



FIAP

ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

DISRUPTIVE ARCHITECTURES: IOT, IOB & IA

05 – Modelos da Aprendizado Supervisionado – Regressão

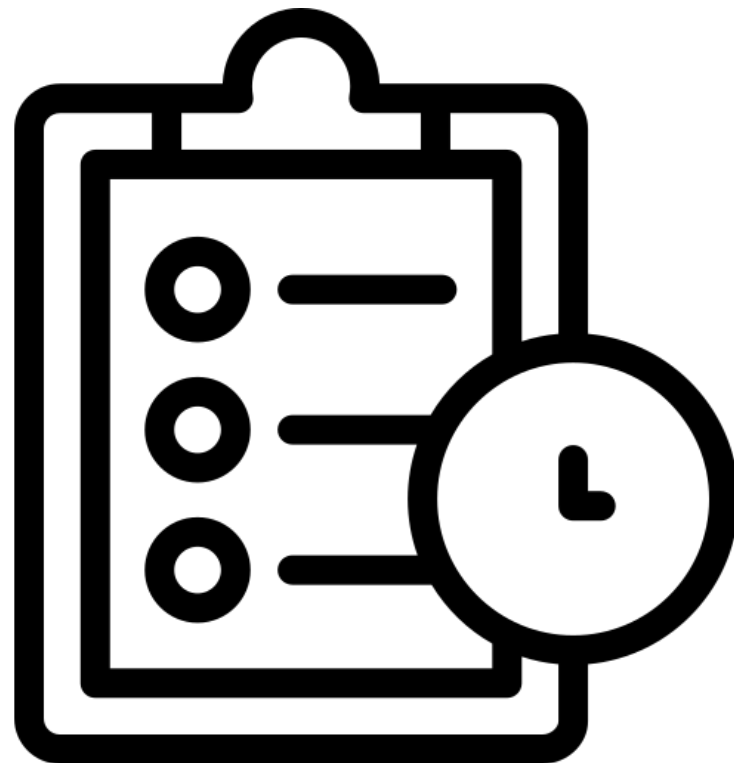


Prof. Airton Y. C. Toyofuku



profairton.toyofuku@fiap.com.br

- Revisão
- O que é Aprendizado Supervisionado
- Tipos de Problemas de Aprendizado Supervisionado
- Algoritmos de Regressão
- Matriz de correlação
- Métricas de Avaliação para Problemas de Regressão



❖ O que é Machine Learning?

- Técnica da Inteligência Artificial que permite que o computador aprenda uma tarefa específica.

❖ Quais os tipos de aprendizado?

- Supervisionado;
- Não Supervisionado;
- Por reforço;

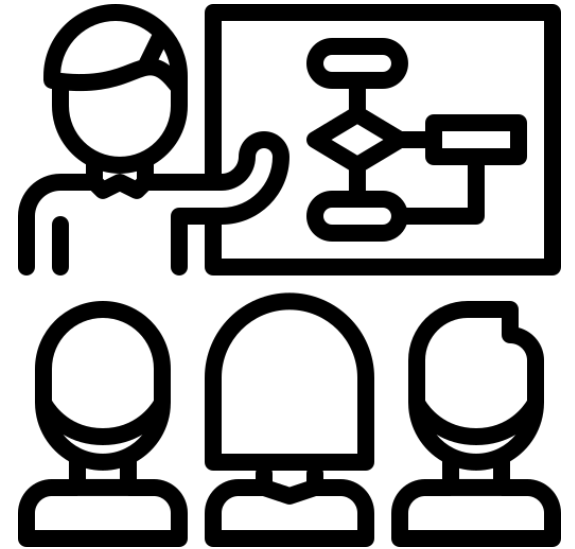
❖ O que precisamos para implementar uma ML?

- Dados!!!!

