



Data Science para Mercado Financeiro

Thyago Carvalho Marques

Agenda

- ☐ Big Data Analytics
- ☐ Ciência de dados
- ☐ Processo em ciência de dados
- ☐ Principais técnicas e ferramentas

Big Data Analytics!!!

□ Big Data

- Quando volume, velocidade ou variedade de dados excede a capacidade de **armazenamento** e/ou **computação**
- Big Data é **relativo**, não absoluto

Analytics

- Aprendizado de máquinas
- Mercado
 - “Relatórios”, painéis, etc



Big Data



❑ A cada segundo¹:

- 100.000 tweets circulam
- 547 websites são criados
- mais de 2 milhões de pesquisas (Google)
- 48h de vídeos são baixadas no YouTube
- 684.478 itens são compartilhados no Facebook...

❑ Em governo (Brasil)²:

- Mais de 7 milhões de notas fiscais eletrônicas (NFe) por dia
- Mais de 16 bilhões de NFe autorizadas...

Big Data



☐ A cada segundo¹:

- 100.000 tweets circulam
- 547 websites são criados
- mais de 2 milhões de pesquisas (Google)
- 48h de vídeos são baixadas no YouTube
- 684.478 itens são compartilhados no Facebook...

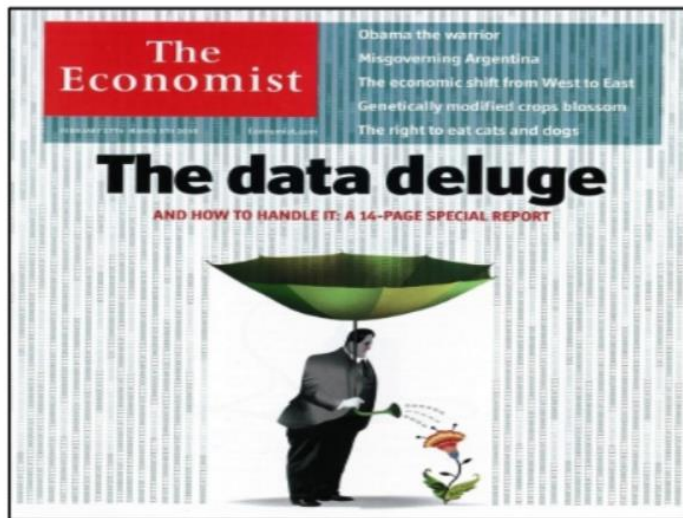
☐ Em governo (Brasil)²:

- Mais de 7 milhões de notas fiscais eletrônicas (NFe) por dia
- Mais de 16 bilhões de NFe autorizadas...

Big Data X Big User

Como lidar com este “dilúvio” de dados?

- ❑ A palavra mais importante no termo “ciência de dados” não é “dados”, mas **ciência**.



Fonte (2010): <http://www.economist.com/node/15579717>

Ciência de Dados

- ❑ A partir da necessidade de análise desse emaranhado de dados surgiu uma “nova” área da ciência, a chamada **ciência de dados**
 - O “quarto paradigma” da ciência¹
 - A profissão mais “sexy” do século 21²
 - Uma nova buzzy word!
- ❑ As atividades executadas pelo “**cientista de dados**”, em menor escala em relação ao volume de dados, são bastante antigas³

1 - Tansley, S.; Tolle, K.M (2009). The Fourth Paradigm: Data-intensive Scientific Discovery. Microsoft Research.

2 - <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

3 - www.forbes.com/sites/gilpress/2013/05/28/a-very-short-history-of-data-science/

Cientista de Dados



WIRED

Tell Your Kids to Be Data Scientists, Not Doctors

BUSINESS

CULTURE

DESIGN

GEAR

SCIENCE

SHARE



SHARE



TWEET



COMMENT



EMAIL

PARTNER CONTENT LINDA BURTCH, BURTCH WORKS

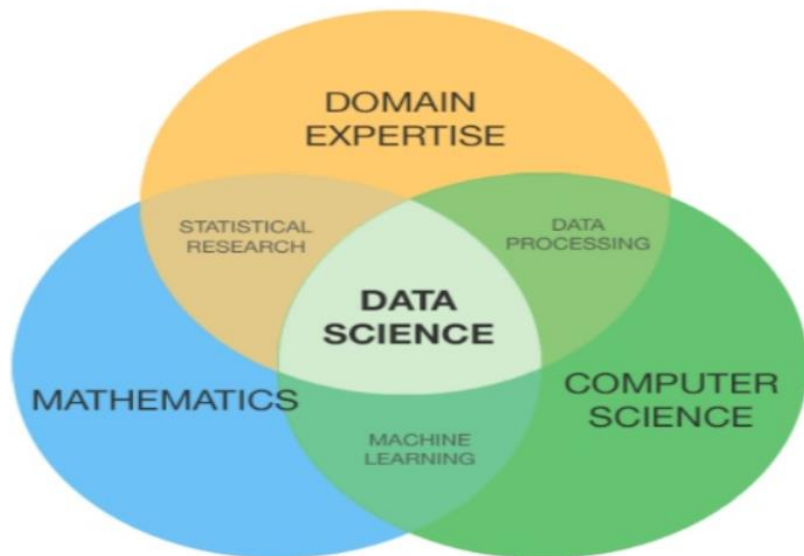
TELL YOUR KIDS TO BE DATA SCIENTISTS, NOT DOCTORS



Image: ifindkarma/Flickr

Fonte (2014): <https://www.wired.com/insights/2014/06/tell-kids-data-scientists-doctors/>

Cientista de Dados



Source: Palmer, Shelly. *Data Science for the C-Suite*.
New York: Digital Living Press, 2015. Print.

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21st century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand what a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- Machine learning
- Statistical modeling
- Experiment design
- Bayesian inference
- Supervised learning: decision trees, random forests, logistic regression
- Unsupervised learning: clustering, dimensionality reduction
- Optimization: gradient descent and variants

DOMAIN KNOWLEDGE & SOFT SKILLS

- Passionate about the business
- Careless about data
- Influence without authority
- Hacker mindset
- Problem solver
- Strategic, proactive, creative, innovative and collaborative

PROGRAMMING & DATABASE

- Computer science fundamentals
- Scripting language e.g. Python
- Statistical computing packages, e.g. R
- Databases: SQL and NoSQL
- Relational algebra
- Parallel databases and parallel query processing
- MapReduce concepts
- Hadoop and Hadoop/Hive/Hig
- Custom reducers
- Experience with Hadoop like AWS

COMMUNICATION & VISUALIZATION

- Able to engage with senior management
- Story telling skills
- Translate data-driven insights into decisions and actions
- Visual art design
- R packages like ggplot2 or lattice
- Knowledge of any of visualization tools e.g. Fream, D3.js, Tableau

MarketingInsider.com is a group of practitioners in the area of a consumer-marketing. Our fields of expertise include marketing strategy and optimization, customer tracking and on-site analytics, predictive analytics and econometrics, data mining and big data systems, marketing channel insights in Paid Search, SEM, Social, CRM and brand.

Marketing
INSIDER.COM
61 King Street, London

O que os cientistas de dados fazem?

- ☐Definem hipóteses e perguntas
- ☐Definem os conjuntos de dados ideais
- ☐Determinam que dados podem ser acessados
- ☐Adquirem os dados
- ☐Preprocessam os dados
- ☐Realizam análise de dados exploratória
- ☐Realizam modelagem estatística dos dados
- ☐Interpretam resultados de análises
- ☐Escrevem relatórios sobre os resultados
- ☐Criam modelos/componentes/códigos reusáveis
- ☐Compartilham modelos e resultados com outras pessoas

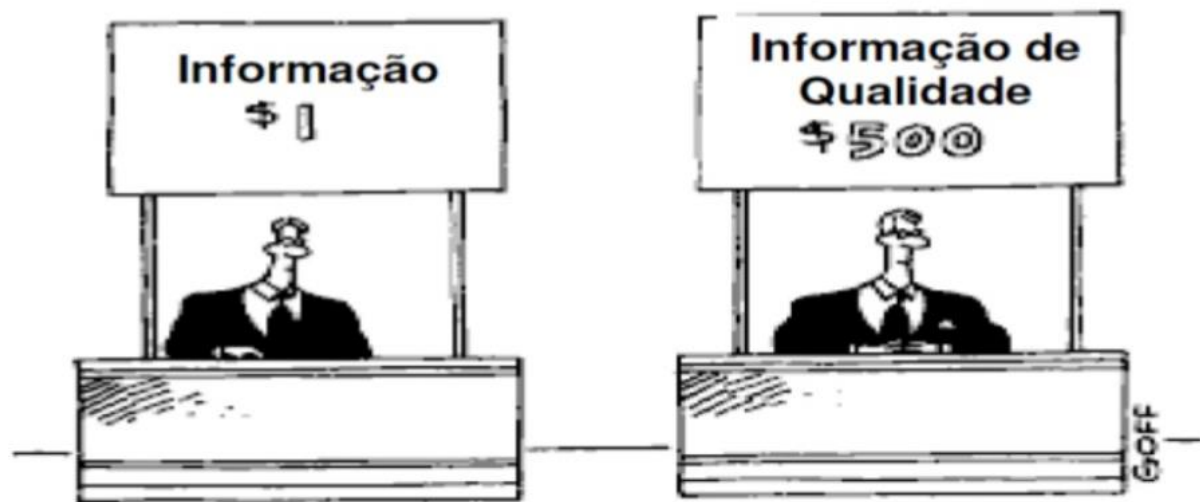
Considerando:

- **Processos**
- **Ambientes**
- **Projetos**



O que os cientistas de dados fazem?

