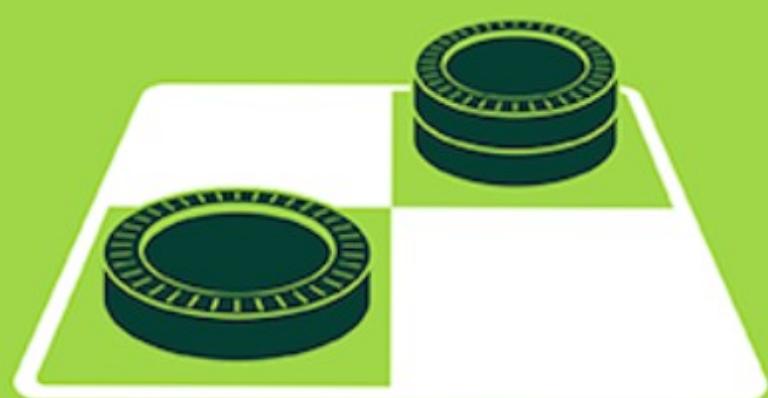


Machine Learning

ARTIFICIAL INTELLIGENCE



1950's

1960's

1970's

1980's

1990's

2000's

2010's

MACHINE LEARNING



DEEP LEARNING



National Museum of Computing, Bletchley Park, UK

Tópicos

- Nearest Neighbors
- Naive Bayes
- Decision Tree
- K-Means
- Random Forest
- Otimização



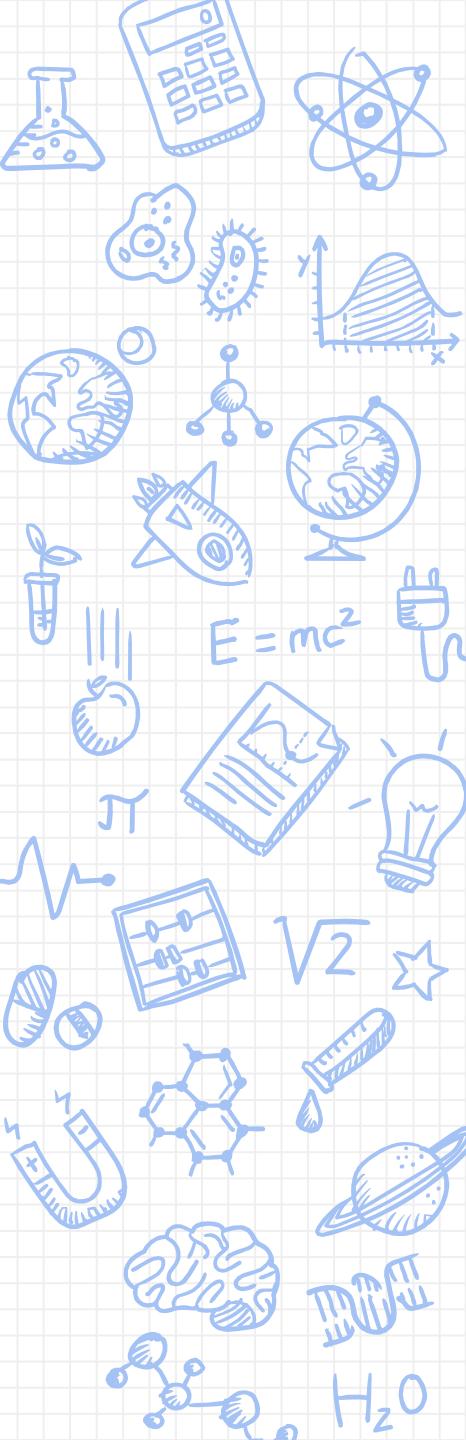
Nearest Neighbors

O que é?

É um método de aprendizado supervisionado e não supervisionado.

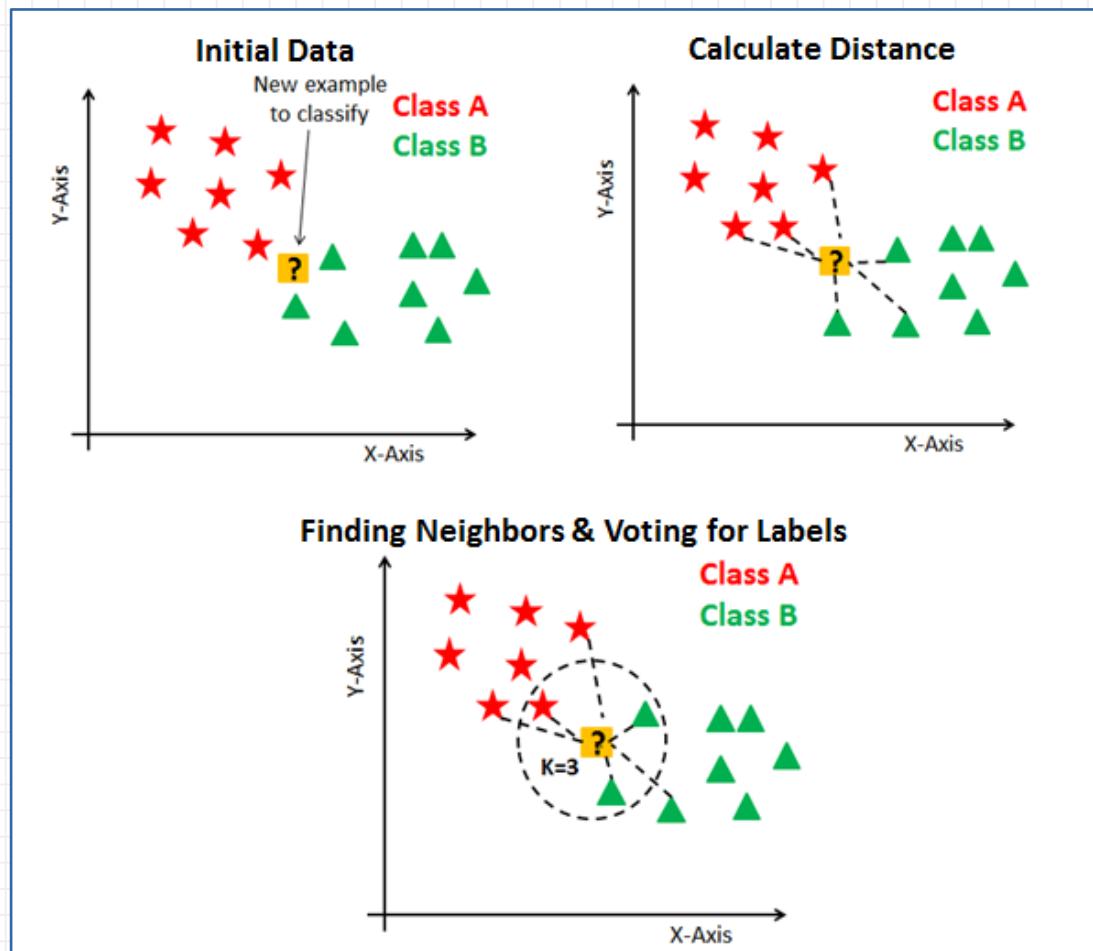
- Classificação (dados discretos);
- Regressão (dados contínuos).

É um dos algoritmos mais simplistas e precisa de menos dados que o convencional.



KNN

- Calcular a distância do novo dado;
- Encontrar os pontos mais próximos (K – vizinhos);
- Votação par definir qual a classe.



KNN

Algorítimos

- Ball Tree;
- Kd Tree;

Calculo da distância:

Euclidean

$$\sqrt{\sum_{i=1}^k (x_i - y_i)^2}$$

Manhattan

$$\sum_{i=1}^k |x_i - y_i|$$

Minkowski

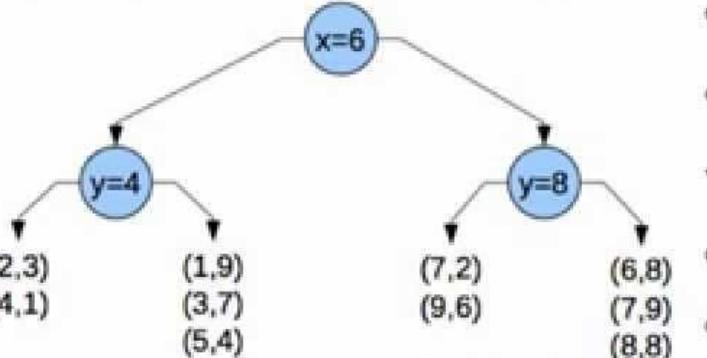
$$\left(\sum_{i=1}^k (|x_i - y_i|)^q \right)^{1/q}$$

KNN

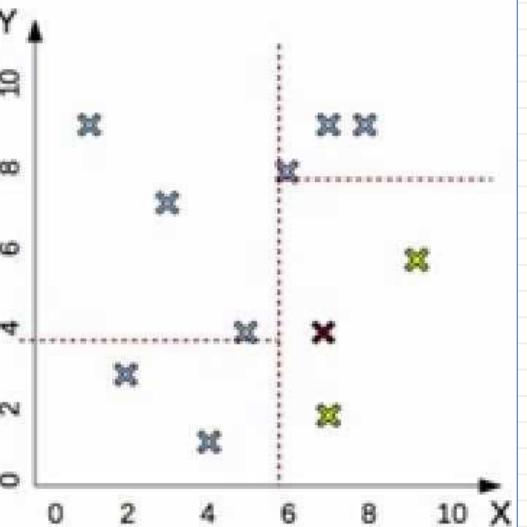
K-D Tree

K-D tree example

- Building a K-D tree from training data:
 - $\{(1,9), (2,3), (4,1), (3,7), (5,4), (6,8), (7,2), (8,8), (7,9), (9,6)\}$
 - pick random dimension, find median, split data, repeat
- Find NNs for new point $(7,4)$
 - find region containing $(7,4)$
 - compare to all points in region

