

O que é classificação?



Maior65Anos: 0



Maior65Anos: 0

Conjuge: Nao



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

MesesDeContrato: 0



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

MesesDeContrato: 0

TelefoneFixo: Sim



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

MesesDeContrato: 0

TelefoneFixo: Sim

VariasLinhasTelefonicas: Nao



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

MesesDeContrato: 0

TelefoneFixo: Sim

VariasLinhasTelefonicas: Nao

ServicoDeInternet: FibraOptica



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

MesesDeContrato: 0

TelefoneFixo: Sim

VariasLinhasTelefonicas: Nao

ServicoDeInternet: FibraOptica

SegurancaOnline: Nao



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

MesesDeContrato: 0

TelefoneFixo: Sim

VariasLinhasTelefonicas: Nao

ServicoDeInternet: FibraOptica

SegurancaOnline: Nao

BackupOnline: Sim



SeguroNoDispositivo: Nao



SeguroNoDispositivo: Nao

SuporteTecnico: Sim



SeguroNoDispositivo: Nao

SuporteTecnico: Sim

TVaCabo: Nao



SeguroNoDispositivo: Nao

SuporteTecnico: Sim

TVaCabo: Nao

StreamingDeFilmes: Sim



SeguroNoDispositivo: Nao

SuporteTecnico: Sim

TVaCabo: Nao

StreamingDeFilmes: Sim

TipoDeContrato: UmAno



SeguroNoDispositivo: Nao

SuporteTecnico: Sim

TVaCabo: Nao

StreamingDeFilmes: Sim

TipoDeContrato: UmAno

PagamentoOnline: Sim



SeguroNoDispositivo: Nao

SuporteTecnico: Sim

TVaCabo: Nao

StreamingDeFilmes: Sim

TipoDeContrato: UmAno

PagamentoOnline: Sim

FormaDePagamento: DebitoEmConta



SeguroNoDispositivo: Nao

SuporteTecnico: Sim

TVaCabo: Nao

StreamingDeFilmes: Sim

TipoDeContrato: UmAno

PagamentoOnline: Sim

FormaDePagamento: DebitoEmConta

ContaMensal: 39.90



SeguroNoDispositivo: Nao

SuporteTecnico: Sim

TVaCabo: Nao

StreamingDeFilmes: Sim

TipoDeContrato: UmAno

PagamentoOnline: Sim

FormaDePagamento: DebitoEmConta

ContaMensal: 39.90

Churn: ?



