



IC 598 - Algoritmos e Estrutura de Dados I

Prof: Marcelo Dib Cruz

Algoritmos e Estrutura de Dados I

Algoritmos?

- Definição
- Para que serve ?

Algoritmos ?

- **Solucionar Problemas** exige a capacidade de criar representações adequadas da realidade (modelos) e, com ajuda delas, encontrar um algoritmo de solução que explique como remover ou superar tais problemas. O algoritmo precisa funcionar na prática e ser rápido o suficiente para solucionar o problema quando este oferece uma solução.

Algoritmos ?

→ Solucionar Problemas

- O que é Modelagem?
- É o processo de produzir um modelo.
- É uma atividade na qual construímos modelos para descrever como um fenômeno ou sistema de interesse se comporta.

Algoritmos ?

→ Solucionar Problemas

- O que é Modelagem?
- É uma representação de algum objeto, fenômeno observado, ou sistema;
- É uma descrição em alguma “linguagem” usada para ajudar a entender ou visualizar alguma coisa.

Modelagem

- Características
 - Um modelo é similar, porém mais simples do que o sistema real que ele representa;
 - Um modelo deve ser uma representação próxima ao sistema real (que ele representa) e incorporar suas características principais

Modelagem

Modelagem

- Por que modelar?
- Para ganhar entendimento: O entendimento de um sistema do mundo real pode ser melhorado através da análise do modelo.
- Para prever (resolver situações): Em muitos casos, experimentar diretamente um sistema real para saber como ele se comportará no futuro pode ser muito caro ou mesmo impossível.

Modelagem

Modelagem

- Por que modelar?
- A construção de um modelo requer:
- Um entendimento claro dos seus objetivos com a modelagem. Exatamente, quais aspectos do sistema você deseja entender ou predizer?
- Um entendimento dos fatores chaves envolvidos no sistema e como eles se relacionam entre si. Isto geralmente requer uma visão muito simplificada do sistema.

Modelagem

Modelagem

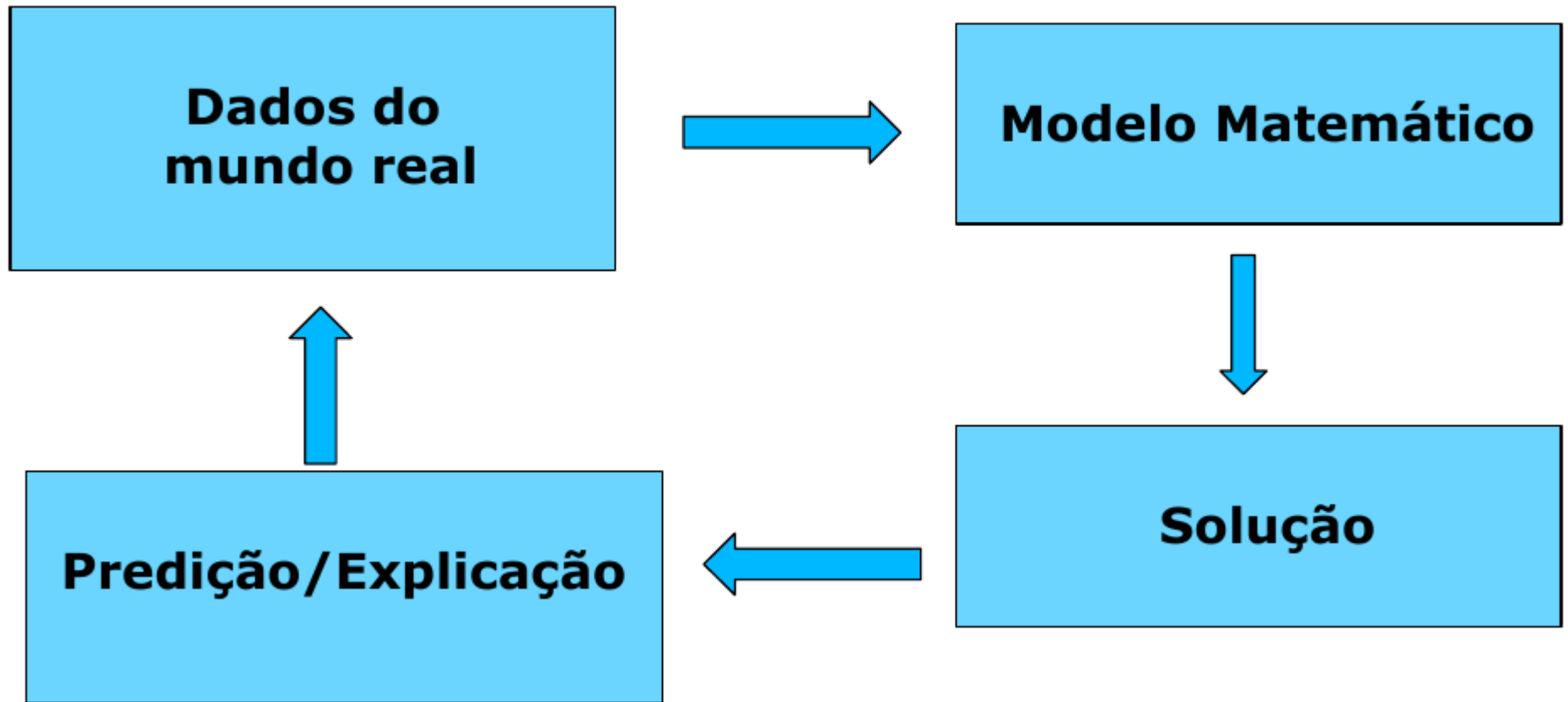
- Por que modelar?
- **O mesmo sistema pode ser modelado de formas diferentes dependendo do objetivo proposto;**

