

UNIDADE 0 - NIVELAMENTO

Marta Noronha

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
CURSO: CIÊNCIA DE DADOS
APRENDIZADO DE MÁQUINA II



- 0.1 - Conceitos básicos do Aprendizado de Máquina
- 0.2 - Definição de tipos de problemas e tipos de aprendizado
- 0.3 - Aplicações

Conceitos

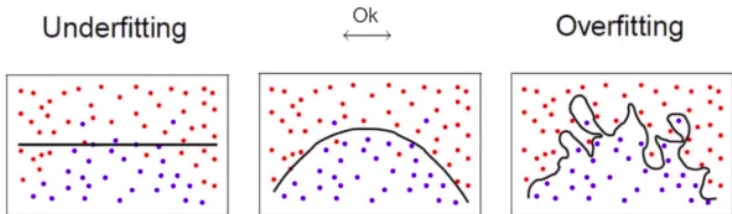
- Dados de Treinamento:
 - Conjunto de dados usados para treinar o modelo.
 - Dados devem ter alta qualidade e possuírem rótulos de qualidade.
 - Exemplos: Treinamento para reconhecimento de imagens como cães e gatos (Dataset disponível no Kaggle).
- Validação Cruzada:
 - Técnica para avaliar a capacidade de generalização de um modelo, a qual reduz o *overfitting* e fornece uma estimativa mais precisa da performance do modelo.
 - Exemplos: K-fold, Leave-One-Out, entre outros.
- Dados de Teste:
 - Conjunto de dados usado para avaliar o modelo após o treinamento.
 - É necessário manter os dados de teste separados dos dados de treinamento para evitar o *overfitting*.

Conceitos

- Modelo de Aprendizado de Máquina:
 - Criado por algoritmos que aprendem padrões a partir dos dados.
 - Exemplos: Redes neurais, árvores de decisão, regressão linear, k-means.
- Características (Atributos):
 - Variáveis independentes usadas para prever a saída.
 - Exemplos: Peso e altura para prever a massa corporal.
- Instâncias (Objetos ou Registros):
 - Cada linha de dados que representa um exemplo.
 - Respondentes de questionários como, por exemplo, para o CENSO.
- Rótulos (Classes):
 - A variável dependente ou a saída que queremos prever.
 - Exemplo: Classificação de câncer em pacientes como benigno ou maligno (Dataset disponível no Kaggle).

Conceitos

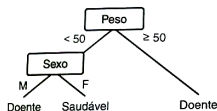
- Indução de Hipóteses
 - Processo de criar uma função a partir dos dados de treinamento.
 - Fundamental para a criação de modelos preditivos eficientes.
 - A indução de hipóteses contribui para a precisão do modelo.
 - Exemplos: Se tivermos dados de clima e vendas de sorvete, a hipótese pode ser "Vendas de sorvete aumentam em dias quentes".
- *Underfitting vs. Overfitting*



Fonte: <https://www.datascienceacademy.com.br/>

Conceitos: Viés indutivo (representação e busca)

- Viés de busca: Suposições feitas pelo modelo para generalizar a partir dos dados de treinamento para dados não vistos.
- O aprendizado prefere modelos mais simples (com menos parâmetros) porque são menos propensos ao *overfitting*.
- Modelos complexos devem ser usados apenas se melhorarem significativamente a performance.
- Viés de representação: linguagem ou representação usada pode restringir os modelos que podem ser encontrados.
- Cada viés pode afetar de forma diferente na performance do modelo, devendo ser controlado para evitar *underfitting* ou *overfitting*.



(a) [Árvore de decisão]

Se $\text{Peso} \geq 50$ então Doente
 Se $\text{Peso} < 50$ e $\text{Sexo} = M$ então Doente
 Se $\text{Peso} < 50$ e $\text{Sexo} = F$ então Saudável

(b) [Conjunto de regras]

$$\begin{pmatrix} 0,45 & -0,40 & 0,54 & 0,12 & 0,98 & 0,37 \\ -0,45 & 0,11 & 0,91 & 0,34 & -0,20 & 0,83 \\ -0,29 & 0,32 & -0,25 & -0,51 & 0,41 & 0,70 \end{pmatrix}$$

(c) [Redes neurais]

FIGURA 1.3 Diferentes vieses de representação.

FONTE: Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina - 2ª ed.

Conceitos: Viés induzido pelo conjunto de dados

- "Pesquisadores descobriram vieses sistemáticos em modelos comerciais de aprendizado de máquina usados para **detecção e rastreamento de rostos, detecção de atributos, justiça criminal, detecção de comentários tóxicos e outras aplicações.**" [Mitchell et al. 2019]

Conceitos: Viés induzido pelo conjunto de dados

- Estudo feito em 2019 pelo Instituto Nacional de Padrões e Tecnologia (“NIST”).
- Análise de desempenho de 189 algoritmos de reconhecimento facial para o reconhecimento de raças.
- 99 desenvolvedores da Microsoft, Intel, Idemia e outras grandes empresas de tecnologia e vigilância.
 - Propensão de 10 a 100 vezes de classificar incorretamente um rosto negro ou do leste asiático como branco por muitos algoritmos.
 - A maioria dos algoritmos identificou mulheres negras com muito menos precisão do que pessoas de outros grupos.