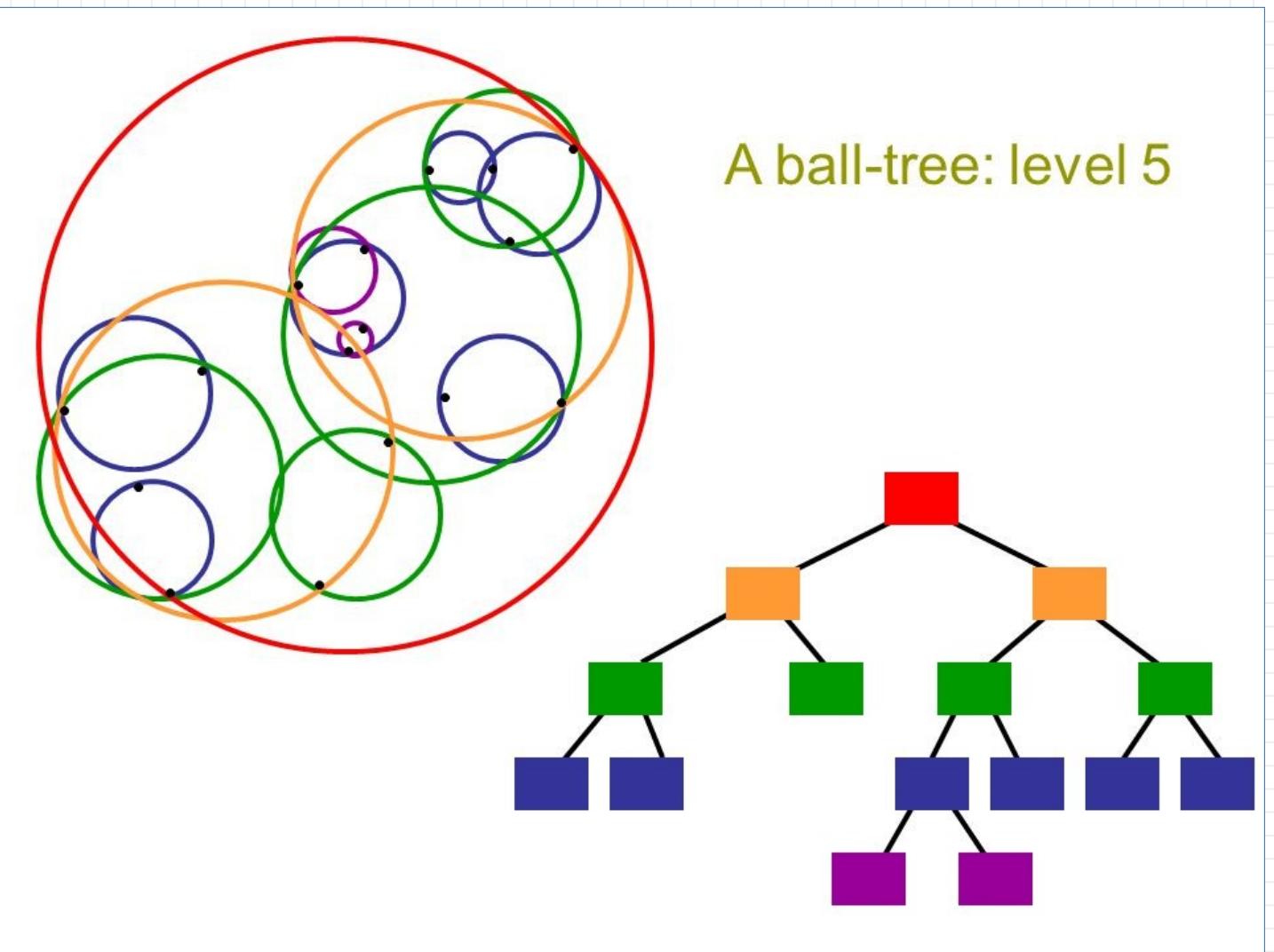


Copyright © 2013 Victor Lavrenko



KNN

Ball Tree



KNN

Comparação

Operati on on Instanc es	Build		Insertion		Deletion		N-NN Search	
	kd-tree	ball-tree	kd-tree	ball-tree	kd-tree	ball-tree	kd-tree	ball-tree
10000	187	1203	188	1187	218	1172	31	47
25000	547	3469	531	3468	516	3454	96	110
50000	1344	7906	1172	8094	1156	8093	118	219
100000	2578	16547	2750	16437	2812	16485	391	422
200000	5937	36672	5688	36703	5640	36375	875	907

KNN

Dados

- Escalonar/padronizar;
- NaN's;
- Funciona melhor para menores dimensões.

KNN



Naive Bayes

O que é?

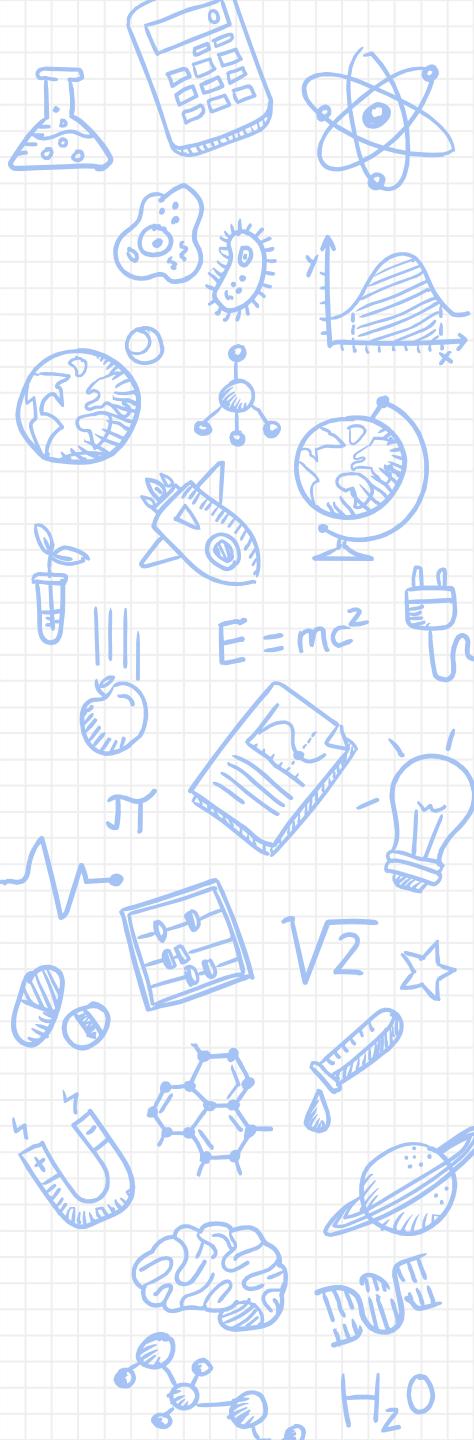
É um método de aprendizado supervisionado simples.

É uma família de algorítimos que se baseiam na aplicação do teorema de Bayes,

Utiliza uma suposição “Naive” - ingênuo.

Ingênuo: supõe que não existe dependência condicional entre os recursos.

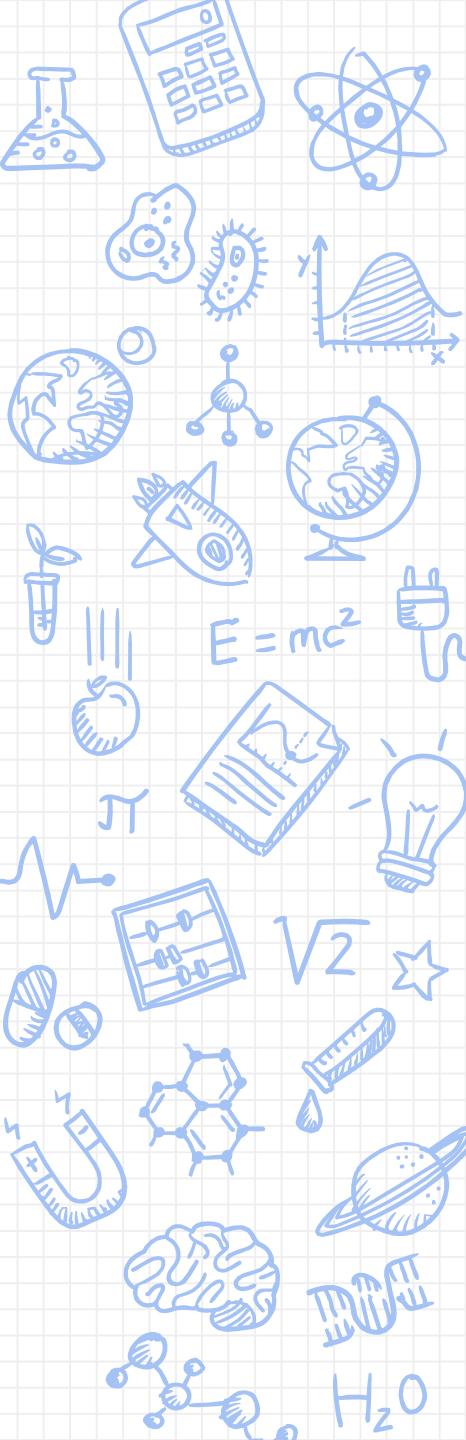
Considera os atributos do espaço probabilístico individualmente,
sendo um independente do outro

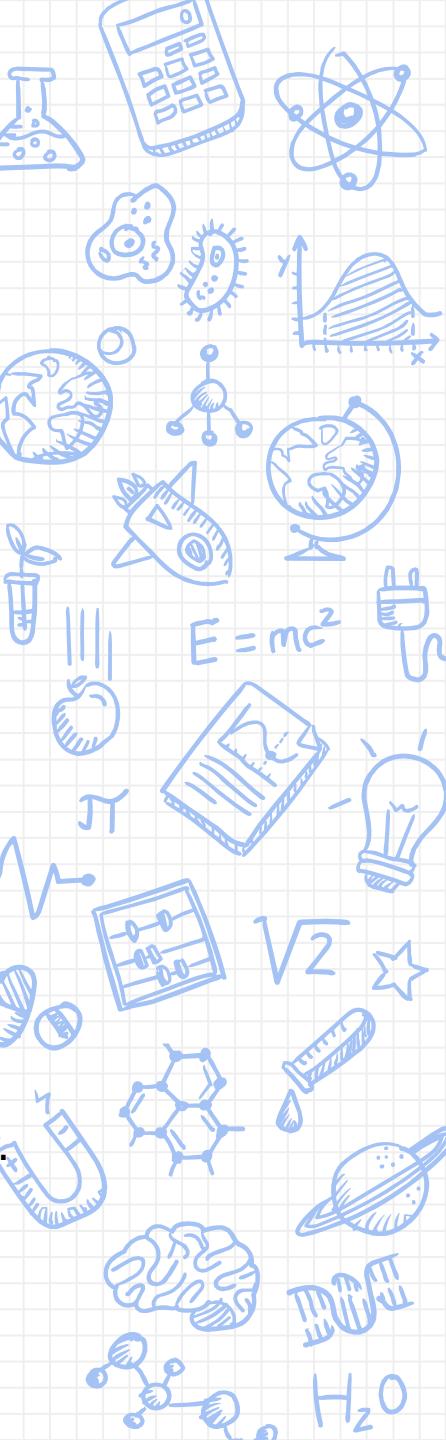


Naive Bayes

Conjunto de Dados “Tempo”