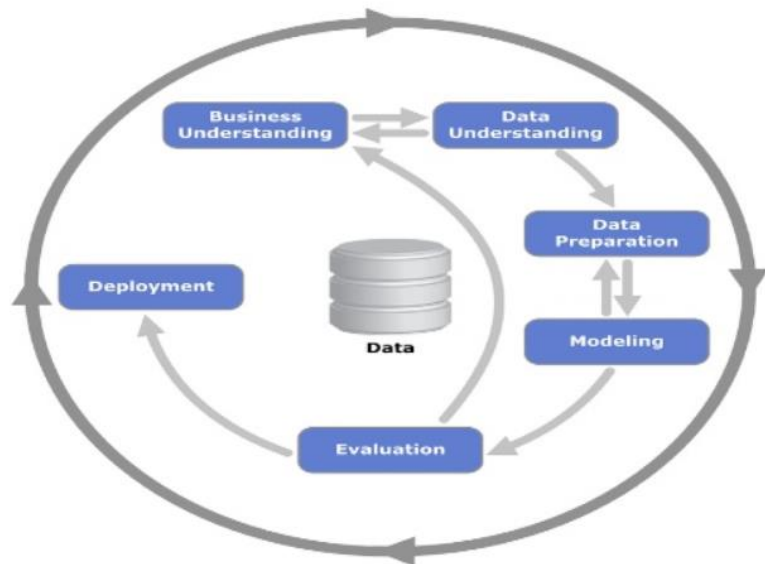
- ❑ Informação com alto valor agregado!



© 1996 Ted Goff

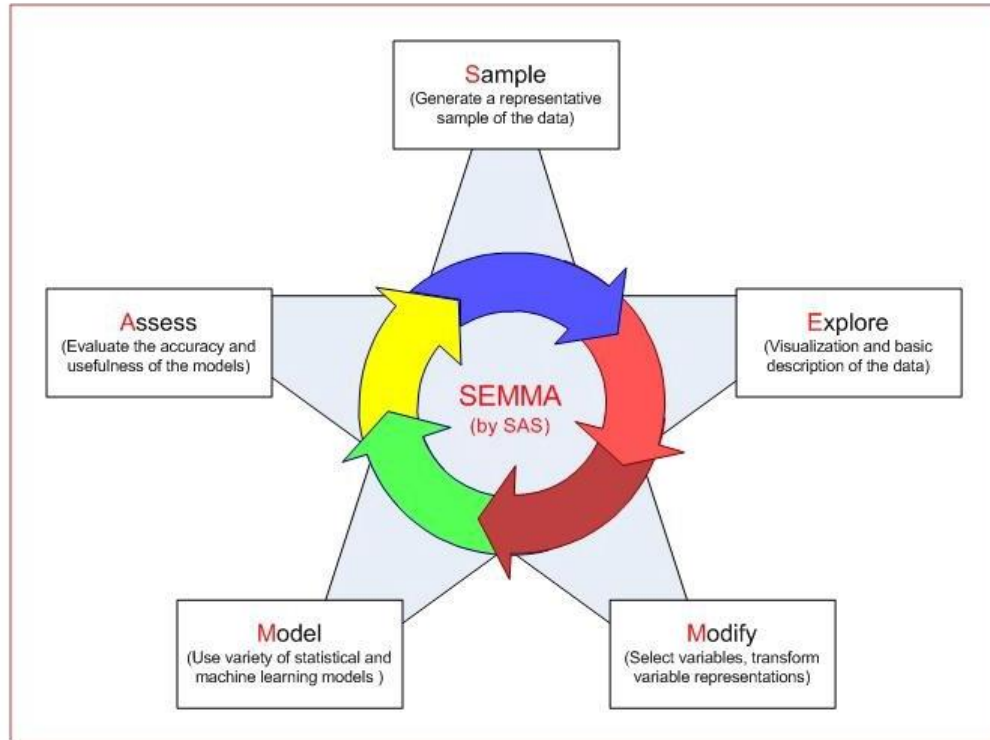
Processo em ciência de dados

- ❑ **CRISP-DM** (CRoss Industry Standard Process for Data Mining)
- ❑ SEMMA (Sample, Explore, Modify, Model and Assess)



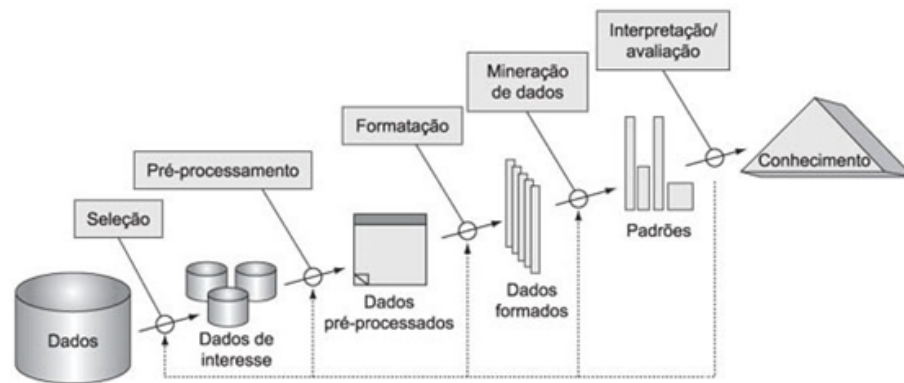
Sample, Explore, Modify, Model e Assess — SEMMA

Primeiro os dados, depois os negócios



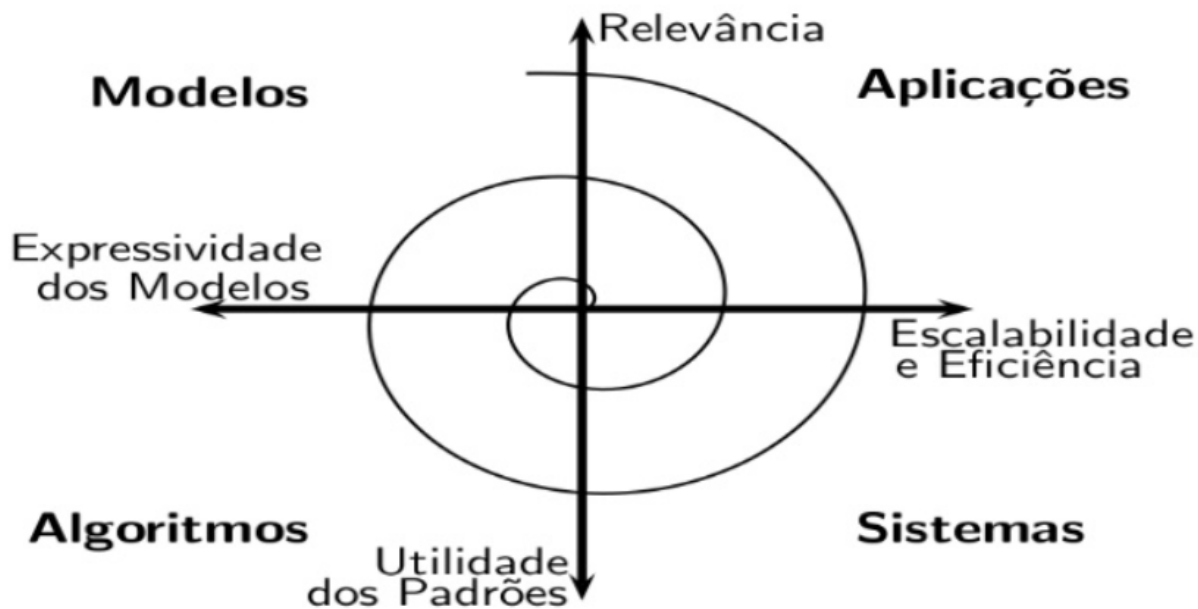
Knowledge Discovery in Databases

Ou só KDD



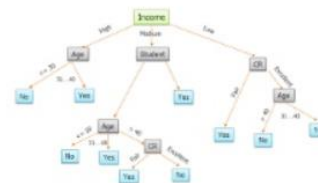
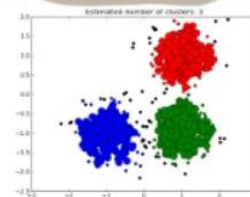
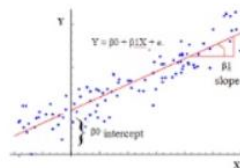
Esse é um dos métodos mais antigos existentes, tendo sua criação feita em 1980. Essa técnica é muito utilizada em Data Mining e, diferente das anteriores, não foca em questões de negócio ou geração de modelos, mas sim na descoberta de conhecimentos a partir dos dados.