Fonte: Why Racial Bias is Prevalent in Facial Recognition Technology?

Conceitos: Viés induzido pelo conjunto de dados

- A raça influencia o desenvolvimento e o desempenho da tecnologia de reconhecimento facial de três maneiras distintas:
 - Imagens de treinamento não diversas: o preconceito humano e a disponibilidade de dados levam à predominância de tons de pele mais claros.
 - Conjunto *Labeled Faces in the Wild*: composto por 83,5% de imagens de pessoas brancas.
 - Conjunto IJB-A: criado pelo NIST para diversidade geográfica. Apresenta 79,6% de imagens de rostos com tons de pele mais claros.
 - Seleção humana de características faciais em algoritmos mais antigos.
 - Problemas de qualidade de imagem que afetam desproporcionalmente os tons de pele mais escuros.
- Coletivamente, esses fatores resultam em um desempenho desigual da tecnologia, geralmente pior para tons de pele mais escuros.

Fonte: Why Racial Bias is Prevalent in Facial Recognition Technology?

Conceitos: Viés induzido pelo conjunto de dados

There Is a Racial Divide in Speech-Recognition Systems, Researchers Say

Technology from Amazon, Apple, Google, IBM and Microsoft misidentified 35 percent of words from people who were black. White people fared much better.



Share full article



Amazon's Echo device is one of many similar gadgets on the market. Researchers say there is a racial divide in the usefulness of speech recognition systems. Grant Hindsley for The New York Times

Fonte: There Is a Racial Divide in Speech-Recognition Systems, Researchers Say

Conceitos: Viés induzido pelo conjunto de dados

Fairness, Accountability, and Transparency in Machine Learning (FAT/ML)

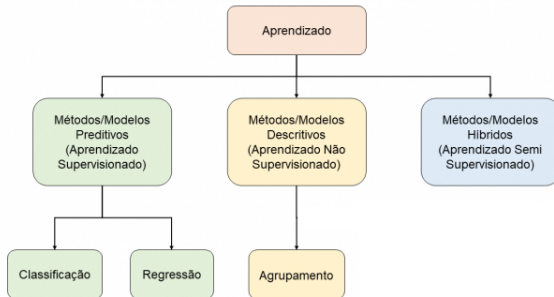
- Foco em ética e justiça:
 - Organização que reúne pesquisadores, profissionais e acadêmicos dedicados a promover justiça, transparência e responsabilidade.
 - Atua na mitigação de vieses e na promoção de tecnologias mais justas.
- Contribuições para a área:
 - Referência em discussões e publicações sobre práticas responsáveis em aprendizado de máquina.
 - Influencia políticas e melhores práticas por meio de conferências e eventos.
- Reconhecimento pela comunidade:
 - Amplamente reconhecida e apoiada pela comunidade acadêmica e industrial.
 - Reforça sua credibilidade e o valor das contribuições éticas na área.

Fonte: Fairness, Accountability, and Transparency in Machine Learning

Conceitos: Viés induzido pelo conjunto de dados

- Razões para a ocorrência do viés:
 - Sub-representação: A sub-representação de grupos ou características criam um modelo com desempenho inferior para aqueles sub-representados.
 - Coleta e amostragem de dados: Métodos de coleta que introduzem distorções (por exemplo, amostras não aleatórias ou enviesadas) podem não refletir a realidade por meio do conjunto de dados.
 - Coleta de características inadequadas: Aprendizado inadequado ou simplificado demais por não capturar informações importantes para a resolução do problema.
- Consequências:
 - Generalização limitada: O modelo tem bom desempenho na classificação de dados de treino, mas falham ao generalizar para dados não vistos que não foram bem representados por amostras do conjunto de treino.
 - Resultados injustos: Pode levar a resultados discriminatórios ou injustos, especialmente em aplicações sensíveis (como em sistemas de recomendação, crédito ou decisões judiciais).