	Outlook	Temperature	Humidity	Windy	Play
D1	overcast	cool	normal	true	yes
D2	overcast	hot	high	false	yes
D3	overcast	hot	normal	false	yes
D4	overcast	mild	high	true	yes
D5	rainy	cool	normal	false	yes
D6	rainy	mild	high	false	yes
D7	rainy	mild	normal	false	yes
D8	sunny	cool	normal	false	yes
D9	sunny	mild	normal	true	yes
D10	rainy	cool	normal	true	no
D11	rainy	mild	high	true	no
D12	sunny	hot	high	false	no
D13	sunny	hot	high	true	no
D14	sunny	mild	high	false	no



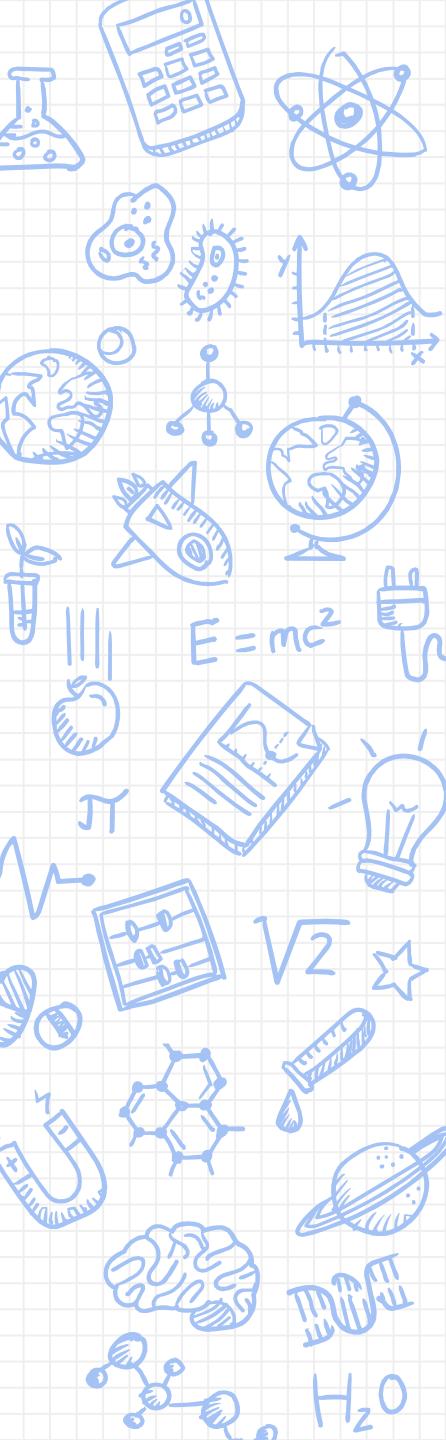


Naive Bayes

Tipos

- Gaussian Naive Bayes;
- Complement Naive Bayes;
- Bernoulli Naive Bayes;
- Categorical Naive Bayes;
- Multinomial Naive Bayes.

Os diferentes classificadores diferem principalmente nos pressupostos feitos em relação a $P(x_i | y)$.



Naive Bayes

Exemplo

CONTRATO DE LOCAÇÃO

Os signatários, que contratam nas qualidades indicadas neste contrato, têm entre si, ajustada a presente locação, mediante as seguintes cláusulas e condições:

I) LOCADOR(ES): _____ CNPJ / CPF _____

II) LOCATÁRIO(S): _____ CNPJ / CPF _____

III) FIADOR(ES) () CAUÇÃO () SEGURO DE FIANÇA LOCATÍCIA () _____ CNPJ / CPF _____

IV) OBJETO DE LOCAÇÃO: _____

FIM A QUE SE DESTINA: _____

V) VALOR MENSAL DA LOCAÇÃO: _____

O aluguel mensal é indicado neste contrato, devendo seu pagamento ser feito até o dia _____ de cada mês subsequente ao vencido, no endereço do LOCADOR ou de seu representante.

VI) PRAZO DA LOCAÇÃO:
INÍCIO: _____ TÉRMINO: _____
REAJUSTE A CADA _____ MESES, COM BASE NO(A)

VII) TRIBUTOS E DEMAIAS ENCARGOS: Obriga-se o LOCATÁRIO além do pagamento de aluguel a satisfazer: ao pagamento, por sua conta exclusiva do consumo de água, luz e esgoto, bem como, ao pagamento, por sua conta exclusiva de todas as despesas de condomínio e que sejam devidas pelo condomínio, ora LOCADOR, de acordo com a convenção do condomínio.

VIII) OBRIGAÇÕES GERAIS: O LOCATÁRIO declara ter procedido a vistoria do imóvel locado recebendo-o em perfeito estado e obrigando-se a:

- a) manter o objeto da locação no mais perfeito estado de conservação e limpeza, para assim o restituir ao LOCADOR, quando finda ou rescindida a locação, correndo por sua conta exclusiva as despesas necessárias para esse fim, notadamente, as que se referem à conservação de pinturas, portas, janelas, fechaduras, trincos, puxadores, vitrais e vidraças, lustres, instalações elétricas, torneiras, aparelhos sanitários, fogão e quaisquer outras, inclusive obrigando-se a pintá-lo novamente em sua desocupação, com tintas e cores iguais as existentes; tudo de acordo com o laudo de vistoria, assinado e anexado a este contrato, fazendo parte integrante do mesmo;
- b) não fazer instalação, adaptação, obra ou benfeitoria, inclusive colocação de luminosos, placas, letreiros e cartazes sem prévia obtenção de autorização, por escrito, do LOCADOR;
- c) não transferir este contrato, não sublocar, não ceder ou emprestar, sob qualquer pretexto e de igual forma alterar a destinação da locação, não constituindo o desuso do tempo, por si só, na demora do LOCADOR reprimir a infração, assentimento à mesma;
- d) encaminhar ao LOCADOR todas as notificações, avisos ou intimações dos poderes públicos que forem entregues no imóvel, sob pena de responder pelas multas, correção monetária e penalidades decorrentes do atraso no pagamento ou satisfação no cumprimento de determinações por aqueles poderes;
- e) no caso de qualquer obra, reforma ou adaptação, devidamente autorizada pelo LOCADOR, reparar por ocasião da entrega efetiva das chaves do imóvel locado, seu estado primitivo, não podendo exigir qualquer indenização;
- f) facultar ao LOCADOR ou ao seu representante legal examinar ou vistoriar o imóvel sempre que for para tanto solicitado, bem como no caso do imóvel ser colocado à venda, permitir que interessados o visitem;
- g) na entrega do prédio, verificando-se infração pelo LOCATÁRIO de quaisquer das cláusulas que se compõe este contrato, e que o prédio necessite de algum conserto ou reparo, ficará o mesmo LOCATÁRIO, pagando o aluguel, até a entrega das chaves;
- h) findo o prazo deste contrato, por ocasião da entrega das chaves, o LOCADOR mandará fazer uma vistoria no prédio locado, a fim de verificar se o mesmo se acha nas condições em que foi recebido, pelo LOCATÁRIO.

Este Contrato de Locação é de uso comum e não pode ser modificado, salvo se houver a anotação de sua respectiva alteração no cartório competente.

COE 10012-2

CONTRATO DE EMPREITADA

IDENTIFICAÇÃO DAS PARTES CONTRATANTES

CONTRATANTE: (Nome da Contratante), com sede em Irecê, na Rua São Lucas, nº 171, bairro Jardim Ouro Branco, Cep 47890-000, no Estado Bahia, inscrita no C.N.P.J. sob o nº 65.780.044/0001-87, e no Cadastro Estadual sob o nº 778.433.679-88, neste ato representada pelo seu diretor Ricardo da Silva Santana, brasileiro, Solteiro, Empresário, Carteira de Identidade nº 793866, C.P.F. nº 048.567.789-02, residente e domiciliado na Rua Castro Alves, nº 157, bairro Jacaré, Cep 47891-002, Cidade Irecê, no Estado Bahia.

CONTRATADA: (Nome da Contratada), com sede em São Paulo, na Rua Ribeiro de Souza, nº 909, bairro Meireles, Cep 85040-036, no Estado São Paulo, inscrita no C.N.P.J. sob o nº 98.667.324/0001-95, e no Cadastro Estadual sob o nº 554.879.099-49, neste ato representada pelo seu diretor Víctor de Castro, brasileiro, Casado, Engenheiro Civil, Carteira de Identidade nº 134317990-03, C.P.F. nº 887.665.443-07, residente e domiciliado na Rua Presidente Prudente, nº 7891, bairro Maracanã, Cep 85040-983, Cidade São Paulo, no Estado São Paulo.

As partes acima identificadas têm, entre si, justo e acertado o presente Contrato de Prestação de Empreitada, que se regerá pelas cláusulas seguintes e pelas condições de pagamento descritas no presente.

DO OBJETO DO CONTRATO

Cláusula 1ª: O presente contrato tem como objeto a realização de obra de restruturação da sede da CONTRATANTE, situada na Rua São Lucas, nº 171, bairro Jardim Ouro Branco, nesta cidade, ajustando-se com a CONTRATADA a sua execução mediante as cláusulas e condições deste instrumento.

Cláusula 2ª: A execução da obra dar-se-á em duas etapas, sendo a primeira de armazenação da ferragem da laje, e a segunda de enchimento da laje.