Modelagem

O Processo de Modelagem

- Para exercer as habilidades para formular ou modelar matematicamente um problema é importante esquematizar alguma metodologia de modelagem.
- Um esquema consensual consiste nas seguintes etapas:
 1. Identificação (reconhecimento);
 2. Modelo (formulação);
 3. Solução

Modelagem



Modelagem

As etapas do esquema básico podem ser detalhadas:

1. Identificar o problema real. Identificar as variáveis do problema;
2. Construir relações apropriadas entre as variáveis;
3. Criar um modelo matemático adequado;
4. Obter a solução;
5. Interpretar a solução matemática;
6. Testar e validar o modelo. Comparar com a realidade.

Modelagem

Em cada uma das etapas algumas perguntas podem ser feitas para auxiliar no processo:

1. Identificar o problema real. Identificar as variáveis do problema;
 1. O que queremos obter?
 2. Quais os objetivos em abordar o problema?
 3. O que pode ser medido ou obtido como dado no problema?
2. Construir relações apropriadas entre as variáveis;
 1. Qual é a relação entre os dados conhecidos e desconhecidos?

Modelagem

3. Criar um modelo matemático adequado;
 1. A técnica escolhida para representar o modelo é adequada?
 2. O modelo possui solução?
4. Obter a solução;
 1. O modelo possui solução analítica?
 2. Qual será o esforço e quais os métodos de solução serão adequados?
 3. Pode-se melhorar o método de solução?
 4. Será melhor desenvolver um método novo?

Modelagem

Exemplos:

- Mercado de ações
- Produção de Petróleo
- Biologia
- Ecologia
- Meteorologia
- Engenharias
- Química
- Medicina
- Etc.

Exemplo

Reconhecimento de Padrões

- **Objetivo : encontrar algum padrão no conjunto de dados**
- **Solução**
 - **Inteligencia Artificial**
 - **Otimização Combinatória**
- **Para que serve ?**
 - **Classificação**
 - **Predição**
 - **Mineração**

Exemplo

Reconhecimento de Padrões

- **Objetivo** : Com rótulos ou classes
- Aprendizado Supervisionado
- Risco de Pacientes com Cancer
 - Alto
 - Medio
 - Baixo

Exemplo

Reconhecimento de Padrões

- **Objetivo** : Sem rótulos ou classes
- Aprendizado Não supervisionado
- RNA não supervisionada
- Aplicação para Encontros

Exemplo

Reconhecimento de Padrões

- **Objetivo** : Sem rótulos ou classes
- Problema de Agrupamento ou Clusterização ou Clustering
- Uma aplicação interessante é quando se deseja encontrar estruturas similares em um conjunto de dados onde não possui classes.
- Agrupamento de dados é uma técnica base para identificar valores anômalos ou outliers.
- K conhecido (numero de grupos)
- K-means

Exemplo

Reconhecimento de Padrões

- **K-means** (MacQuenn, 1967) é um método de clusterização que objetiva particionar um conjunto com n observações em um número k de grupos ou clusters, fixado a priori. O algoritmo visa minimizar uma função objetivo, neste caso uma função do erro quadrado.

