1

Machine Learning - Aprendizado de máquina.

É uma área da inteligência artificial que treina um algoritmo para solucionar problemas complexos.

Features modelos que eu texto e treino

2 Como trabalha o maching learning

ETL, Analise exploratória, definição do modelo, treino, teste, predição e decisão;

3 Etapas de desenvolvimento

4

Um algoritmo é uma sequência lógica, finita e definida de instruções que devem ser seguidas para

resolver um problema ou executar uma tarefa.

Não existe soluções simples para problemas complexos.

5

Necessidade do negócio, dados, modelos, treinamento, testes, implementação, conecxão e monitoramento

6

Os algoritmos normalmente não são desenvolvidos pelo analista de dados.

Para solucionar um problema, utilizamos mais de um algoritmo pronto, treinamos e testamos.

7 Algoritmo é milemar

O Ábaco já utilizava algoritmos(procedimentos para obter um resultado específico).

9

Arthur Lee Samuel pioneiro em maching learning criou esse termo na década de 1950.

10

Teste de Turing – Foi criado por Alan Turing em seu artigo de 1950.

“Computing Machinery and Intelligence”

O teste inicia com as palavras: Eu proponho considerar a questão

“As máquinas podem pensar?”

O objetivo do teste é determinar se uma máquina é capaz de exibir comportamento inteligente

Indistinguível do ser humano.

11

Obter aprendizado sem ser programado especificamente para ele, Arthur Samuel.

Linguagem Natural - Processamento de [**Linguagem Natural (PLN)** é uma área que combina **linguística computacional com algoritmos de aprendizado de máquina**](https://www.bing.com/ck/a?!&&p=fa6d4047bf65613ac9d81d0fbb87cc3a391bf2fc5cc9795b7b33b9f9a2f353c6JmltdHM9MTczMzE4NDAwMA&ptn=3&ver=2&hsh=4&fclid=3d77ace7-f03f-66c3-17c9-b820f146671e&psq=linguagem+natural+algoritmo&u=a1aHR0cHM6Ly93d3cuaWJtLmNvbS9ici1wdC90b3BpY3MvbmF0dXJhbC1sYW5ndWFnZS1wcm9jZXNzaW5n&ntb=1) . Utiliza quatro principais tipos de modelo:

12

Modelo matemático: Nós acreditamos que o modelo tem uma porção de if´s.

A matemática e muito utilizada em modelos de dados,

Especificação de um conjunto de regras de relação matemática entre variáveis.

13 Fórmulas matemáticas, quantitavivas e qualificativas e estatidticas

14 algebra linear, vetores, matrizes e determinantes

15 Diversos tipos de calculos, calculo provabilidade

16 conseitos de estatistica, amostra, experimento, população finita e infinita,estatistica descritiva.

17 tipos de maching learning, supervisionado, Não supervisionad e de reforço

18 Não supervisionado

Agentes inteligentes,perceptions e ambiente.

Perception

18 tecnicas de machine learning

Supervisionado Regreção, classificação

Não supervisionado Agrupamento

19 Representação gráfica

Classificação(Tem um rotulo)

Regressão

Clusterização(Não tenho rotulo)

20 Classificação

Método de Classificação – Encontrar uma classe de um item.

Dados com Label, rotulados e bem estabelecidos.

21 Tipos de regreção

Método de Regressão – Prever uma informação em determinado tempo.

Predição – Os dados são rotulados

22 Modelo supervisionado

Arvore de decisão

Random Florest

23 Modelo não supervisionado –

23ª Reforço, Tentativa e erro, recompensa posivitva, negativa

24 Clusterização

Clusterização ex. separar em várias categorias sem rotular(grupos).

25 Temos também o modelo Semi supervisionado

27 modelo gerador de recomendação

28 modelo de reforço

Modelo de Reforço

Agente\Ambiente

29 Alguns Algoritmos utilizados

KNN – Baseado em instancias

30

Detecção de anomalias out layers - Valores muito fora da curva

31

Treinamento e teste do modelo(Dados de teste e dados de treino).

Underfitting - Sub ajuste algoritmo muito fraco no treinamento e no teste)

Overfitting – Super ajuste (resultado bons na base de treinamento, ruins na do teste).