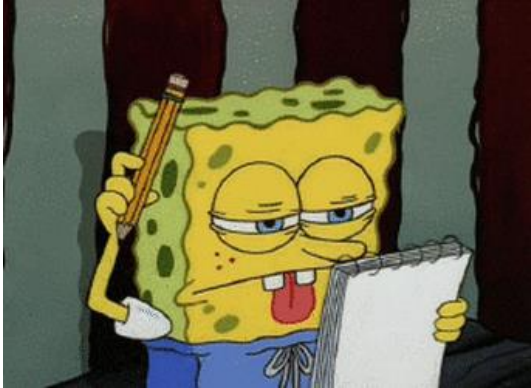


O que é Aprendizado Supervisionado

- ❖ Consiste em treinar um modelo com um grupo de dados já rotulados;
- ❖ Ou seja, nós já sabemos qual é a resposta ou correlação entre os dados;
- ❖ O Objetivo é fazer o modelo entender essa correlação e generalizar para novos dados que ele não conhece;



Tipos de Problemas de Aprendizado Supervisionado



❖ Existem dois tipos de problemas em que Podemos aplicar Aprendizado Supervisionado:

❖ De Classificação:

- Sim ou Não;
- Vai comprar ou Não vai comprar;
- É do tipo A, do tipo B ou do tipo C;

❖ De Regressão:

- Prever um determinado número;
- Quanto custa uma casa;
- Qual vai ser a quilometragem;

❖ Regressão Linear ($Y = A + BX$)

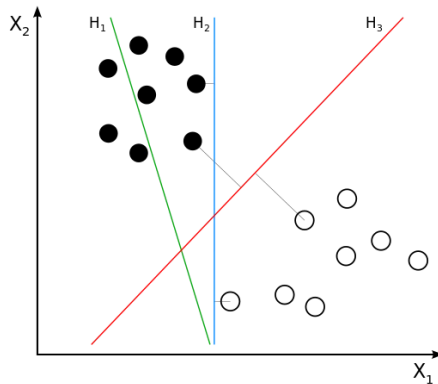
- modelo de Regressão simples e popular que se ajusta a um conjunto de dados com uma linha reta. Ele é útil quando há uma relação linear entre as variáveis dependentes e independentes.

❖ Regressão Polinomial ($Y = A_0 + A_1X + A_2X^2 + \dots + A_nX^n$)

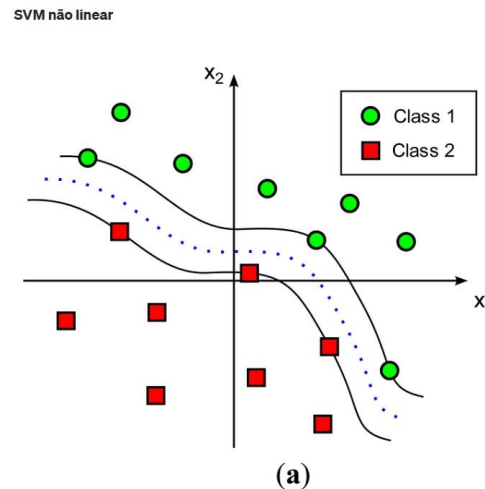
- modelo de Regressão que se ajusta a um conjunto de dados com uma curva polinomial. Ele é útil quando há uma relação não linear entre as variáveis dependentes e independentes.

❖ Support Vector Machine

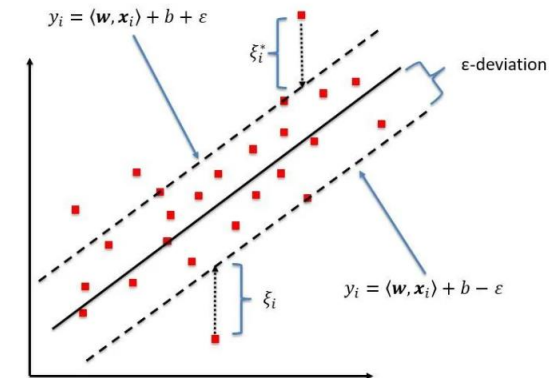
- Pode ser usado tanto para classificação, quanto para regressão. Ele pega os dados e traça uma divisão entre eles, considerando a inclinação que dá a maior distância entre os grupos, mas ao mesmo tempo, a menor distância entre os pontos da reta



A reta ótima (H3) é a mais distante dos dois grupos, considerando apenas os pontos de cada grupo mais próximos à reta (como indicado pelas linhas cinzas. Fonte: Wikimedia Commons.