Churn

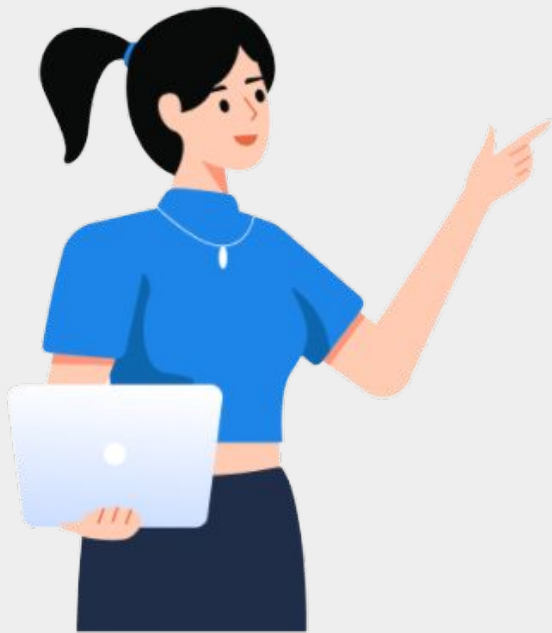




Churn



**Continua
Cliente (0)**



Churn

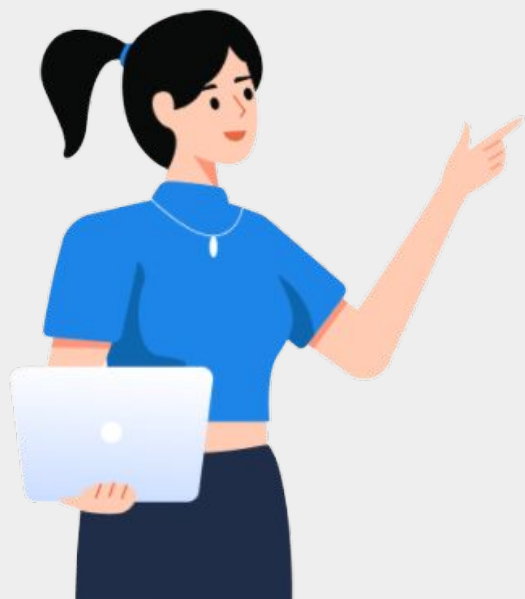


**Continua
Cliente (0)**



**Deixa de ser
Cliente (1)**

**O que é K-Nearest
Neighbors?**



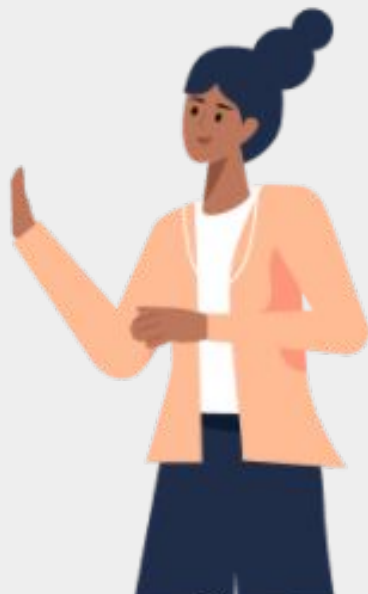
Maria

Pedro





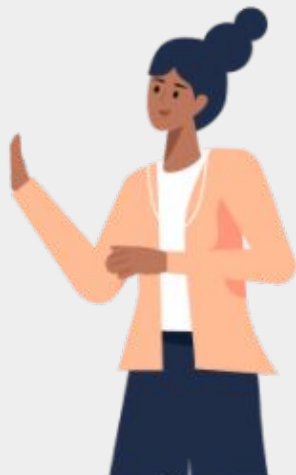
Pedro



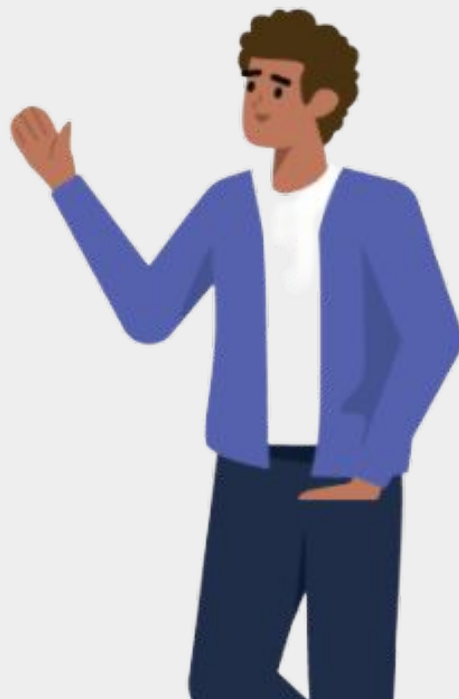
Júlia



Pedro



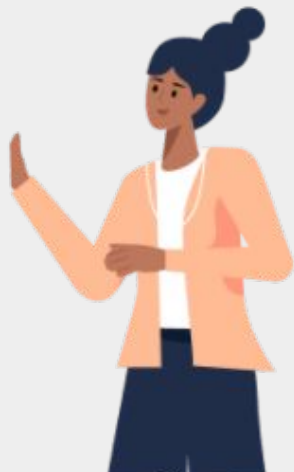
Júlia



João



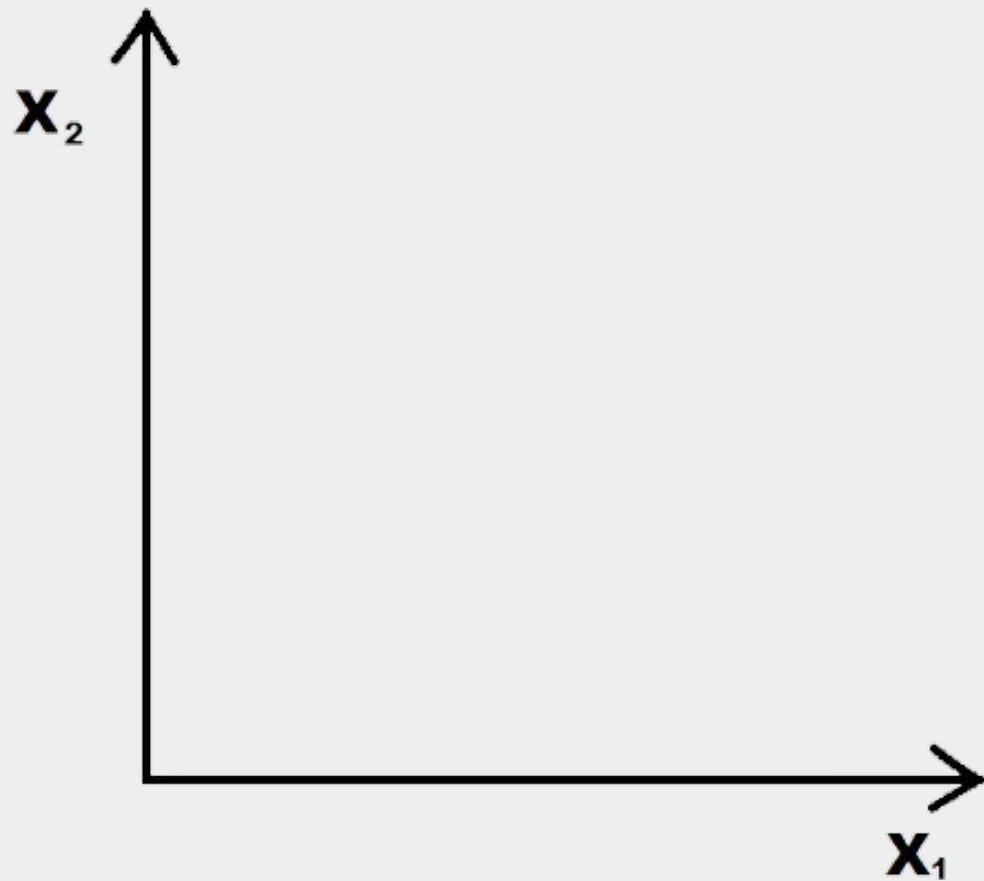
Pedro

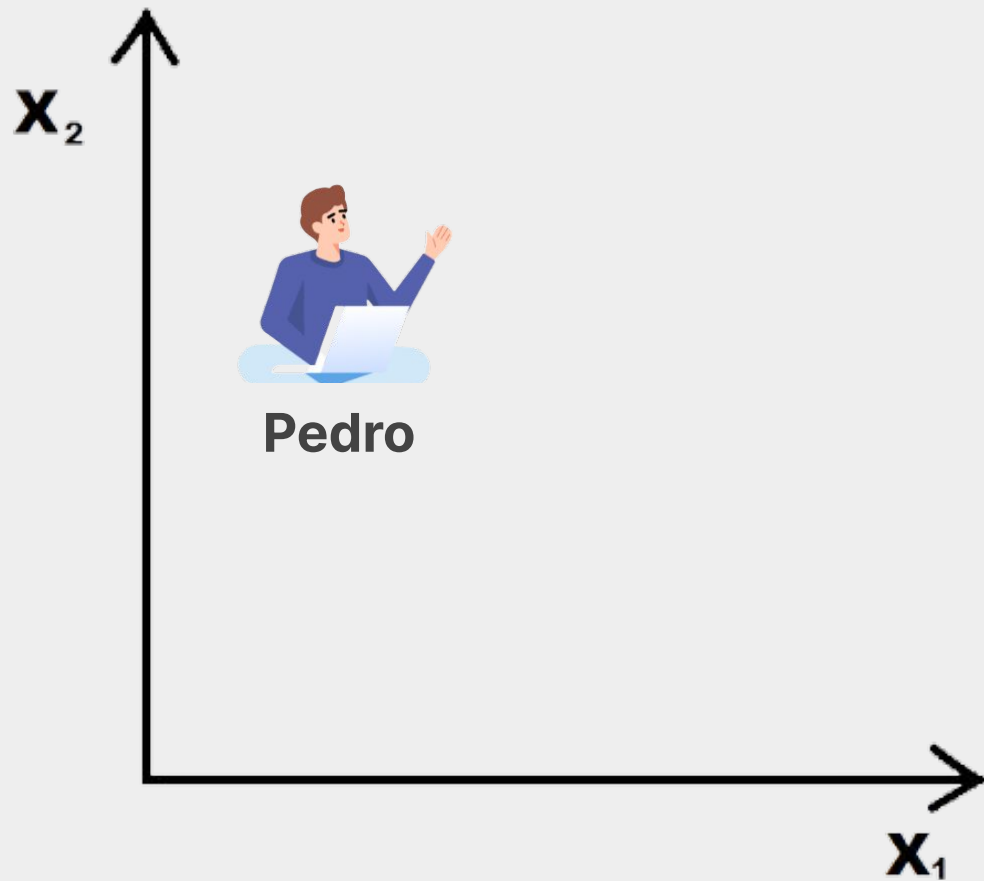


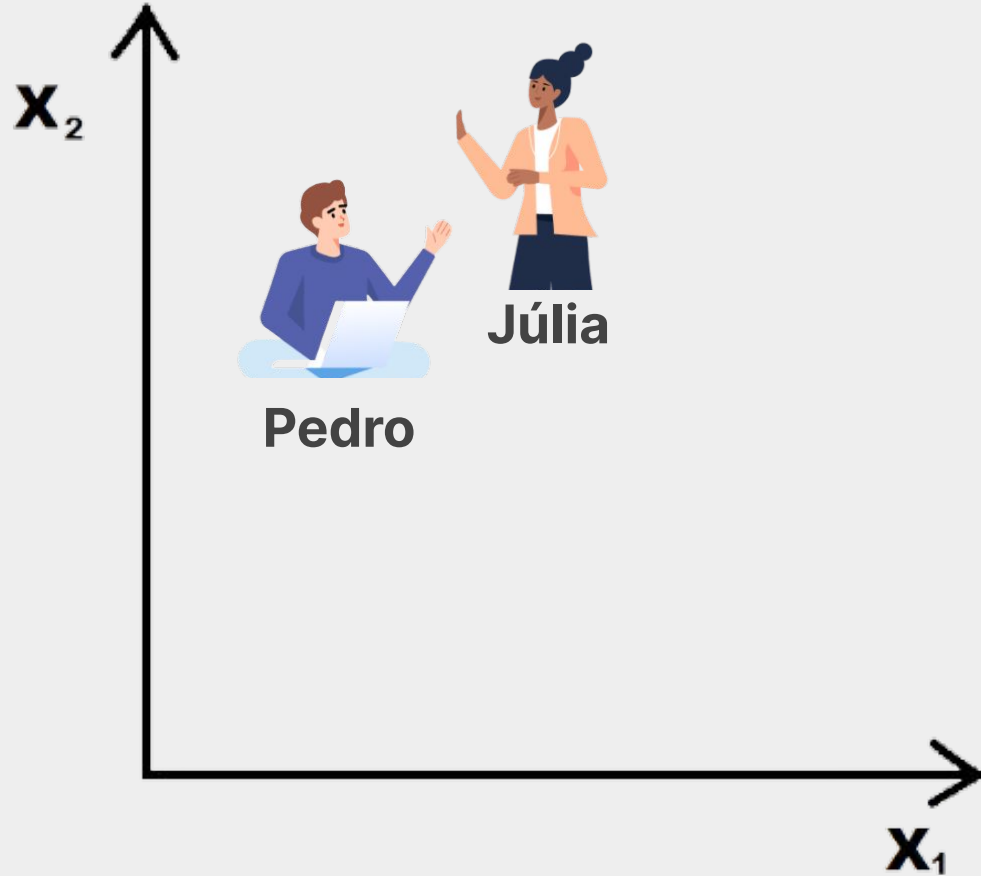
Júlia

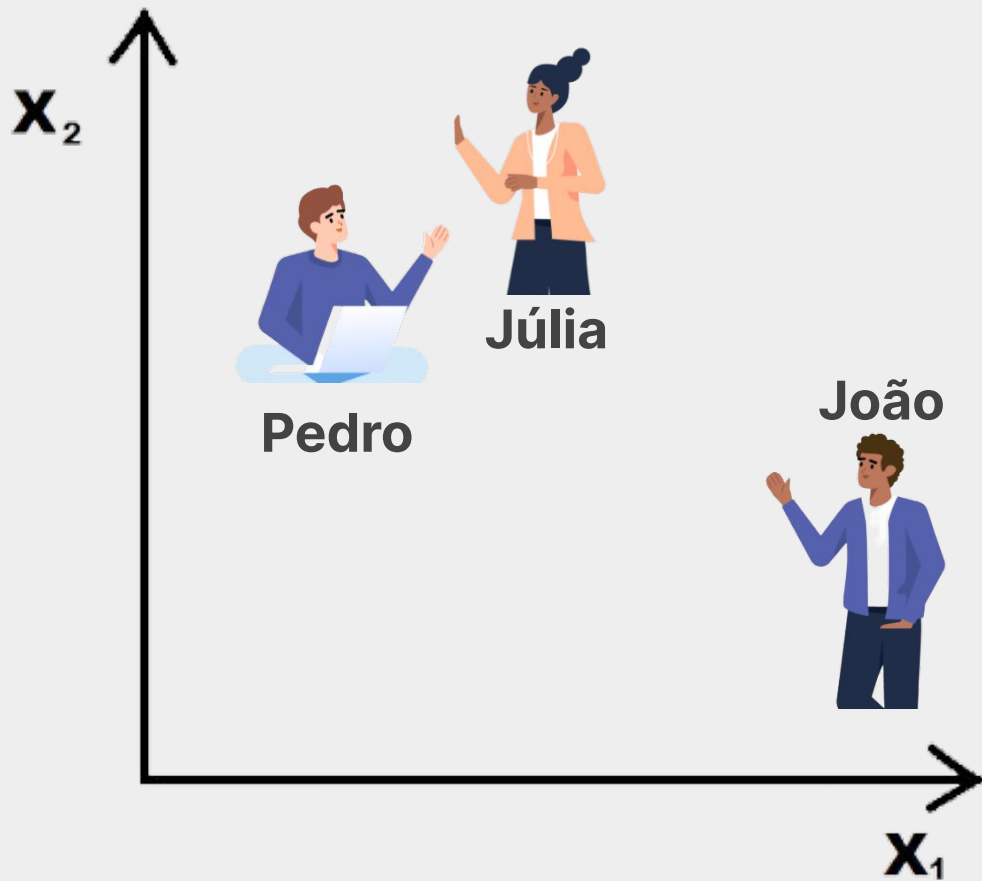


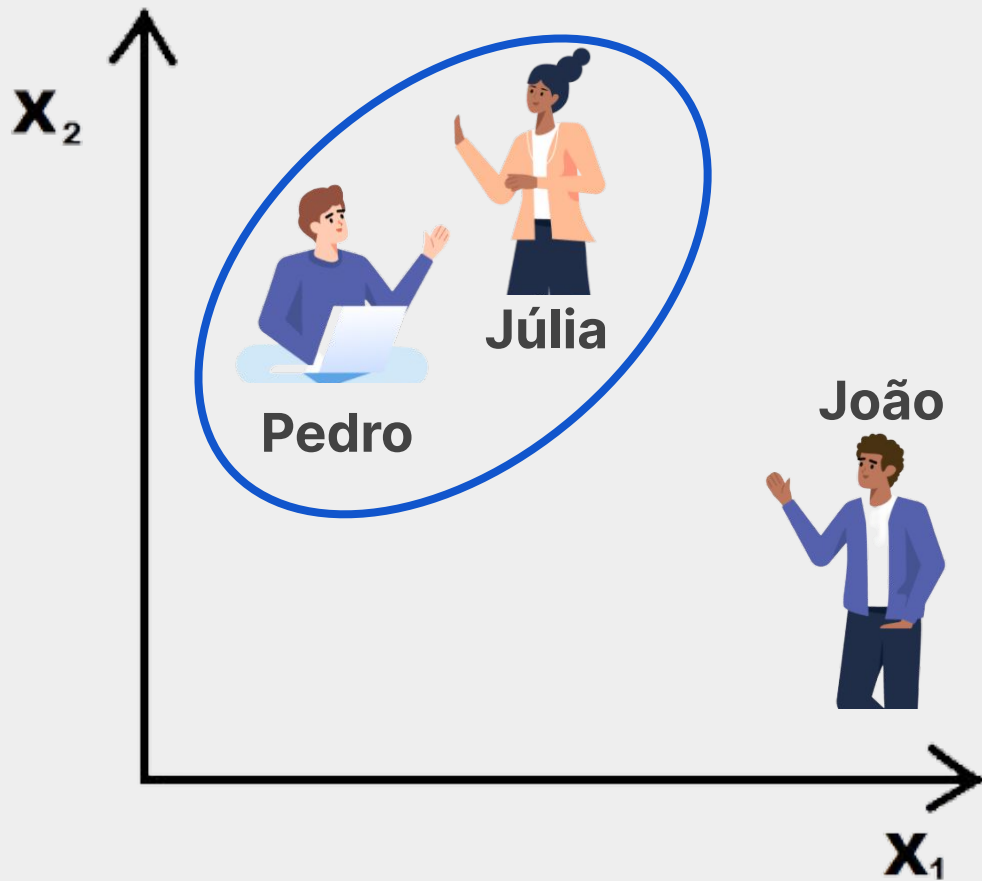
João

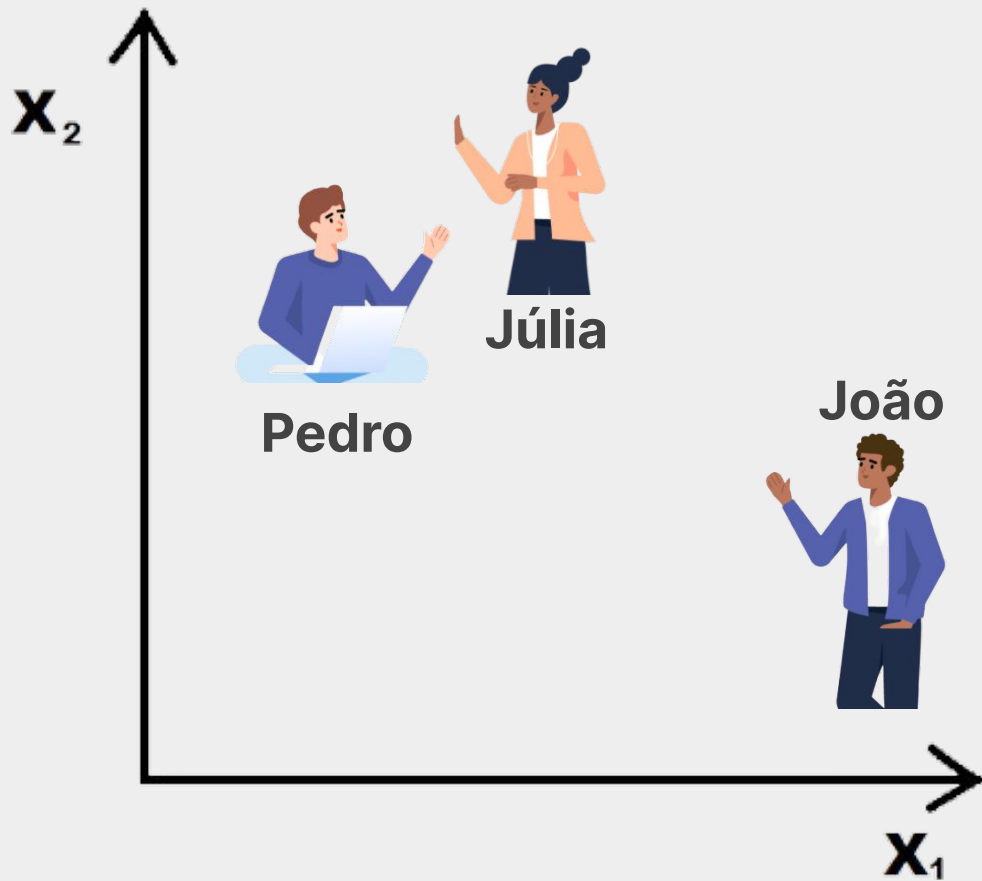


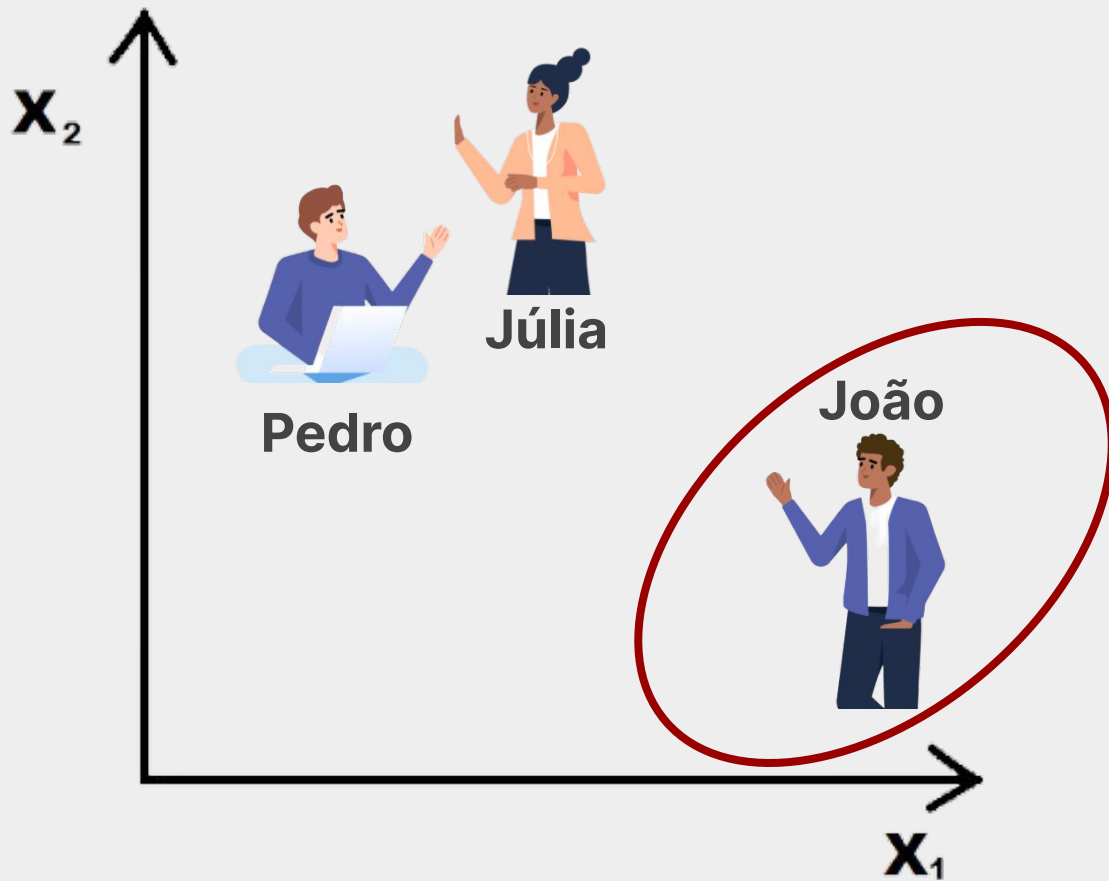


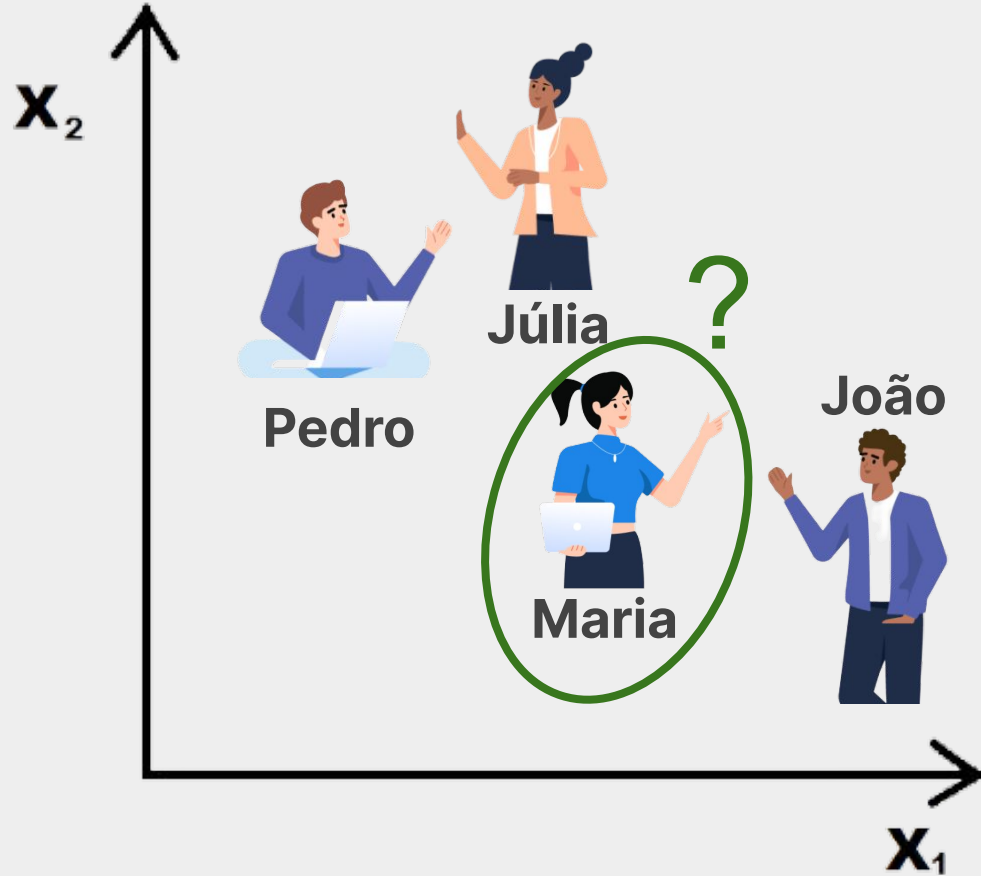