Processo em ciência de dados



Principais técnicas

- ❑ Algumas classes de problemas
 - Modelos de regressão
 - Descoberta de conjuntos frequentes
 - Agrupamento
 - Classificação
- ❑ Existem algoritmos diferentes para cada classe
- ❑ Existem outras classe de problemas
 - Análise de imagem
 - Redes sociais
 - Mineração de texto
 - etc





Tipos de Aprendizado



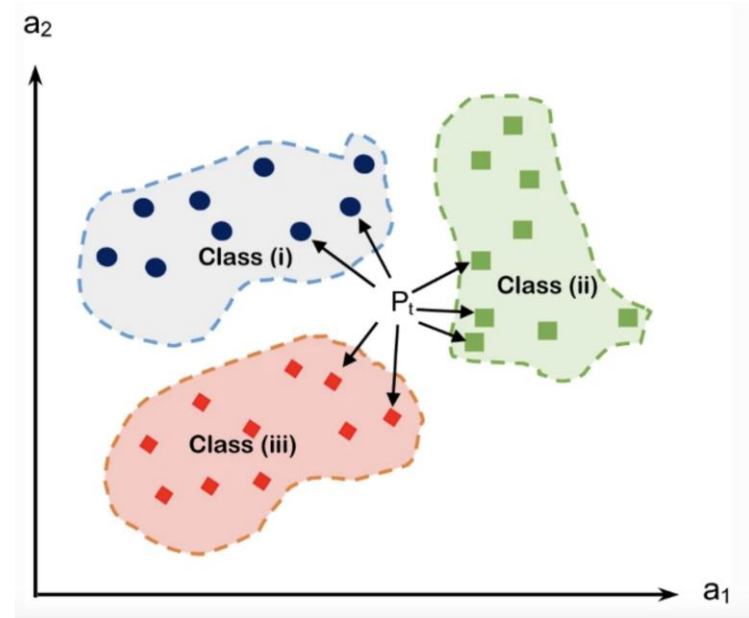
APRENDIZADO SUPERVISIONADO

O aprendizado supervisionado ocorre quando o **modelo aprende a partir de resultados pré-definidos**, utilizando os valores passados da variável alvo para aprender quais devem ser seus resultados de saída.

Estes mesmos valores servem como “supervisão” destas previsões, **permitindo o ajuste nas previsões com base nos erros**, ou seja, o modelo possui uma referência daquilo que está certo e daquilo que está errado.

KNN (K-Nearest Neighbors)

- Algoritmo do tipo supervisionado, tanto para tarefas de classificação como de regressão
- Usado para classificar objetos com base em exemplos de treinamento que estão mais próximos no espaço de características
- *'diga-me com quem tu andas que direi quem tu és'*



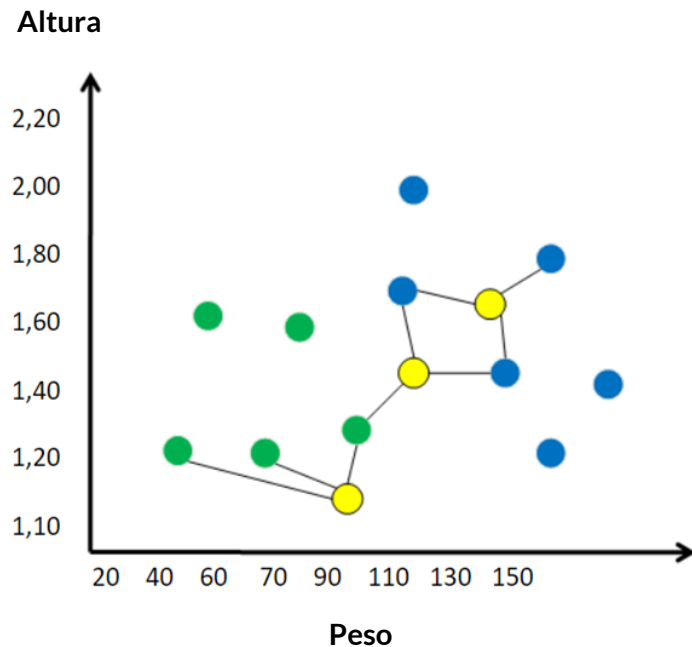
KNN (K-Nearest Neighbors)

Para utilizar o kNN, necessita-se de:

1. Um conjunto de exemplos de treinamento
2. Definir uma métrica para calcular a distância entre os exemplos de treinamento
3. Definir o valor de K (número de vizinhos próximos que serão considerados pelo algoritmo)

Para classificar um exemplo desconhecido:

1. Calcula-se a distância entre o exemplo desconhecido e os outros pontos de dados
2. Identifica-se aqueles mais próximos (distância euclidiana por exemplo)
3. Definir a classe do exemplo desconhecido pela votação majoritária de cada classe dos vizinhos mais próximos



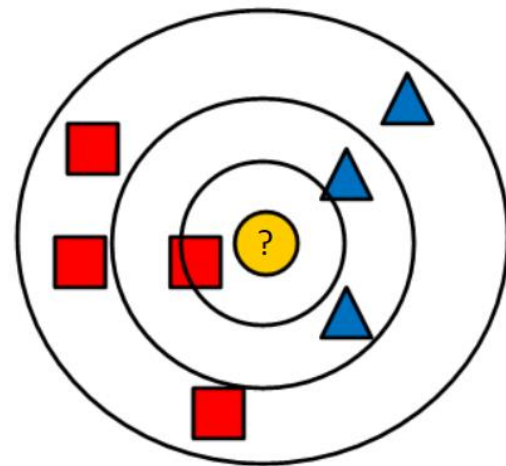
KNN (K-Nearest Neighbors)

Vantagens:

- Técnica simples e facilmente implementada, bastante flexível
- Funciona bem para muitas classes
- Apresenta resultados ótimos em alguns casos

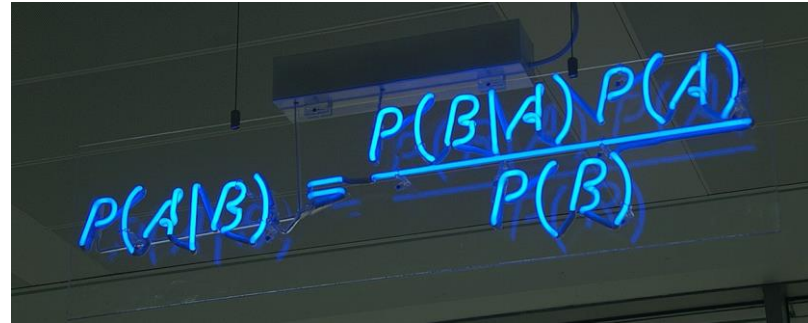
Desvantagens:

- Pode consumir muito tempo quando o conjunto de treinamento é muito grande, já que requer um cálculo de distância para cada exemplo de treinamento
- A precisão da classificação pode ser degradada pela presença de ruído
- é preciso escolher um número k ótimo para que os problemas sejam bem solucionados



Naive Bayes

- Algoritmo probabilístico do tipo supervisionado para tarefas de classificação
- Mostra como determinar a probabilidade de um evento condicional através da probabilidade inversa
- Assume que a presença ou ausência de um atributo não tem relação alguma com qualquer outro atributo (*naive*)

A photograph of a chalkboard with the formula for Bayes' Theorem written in blue chalk. The formula is $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$. The chalkboard is dark, and the blue chalk is clearly visible. The formula is written in a slightly slanted, handwritten style. The background is dark, and the lighting is focused on the chalkboard.
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Teorema de Bayes

Naive Bayes

Partindo do Teorema de Bayes:

$$P(classe|atributos) = \frac{P(classe) \cdot P(atributos|classe)}{P(atributos)}$$

Porém, como se assume que os atributos são independentes, conclui-se que:

$$P(classe|atributos) = \frac{P(classe) \cdot P(a_1|classe) \cdot ... \cdot P(a_n|classe)}{P(atributos)}$$

O algoritmo portanto calcula apenas o denominador para cada classe e o normaliza para que a soma seja 1 (evidência):

$$P(classe|atributos) = \frac{P(classe) \cdot P(a_1|classe) \cdot ... \cdot P(a_n|classe)}{evidência}$$