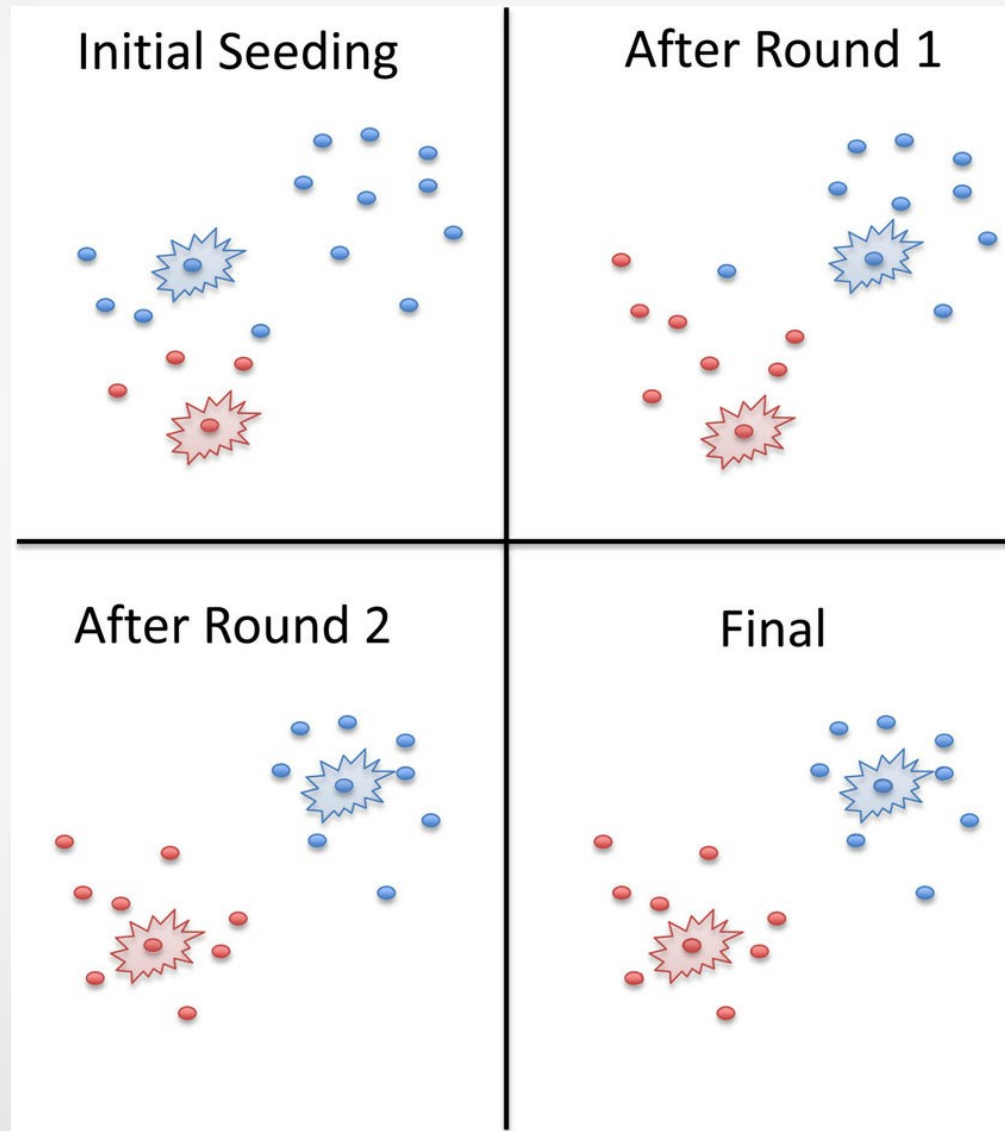
Exemplo

K-Means

- 1) Escolher aleatoriamente um número k de centros para os clusters;
- 2) Atribuir cada objeto para o cluster de centro mais próximo (ex. usando a distância Euclidiana)
- 3) Mover cada centro para a média dos objetos atribuídos;
- 4) Repetir os passos 2 e 3 até que algum critério de convergência seja obtido (número de iterações, tolerância em relação às mudanças nos centróides).

Exemplo

K-Means



Exemplo

Reconhecimento de Padrões

- **Objetivo** : Sem rótulos ou classes
- Problema de Agrupamento Automático ou Clusterização Automática ou Automatic Clustering
- Numero de grupos não conhecido

Exemplo

Problema de Agrupamento Automático ou Automatic Clustering

K não conhecido (numero de grupos)

```
from sklearn.cluster import DBSCAN
```

```
from sklearn.cluster import AgglomerativeClustering
```

```
from sklearn.cluster import OPTICS
```

```
from sklearn.cluster import MiniBatchKMeans
```

```
from sklearn.cluster import AffinityPropagation
```

```
from sklearn.cluster import MeanShift
```

```
from sklearn.cluster import SpectralClustering
```

```
from sklearn.cluster import Birch
```


Exemplo

Problema de Agrupamento Automático ou Automatic Clustering

K não conhecido (numero de grupos)