Tipos de aprendizado



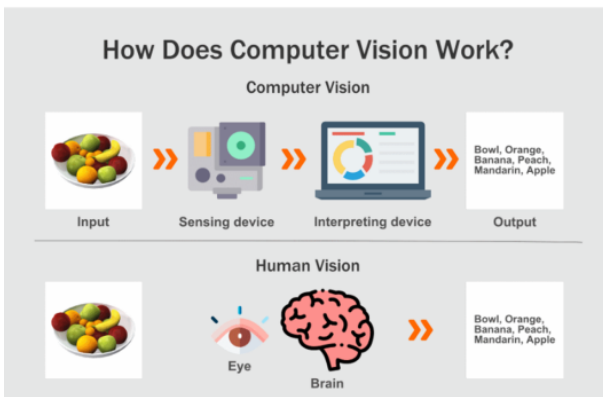
Fonte: Adaptado de
<https://embarcados.com.br/classificacao-multirrotulo-hierarquica-intro/>

Tipos de problemas e tipos de aprendizado

	Supervisionado	Não Supervisionado	Reforço
Dados de Treino	Rotulados	Não rotulados	Recompensas e punições com base em ações
Objetivo	Prever saídas baseadas em entradas	Encontrar padrões ocultos ou agrupamentos nos dados	Maximizar a recompensa total através de ações sequenciais
Exemplos	Classificação de e-mails, previsão de vendas	Segmentação de clientes, detecção de anomalias	Jogos de tabuleiro, controle de robôs, veículos autônomos
Modelos Comuns	Regressão Linear, SVM, Redes Neurais	K-means, PCA, Algoritmo de Floresta de Isolamento	Q-learning, Deep Q-Networks (DQN), Algoritmo AlphaGo

Aplicações

- Visão computacional
 - Exemplos: Reconhecimento facial, detecção de objetos.
 - Tecnologias usadas: CNNs, RNNs.




Fonte: <https://algorithmxlab.com/blog/computer-vision/>

Aplicações

- Recomendação Personalizada
 - Exemplos: Netflix, Amazon.
 - Técnicas: Filtragem colaborativa, sistemas baseados em conteúdo.

Os clientes que viram os produtos em seu histórico de navegação também viram





Data science para negócios: O que você precisa saber sobre...
 > Foster Provost
 ★★★★★ 1.142
 Capa comum
 R\$77,52
 Receba até **quinta-feira, agosto 3**
 Frete GRÁTIS em pedidos acima de R\$ 129,00 enviados pela Amazon



Python Para Análise de...
 > Wes McKinney
 ★★★★★ 20
 Capa comum
1º mais vendido em Livros de Processamento de Dados
29% off Oferta
 R\$120,00
 De: ~~R\$169,00~~
 Receba até **quinta-feira, agosto 3**
 Frete GRÁTIS em pedidos acima de R\$ 129,00 enviados



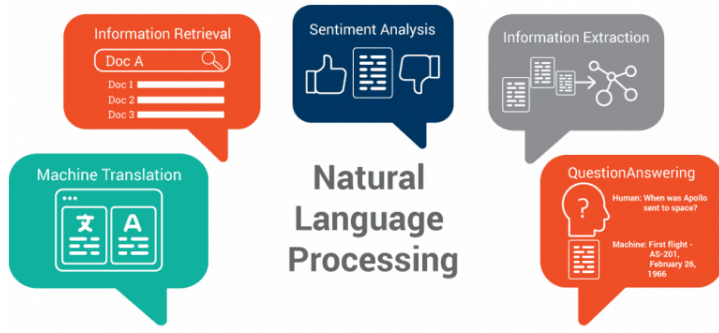
Machine Learning – Guia de Referência Rápida: Trabalhando com Dad...
 Matt Harrison
 ★★★★★ 400
 Capa comum
 R\$67,00
 Receba até **quinta-feira, agosto 3**
 Frete GRÁTIS em pedidos acima de R\$ 129,00 enviados pela Amazon



Estatística e Ciência de Dados
 Pedro Alberto Morettin
 ★★★★★ 111
 Capa comum
 R\$173,61
 Receba até **quinta-feira, agosto 3**
 Frete GRÁTIS em pedidos acima de R\$ 129,00 enviados pela Amazon

Aplicações

- Processamento de Linguagem Natural
 - Exemplos: Análise de sentimentos, chatbots.
 - Técnicas: BERT, GPT (Artigo).



Fonte : <https://legaltechnobrasil.com.br/legaltechs/10786-por-que-o-processamento-de-linguagem-natural-pln-impulsiona-a-gestao-de-contratos-gerenciamento-de-contratos/>

Outras Aplicações

- Medicina e Saúde
 - Exemplos: Diagnóstico assistido por IA, previsão de surtos.
- Previsões Financeiras
 - Exemplos: Previsão de mercado, análise de risco.
- Agricultura Inteligente
 - Exemplos: Detecção de doenças em plantas, otimização de colheitas.
- Jogos e Entretenimento(NPCs)
 - Exemplos: NPCs inteligentes, geração procedural de conteúdo.

Referências



MITCHELL, M. et al. Model cards for model reporting. In: . New York, NY, USA: Association for Computing Machinery, 2019. (FAT* '19), p. 220–229. ISBN 9781450361255.