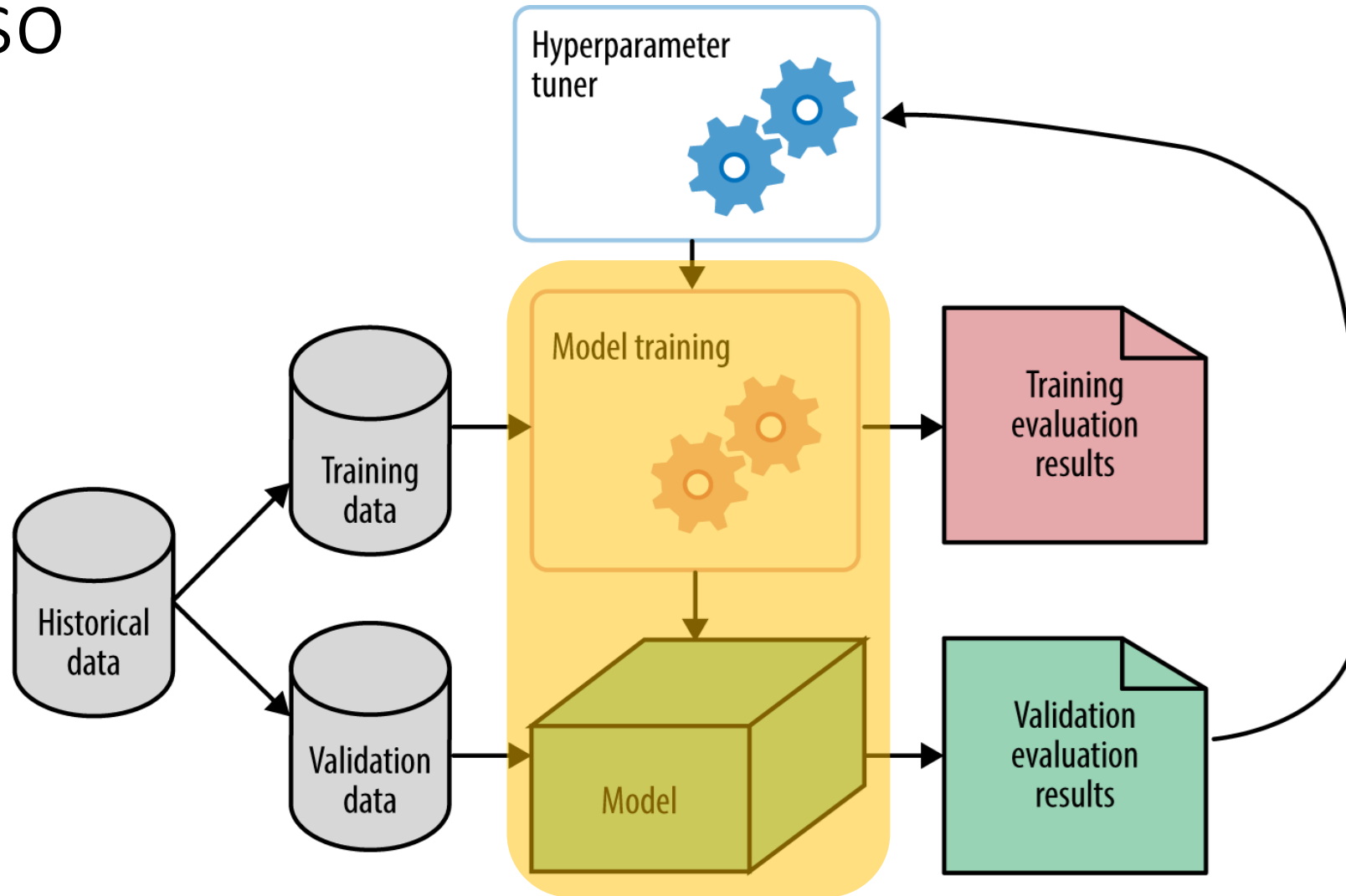


# Aprendizado de Máquina e Deep Learning

## Regressão

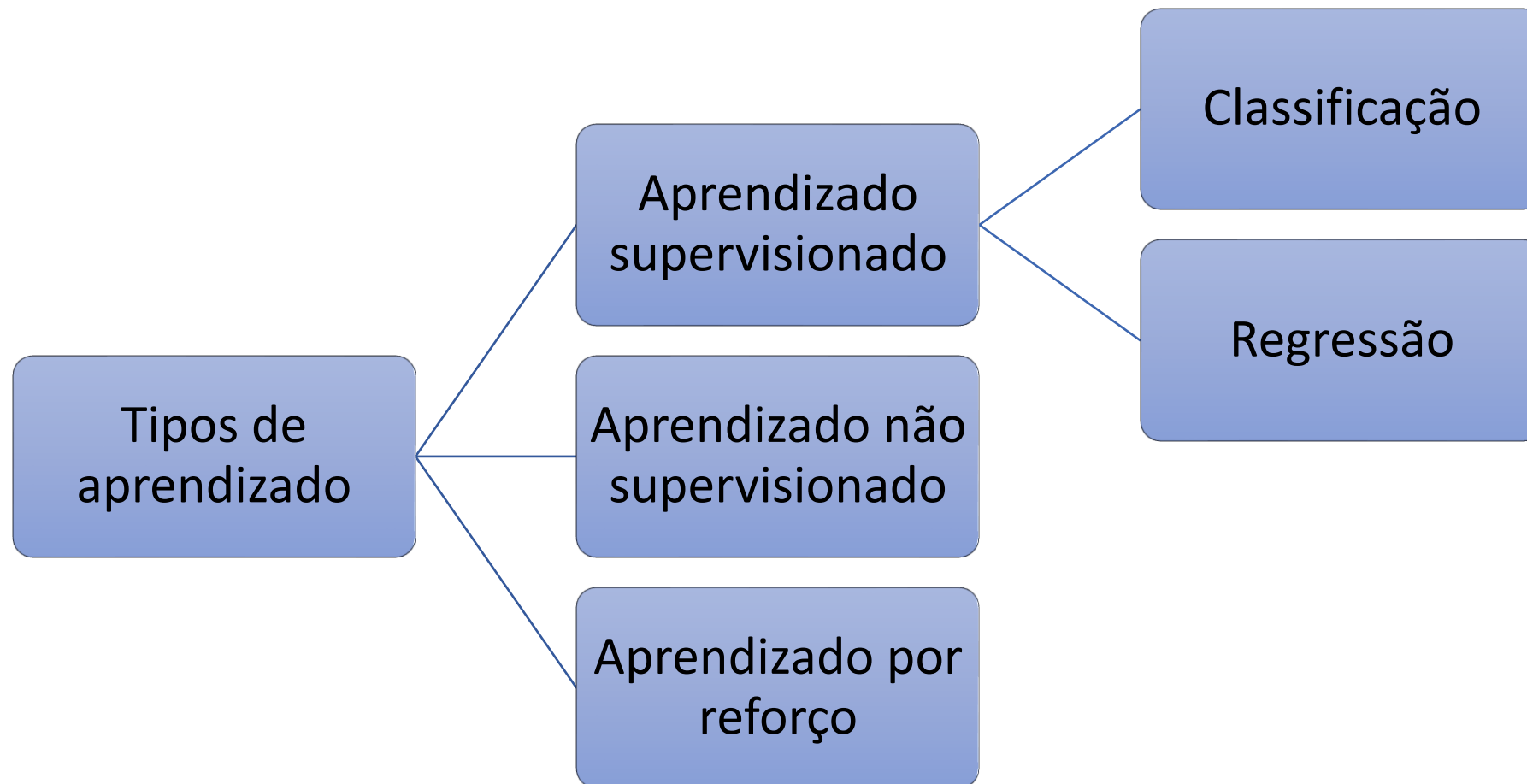
Prof. Dr. Thiago Meirelles Ventura