Algoritmos Genéticos

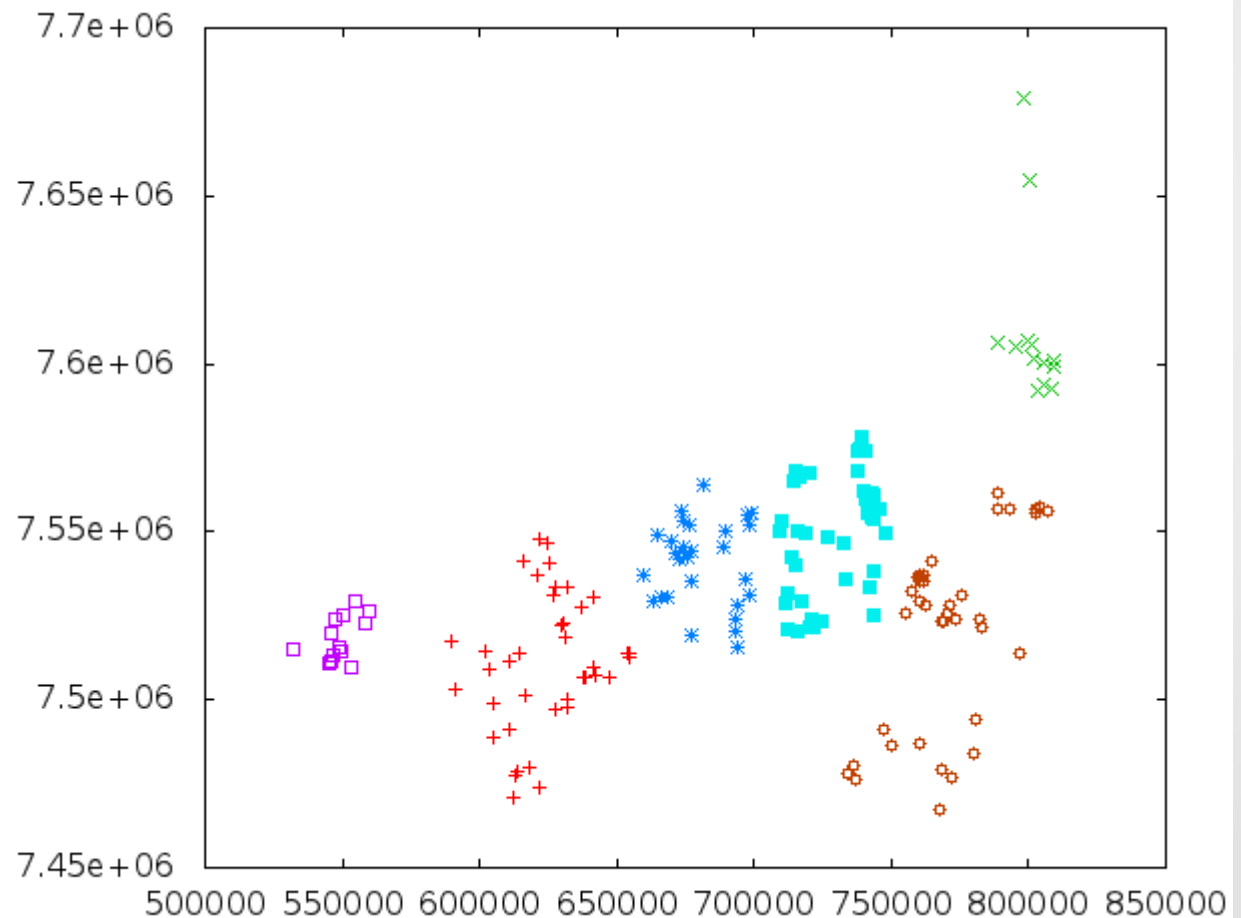
Exemplo

Problema de Agrupamento Automático

K não conhecido

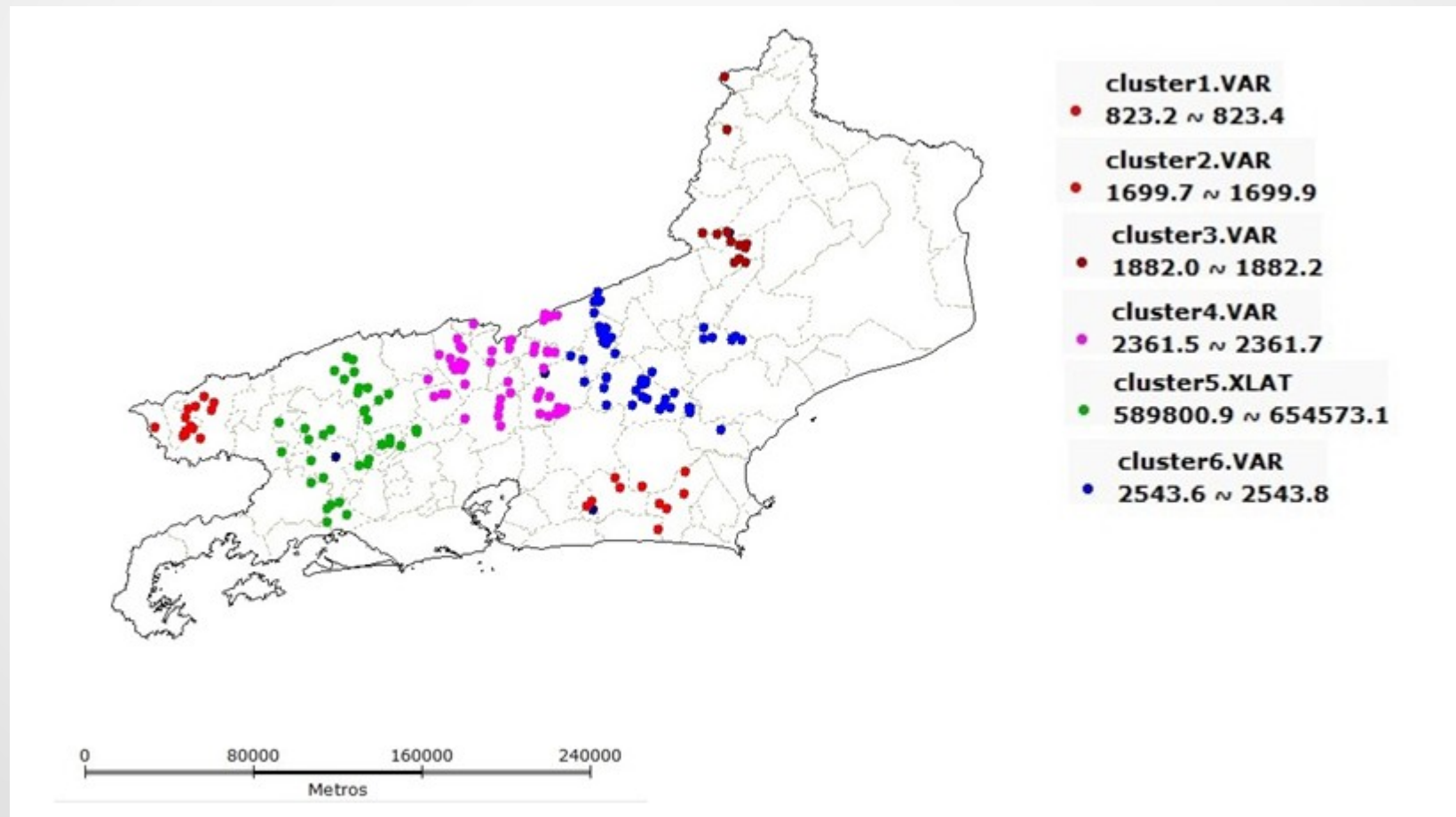
AG

Abelhas



Exemplo

Problema de Agrupamento Automático



Exemplo

Problema de Agrupamento Automático

Aglomerado	Produção anual mel/colmeia (kg)	Prevalência de apiários doentes (%)	Raio médio do aglomerado (km)	Distâncias médias dos apiários (km)
1	13.723	13,3	17,61	0,82