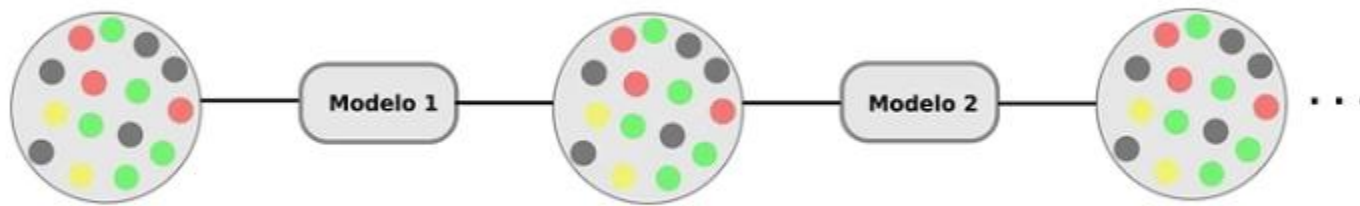
SVM não linear. Fonte: Ruben Ruiz-Gonzalez et al. (adaptado)



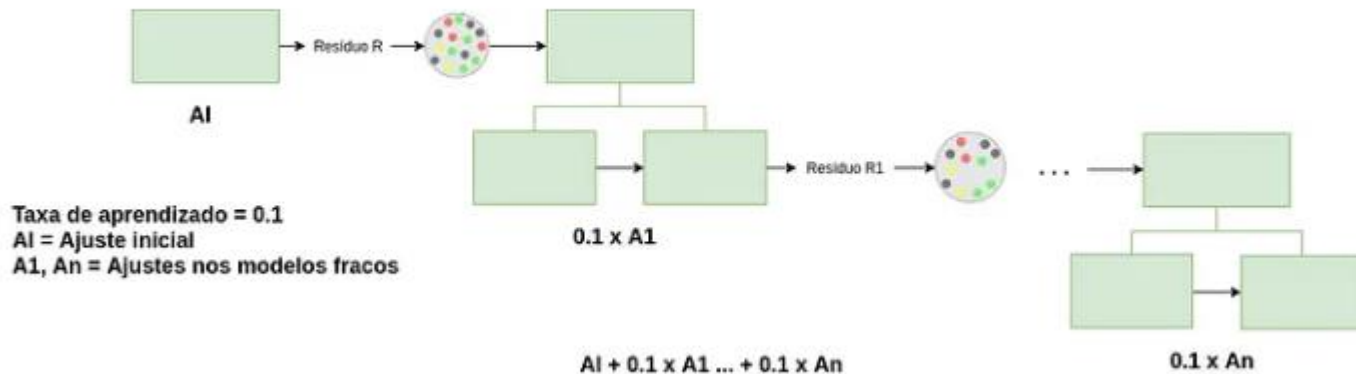
Regressão por SVM. Fonte: Tania Kleynhans et al.