DO PRAZO

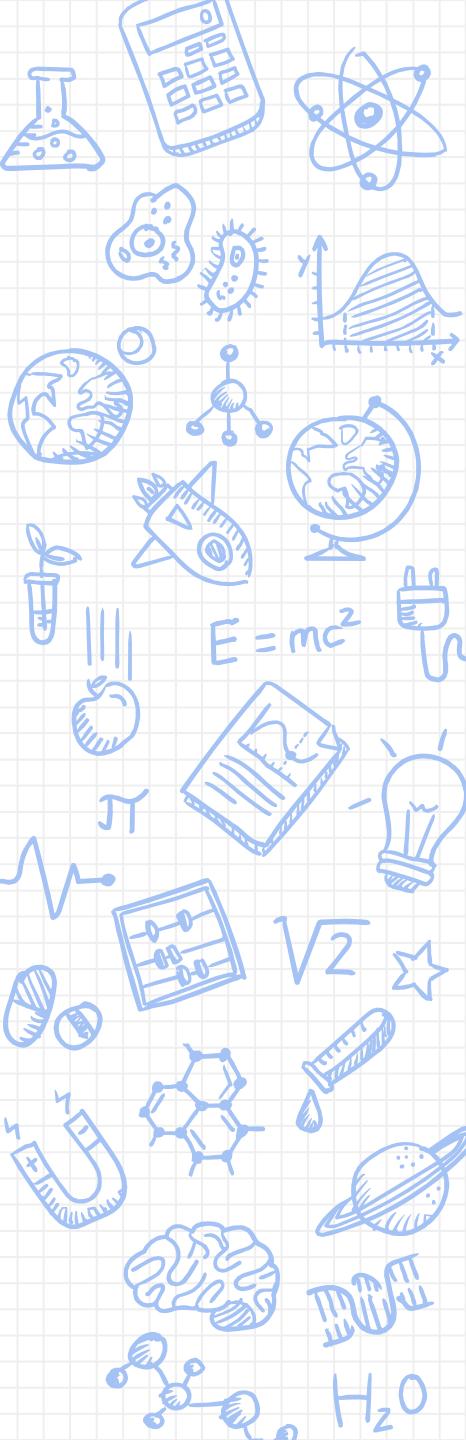
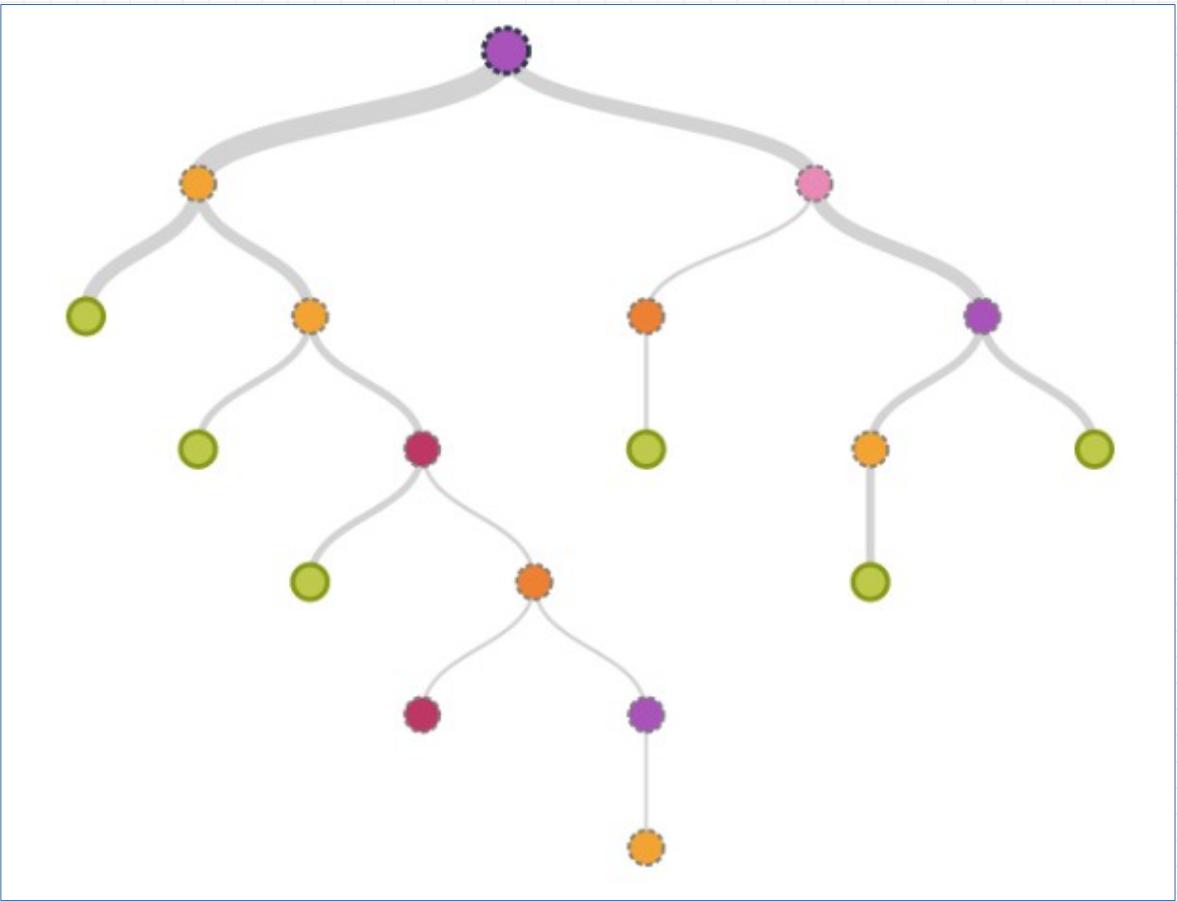
Cláusula 3ª: O presente contrato de empreitada terá início na data de 17/06/2011, com seu término previsto para até o dia (xxx).

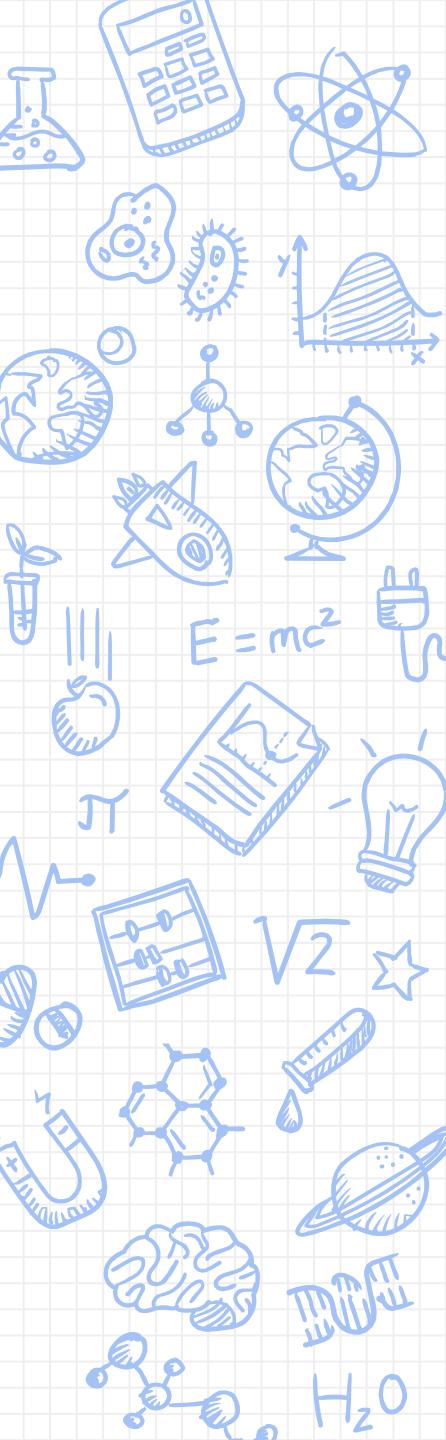
DAS OBRIGAÇÕES DA CONTRATADA

Cláusula 4ª: A CONTRATADA fornecerá por sua própria conta, além do seu trabalho, as ferramentas necessárias à execução da obra, incumbindo-se da limpeza da laje e da remoção necessária de entulhos, responsabilizando-se por qualquer contratação de terceiros, eximindo