32

Acurácia de precisão:

Entropia, desordem, percentual de acerto

33 Metodos de coleta de dados

34 Principais Perguntas utilizadas na coleta de dados

35 ETL

36 Exploração de dados

Mitigar dados inválidos

36 Tipos de variáveis(Estatísticas)

38 Modelagem de Dados

39 Dicionário de dados

PowerBI – Temos curso, link nos comentarios

JuPyter – Jlia, Python e R

Download Anaconda

Sem instalação:

[Project Jupyter | Try Jupyter](https://jupyter.org/try)

Grafana, Dashboard de Monitoramento

Download grafana

Localhost:3000

[A **Inteligência Artificial Supervisionada** é um ramo da inteligência artificial que utiliza algoritmos para aprender a partir de dados rotulados](https://www.bing.com/ck/a?!&&p=285064b27728ba4f464becf9f71a7a06c9a814ab081f227f016e9bba04e22aceJmltdHM9MTczMzE4NDAwMA&ptn=3&ver=2&hsh=4&fclid=3d77ace7-f03f-66c3-17c9-b820f146671e&u=a1aHR0cHM6Ly90ZWNub2xvZ2lhLmNyZWRpdGVkLmNvbS5ici9nbG9zc2FyaW8vby1xdWUtZS1pYS1zdXBlcnZpc2lvbmFkYS1lbnRlbmRhLW8tY29uY2VpdG8v&ntb=1).

Dados rotulados.

Fatos e dados qualitativos e quantitativos.

Classificação

Regressão

[O **aprendizado não supervisionado** é uma técnica de **aprendizado de máquina** em que o algoritmo aprende a partir de **dados sem que haja um rótulo ou resposta definida**](https://www.bing.com/ck/a?!&&p=6505d56eabe91c504d97b5999421eb0d4ea1b6d2574cdfb9d6ef0ad27743b8ebJmltdHM9MTczMzE4NDAwMA&ptn=3&ver=2&hsh=4&fclid=3d77ace7-f03f-66c3-17c9-b820f146671e&u=a1aHR0cHM6Ly93d3cuaWJtLmNvbS9ici1wdC90b3BpY3MvdW5zdXBlcnZpc2VkLWxlYXJuaW5n&ntb=1).

Agrupamento de dados. Identificar padrões ocultos.

Associação utiliza padrões frequentes.

Classificação, regressão, clusterização e predição.

Algoritmo preditivo.

A Aprendizagem Por Reforço é o treinamento de modelos de aprendizado de máquina para tomar uma sequência de decisões. O agente aprende a atingir uma meta em um ambiente incerto e potencialmente complexo.

Aprendizado de reforço. Recompensa positiva, Recompensa negativa(Castigo, Penalidade).

Tentativa e erro:

Equação de Bellman, MDP Markov Decision Process, Q-Learning e Diferença temporal.

Semi supervisionado.

A IA generativa, às vezes chamada de IA gen, é a [inteligência artificial](https://www.ibm.com/br-pt/topics/artificial-intelligence) (IA) que pode criar conteúdo original, como texto, imagens, vídeo, áudio ou código de software, em resposta a um prompt ou solicitação do usuário.

Dados de entrada, mapeamento de dados(informações,regras),tomada de decisão.

Estruturar categorizar dados transformar em informações.

# Kasparov vs. Deep Blue 1996.

Veículo autônomo :[**Um veículo autônomo é um veículo terrestre que pode transportar pessoas ou bens sem a utilização de um condutor human**](https://www.bing.com/ck/a?!&&p=64a67c7653b1a5b6ecae0f6d9fe597ec514c676e476f0f878217fb17f24cacf8JmltdHM9MTczMzI3MDQwMA&ptn=3&ver=2&hsh=4&fclid=3d77ace7-f03f-66c3-17c9-b820f146671e&u=a1aHR0cHM6Ly9wdC53aWtpcGVkaWEub3JnL3dpa2kvVmUlQzMlQURjdWxvX2F1dCVDMyVCNG5vbW8&ntb=1)**o.**

