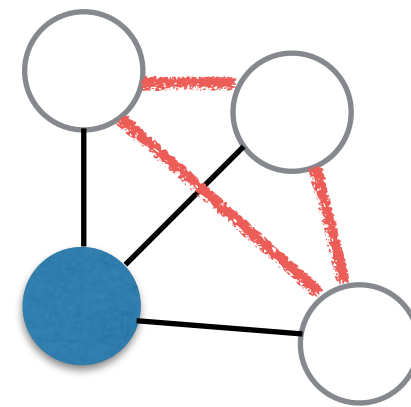
---

## PROPRIEDADE DA REDE

- Coeficiente de Agrupamento Local
- Número de conexões entre vizinhos (número de triângulos)



*Grafo com 4 nós*

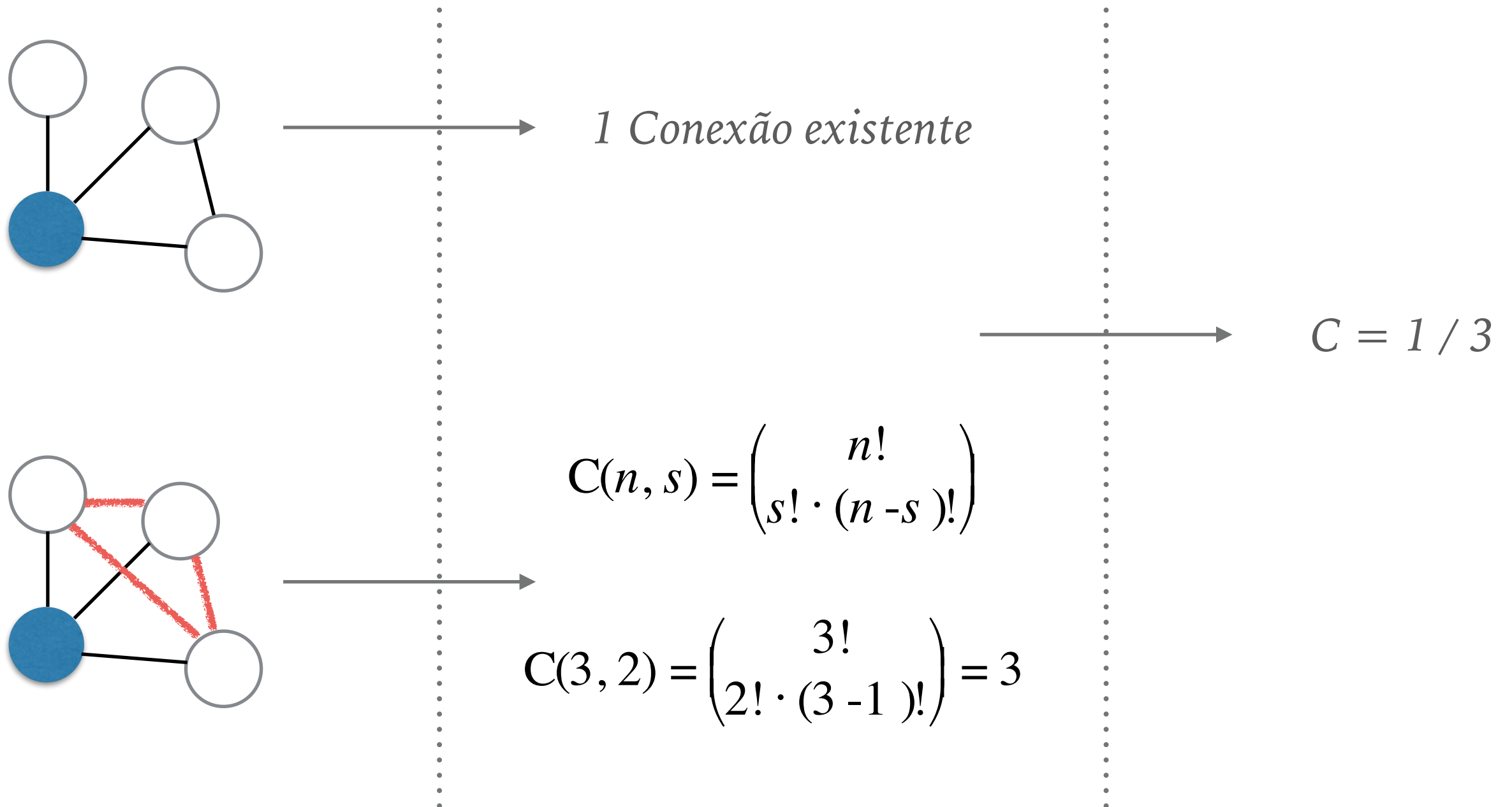


*Em vermelho total de conexões possíveis*

# TÉCNICAS UTILIZADAS E PREDITORES

---

## PROPRIEDADE DA REDE – COEFICIENTE DE AGRUPAMENTO





# TÉCNICAS UTILIZADAS E PREDITORES

---

## COMMON NEIGHBOURS E JACCARD

- Métodos baseados no cálculo de propriedades dos vizinhos do nó avaliado

### COMMON NEIGHBOURS

$$y = | \tau(x) \cap \tau(y) |$$

### JACCARD

$$y = \frac{| \tau(x) \cap \tau(y) |}{| \tau(x) \cup \tau(y) |}$$

# TÉCNICAS UTILIZADAS E PREDITORES

---

## COMMON NEIGHBOURS E JACCARD

- Métodos baseados no cálculo de propriedades dos vizinhos do nó avaliado

### COMMON NEIGHBOURS

$$y = | \tau(x) \cap \tau(y) |$$

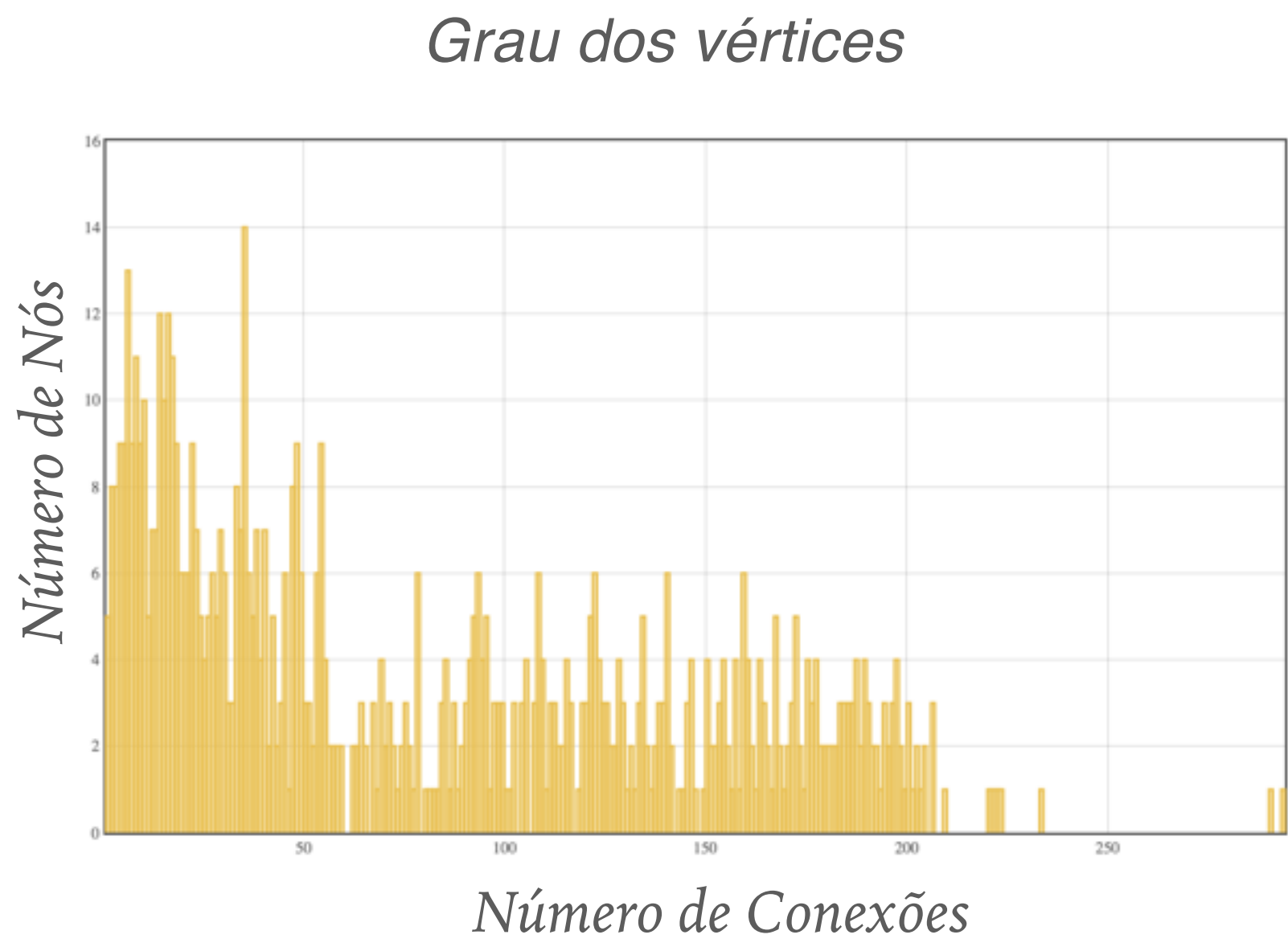
### JACCARD

$$y = \frac{| \tau(x) \cap \tau(y) |}{| \tau(x) \cup \tau(y) |}$$

# RESULTADOS

## MÉTRICAS GERAIS

Métricas	Valores
Número de Nós da Rede	748
Número de Conexões	30025
Média de grau da rede	80.2820