2	10.662	66,6	27,01	1,70
3	13.317	8,6	46,79	1,88
4	7.105	40,0	42,50	2,36
5	9.407	28,7	40,50	2,47
6	7.175	52,0	11,59	2,54

Exemplo

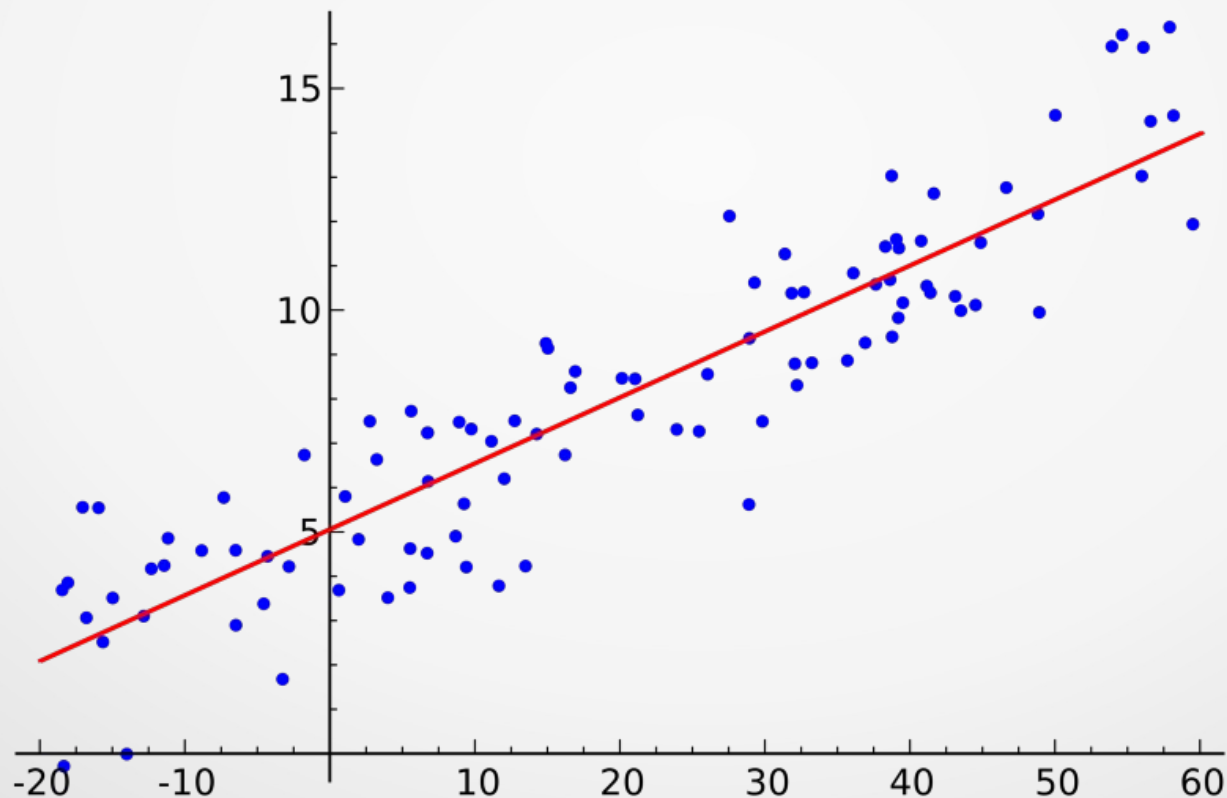
Problema de classificação/Regressão

- Regressão Linear
- Regressão Logística
- K-NN Classified (Vizinho mais próximo)
- Rede Neural Artificial
- Support Vector Machine(SVM)
- Árvore de decisão
- Random Forests
- Ensemble Methods
- Gradiente Tree Boosting