**Robos humanóides,** são construídos para imitar o comportamento e corpo humano.

ETL, extração, transformação e load

Web Scfaping

[A análise exploratória de dados é uma abordagem usada para **analisar e investigar dados**](https://www.bing.com/ck/a?!&&p=d179d0b00d68f21ff7f1a99a2823c2eb5546987bfd5bd0b5ffd128a78e42f3daJmltdHM9MTczMzQ0MzIwMA&ptn=3&ver=2&hsh=4&fclid=3d77ace7-f03f-66c3-17c9-b820f146671e&psq=an%c3%a3lise+exploratoria&u=a1aHR0cHM6Ly93d3cuYWx1cmEuY29tLmJyL2FydGlnb3MvYW5hbGlzZS1leHBsb3JhdG9yaWE&ntb=1). [Ela permite que o cientista de dados tenha uma visão panorâmica dos dados e extraia conhecimento a partir deles](https://www.bing.com/ck/a?!&&p=d179d0b00d68f21ff7f1a99a2823c2eb5546987bfd5bd0b5ffd128a78e42f3daJmltdHM9MTczMzQ0MzIwMA&ptn=3&ver=2&hsh=4&fclid=3d77ace7-f03f-66c3-17c9-b820f146671e&psq=an%c3%a3lise+exploratoria&u=a1aHR0cHM6Ly93d3cuYWx1cmEuY29tLmJyL2FydGlnb3MvYW5hbGlzZS1leHBsb3JhdG9yaWE&ntb=1)

* Coletar os dados relevantes para a análise.
* Limpar os dados, removendo valores ausentes, tratando outliers e corrigindo possíveis erros.
* Explorar as variáveis.
* Identificar outliers e tratar valores extremos.
* Realizar análise de correlação.

Netflix, google

Chatbots

Regressão Linear.

[O que é regressão linear? – Explicação sobre o modelo de …](https://www.bing.com/ck/a?!&&p=65a122f1a7a4eabe84e5973840196251ee9411ee070cae3783b63131774670d4JmltdHM9MTczMzE4NDAwMA&ptn=3&ver=2&hsh=4&fclid=3d77ace7-f03f-66c3-17c9-b820f146671e&psq=regress%c3%a3o+linear&u=a1aHR0cHM6Ly9hd3MuYW1hem9uLmNvbS9wdC93aGF0LWlzL2xpbmVhci1yZWdyZXNzaW9uLw&ntb=1)

A regressão linear é uma técnica de análise de dados que prevê o valor de dados desconhecidos usando outro valor de dados relacionado e conhecido.

Arvore de decisão.

Python é uma das linguagens mais utilizadas em maching learning.

As bibliotecas mais utilizadas em Python são:

Scikit-learn – Algoritmos e ferramentas para Maching Learning

Pandas - Leitura e manipulação de dados

Matplotlib – Visualização de dados, graficos

Seaborn – Visualização de dados, graficos

Jupyter Notebooks – Ambiente para desenvolvimento(consorcio entre Julia, Python e R).

Anaconda - gerenciador de pacotes, para evitar conflitos de dependências.

Dados podem ser imagens, textos, vídeos, sons e outros.

Redes neurais.

Arquitetura de redes neurais é a tentativa de imitar o cérebro humano, mas não tem neurônios.

Utilizam várias fórmulas matemáticas que direcionam a um resultado.

Deep Learning(utiliza arquitetura de redes neurais, simulando o comportamento do cérebro humano)

Escolha do modelo.

Variáveis: Objetivo, tempo, custo, indicadores, critérios, métricas, percentual de erro,

Computação Quântica,

[O desenvolvimento "no-code" é uma abordagem que permite que usuários criem aplicativos, sites e workflows automatizados usando ferramentas visuais de arrastar e soltar, sem a necessidade de codificação](https://www.bing.com/ck/a?!&&p=64e16641b68f231bf2ed73cf8bfb3f698c3f2cc883b179c21713c78330240fa3JmltdHM9MTczMzE4NDAwMA&ptn=3&ver=2&hsh=4&fclid=3d77ace7-f03f-66c3-17c9-b820f146671e&u=a1aHR0cHM6Ly93d3cuc2FwLmNvbS9icmF6aWwvcHJvZHVjdHMvdGVjaG5vbG9neS1wbGF0Zm9ybS9idWlsZC9ndWlkZS10by1uby1jb2RlLWRldmVsb3BtZW50Lmh0bWw&ntb=1)

Melhores ferramentas de No-code:

FlutterFlow

.bubble

Pensamento algorítmico, capacidade de efetuar cálculos e conhecer sobre matemática e vetores, saber agregar valor.

Reconhecimento facial, reconhecimento de voz, reconhecimento de imagens.