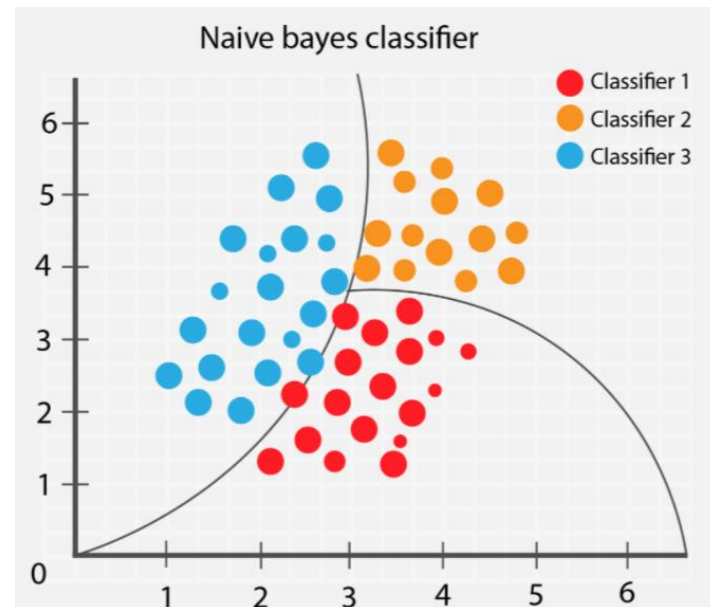
Naive Bayes

Vantagens:

- É um algoritmo super rápido, utilizado em previsões em tempo real
- Funciona bem para dados multidimensionais como textos, sendo muito utilizado na tarefa de classificar emails como spam ou não

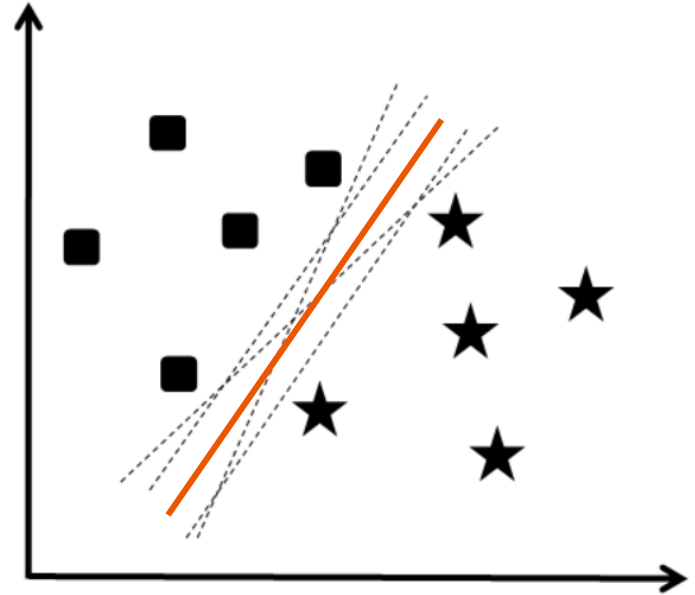
Desvantagens:

- Determinados problemas não podem ser bem modelados para este algoritmo, já que ele não considera a correlação entre as variáveis, algo bastante difícil de aparecer no mundo real



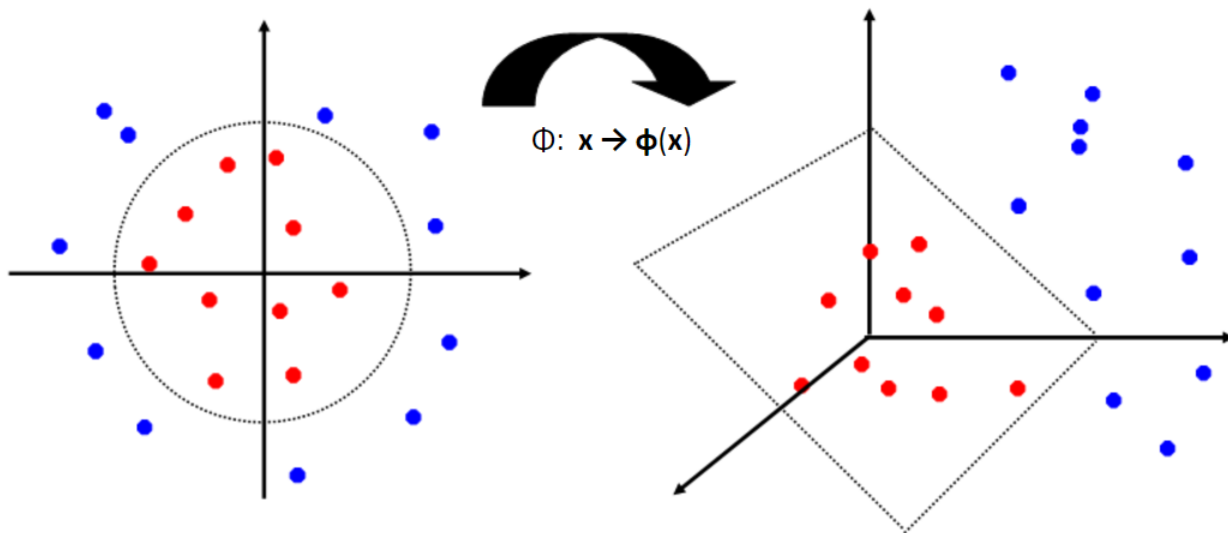
Support Vector Machines (SVM)

- Algoritmo do tipo supervisionado para tarefas de regressão e classificação, sendo mais utilizado para classificação
- Consiste em um método que busca encontrar um hiperplano com a maior margem que separa diferentes classes com menor erro de classificação



Support Vector Machines

- Em determinados problemas não é possível separar linearmente o conjunto de dados
- Para contornar isso, utiliza-se um 'truque' de mapear os dados para um espaço de dimensão maior através de uma função e distribuir esses pontos usando o critério de similaridade entre eles (kernel trick)
- Neste novo espaço, os dados são linearmente separáveis e assim o algoritmo atua para encontrar um hiperplano que melhor os classifica



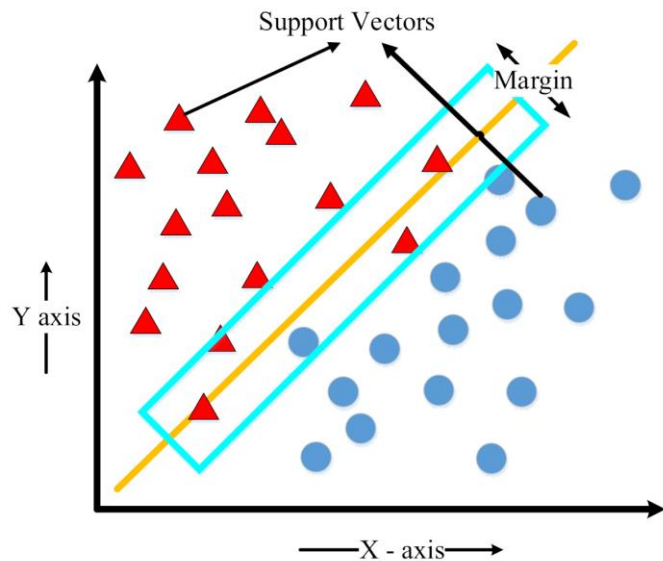
Support Vector Machines

Vantagens:

- Pode ser utilizado em situações em que o número de parâmetros é maior que o número de dados de exemplo para alimentar o modelo
- Consegue lidar bem com grandes conjuntos de exemplo, pois apenas os mais importantes são considerados

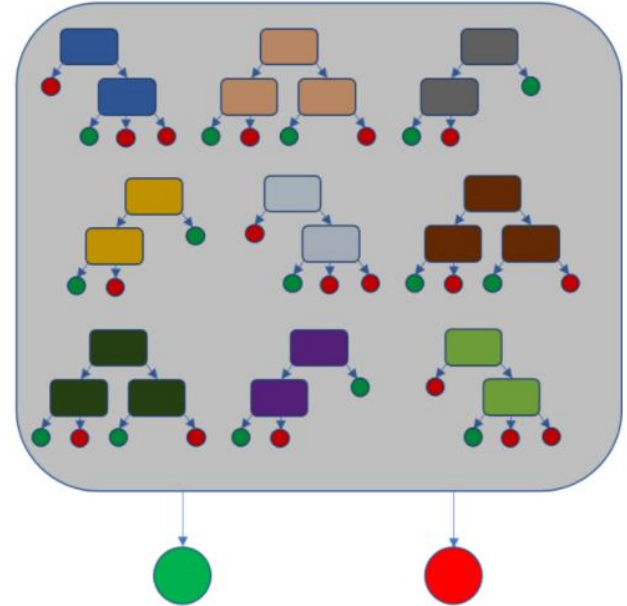
Desvantagens:

- É necessário definir um bom kernel para realizar a classificação de forma eficiente
- O algoritmo não é muito claro em como ele está classificando os dados, por conta disso é conhecido como 'caixa-preta'



Random Forest

- Técnica de machine learning utilizada tanto para tarefas de classificação como de regressão, combinando vários classificadores para solucionar problemas complexos
- É utilizada por **bancos** para determinar se uma pessoa terá seu empréstimo aceito ou não a depender da qualidade do seu crédito, na **área da saúde** para diagnosticar pacientes dado seu histórico clínico, em **e-commerces** para prever as preferências dos consumidores pelo seu comportamento histórico, entre outros.