Exemplo

Problema de classificação/Regressão

- Regressão Linear



Exemplo

Problema de classificação/Regressão

- Regressão Logística

$$P(Y = 1) = \frac{1}{1 + e^{-g(x)}}$$

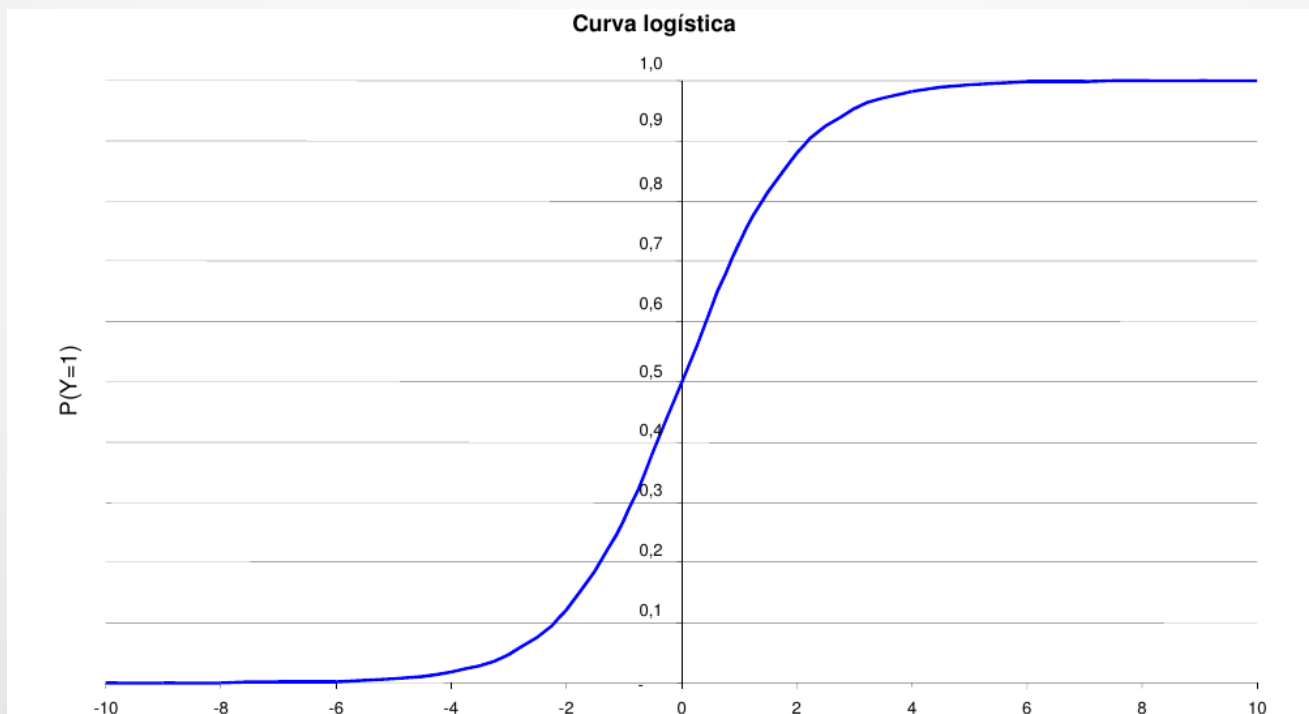
onde,

$$g(x) = B_0 + B_1X_1 + \dots + B_pX_p$$

Exemplo

Problema de classificação/Regressão

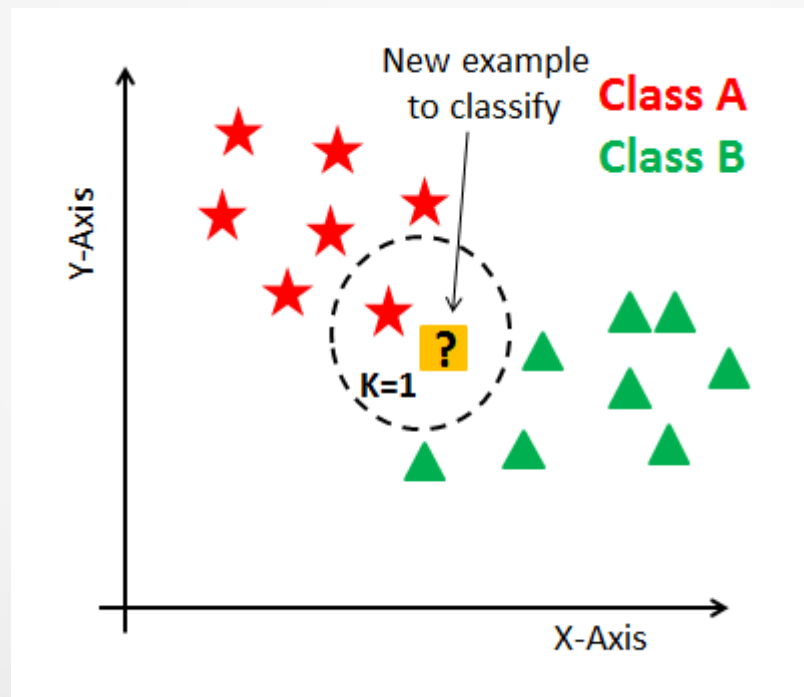
- Regressão Logística



Exemplo

Problema de classificação/Regressão

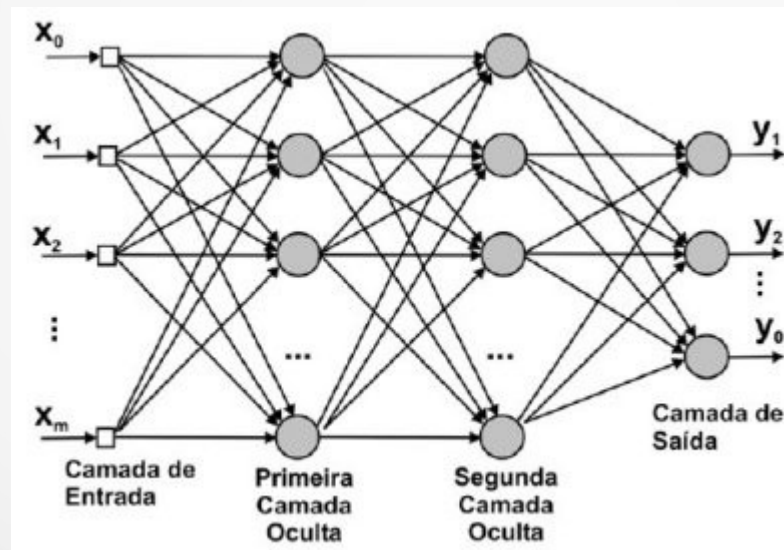
- K-NN Classified (Vizinho mais próximo)