❖ Gradient Boosting

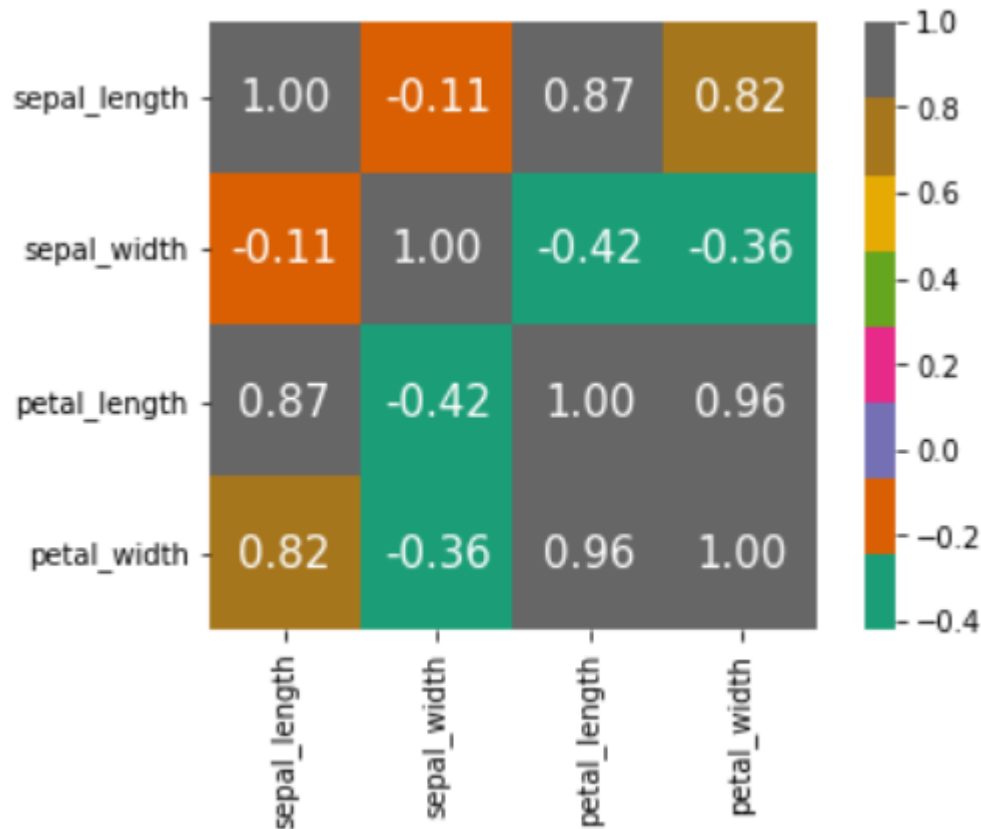
- Algoritmo que emprega a técnica de boosting, visando tentar minimizar o erro residual proveniente do modelo fraco anterior.



Exemplo de funcionamento do Algoritmo Boosting



Matriz de correlação



- a matriz de correlação é uma ferramenta utilizada na análise exploratória de dados, pois ajuda a identificar as relações mais fortes e mais fracas entre as variáveis.
- Pode ser usada para selecionar as variáveis mais importantes para a modelagem e pode ajudar a evitar problemas de multicolinearidade

Métricas de Avaliação para Problemas de Regressão

❖ MSE

- Erro Quadrático Médio : É a média do erro das previsões ao quadrado:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

❖ RMSE

- Raiz do Erro Quadrático Médio: É a raiz quadrada do MSE

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$

❖ R²

- R-Quadrado: É a porcentagem da variância que pode ser prevista pelo modelo.

$$R^2 = 1 - \frac{\text{Variança Residual}}{\text{Variança Total}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Métricas de Avaliação para Problemas de Regressão

❖ R^2 Score

- É a porcentagem da variância que pode ser prevista pelo modelo, mas sem viés

$$R_a^2 = 1 - \frac{(1 - R^2)(N - 1)}{N - p - 1}$$

❖ MAE

- Erro Absoluto Médio: É a media das distâncias dos valores reais e do preditos.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

Copyright © 2023 Prof. Airton Y. C. Toyofuku

Todos direitos reservados. Reprodução ou divulgação total ou parcial deste documento é expressamente proibido sem o consentimento formal, por escrito, do Professor (autor).