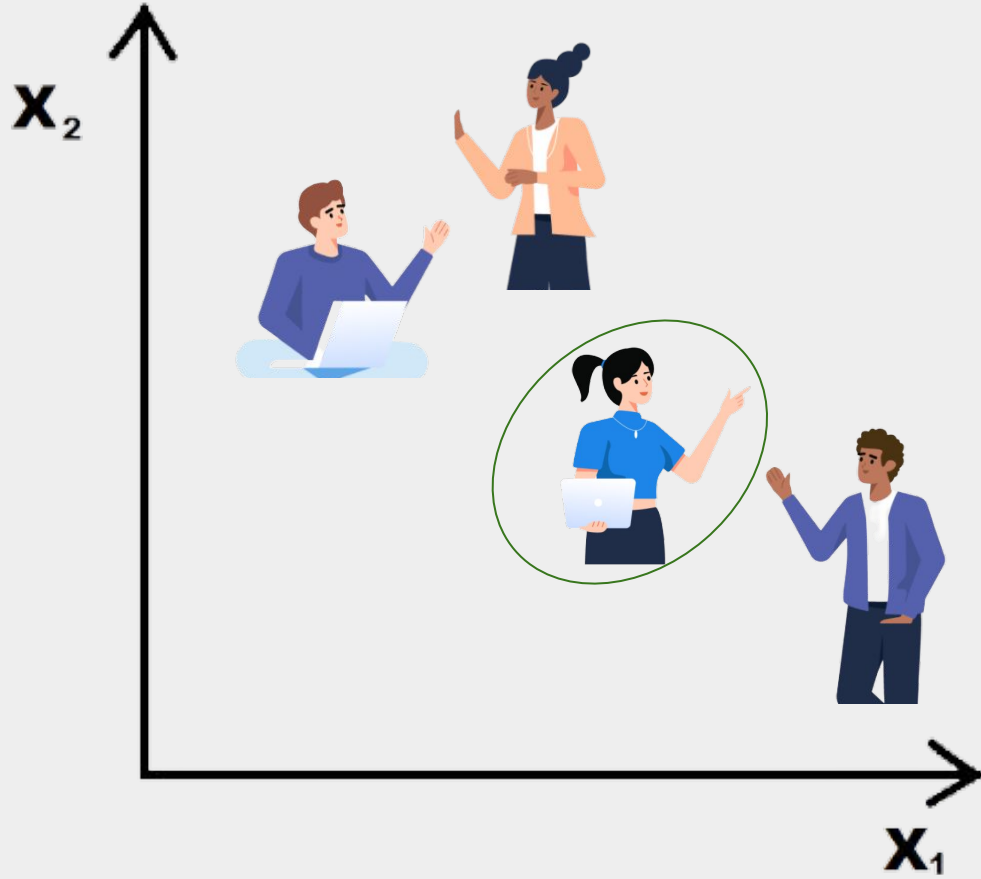


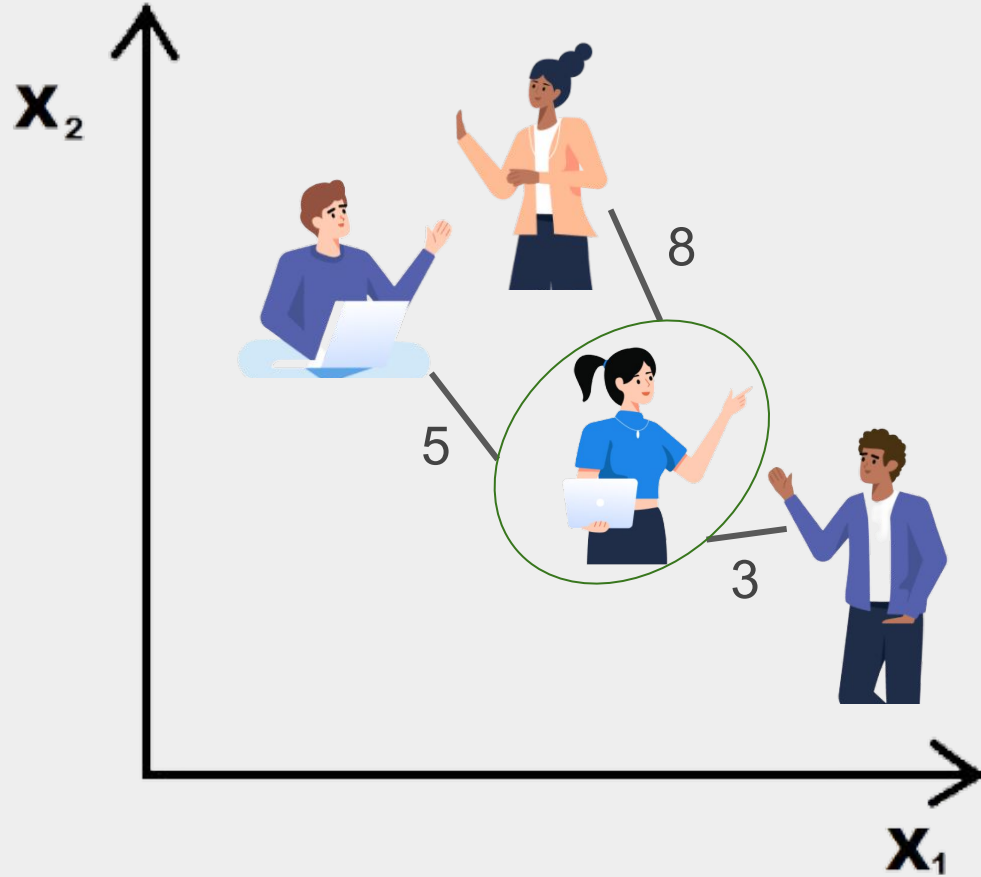


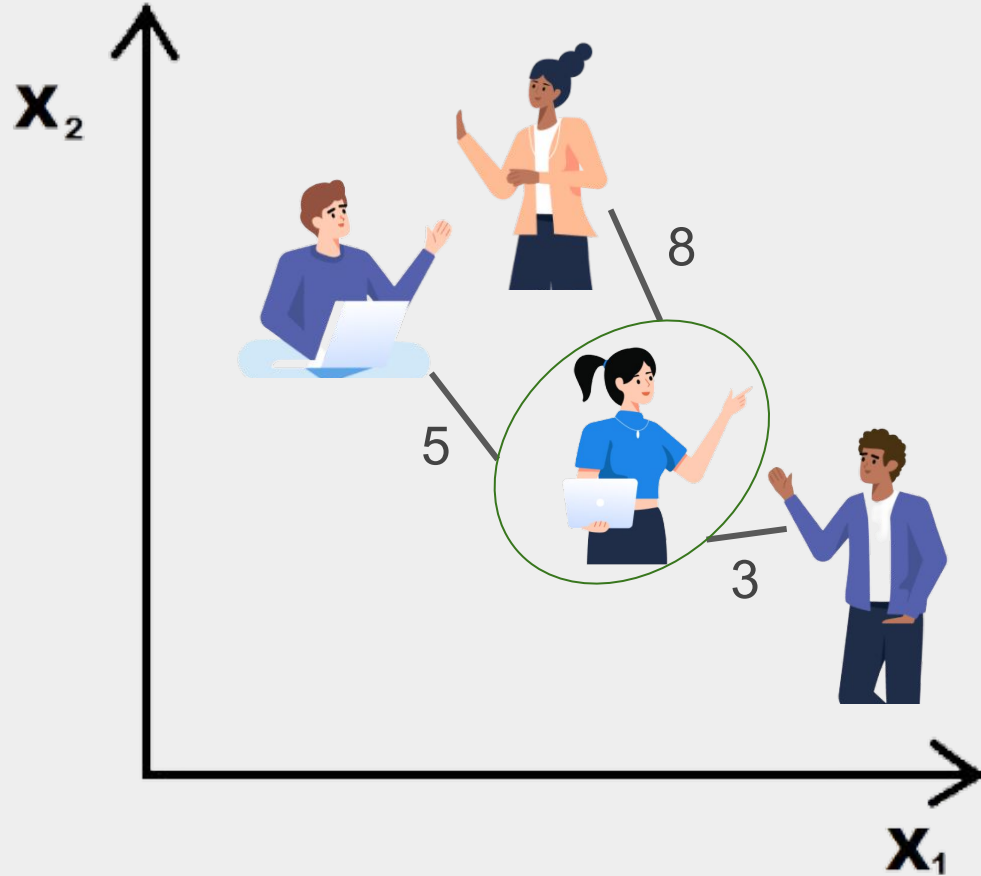




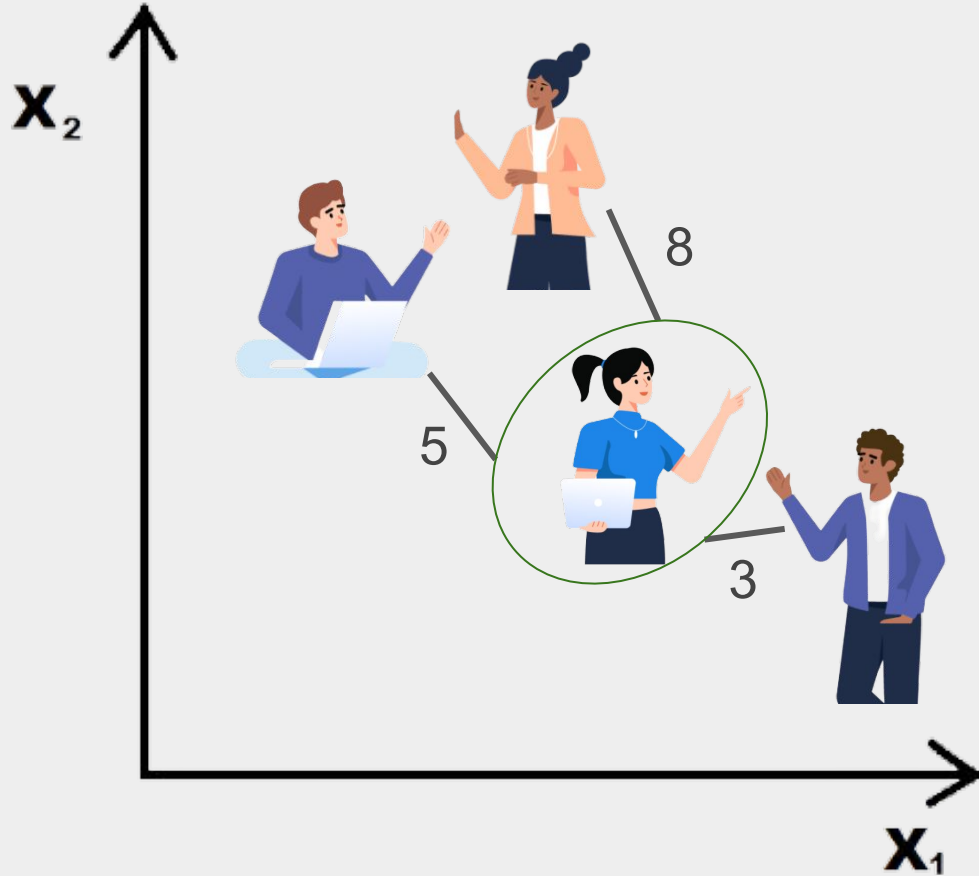






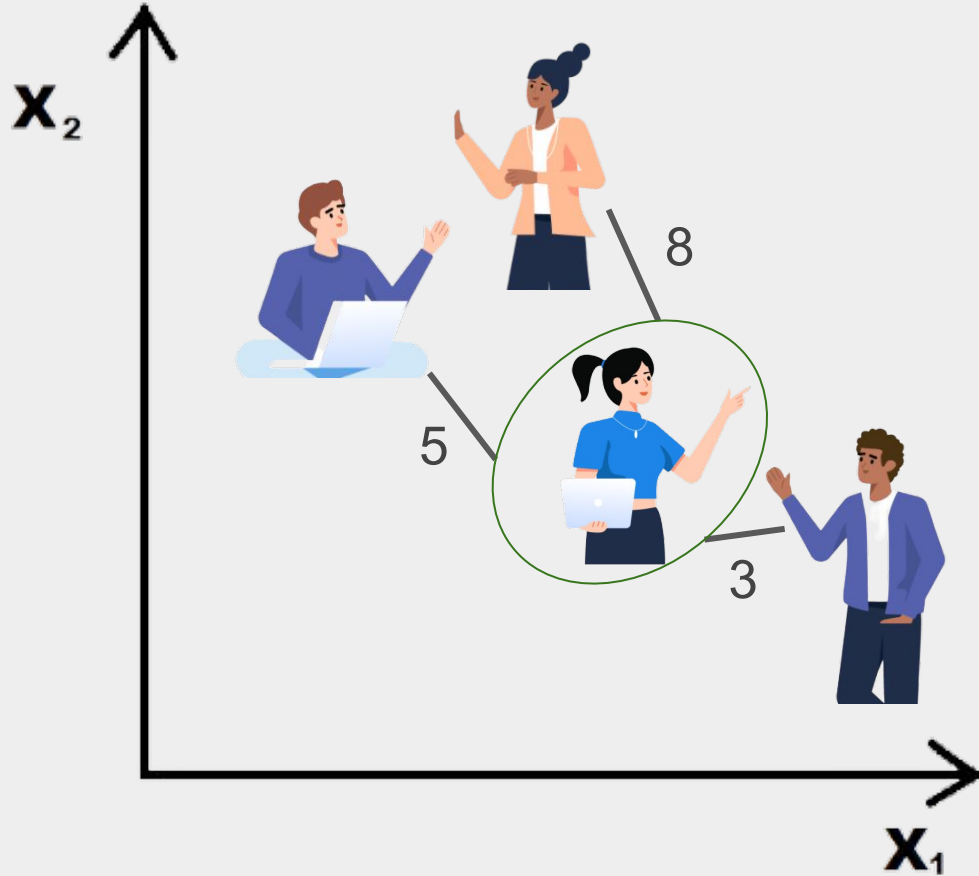


João - 3



João - 3

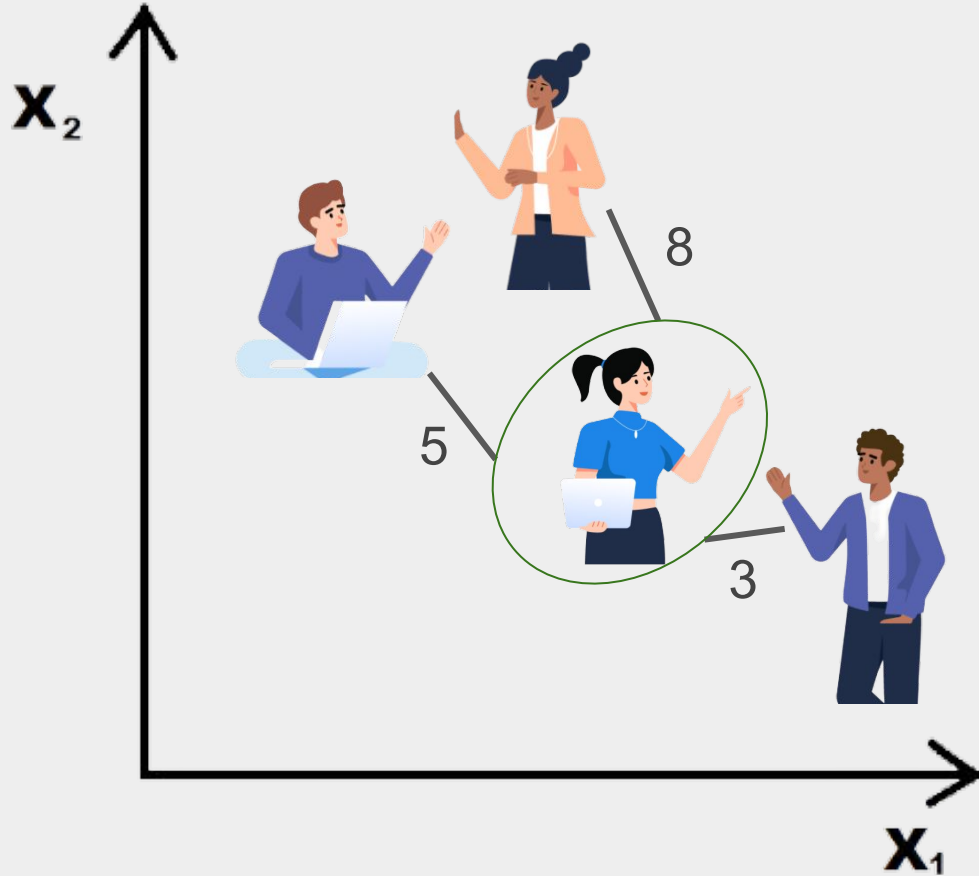
Pedro - 5



João - 3

Pedro - 5

Júlia - 8

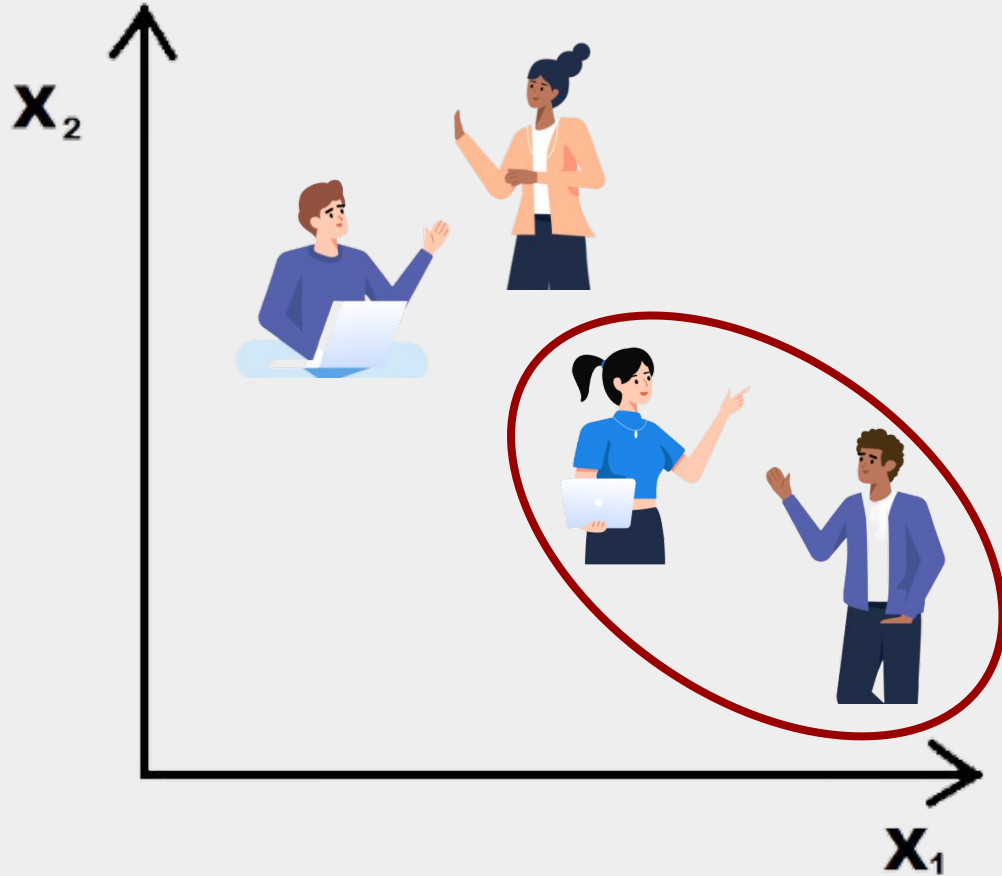


$$K = 1$$

João - 3

Pedro - 5

Júlia - 8

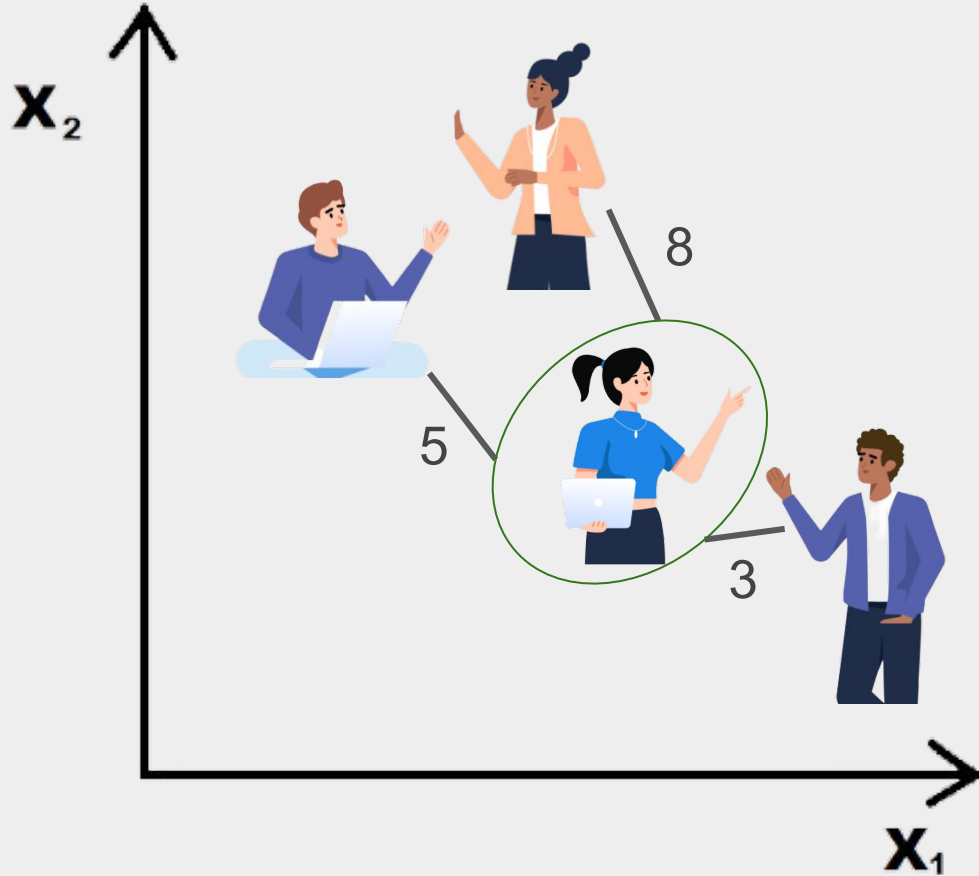


$$K = 1$$

João - 3

Pedro - 5

Júlia - 8

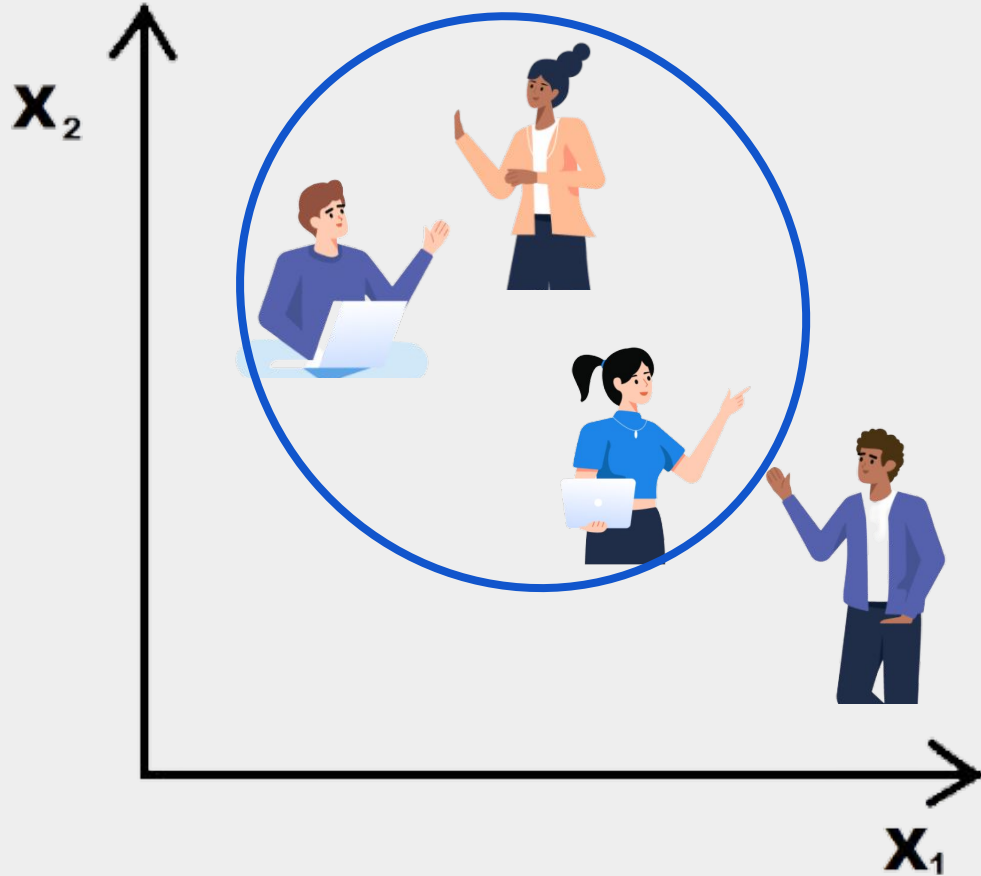


$$K = 3$$

João - 3

Pedro - 5

Júlia - 8



$$K = 3$$

João - 3

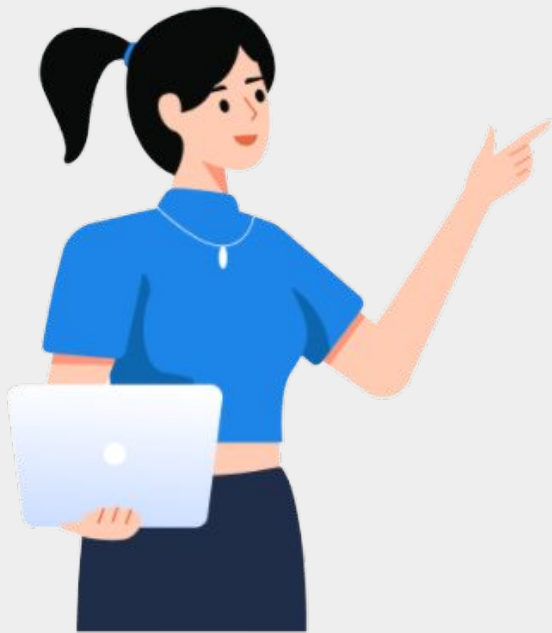
Pedro - 5

Júlia - 8

O que é Naive Bayes?

Churn





Churn



**Continua
Cliente (0)**



Churn



**Continua
Cliente (0)**



**Deixa de ser
Cliente (1)**



