

Métodos para problemas reais

Carlos Diego Rodrigues

21 de outubro de 2021

Universidade Federal do Ceará

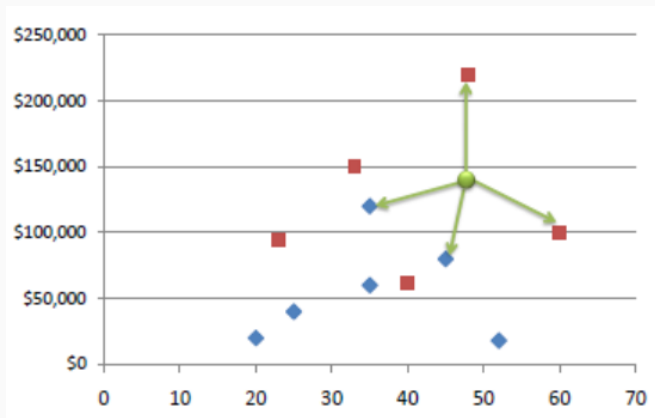
Métodos para problemas reais: ferramentas avançadas.

- Quais ferramentas avançadas são utilizadas atualmente para problemas de classificação?
- Métodos estatísticos e de otimização:
 - K Vizinhos mais próximos (KNN)
 - Regressão Logística
 - *Support Vector Machine* (SVM)
 - Redes neurais artificiais (ANN)

K Vizinhos mais próximos (KNN)

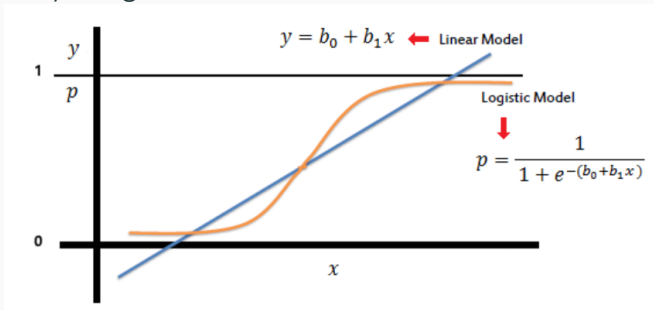
- Método simples que utiliza todos os dados de treino.
- Classificador "lazy".
- Baseado em alguma medida de distância:
 - Linear: $\sum_{i=1}^k |x_i - y_i|$
 - Euclidiana: $\sqrt{\sum_{i=1}^k (x_i - y_i)^2}$
 - Minkowski: $\left(\sum_{i=1}^k (|x_i - y_i|)^q \right)^{\left(\frac{1}{q}\right)}$
- Em geral utiliza-se atributos padronizados

Exemplo: KNN



Regressão Logística

- Utiliza uma ideia similar à regressão.
- Função logística:



- Método da Máxima Verossimilhança para encontrar os b .
- Repetido iterativamente até que não haja mais uma mudança significativa no modelo.

Desenvolvendo o modelo logístico

- Podemos escrever o modelo logístico da seguinte forma:

$$p = \frac{1}{1 + e^{-(b_0 + b_1 x)}}$$

O que pode ser escrito como:

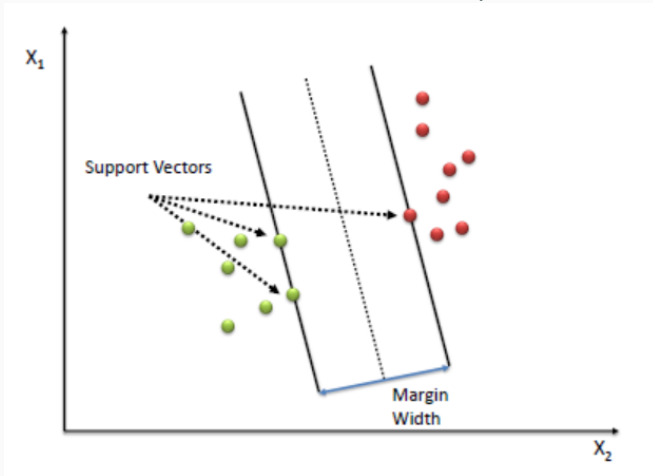
$$\frac{p}{1 - p} = e^{-(b_0 + b_1 x)}$$

- Pode ser aplicado a um número qualquer de dimensões:

$$p = \frac{1}{e^{-(b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n)}}$$

Support Vector Machine (SVM)

- Uma máquina de vetor de suporte realiza uma classificação baseada numa margem definida entre duas classes.
- Os casos extremos são os vetores de suporte.



Encontrando o melhor SVM

- Nós encontramos os valores que definem o melhor vetor resolvendo o problema:

$$\begin{aligned} \min & \frac{1}{2} ||w||^2 \\ \text{s.a.} & : y_i(w^T x_i + b) \geq 1, \quad \forall x_i \end{aligned}$$

- Penalizações para casos não separáveis
- *Kernel trick*

Redes neurais artificiais (ANN)

- Uma rede neural artificial é um sistema inspirado nas redes neurais biológicas.
- A conexão de informações repassadas a partir de sensores é trabalhada em camadas através de impulsos que vão dos sensores (entrada) ao sistema central (saída).
- Matematicamente se configura como uma rede (ou grafo) onde cada neurônio é um nó do grafo que está conectado e passando ou recebendo informação dos outros nós.
- Podemos dividi-los em três categorias: entrada, internos e saída.

Exemplo: ANN