Coeficiente Agrupamento Médio	0.63540



# RESULTADOS

## ..... AVALIAÇÃO DOS ALGORITMOS DE PREDIÇÃO

➤ 90% dos links para treino

➤ Links preditos são comparados com os 10% removidos

Top 10	Common Neighbour	Rede Futura Contem Link	Jaccard	Rede Futura Contem Link
1	[1983,2266] = 179	NAO	[1961,2487] = 1	NAO
2	[2244,2464] = 174	SIM	[2080,2358] = 1	NAO
3	[2123,2324] = 173	SIM	[2244,2464] = 0.8405	SIM
4	[2218,2244] = 173	SIM	[2218,2244] = 0.8277	SIM
5	[2150,2206] = 171	SIM	[2150,2206] = 0.8181	SIM
6	[2206,2324] = 170	NAO	[2078,2593] = 0.8155	NAO
7	[2201,2266] = 169	NAO	[2123,2324] = 0.81220	SIM
8	[2078,2593] = 168	NAO	[2201,2206] = 0.81159	NAO
9	[2201,2206] = 168	NAO	[2590,2607] = 0.8029	SIM
10	[2088,2369] = 167	NAO	[2059,2131] = 0.7990	SIM

## ACERTOS

CN: 40%

Jaccard: 60%

# IDEIA PARA PRE-PROCESSAMENTO

---

- Remover hubs
  - Alteram comportamento da rede
- Remover nós com quantidade com desvio alto da média
  - Nós com poucas e muitas conexões

# CONCLUSÃO

---

- Possível avaliar surgimento de novos links utilizando medidas simples
- Custo computacional alto
- Área importante, principalmente com predição de redes terroristas