Churn



**Continua
Cliente (0)**



**Deixa de ser
Cliente (1)**

Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

...

FormaDePagamento:
DebitoEmConta

ContaMensal: 39.90

Churn = Não



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

...

FormaDePagamento:
DebitoEmConta

ContaMensal: 39.90

Churn = Não



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

...

FormaDePagamento:
DebitoEmConta

ContaMensal: 39.90

Churn = Não



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

...

FormaDePagamento:
DebitoEmConta

ContaMensal: 39.90

Churn = Não



$$P(y|X) = \frac{P(X|y) * P(y)}{P(X)}$$



Verossimilhança

$$P(y|X) = \frac{P(X|y) * P(y)}{P(X)}$$



Probabilidade a priori da classe

$$P(y|X) = \frac{P(X|y) * P(y)}{P(X)}$$



$$P(y|X) = \frac{P(X|y) * P(y)}{P(X)}$$

**Probabilidade a
priori da evidência**



Probabilidade a posteriori

$$P(y|X) = \frac{P(X|y) * P(y)}{P(X)}$$



$$P(y|X) = \frac{P(X|y) * P(y)}{P(X)}$$



Exemplo: Maria

$$P(X) = 0.015$$



Exemplo: Maria

$$P(X) = 0.015$$

$$P(y) = 0.14$$



Exemplo: Maria

$$P(X) = 0.015$$

$$P(y) = 0.14$$

$$P(X|y) = 0.006$$



Exemplo: Maria

$$P(X) = 0.015$$

$$P(y) = 0.14$$

$$P(X|y) = 0.006$$

$$P(y|X) = ?$$



Exemplo: Maria

$$P(y|X) = \frac{0.006 * 0.14}{0.015}$$



Exemplo: Maria

$$P(y|X) = 0.056 \text{ ou } 5,6\%$$



E o Bernoulli Naive Bayes?

$$P(y|X) = \frac{P(X|y) * P(y)}{P(X)}$$



Verossimilhança

$$P(y|X) = \frac{P(X|y) * P(y)}{P(X)}$$



$$P(x_i|y) = P(i|y)x_i + (1 - P(i|y))(1 - x_i)$$



Parâmetro p da Distribuição de Bernoulli

$$P(x_i|y) = P(i|y)x_i + (1 - P(i|y))(1 - x_i)$$

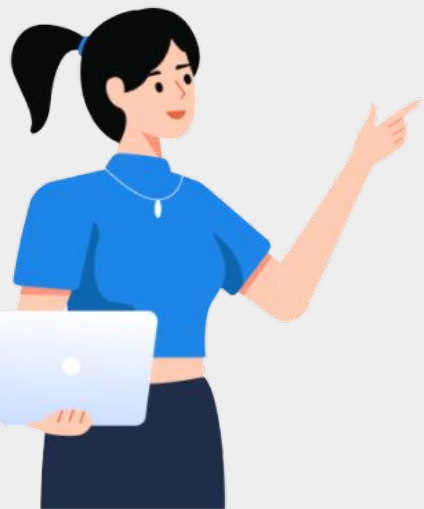


$$P(x_i|y) = P(i|y)x_i + (1 - P(i|y))(1 - x_i)$$



Verossimilhança

$$P(x_i|y) = P(i|y)x_i + (1 - P(i|y))(1 - x_i)$$



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

...

FormaDePagamento:
DebitoEmConta

ContaMensal: 39.90

Churn = Sim
ou
Churn = Não



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

...

**FormaDePagamento:
DebitoEmConta**

ContaMensal: 39.90

Churn = Sim
ou
Churn = Não



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

...

FormaDePagamento:
DebitoEmConta

ContaMensal: 39.90

Churn = Sim

ou

Churn = Não



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

...

FormaDePagamento:
DebitoEmConta

ContaMensal: 39.90

Churn = Sim

ou

Churn = Não



Maior65Anos: 0

Conjuge: Nao

Dependentes: Nao

...

FormaDePagamento:
DebitoEmConta

ContaMensal: 39.90

?