Decision Tree

São árvores com regras nos nós



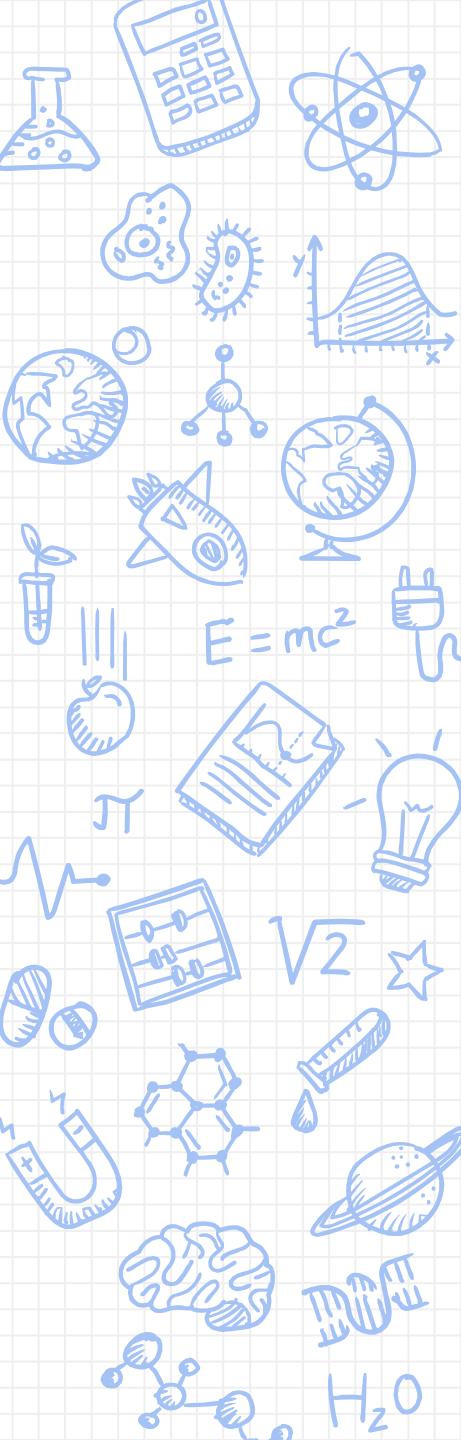
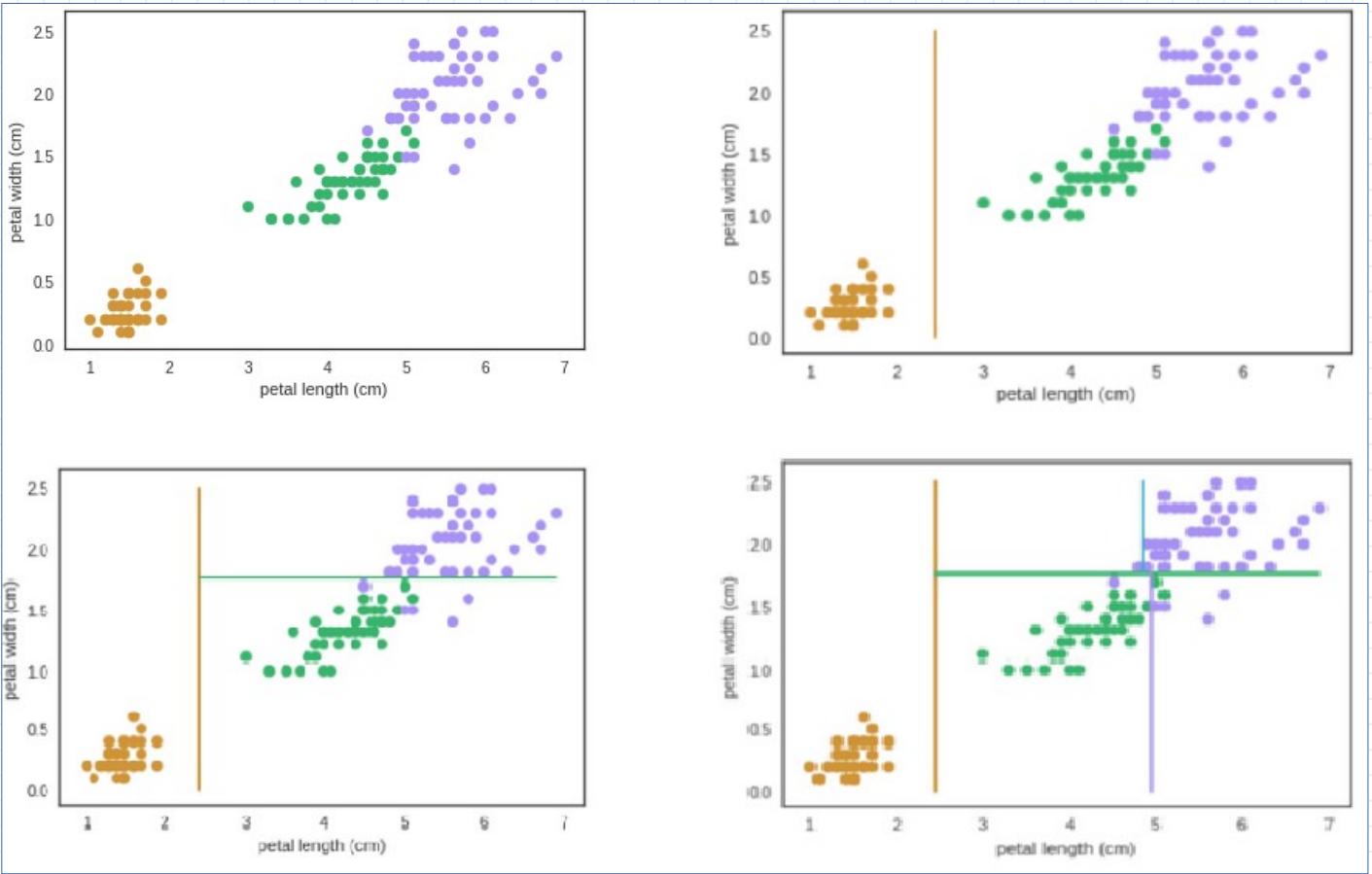


Decision Tree

O que é?

- Supervisionado;
- Dividir para conquistar (se/então);
- Separa baseado na divisão do espaço em outros sub-espacos;
- O critério de separação é referente a utilidade do atributo.

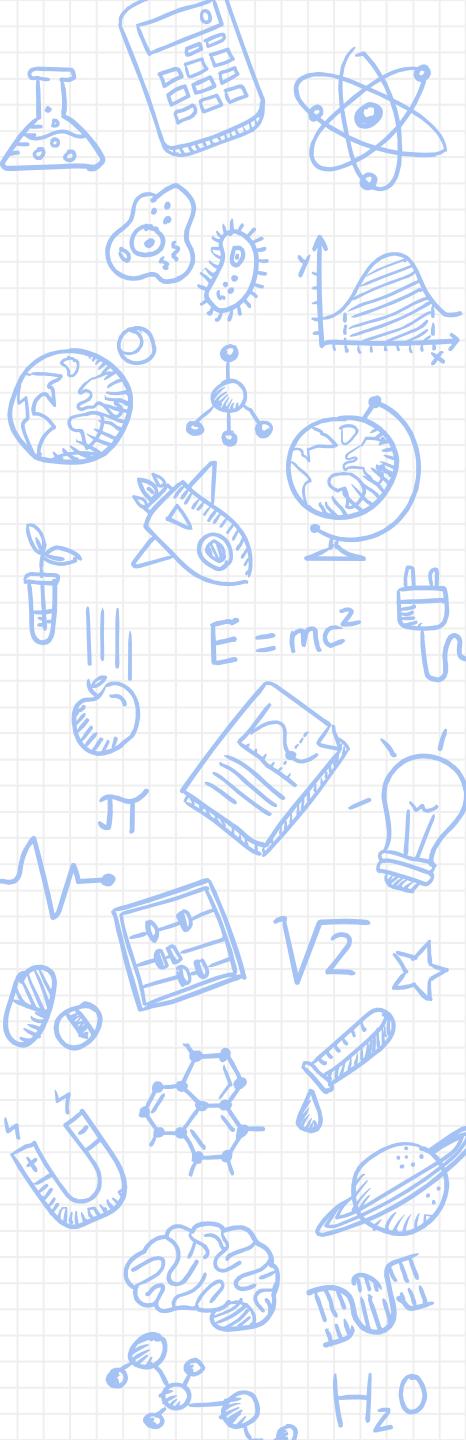
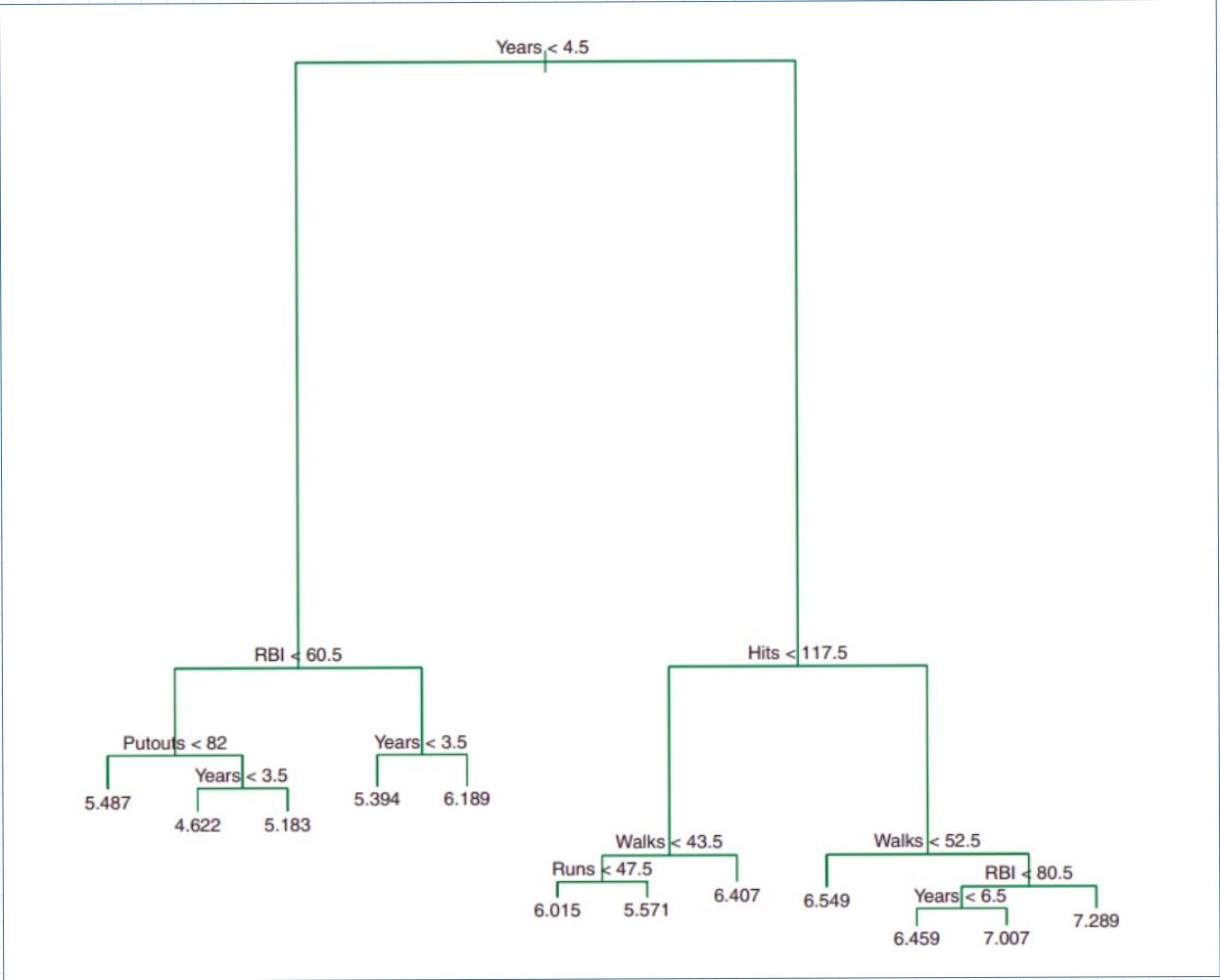
Decision Tree