# Processo

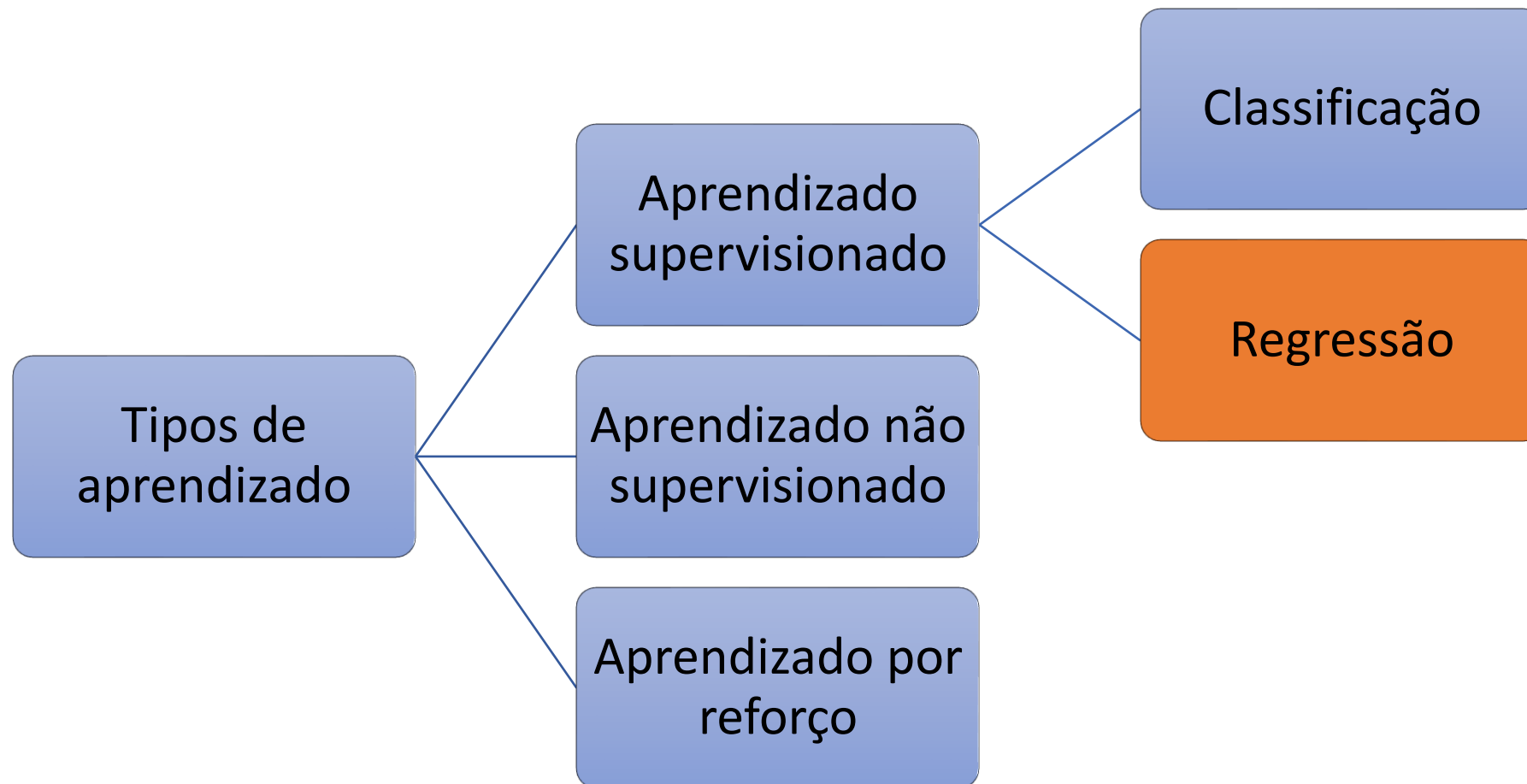


Deep Learning Bible - A. End-to-End Pipelines

# Tipos de aprendizado



# Tipos de aprendizado



# Regressão

- O objetivo é estimar valores contínuos
- Possibilita predição de valores desconhecidos ou ausentes
- Aplicações de exemplos de regressão são as que estimam o dado de uma temperatura ou do valor da bolsa de valores