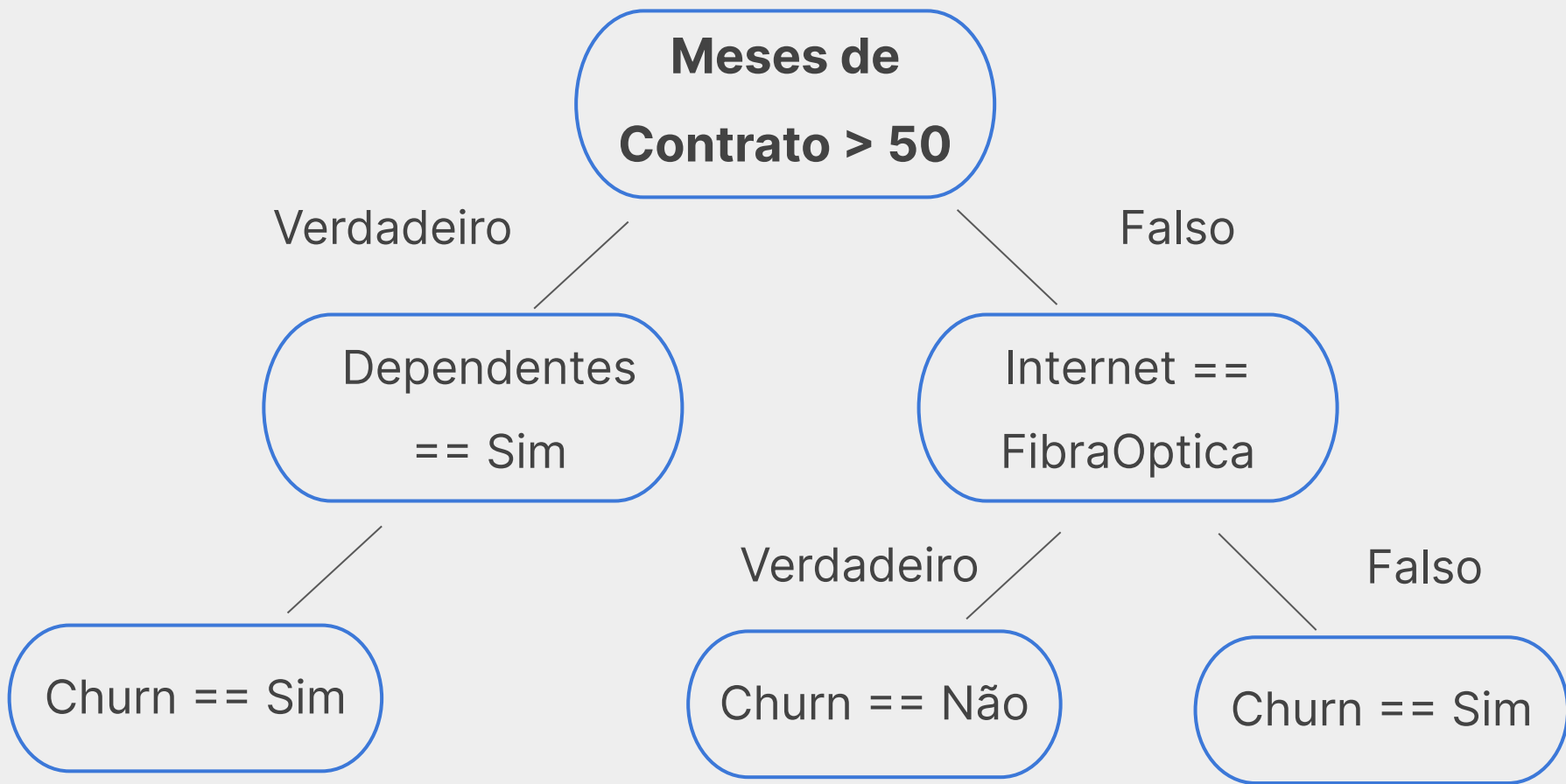
Churn = Sim

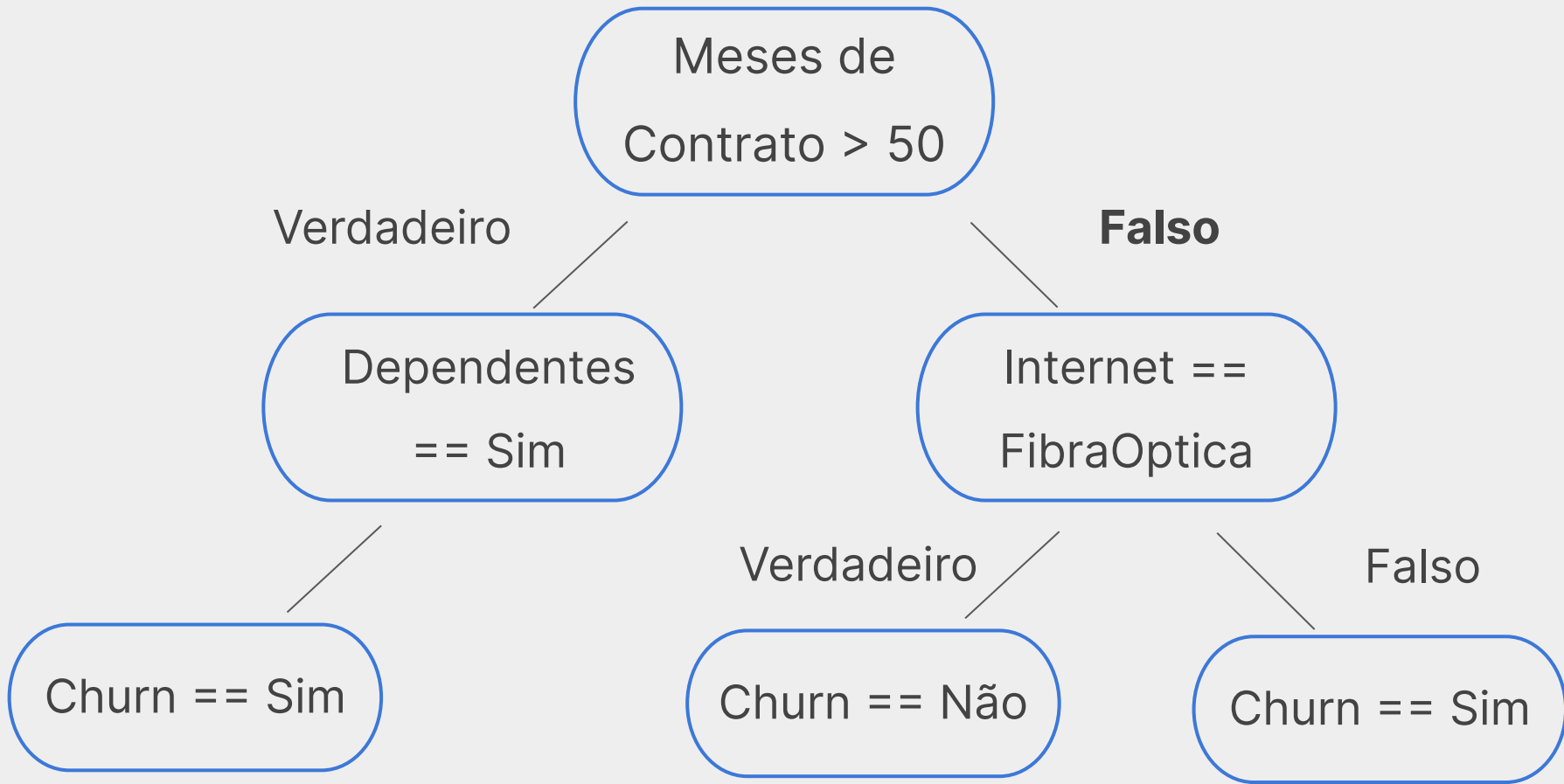
ou

Churn = Não



**O que são árvores de
decisão?**





Meses de
Contrato > 50

Verdadeiro

Falso

Dependentes
== Sim

**Internet ==
FibraOptica**

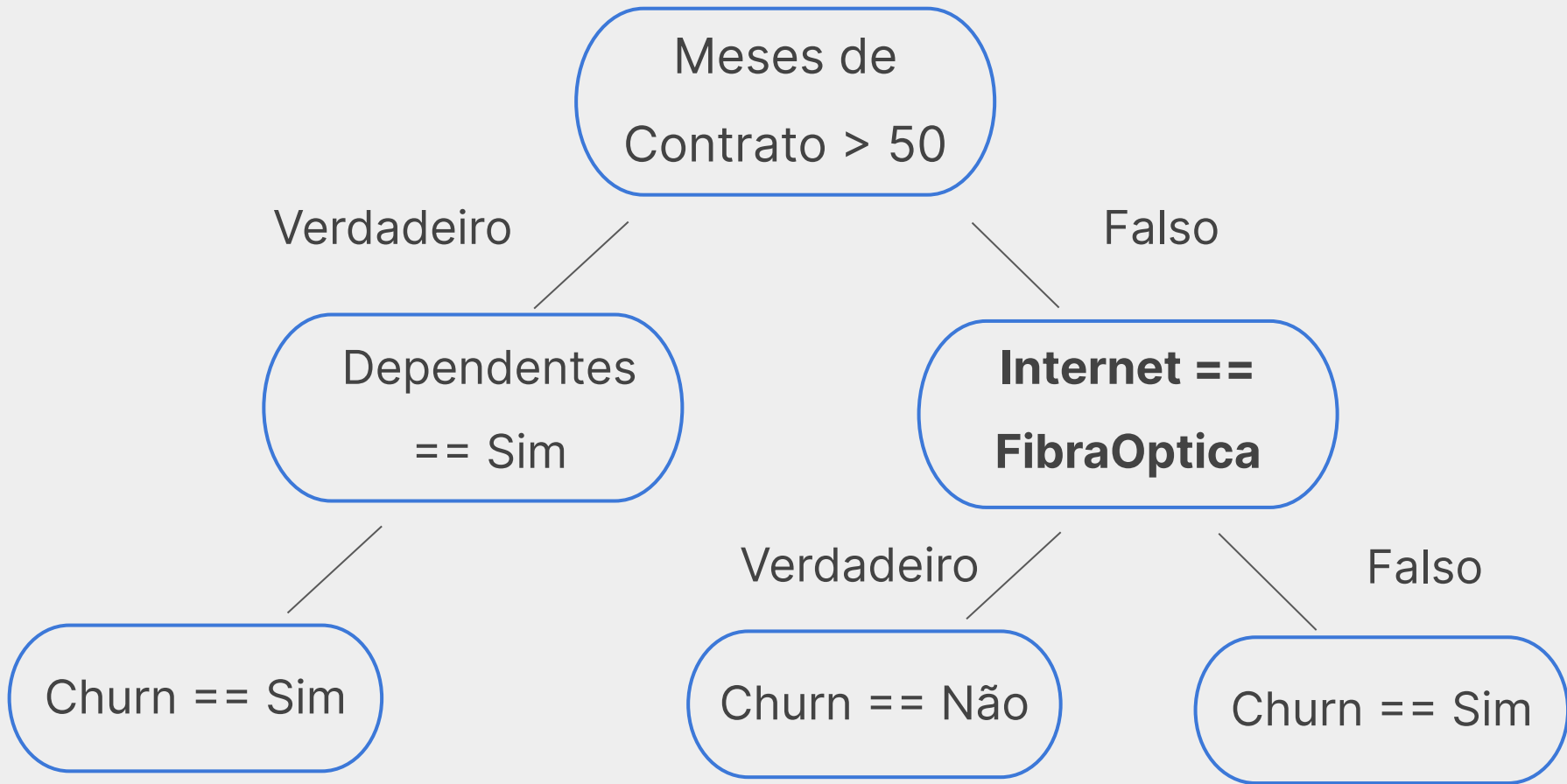
Churn == Sim

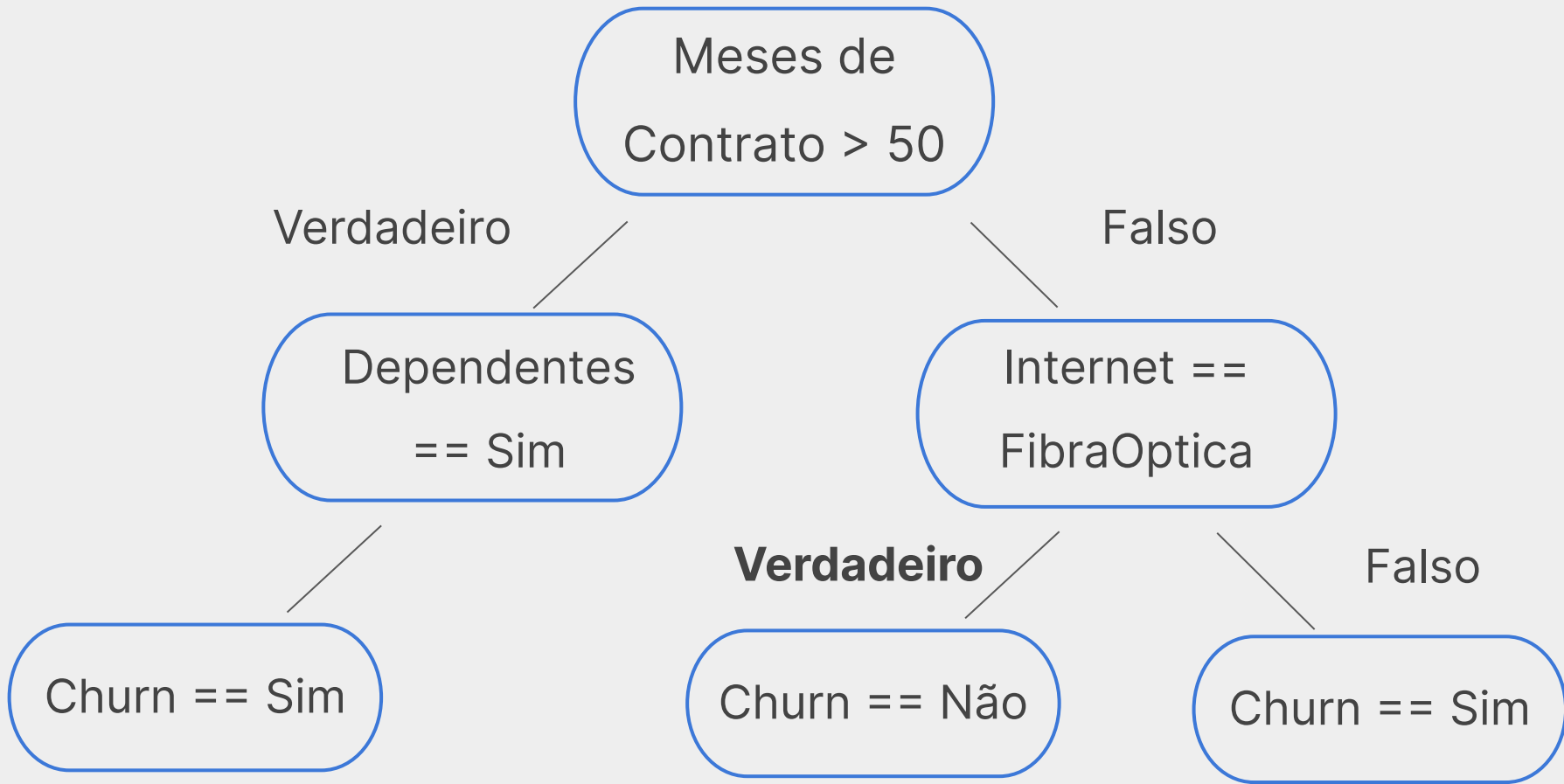
Verdadeiro

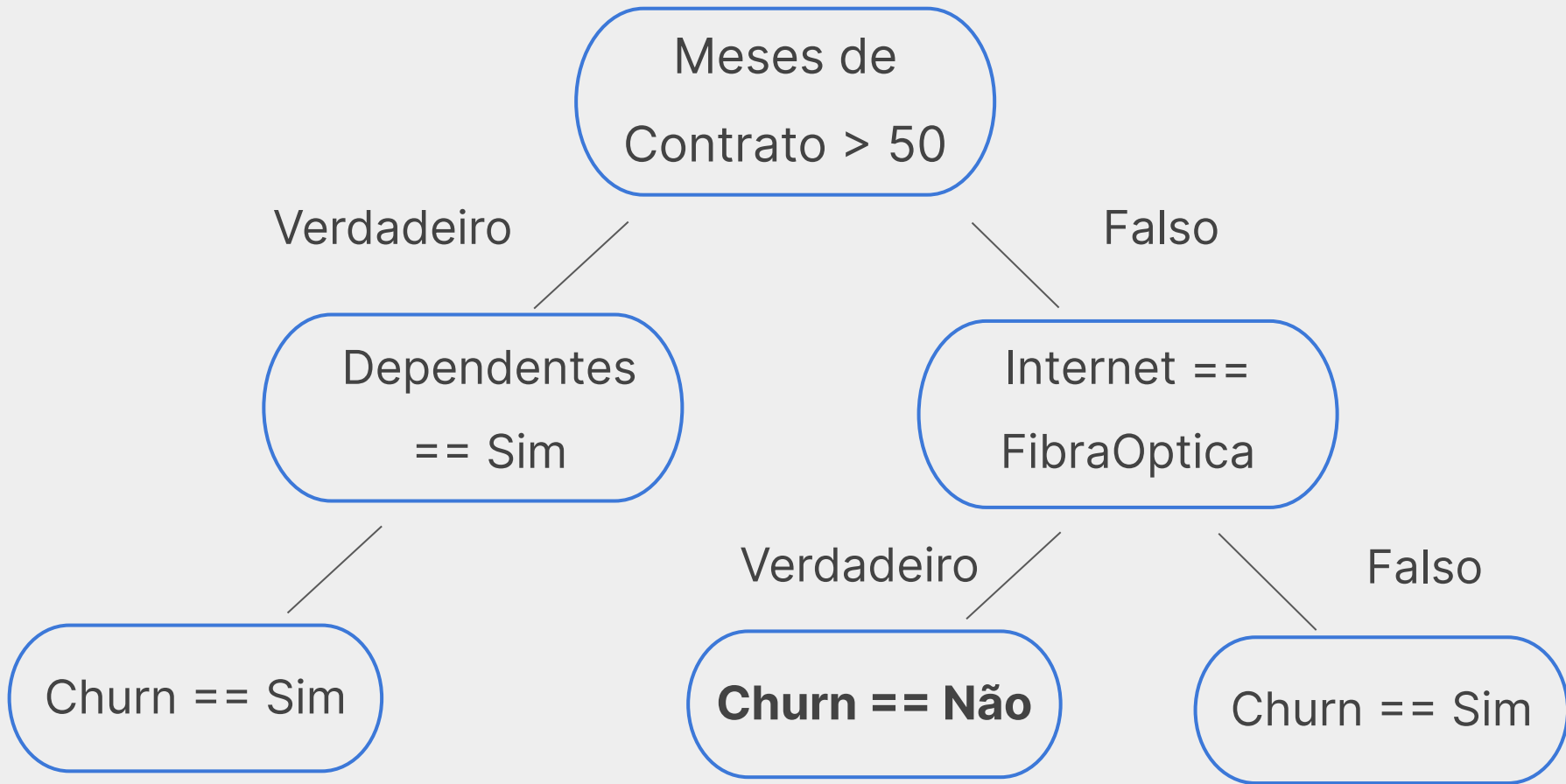
Falso

Churn == Não

Churn == Sim







Nó raiz

Verdadeiro

Falso

Nó de decisão

Nó de decisão

Verdadeiro

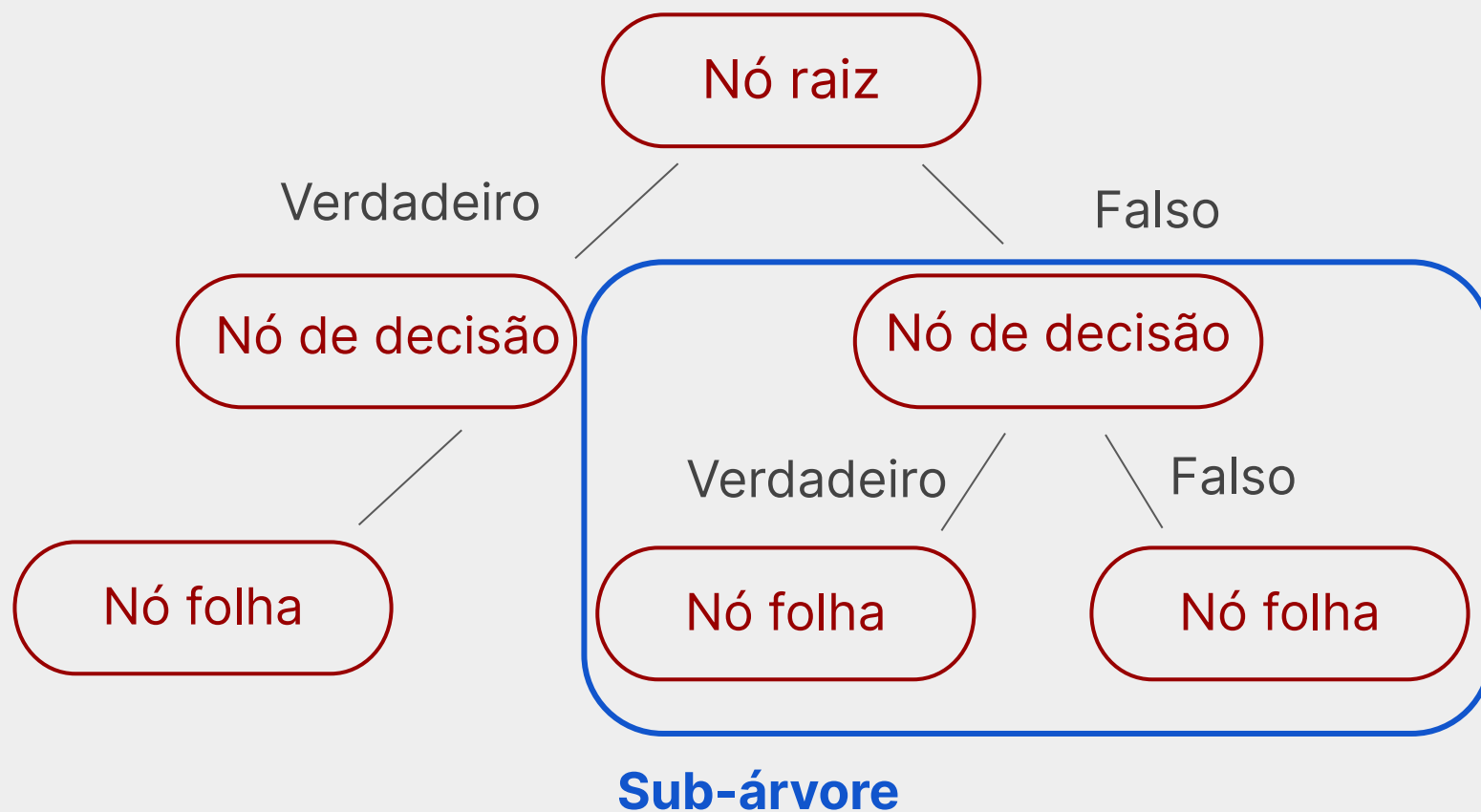
Falso

Nó folha

Nó folha

Nó folha





Por trás da Árvore de decisão

Critérios de divisão de nós

Critérios de divisão de nós

Índice Gini

Critérios de divisão de nós

Índice Gini

Qui-Quadrado

Ganho de informação (entropia)

Cr terios de divis o de n s

 ndice Gini

Qui-Quadrado

Ganho de informa  o (entropia)

Redu  o na vari ncia

Critérios de divisão de nós

Índice Gini

Qui-Quadrado

Ganho de informação (entropia)

Redução na variância

Índice Gini

$$G(\text{node}) = \sum_{k=1}^c p_k (1 - p_k)$$

Probabilidade de ocorrência de um dado da classe k

$$G(\text{node}) = \sum_{k=1}^c p_k (1 - p_k)$$

**Probabilidade de não
ocorrência de um
dato da classe k**

$$G(\text{node}) = \sum_{k=1}^c p_k (1 - p_k)$$

Entropia

$$\text{Entropy}(\text{node}) = - \sum_{i=1}^c p_k \log(p_k)$$

Entropia

$$\text{Entropy}(\text{node}) = - \sum_{i=1}^c p_k \log(p_k)$$

Créditos

Alura - Instrutora Ana Clara Mioto