Exemplo

Problema de classificação/Regressão

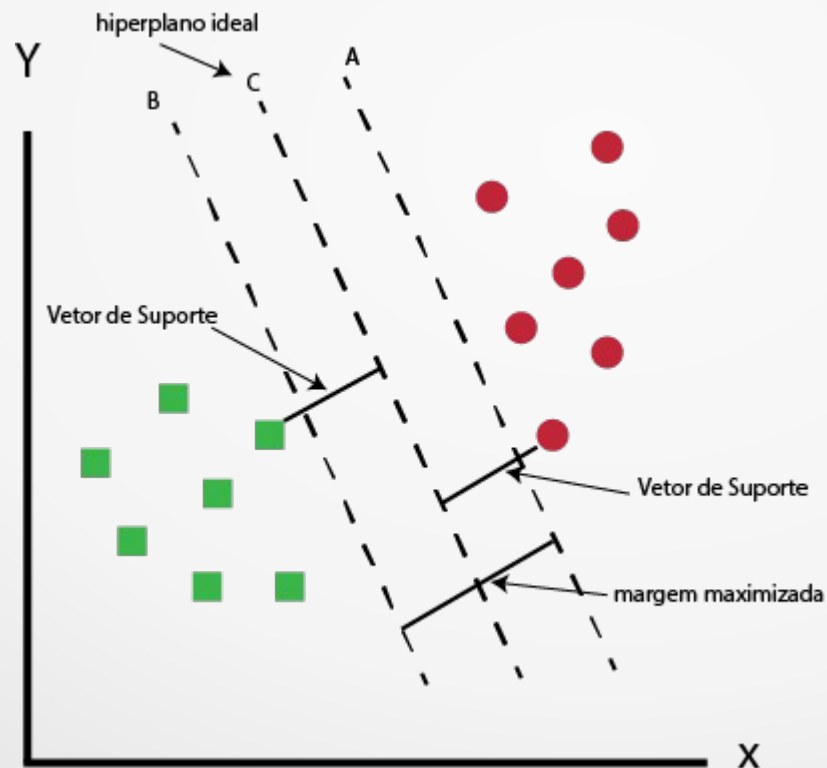
- Rede Neural Artificial



Exemplo

Problema de classificação/Regressão

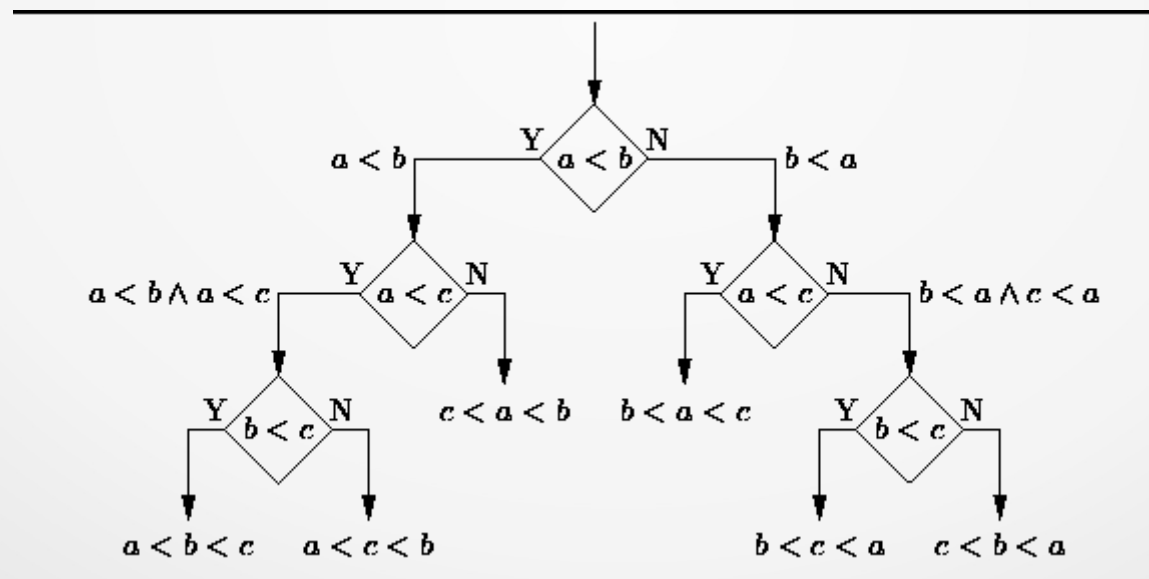
- Support Vector Machine



Exemplo

Problema de classificação/Regressão

- Árvore de decisão

