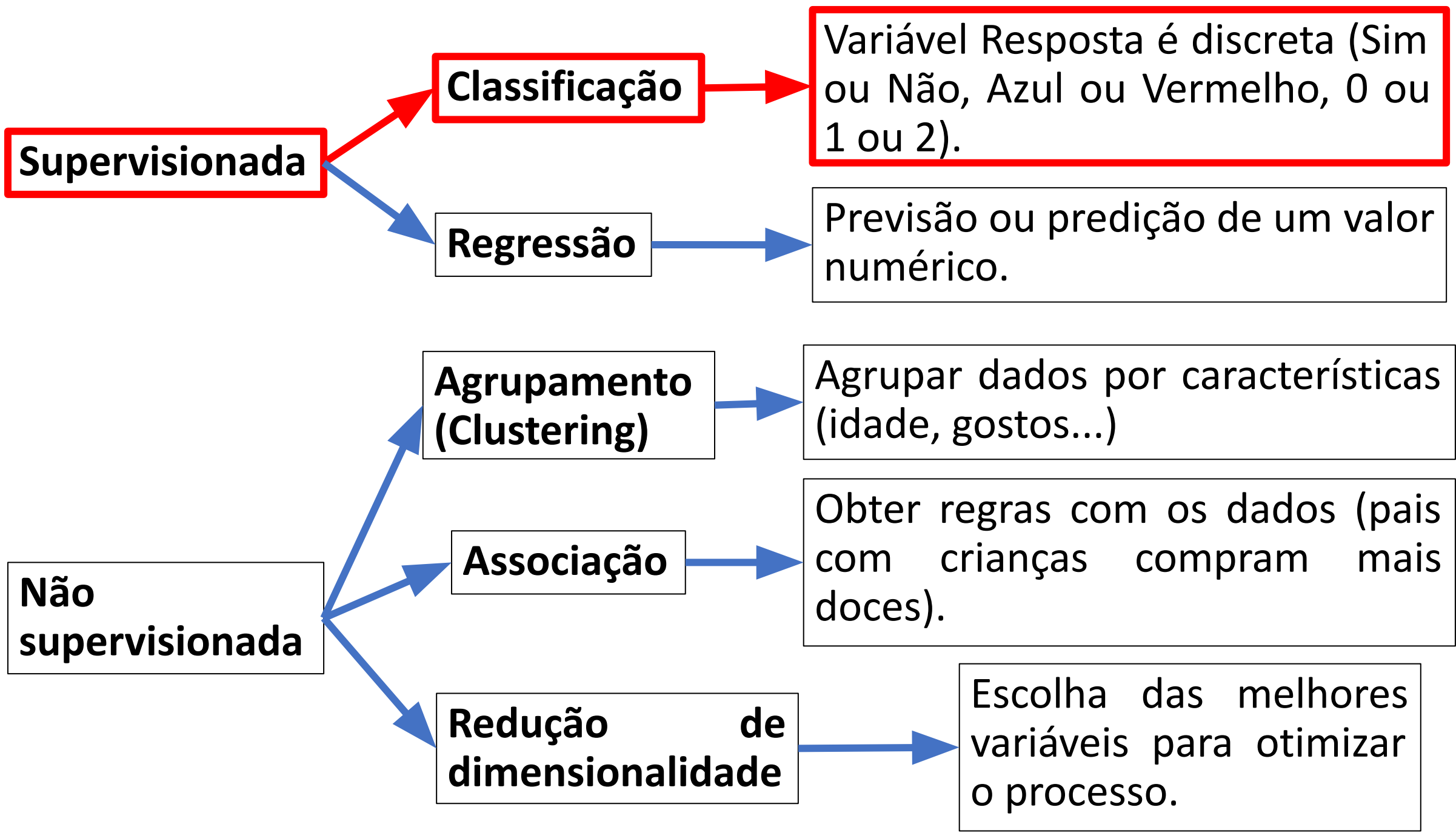


Machine Learning com Python

Prof. Luciano Galdino



Com base nos dados de entrada estima-se um “classificador” que gera como saída uma classificação qualitativa de um dado não observado (Ex.: análise de crédito, chances de desenvolver doenças).