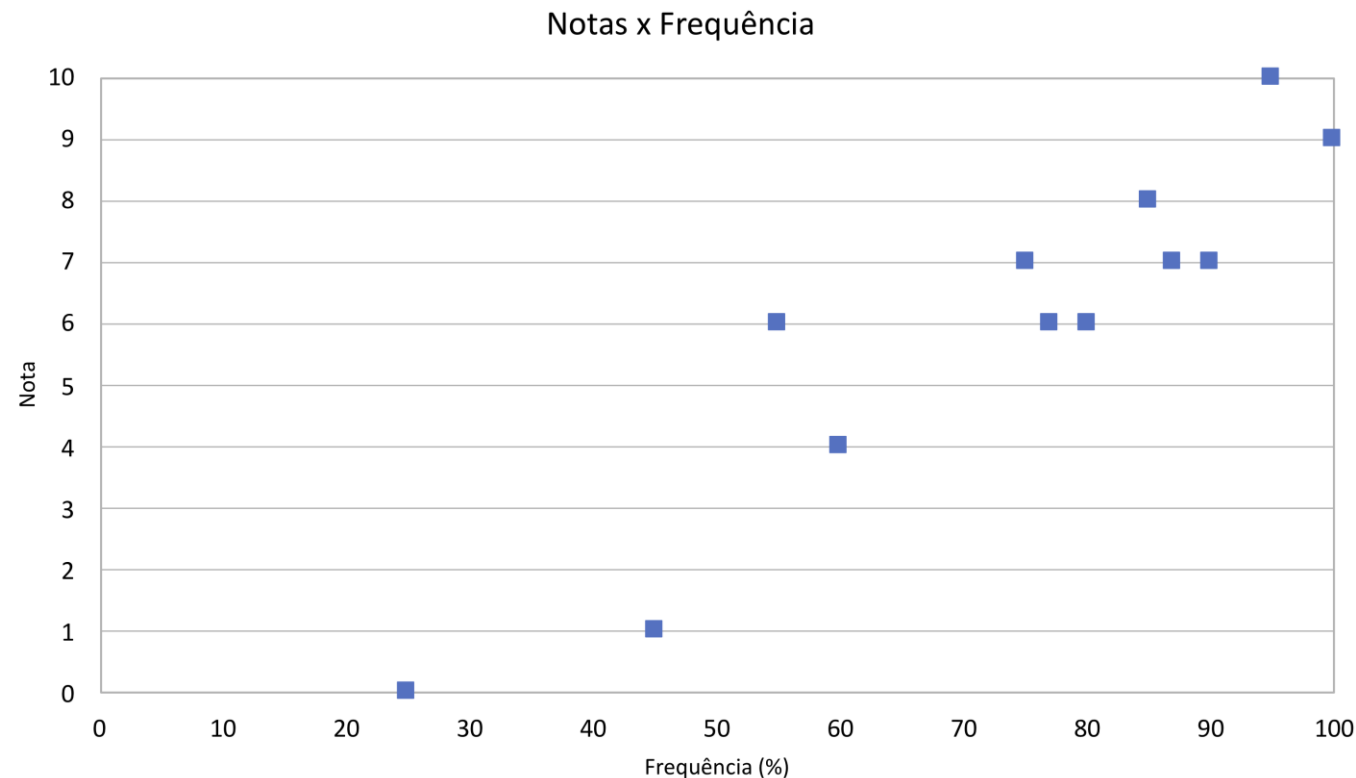
# Regressão

- Vamos tentar estimar a nota de um aluno apenas sabendo a sua frequência nas aulas
- Dados: frequência e nota final