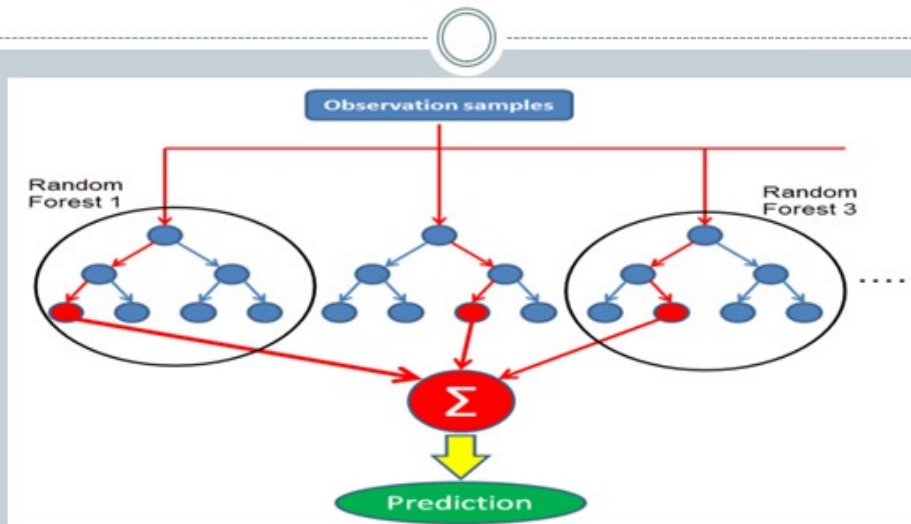


Exemplo

Problema de classificação/Regressão

- Random Forests

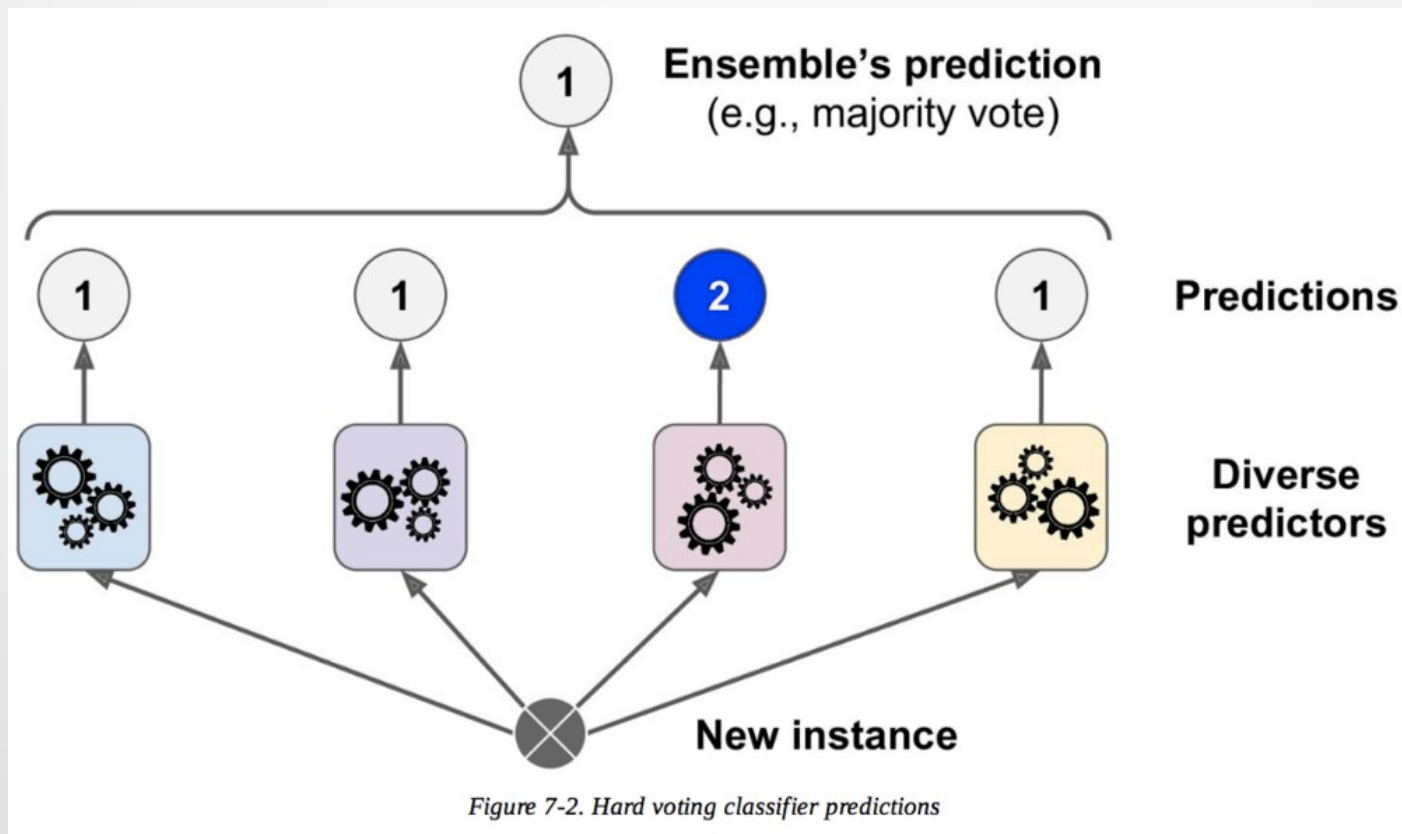
Chapter 5 : Random Forest Algorithm



Exemplo

Problema de classificação/Regressão

- Ensemble Methods



Exemplo

Problema de classificação/Regressão

Gradiente Tree Boosting