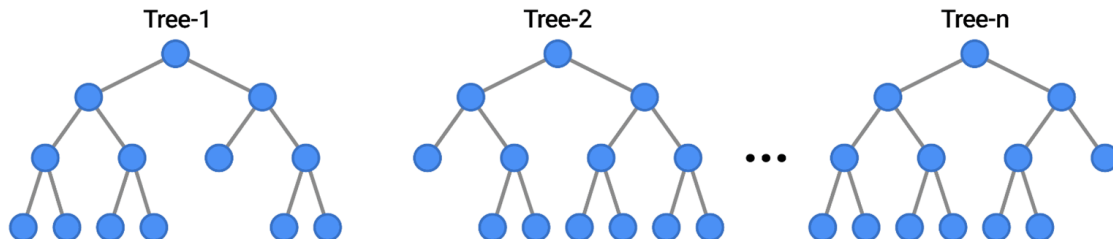


Random Forest

A técnica funciona da seguinte maneira:

1. O algoritmo seleciona amostras aleatórias do dataset
2. Então, cria-se uma árvore de decisão para cada amostra selecionada. Cada árvore terá uma predição para aqueles dados
3. Com essas predições, utiliza-se a técnica do voto majoritário para chegar na conclusão da floresta: para tarefas de classificação será a moda dos valores, enquanto que para tarefas de regressão será a média dos valores
4. O algoritmo portanto seleciona o resultado mais votado como predição final

EXAMPLES



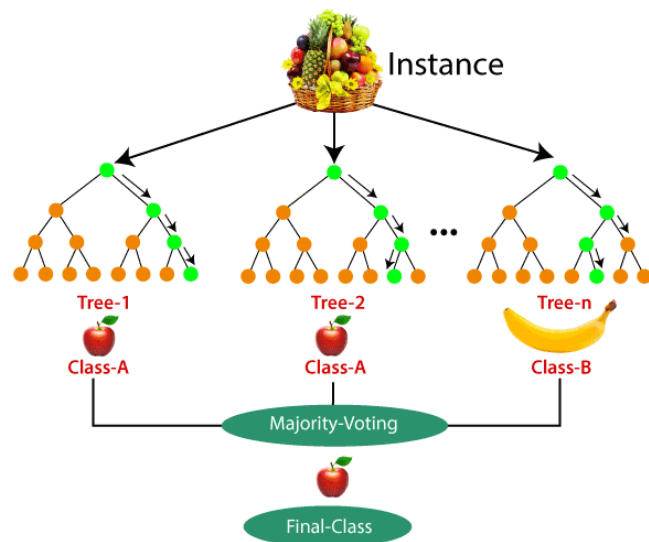
Random Forest

Vantagens:

- Não é necessário normalizar as variáveis como em outras técnicas de machine learning, é um modelo super robusto, performático mesmo com grandes datasets e com grande acurácia em diferentes tipos de problema
- A natureza aleatória de construção de cada árvore minimiza o sobreajuste
- Possibilita método efetivo de substituição de dados ausentes

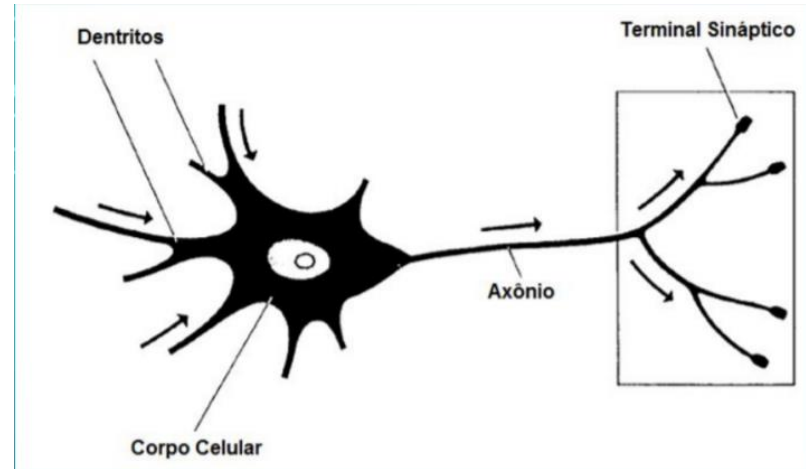
Desvantagens:

- Problemas que possuem dimensionalidade elevada como a classificação de textos não são bem trabalhados com as random forests
- Assim como o Support Vector Machines, é uma 'caixa-preta'



Redes Neurais Artificiais (RNA) - Neural Networks

- Tem habilidades de adquirir e armazenar conhecimento para realizar uma tarefa.
- Consistem em modelos inspirados no cérebro humano.
- O neurônio artificial imita o neurônio biológico.

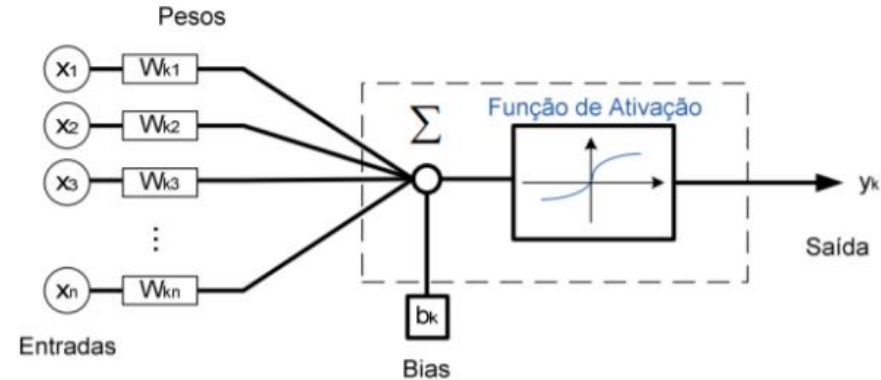


Neurônio Biológico

Redes Neurais Artificiais (RNA) - Neural Networks

São Capazes de:

- APRENDER a partir de amostras de treinamento.
- GENERALIZAR a partir do conhecimento adquirido.
- SE ADAPTAR ajustando-se a uma nova realidade.



Redes Neurais Artificiais (RNA) - Neural Networks

Vantagens:

- **extremamente simples**, uma vez que tenhamos entendido os modelos lineares
- **bastante intuitivas**, pois permitem a interpretação de aprendizado de níveis de abstrações hierárquicos
- são **muito flexíveis**, o que as torna ideais para resolver os mais diversos tipos de problemas

Desvantagens:

- Dificuldade de configuração das redes em relação à sua estrutura inicial e também no que se refere aos parâmetros dos algoritmos de aprendizado;
- Dificuldade de convergência (bloqueios) e instabilidade, inerentes aos algoritmos de otimização empregados;
- Lentidão do processo de aprendizado / adaptação.



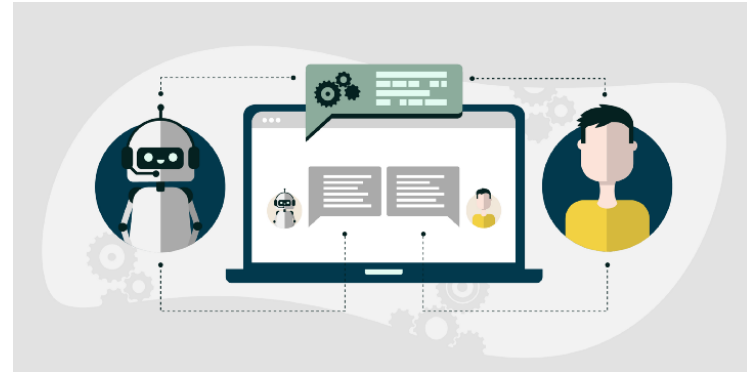
APRENDIZADO NÃO SUPERVISIONADO

O aprendizado supervisionado ocorre quando o **modelo aprende a partir de resultados pré-definidos**, utilizando os valores passados da variável alvo para aprender quais devem ser seus resultados de saída.

Estes mesmos valores servem como “supervisão” destas previsões, **permitindo o ajuste nas previsões com base nos erros**, ou seja, o modelo possui uma referência daquilo que está certo e daquilo que está errado.