Frequência	Nota
100	9
95	10
90	7
85	8
87	7
75	7
80	6
77	6
55	6
60	4
45	1
25	0

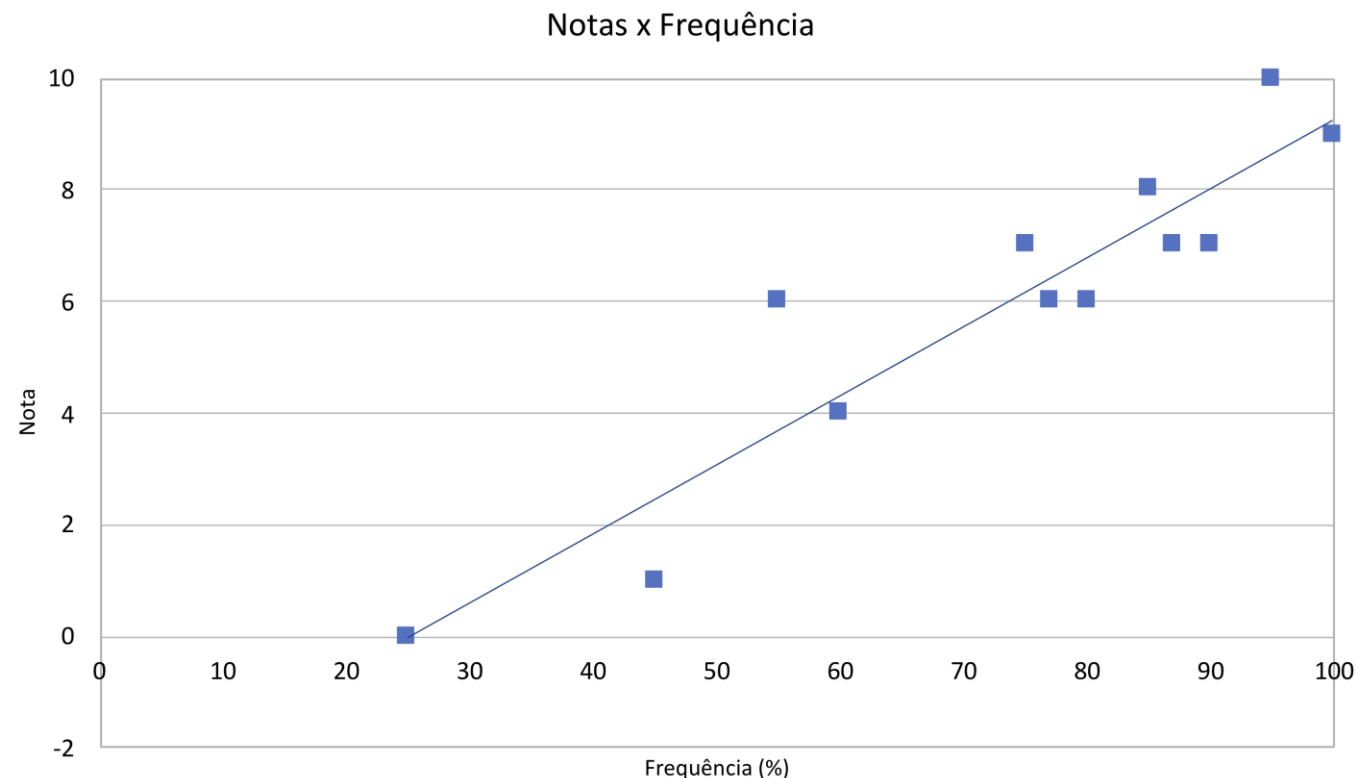
# Regressão

- Baseado nos dados, é possível criar um gráfico que relaciona a frequência com a nota



# Regressão

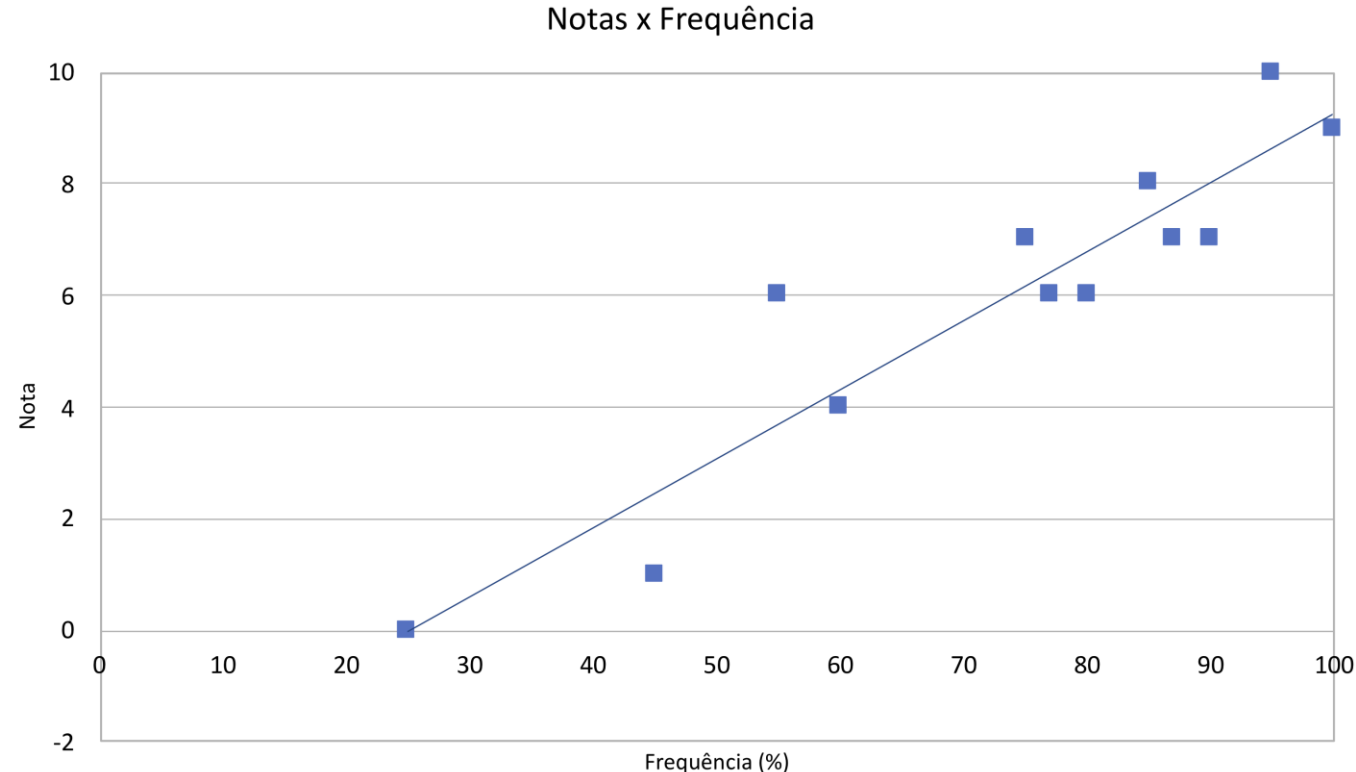
- É possível criar uma equação que represente esses dados, por meio de um método de regressão