Paciente	Pressão Alta	Colesterol Alto	Triglicérides Alto	Pratica Esporte	TEVE AVC?
1	SIM	SIM	SIM	NÃO	SIM
2	NÃO	SIM	NÃO	SIM	NÃO
3	NÃO	NÃO	SIM	NÃO	SIM
4	SIM	NÃO	NÃO	SIM	NÃO
5	SIM	SIM	NÃO	NÃO	NÃO

Paciente	Pressão Alta	Colesterol Alto	Triglicérides Alto	Pratica Esporte	PODERÁ TER AVC?
A	NÃO	SIM	SIM	SIM	?????

Análise do desempenho: Matriz de confusão

	NEGATIVO	POSITIVO
NEGATIVO	VERDADEIRO NEGATIVO	FALSO POSITIVO
POSITIVO	FALSO NEGATIVO	VERDADEIRO POSITIVO

Análise do desempenho: Matriz de confusão

Gênero	Estado Civil	Grau de instrução	Produto
Masculino	Casado	Básico	A
Feminino	Casado	Superior	B
Feminino	Viúva	Básico	A
Masculino	Solteiro	Superior	B
Feminino	Solteira	Superior	B

Previsores
(atributos)

Classe
(target)

Gênero	Estado Civil	Grau de instrução	Produto
0	0	0	0
1	0	1	1
1	2	0	0
0	1	1	1
1	1	1	1

Análise do desempenho: Matriz de confusão

Gênero	Estado Civil	Grau de instrução	Produto
Masculino	Casado	Básico	A
Feminino	Casado	Superior	B
Feminino	Viúva	Básico	A
Masculino	Solteiro	Superior	B
Feminino	Solteira	Superior	B

Previsores
(atributos)

Classe
(target)

Gênero	Estado Civil	Grau de instrução	Produto	Saída
0	0	0	0	0
1	0	1	1	0
1	2	0	0	1
0	1	1	1	1
1	1	1	1	0

Matriz de Confusão

	Referência		
Saída		0	1
	0	1	2
	1	1	1

40% acerto

Análise do desempenho: Matriz de confusão

Gênero	Estado Civil	Grau de instrução	Produto
Masculino	Casado	Básico	A
Feminino	Casado	Superior	B
Feminino	Viúva	Básico	A
Masculino	Solteiro	Superior	B
Feminino	Solteira	Superior	B

Previsores
(atributos)

Classe
(target)

Gênero	Estado Civil	Grau de instrução	Produto	Saída
0	0	0	0	0
1	0	1	1	0
1	2	0	0	0
0	1	1	1	1
1	1	1	1	0

Matriz de Confusão

	Referência		
Saída		0	1
	0	2	2
	1	0	1

60% acerto

Análise do desempenho: Matriz de confusão

Gênero	Estado Civil	Grau de instrução	Produto
Masculino	Casado	Básico	A
Feminino	Casado	Superior	B
Feminino	Viúva	Básico	A
Masculino	Solteiro	Superior	B
Feminino	Solteira	Superior	B

Previsores
(atributos)

Classe
(target)

Gênero	Estado Civil	Grau de instrução	Produto	Saída
0	0	0	0	0
1	0	1	1	0
1	2	0	0	0
0	1	1	1	1
1	1	1	0	0

Matriz de Confusão

	Referência		
Saída		0	1
	0	3	1
	1	0	1

80% acerto

Análise do desempenho: Matriz de confusão

$$Accuracy = \frac{VN + VP}{VN + VP + FN + FP}$$

$$Precision = \frac{VP}{VP + FP}$$

$$Recall = \frac{VP}{VP + FN}$$

$$F1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

	NEGATIVO	POSITIVO
NEGATIVO	VERDADEIRO NEGATIVO	FALSO POSITIVO
POSITIVO	FALSO NEGATIVO	VERDADEIRO POSITIVO

Algoritmos de Classificação

❖ **Regressão Logística**

❖ **Máquinas de Vetor de Suporte**

❖ **Naive Bayes**

❖ **XGBoost**

❖ **Árvore de decisão**

❖ **LightGBM**

❖ **Random Forest**

❖ **CatBoost**

❖ **KNN**

❖ **Redes Neurais Artificiais**