



INSTITUTO FEDERAL  
SÃO PAULO  
Câmpus Presidente Epitácio

# NAIVE BAYES

---

Compreensão do Algoritmo, Análise e  
Codificação de um Exemplo

Leandro R. de Souza  
r.leandro@aluno.ifsp.edu.br

Discente do Curso Bacharelado a Ciência da  
Computação – IFSP – Campus Presidente Epitácio

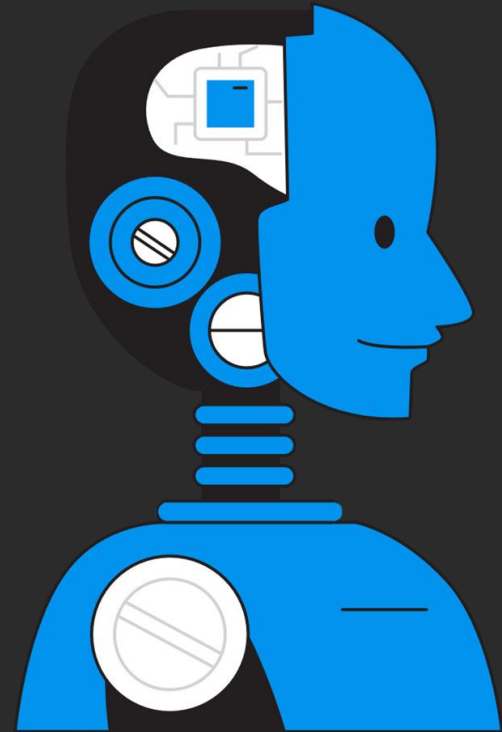
# > Roteiro

- Naive Bayes
  - Conceito.
  - Aplicação.
  - Vantagens.
- Exemplo prático - Manual.
  - Conhecimento da Base de Dados
  - Construção da Tabela de Probabilidade
  - Classificação de um Novo Registro
  - Vantagens e Desvantagens
- Exemplo prático - Linguagem R.



# > Naive Bayes

- Classificador Probabilístico.
- Ingênuo
  - Premissa Central do Algoritmo.
- Utilidade
  - Risco de Crédito
  - Spam de e-mail.
- Vantagens
  - Desempenho (Tempo de Treinamento Linear).
  - Baixa Variação.
  - Aprendizagem Incremental.



# > Base de Dados

História do Crédito	Dívida	Garantias	Renda Anual	Risco
Ruim	Alta	Nenhuma	< 15.000	Alto
Desconhecida	Alta	Nenhuma	>= 15.000 a <= 35.000	Alto
Desconhecida	Baixa	Nenhuma	>= 15.000 a <= 35.000	Moderado
Desconhecida	Baixa	Nenhuma	> 35.000	Alto
Desconhecida	Baixa	Nenhuma	> 35.000	Baixo
Desconhecida	Baixa	Adequada	> 35.000	Baixo
Ruim	Baixa	Nenhuma	< 15.000	Alto
Ruim	Baixa	Adequada	> 35.000	Moderado
Boa	Baixa	Nenhuma	> 35.000	Baixo
Boa	Alta	Adequada	> 35.000	Baixo
Boa	Alta	Nenhuma	< 15.000	Alto
Boa	Alta	Nenhuma	>= 15.000 a <= 35.000	Moderado
Boa	Alta	Nenhuma	> 35.000	Baixo
Ruim	Alta	Nenhuma	>= 15.000 a <= 35.000	Alto



# > Extração de Dados

	História do Crédito			Dívida		Garantias		Renda Anual		
Risco de Crédito	Boa	Desconhecida	Ruim	Alta	Baixa	Nenhuma	Adequada	< 15	>= 15 <= 35	> 35
Alto										
Moderado										
Baixo										



# > Extração de Dados

	História do Crédito		
Risco de Crédito	Boa	Desconhecida	Ruim
Alto			
Moderado			
Baixo			

História do Crédito	Risco
Desconhecida	Baixo
Desconhecida	Baixo
Boa	Baixo
Boa	Baixo
Boa	Baixo