Processamento de Linguagem Natural

- Técnica de Inteligência Artificial com objetivo de permitir que os computadores compreendam a fala humana e com isso possa realizar diferentes ações com essa informação, como tradução automática, análise de sentimento e pesquisa por voz
- Buscadores como o Google e o Bing utilizam essa técnica para prever o que o usuário está desejando procurar. Permite a criação de produtos como chatbots e assistentes virtuais



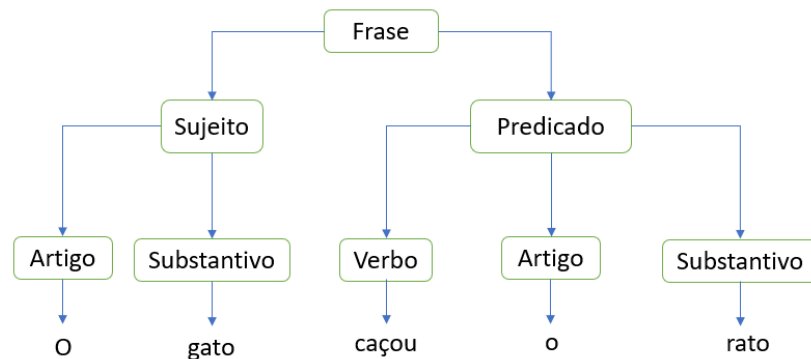
Processamento de Linguagem Natural

1. Encontra-se **sentenças únicas** dentro deste texto
2. Para cada sentença, obtém-se uma **lista de palavras** dessas sentenças (tokens)
3. Realiza-se a **classificação morfológica** das palavras. Para isso, utiliza-se um modelo pré-treinado
4. Realiza-se uma operação para identificar qual a **forma básica de determinado token**, o chamado lemma

Quanto	vai	me	custar	para	ligar	para	a	França
pronome	verbo	pronome	verbo	preposição	verbo	preposição	artigo	substantivo

Processamento de Linguagem Natural

5. Algumas palavras em português são utilizadas frequentemente para dar sentido ao texto, (e, como, o), logo, ao alimentar um modelo com determinado texto, faz sentido filtrar estas palavras (**stop words**)
6. Analisa-se como as palavras na sentença se relacionam entre si. O objetivo é construir uma árvore que atribua uma única palavra pai a cada palavra da frase, com isso a **raiz da árvore será o verbo principal da frase**
7. Utiliza-se as informações da árvore de análise de dependência para **agrupar palavras que falam sobre a mesma coisa**
8. Por fim, utiliza REM para **detectar e rotular substantivos das frases** com os conceitos do mundo real que eles representam (GATE, OpenNLP e SpaCy)



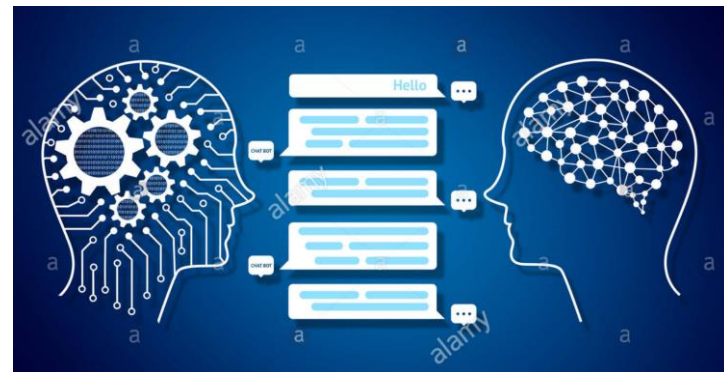
Processamento de Linguagem Natural

Vantagens:

- Permite que se realize em larga escala a análise de documentos, e-mails, dados de mídias sociais, pesquisa por voz, atendimento por chatbots, trazendo agilidade para as pessoas e redução de custos para as empresas

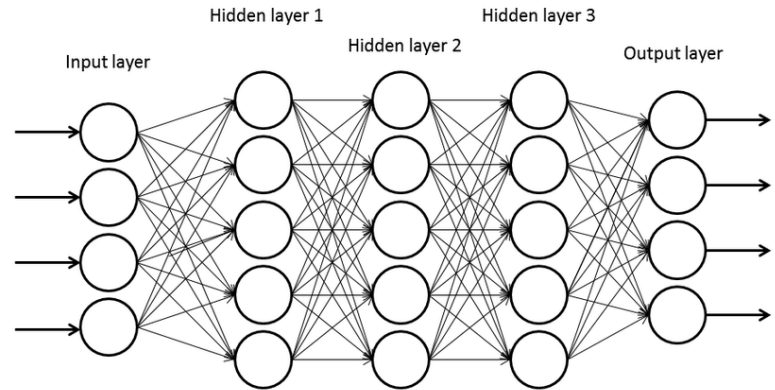
Desvantagens:

- Campo muito complexo, já que na linguagem humana características como ironia e sarcasmo não são identificadas por estes modelos de classificação de texto, o que poderia resultar em respostas completamente destoantes do conteúdo original



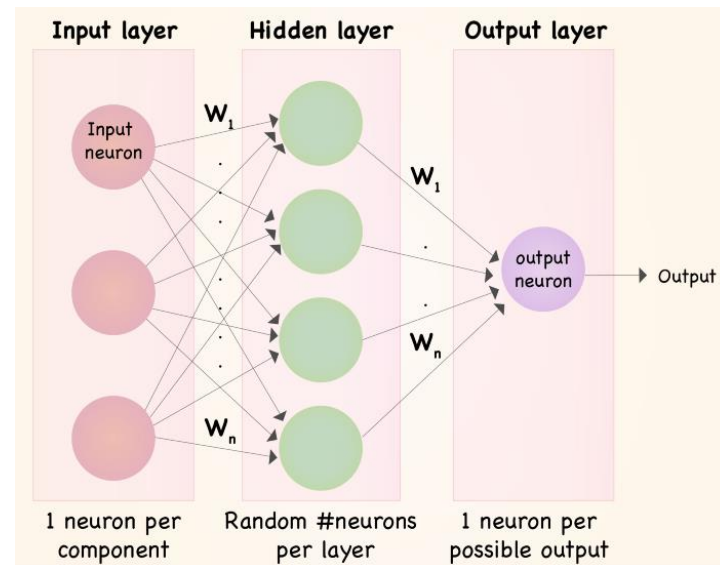
Redes Neurais Profundas

- Tipo de rede que usa camadas de neurônios matemáticos para processar dados e é utilizada pelos mais variados tipos de aplicações.
- A informação é passada através de cada camada, com a saída da camada anterior fornecendo entrada para a próxima camada.
- Cada camada é tipicamente um algoritmo simples e uniforme contendo um tipo de função de ativação



Redes Neurais Profundas

- Na estrutura de Redes Neurais Profundas, os padrões são introduzidos na rede pela camada de entrada e é comunicada a uma ou mais camadas ocultas presente na rede.
- As camadas ocultas recebem este nome somente por não constituírem a camada de entrada ou saída, são como camadas intermediárias. São nestas camadas que todo o processamento acontece por meio de um sistema de conexões dos chamados pesos e vieses:
- A entrada é recebida, o neurônio calcula uma soma ponderada adicionando também o viés e de acordo com o resultado e uma função de ativação predefinida, ele decide se deve ser 'disparado' ou ativado.
- Posteriormente, o neurônio transmite a informação para outros neurônios conectados em um processo chamado “forward pass”. Ao final desse processo, a última camada oculta é vinculada à camada de saída que possui um neurônio para cada saída possível desejada, tornando, assim, a previsão possível.



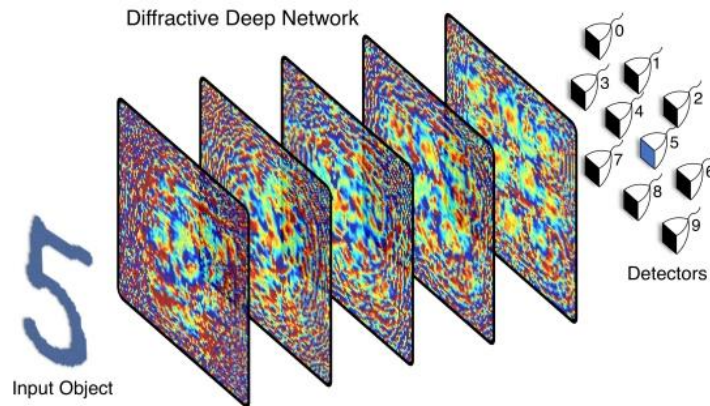
Redes Neurais Profundas

Vantagens:

- As Redes Neurais são capazes de processar e otimizar altos índices de dados, de forma eficiente e muito rápida.
- É capaz de se corrigir e melhorar, assim consegue alcançar índices de acerto super elevados, com taxas de erro inferiores a 10^{-5} . Isso significa ter um erro a cada 100.000 amostras.

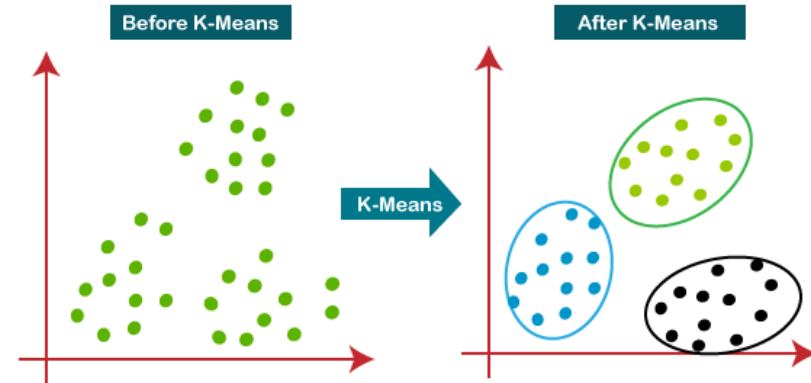
Desvantagens:

- A desvantagem da utilização de aprendizagem profunda ao invés de outras técnicas de Machine Learning, é que o aprendizado profundo é mais complexo, portanto, necessita de maior poder de processamento e quantidade de dados, além de precisar de um período maior de tempo para treinamento.