Decision Tree

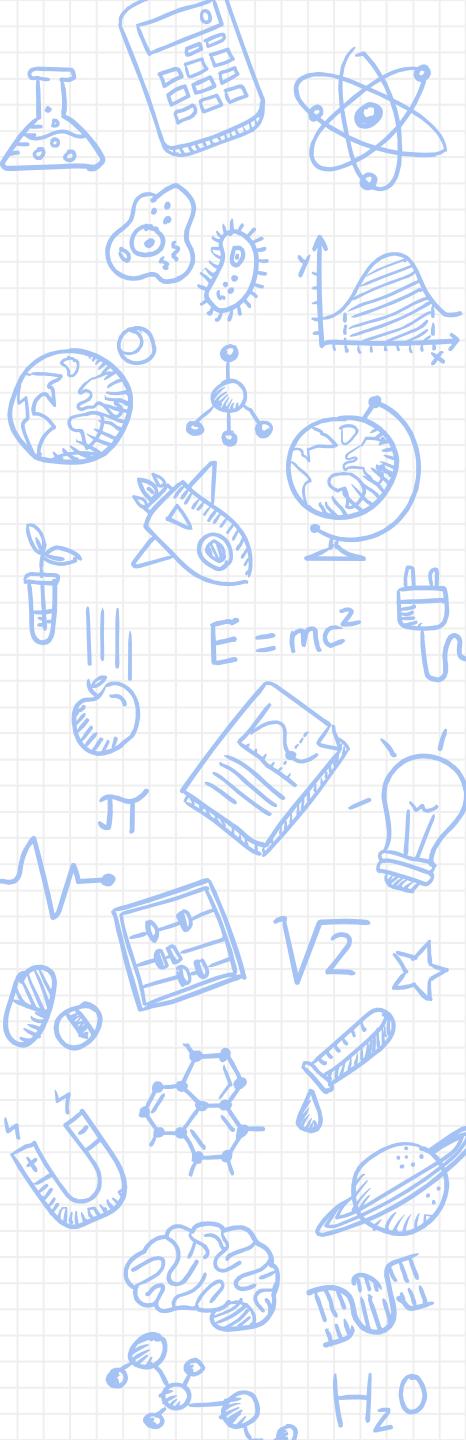
O que é?

Cada nó é um ponto de divisão



Decision Tree

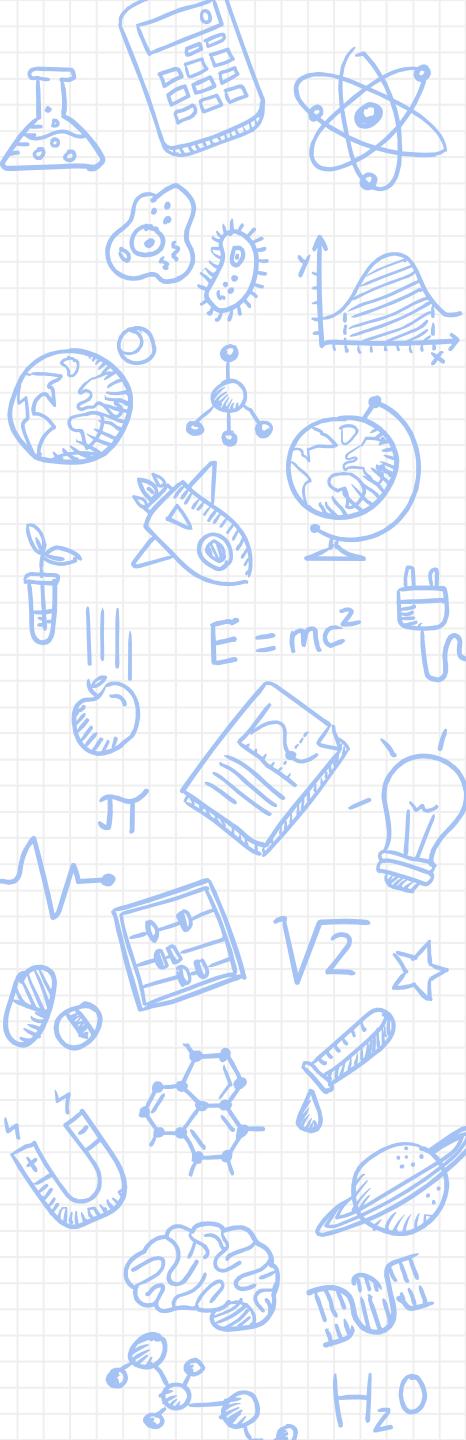
- Causa muito overfitting;
- Se tiver um parâmetro muito forte, o resultado fica tendencioso;
- Algorítimo ganancioso;
- Não paramétrico, aumenta a correção entre as árvores.



Random Forest

O que é?

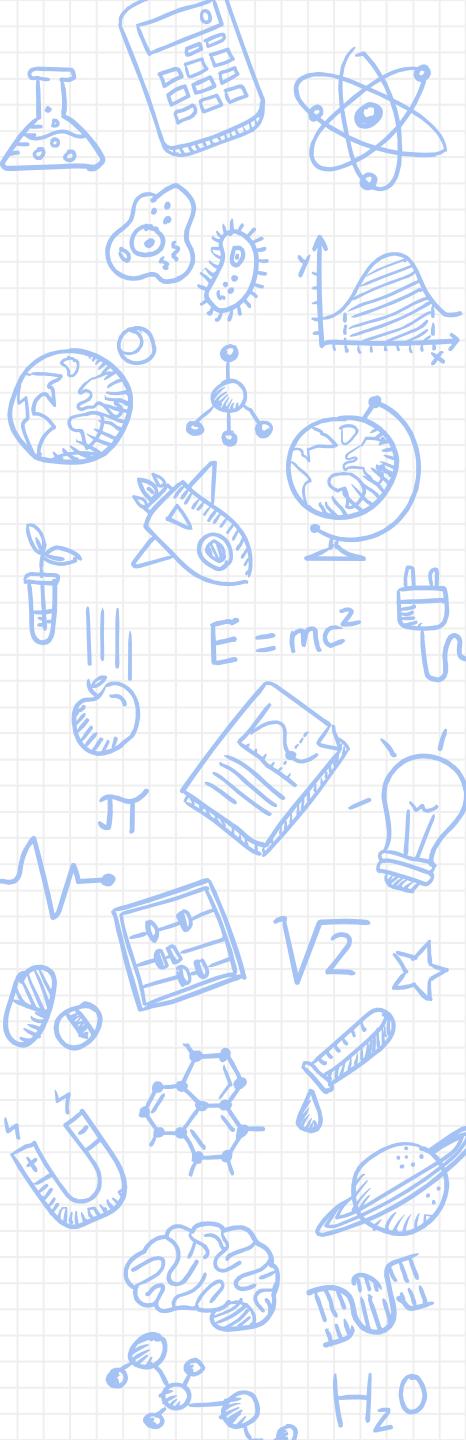
- Random Forest é um algoritmo supervisionado;
- É basicamente um *Ensemble* de árvores de decisão;
- Cria árvores aleatórias com parâmetros e divisões aleatórias;



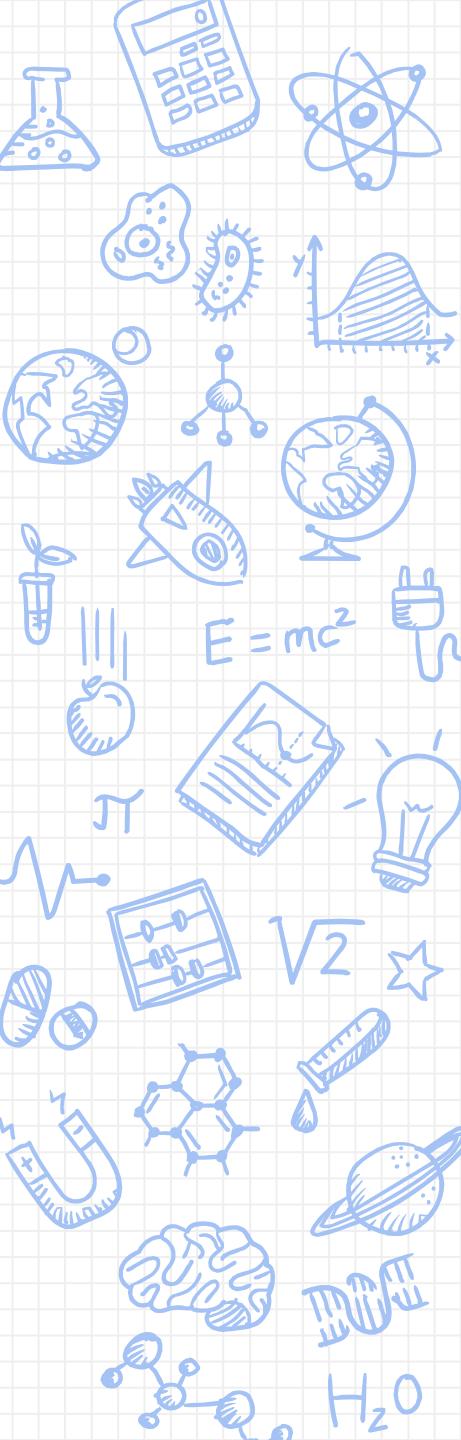
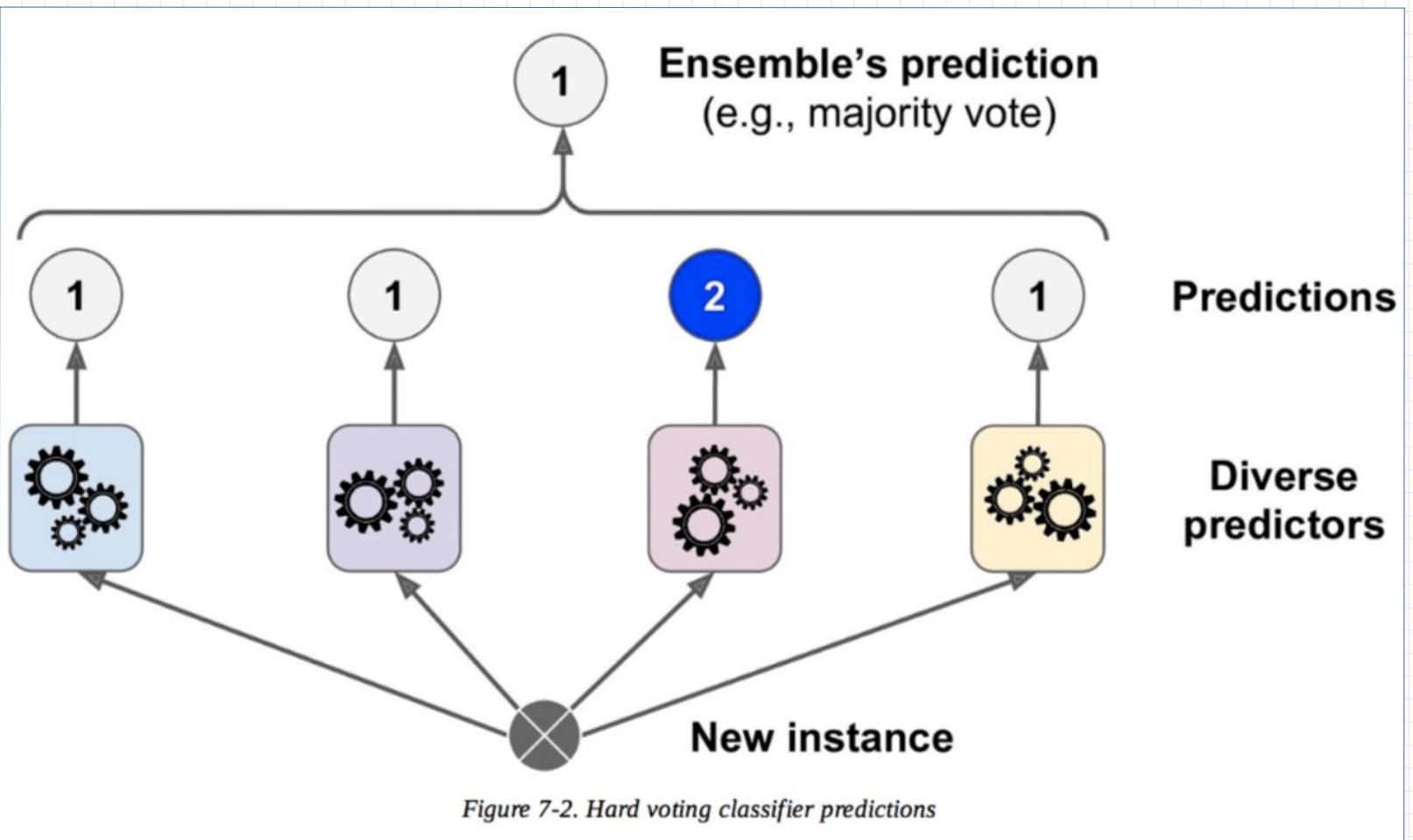
Ensemble