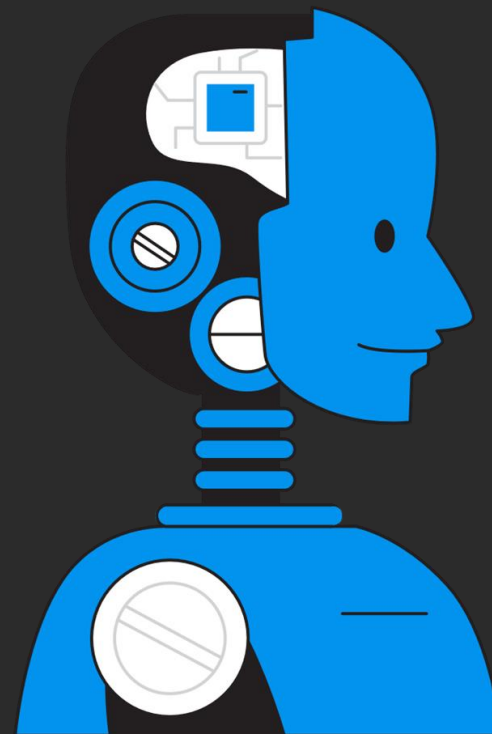
# > Após Extração

	História do Crédito			Dívida		Garantias		Renda Anual		
Risco de Crédito	Boa 5	Desconhecida 5	Ruim 4	Alta 7	Baixa 7	Nenhuma 11	Adequada 3	< 15 3	>= 15 <= 35 4	35 7
Alto 6 14	1 6	2 6	3 6	4 6	2 6	6 6	0	3 6	2 6	1 6
Moderado 3 14	1 3	1 3	1 3	1 3	2 3	2 3	1 3	0	2 3	1 3
Baixo 5 14	3 5	2 5	0	2 5	3 5	3 5	2 5	0	0	5 5



## > Classificando um Registro

- Um novo registro:
  - História: Boa.
  - Dívida: Alta.
  - Garantias: Nenhuma.
  - Renda:  $> 35$ .
- Calcular a Probabilidade Individual.
- Selecionar o maior “valor”.





# > Classificando um Registro

História: Boa.

Dívida: Alta.

Garantias: Nenhuma.

Renda: > 35.

	História do Crédito			Dívida		Garantias		Renda Anual		
Risco de Crédito	Boa 5			Alta 7		Nenhuma 11				> 35 7
Alto 6 14	1 6			4 6		6 6				1 6
Moderado 3 14	1 3			1 3		2 3				1 3
Baixo 5 14	3 5			2 5		3 5				5 5



# > Classificando um Registro

História: Boa.

Dívida: Alta.

Garantias: Nenhuma.

Renda: > 35.

	História do Crédito			Dívida		Garantias		Renda Anual		
Risco de Crédito	Boa 5			Alta 7		Nenhuma 11				> 35 7
Alto 6 14	1 6			4 6		6 6				1 6
Moderado 3 14	1 3			1 3		2 3				1 3
Baixo 5 14	3 5			2 5		3 5				5 5

$$P(Alto) = \frac{6}{14} * \frac{1}{6} * \frac{4}{6} * \frac{6}{6} * \frac{1}{6}$$

$$P(Alto) = 0,0079$$

$$P(Moderado) = \frac{3}{14} * \frac{1}{3} * \frac{1}{3} * \frac{2}{3} * \frac{1}{3}$$

$$P(Moderado) = 0,0052$$

$$P(Baixo) = \frac{5}{14} * \frac{3}{5} * \frac{2}{5} * \frac{3}{5} * \frac{5}{5}$$

$$P(Baixo) = 0,0514$$





INSTITUTO FEDERAL  
SÃO PAULO  
Câmpus Presidente Epitácio

# OBRIGADO PELA ATENÇÃO

---

Leandro R. de Souza  
`r.leandro@aluno.ifsp.edu.br`

Discente do Curso Bacharelado a Ciência da  
Computação – IFSP – Campus Presidente Epitácio