# Regressão

- Este método simples já permite a estimativas de notas
- Qual a nota estimada que um aluno terá comparecendo em 50% das aulas?



# Regressão

- Mesmo não tendo essa frequência na tabela, o método utilizado permite realizar a estimativa
- Isso foi feito graças aos exemplos passados

# Regressão

- A Regressão Linear gera uma hipótese no formato
  - $h(x) = \alpha x + \beta$
- Onde
  - $x$  é a frequência
  - $\alpha$  e  $\beta$  devem ser estimados baseado no conjunto de treinamento

# Regressão

- A Regressão Linear gera uma hipótese no formato
  - $h(x) = \alpha x + \beta$
- $\alpha$  e  $\beta$  devem ter valores que consigam estimar com o menor erro possível todos os dados do conjunto de treinamento

# Exercício 3

- Das possibilidades abaixo, qual seria a melhor opção para os dados que foram passados?
  - a)  $\alpha = 1$ ;  $\beta = 2$
  - b)  $\alpha = 0$ ;  $\beta = 0$
  - c)  $\alpha = 0,5$ ;  $\beta = -41$
  - d)  $\alpha = 0,1$ ;  $\beta = -3$
  - e)  $\alpha = 0$ ;  $\beta = 1,5$
  - f)  $\alpha = 3$ ;  $\beta = -1$

Frequência	Nota
100	9
95	10
90	7
85	8
87	7
75	7
80	6
77	6
55	6
60	4
45	1
25	0

# Exercício 4

- Considerando esses valores
  - $\alpha = 0,1$ ;  $\beta = -3$
- Quais as estimativas e erros desses exemplos?
  - Frequência = 91; Nota = 6
  - Frequência = 74; Nota = 10
  - Frequência = 57; Nota = 5
  - Frequência = 26; Nota = 8

Frequência	Nota
100	9
95	10
90	7
85	8
87	7
75	7
80	6
77	6
55	6
60	4
45	1
25	0

# Exercício 5

- Considerando os valores anteriores, qual o Erro Médio Absoluto do modelo?

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - d_i|$$

# Programação

- Preparação dos dados
- Construção do modelo
- Avaliação dos resultados



# Programação – preparação dos dados

```
import pandas as pd
```

```
df = pd.read_csv('matogrosso.csv')
```

```
x = df[['ESCOLARIZACAO']]
```

```
y = df['IDHM']
```

# Programação – preparação dos dados

```
from sklearn.model_selection import train_test_split
```

```
X_train, X_test, y_train, y_test =  
    train_test_split(  
        x, y,  
        test_size=0.3,  
        random_state=18)
```

# Programação – construção do modelo

```
from sklearn.linear_model import LinearRegression
```

```
regressao = LinearRegression()
```

```
regressao.fit(X_train, y_train)
```

# Programação – avaliação dos resultados

```
predicao = regressao.predict(X_train)
```

```
plt.scatter(predicao, y_train)
```

```
plt.xlabel('IDHM real')
```

```
plt.ylabel('IDHM estimado')
```

# Programação – avaliação dos resultados

```
from sklearn import metrics
```

```
print('r2:', metrics.r2_score(y_train, predicacao))  
print('MAE:', metrics.mean_absolute_error(y_train, predicacao))  
print('MSE:', metrics.mean_squared_error(y_train, predicacao))  
print('RMSE:', np.sqrt(metrics.mean_squared_error(y_train,  
predicacao)))
```

script regressao.ipynb

# Exercício 6

- Melhore o desempenho do modelo
  - Teste novos atributos
  - Use mais de 1 atributo
- Qual foi o seu melhor resultado?
  - Escolha uma métrica
  - Informe o valor da métrica escolhida
  - Compare com os outros resultados
  - Informe o(s) atributo(s) utilizado(s)