O que é?

- Conjuntos de modelos;
- Funciona melhor com algorítimos diferentes.

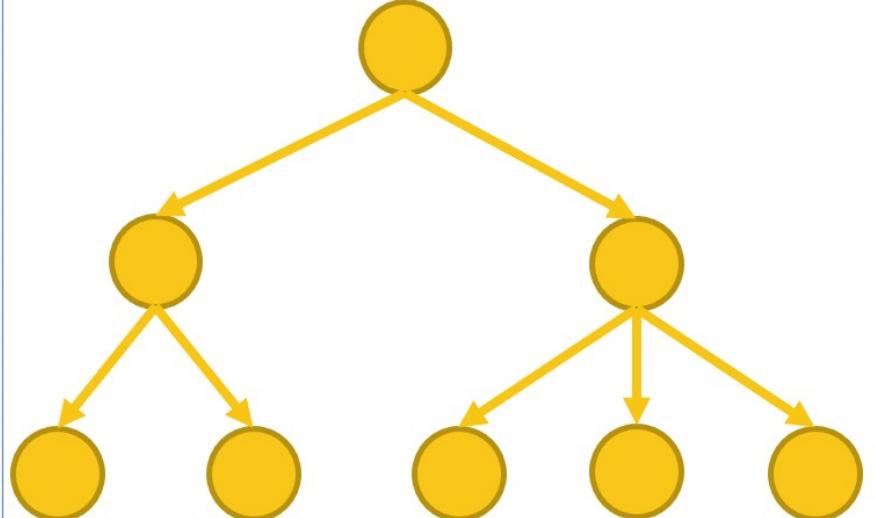


Ensemble

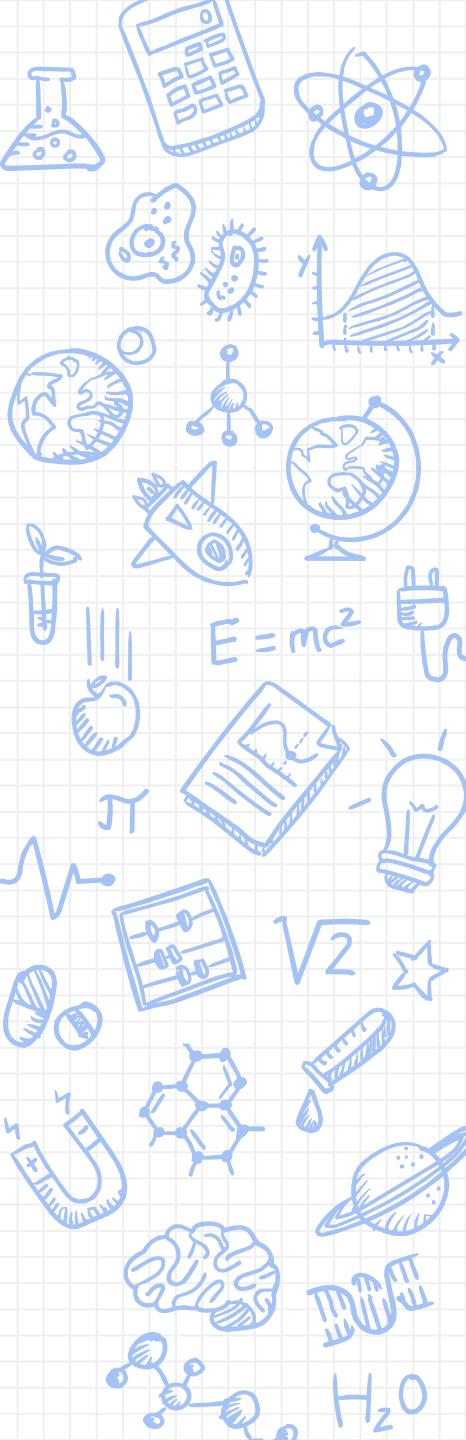
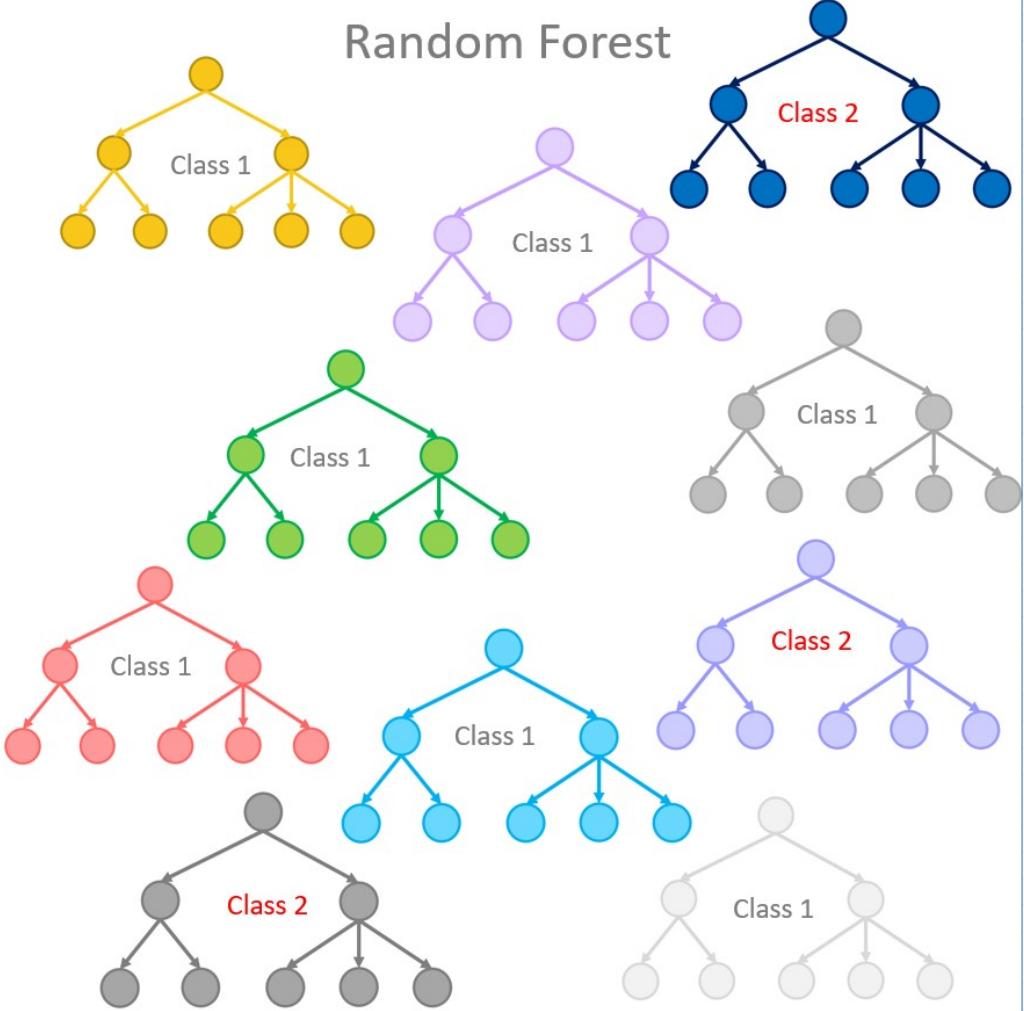