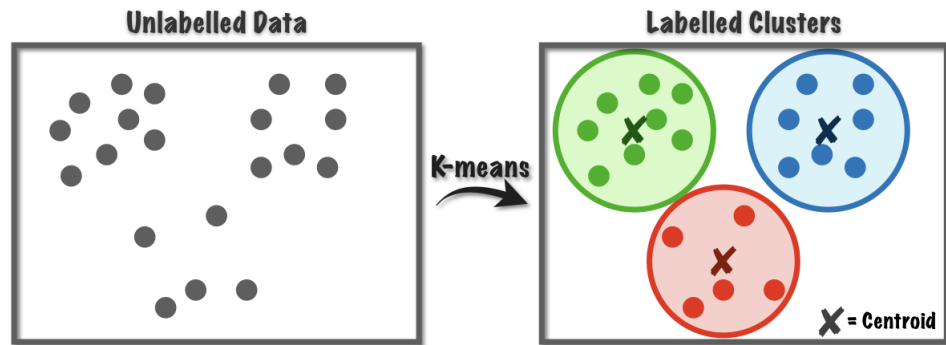
K-Means Clustering

- Um dos algoritmos de aprendizado de máquina não supervisionados mais simples e populares.
- O objetivo é agrupar dados similares e descobrir padrões desconhecidos.
- Para isso, o algoritmo busca um número fixo (chamado de k) de clusters (coleção de pontos de dados agregados devido a certas semelhanças) em um conjunto de dados.
- Em outras palavras, o algoritmo K-means identifica o número k de centróides e, em seguida, aloca cada ponto de dados para o cluster mais próximo, mantendo os centróides tão pequenos quanto possível.



K-Means Clustering

- Para realizar o processo de agrupamento dos dados, o algoritmo de K-Means Clustering começa com um primeiro grupo de centróides selecionados de forma aleatória, que são usados como pontos iniciais para cada cluster.
- Em seguida, o algoritmo realiza cálculos iterativos, buscando sempre otimizar as posições dos centróides.
- A execução é interrompida quando ou os centróides se estabilizam (o agrupamento é bem sucedido) ou quando o número definido de iterações foi alcançado.



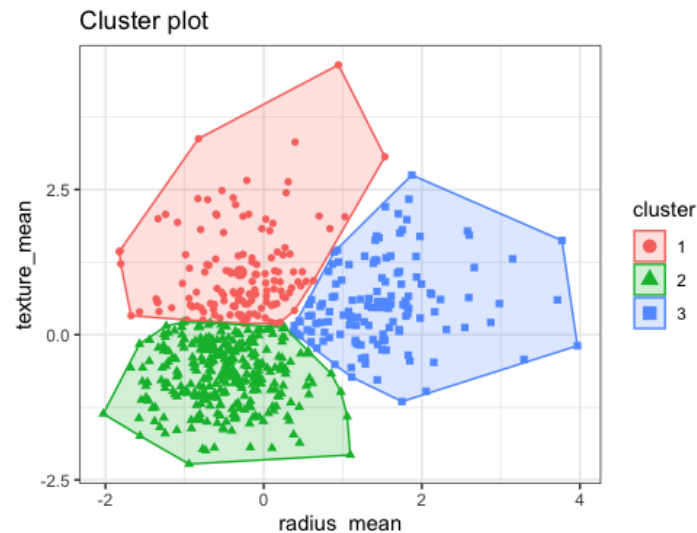
K-Means Clustering

Vantagens:

- Simplicidade de implementá-lo
- Escala obtida para grandes conjuntos de dados, já que o algoritmo consegue boa performance.
- Convergência da solução, pois o algoritmo pode inicializar de forma aleatória as posições dos centróides.
- Facilmente de adaptação a novos exemplos.

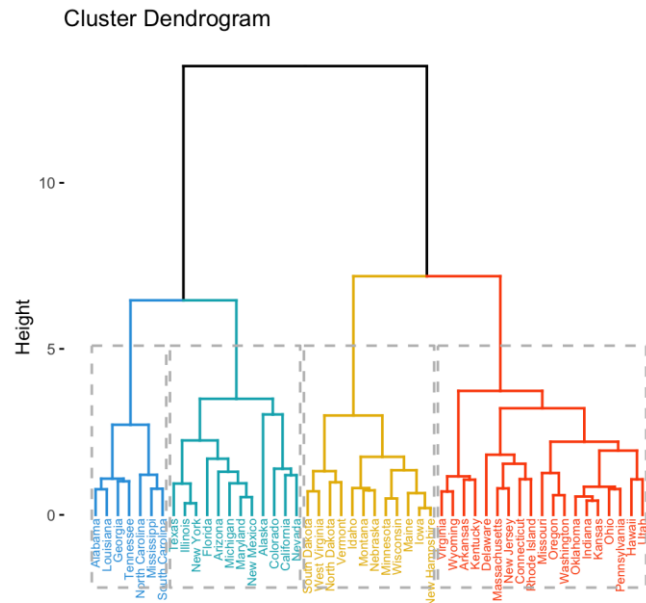
Desvantagens:

- Ocorrência de outliers nos clusters, portanto é preciso ter um olhar especialista sobre os grupos formados.
- o algoritmo pode acabar agrupando dados de densidades diferentes, pois não leva isso em consideração no momento de criar os cluster.
- É preciso setar um número k inicial.



Agglomerative Clustering

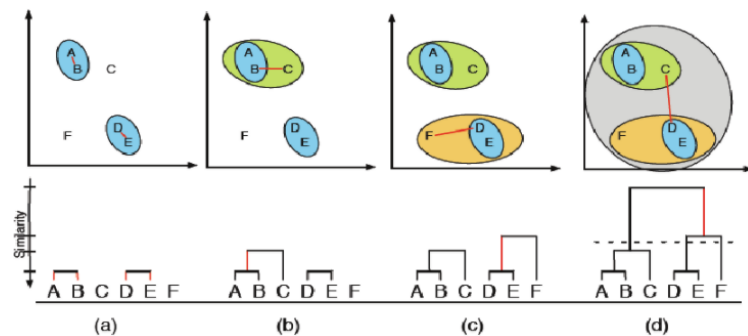
- É um método de análise de cluster que busca construir uma hierarquia de clusters.
- O Agrupamento hierárquico é o tipo mais comum de clustering hierárquico usado para agrupar objetos em clusters com base em sua similaridade.
- É também conhecido como AGNES (Agglomerative Nesting).



Agglomerative Clustering

- O agrupamento hierárquico começa com N grupos, cada um contendo inicialmente uma entidade e, em seguida, os dois grupos mais semelhantes se fundem em cada estágio até que haja um único grupo contendo todos os dados.
- Uma heurística típica para N grande é executar primeiro k -means e depois aplicar o agrupamento hierárquico aos centros do cluster estimados.
- Uma árvore binária chamada dendrograma representará o processo de fusão.
- Os grupos iniciais (objetos) estão nas folhas (na parte inferior da figura), e nós os juntamos na árvore cada vez que dois grupos são fundidos.
- A raiz da árvore (que está no topo) é uma categoria com todos os dados.
- Produzimos um agrupamento de determinado tamanho se cortarmos a árvore em qualquer altura.

Example: Hierarchical Agglomerative Clustering



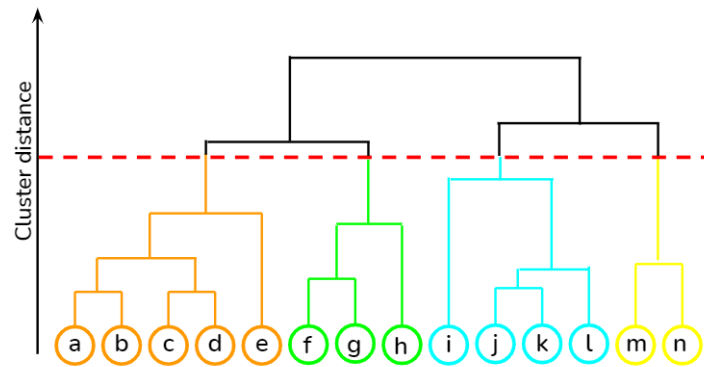
Agglomerative Clustering

Vantagens:

- O agrupamento hierárquico produz uma hierarquia, ou seja, uma estrutura que é mais informativa do que o conjunto não estruturado de aglomerados planos retornado por k-means.
- Portanto, é mais fácil decidir sobre o número de clusters, olhando para o dendrograma.
- Fácil de implementar

Desvantagens:

- Não é possível desfazer a etapa anterior: uma vez que as instâncias foram atribuídas a um cluster, eles não podem mais ser movidos.
- Complexidade de tempo: não adequado para grandes conjuntos de dados.
- As sementes iniciais têm um forte impacto nos resultados.
- A ordem dos dados tem impacto nos resultados.
- Muito sensível a outliers.



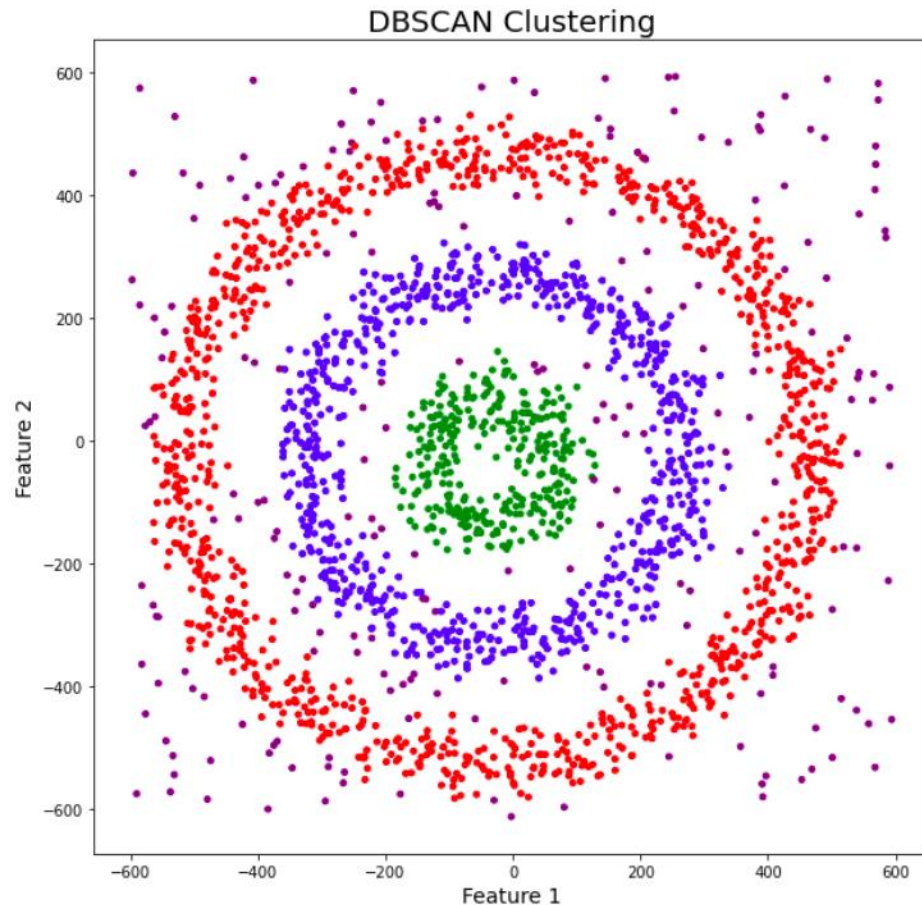
DBSCAN

- É um algoritmo de clustering (classificação) baseado em densidade, que pode ser usado para identificar clusters de qualquer forma em um conjunto de dados contendo ruídos e outliers.
- Funciona para cada ponto de um cluster, a vizinhança de um determinado raio deve conter pelo menos um número mínimo de pontos.
- Não é necessário definir o número de clusters.
- DBSCAN funciona melhor para problemas de classificação, principalmente quando se tem ruído nos dados, é ótimo para dados que contém classes com densidades parecidas.



DBSCAN

- É definida uma distância (raio) dos pontos no espaço e define o cluster baseado nessa distância.
- Caso não tenha nenhum ponto próximo a um conjunto de pontos próximos, é categorizado como um novo cluster.
- O próprio algoritmo define o número de clusters de acordo com a quantidade de registros.
- Dependendo da forma como esse algoritmo é iniciado pode se obter clusters diferentes.



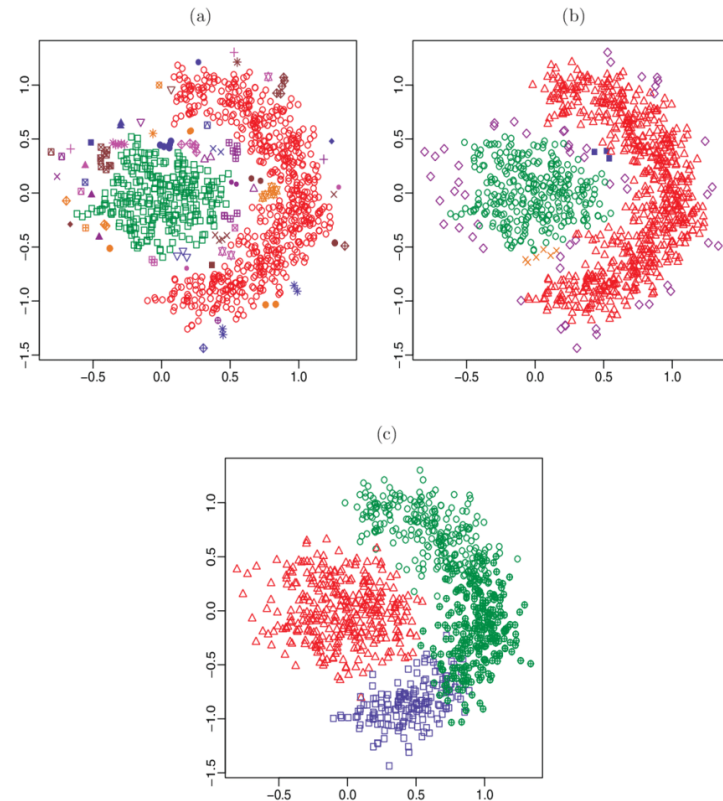
DBSCAN

Vantagens:

- Encontra padrões não lineares.
- Robusto contra outliers.
- Resultado pode ser mais consistente que o k-means pois a inicialização dos centróides não afeta tanto o algoritmo.

Desvantagens:

- Dependendo da inicialização, um ponto pode pertencer ao cluster diferente.
- Difícil encontrar um bom valor para o parâmetro da distância.



Principais Ferramentas e Bibliotecas

1. Jupiter Notebook
2. NumPy
3. Pandas
4. SciPy
5. Keras
6. PyTorch

Problemas Propostos

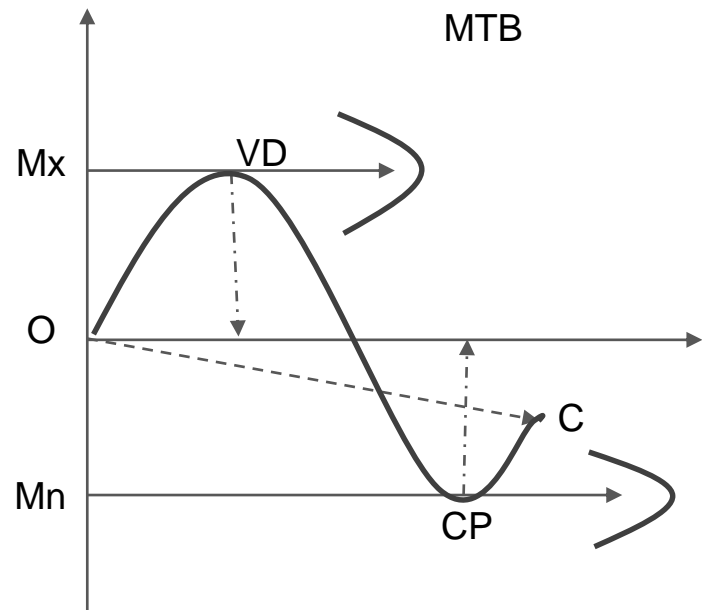
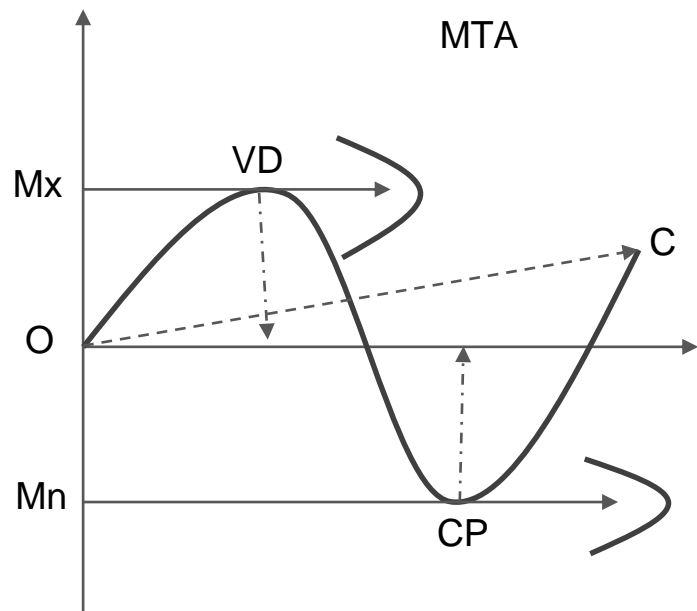
1. Trade de Abertura

1. Estatístico
2. E relação com as bolsas de outros países, naquele momento de abertura. Podendo gerar algum modelo de ML

2. Identificar e acompanhar a movimentação do Big Player

1. Baseado em fortes movimentações do volume e do preço

Trade de Abertura



Curso Nivelamento Trader

1. <https://scalpertrader.com.br/>
2. www.b3.com.br
3. http://www.b3.com.br/pt_br/b3/educacao/cursos/
4. Ferramentas de Noticias e Gráficas
 - a) <https://br.tradingview.com/>
 - b) www.infomoney.com.br