Random Forest

Single Decision Tree



Random Forest

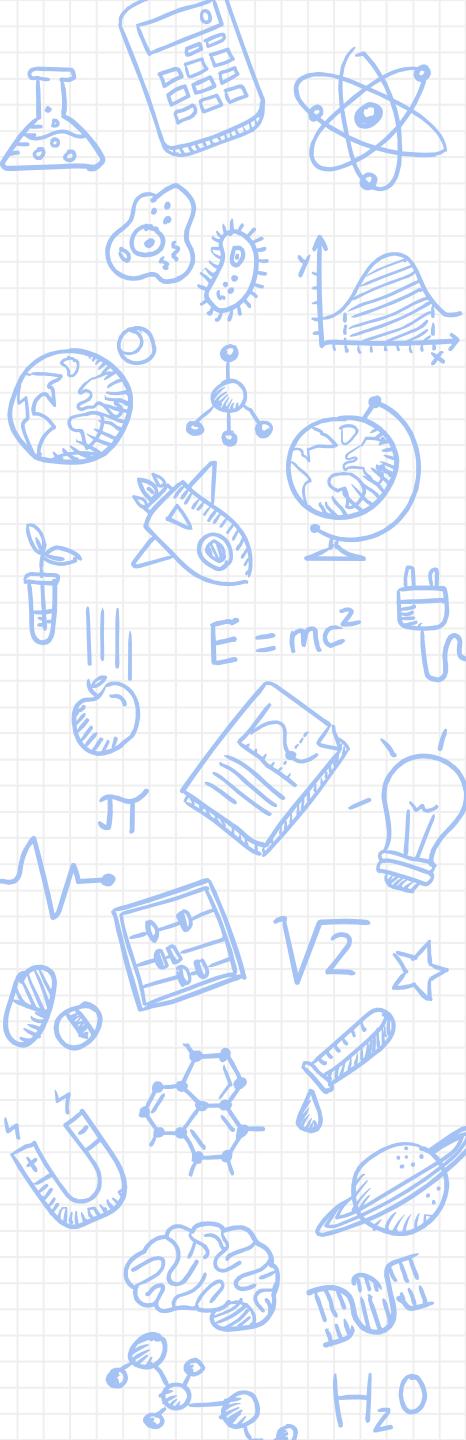


Grid Search

Otimização de algoritmos

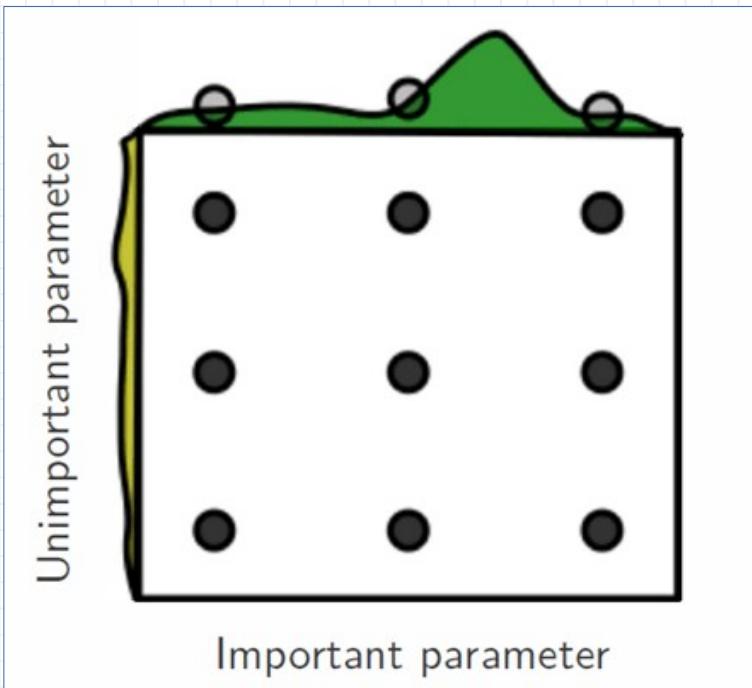
Parâmetros: Propriedade que são aprendidas por conta própria no treinamento.
Característica interna do modelo e seu valor pode ser estimado pelos dados;

Hiper-parâmetros: Configurações externas ao modelo e o valor pode ser estimado
à partir dos dados. Valor definido no início do processo de aprendizado.

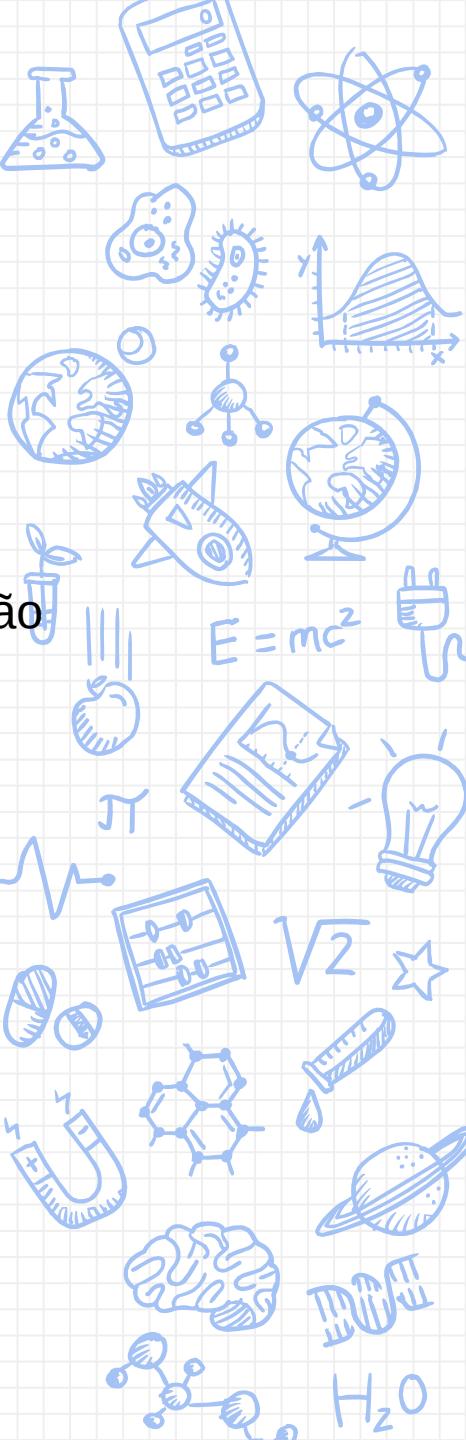


Grid Search

- Abordagem ingênuo e de força bruta que tenta todas as configurações possíveis dentro de um subconjunto de possibilidades;
- Procura no espaço de hiper-parâmetros a melhor pontuação de a melhor pontuação na validação cruzada;

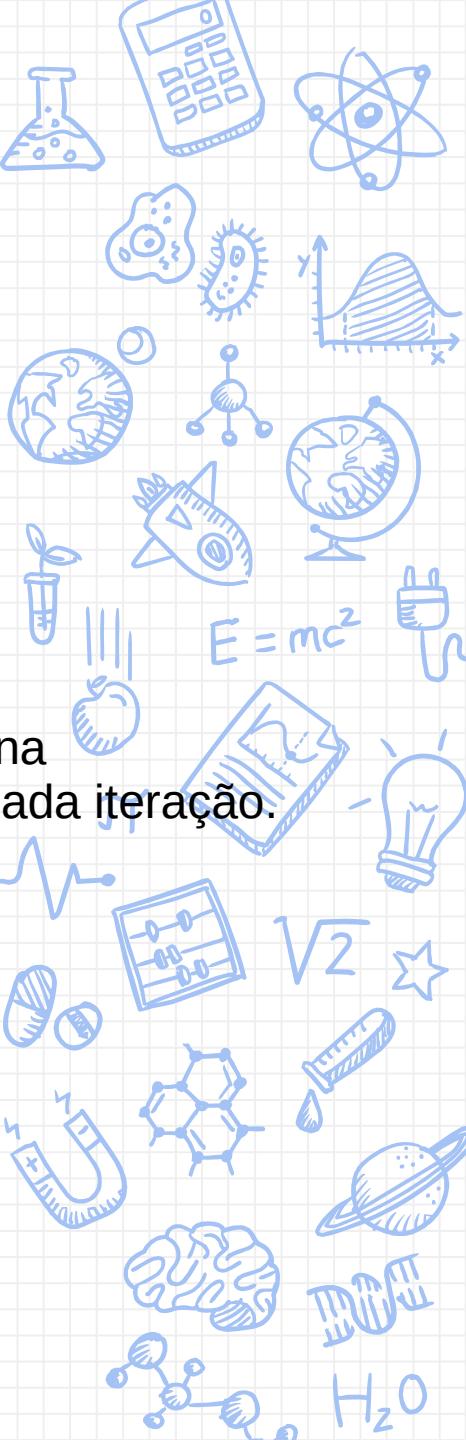
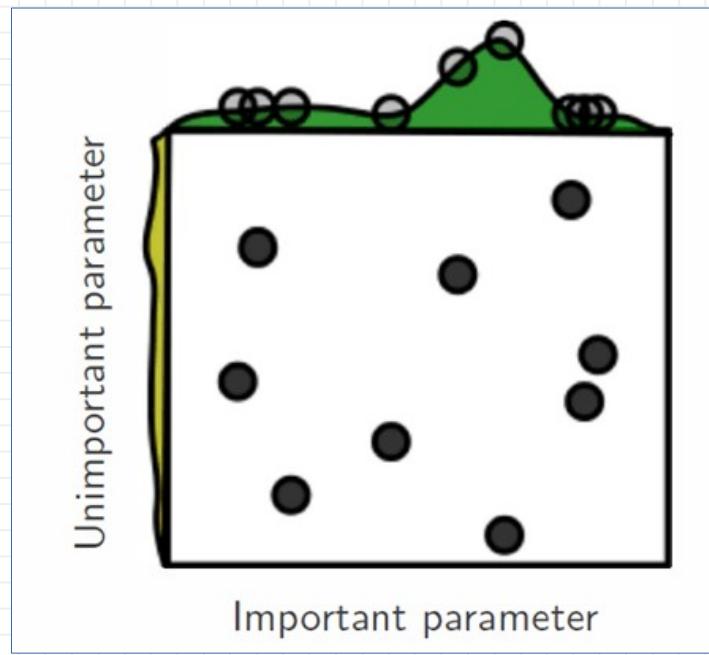


- Desvantagem do Grid Search é a quantidade de dimensões.
- Testa todas as combinações possíveis.



Random Search

- Usa combinações aleatórias de hiper-parâmetros são usados para encontrar
 - a melhor solução;
 - Procura amostras aleatórias, avalia o conjunto e assume que um hiper-parâmetro é melhor que outro.
 - Ao invés de testar todas as combinações ele seleciona um valor aleatório para cada hiperparâmetro em cada iteração.



Obrigado!!!
Dúvidas?

