

Como desenvolver e utilizar modelos de previsão de dados



Superintendência de
Inovação da Gestão
Pública

Subsecretaria de
Inovação da Gestão e
dos Serviços Públicos

SEAD
Secretaria de Estado de
Administração



GOVERNADOR

Ronaldo Ramos Caiado

SECRETÁRIO DA ADMINISTRAÇÃO

Alan Farias Tavares

SUBSECRETÁRIO DE INOVAÇÃO DA GESTÃO E DOS SERVIÇOS PÚBLICOS

Rômulo Mendonça Bailão

SUPERINTENDENTE DE INOVAÇÃO DA GESTÃO PÚBLICA

Daniel Soares Santana

GERENTE DE INTELIGÊNCIA DE NEGÓCIOS PÚBLICOS

Camila do Nascimento Freitas

EQUIPE TÉCNICA E EDIÇÃO

Carlos Alexandre Alencar de Assis

Giovanna Marques Grassini

Jessé Daniel de Souza

Rafaela Morais Rosa

REVISÃO

Mirielly dos Santos Dourado



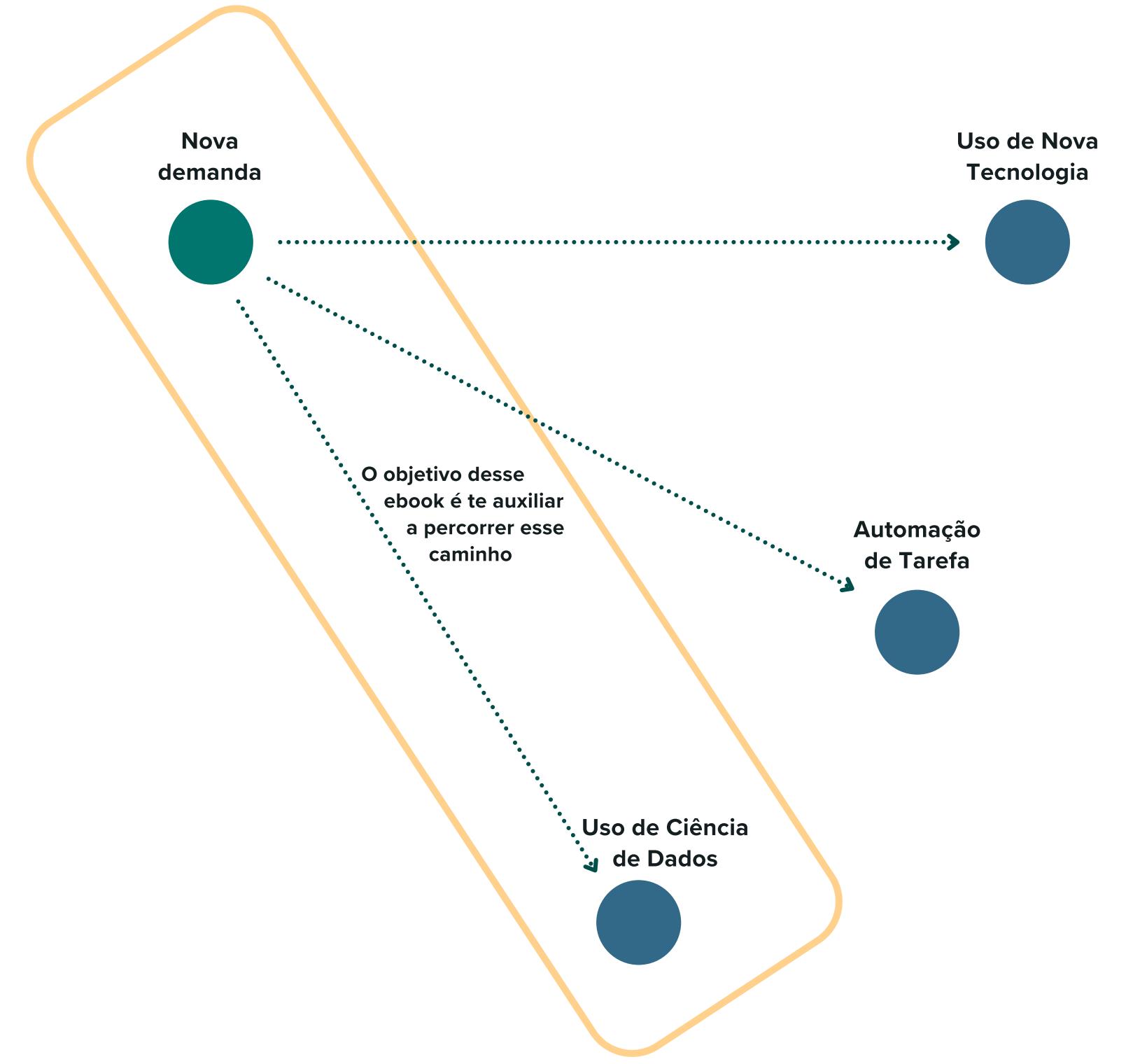
Conheça a GEINP

Somos a **Gerência de Inteligência de Negócios Públicos (GEINP)**, uma equipe multidisciplinar e colaborativa dedicada a buscar soluções eficazes e inovadoras para a administração pública nas áreas de novas tecnologias, automatização de tarefas e ciência de dados.

Nosso propósito é garantir que a gestão pública esteja sempre à frente, antecipando problemas e identificando oportunidades, para oferecer soluções eficazes e transparentes que atendam às necessidades da população de maneira exemplar.

Em nossa Jornada, atuamos em **três frentes**:

- Prospecção e adoção de novas tecnologias;
- Automação de tarefas manuais e repetitivas; e
- Aplicação de Ciência de Dados.



Índice

| | |
|----------------------------------|----|
| Introdução à modelos de previsão | 5 |
| Ciclo de vida de desenvolvimento | 7 |
| Modelos de previsão | 10 |
| Ferramentas e Bibliotecas | 15 |
| Dicas finais | 16 |
| Referências | 17 |



Introdução

Os modelos de previsão são algoritmos ou técnicas matemáticas utilizadas para prever valores futuros com base em dados históricos.

Em outras palavras, é a **arte de "adivinar o futuro" com base no passado.**

Imagine-se tendo que **"adivinar"** as respostas para as seguintes perguntas:

- »»»» Será que chove amanhã?
- »»»» Quantos professores precisaremos no próximo ano letivo?
- »»»» Quantos atendimentos presenciais teremos no mês que vêm?
- »»»» De quanto será a arrecadação de tributos nos próximos 2 anos?
- »»»» Qual será a demanda por programas sociais no próximo semestre?





A previsão atua como uma **ferramenta de planejamento** para ajudar os órgãos a se prepararem para a incerteza que pode ocorrer no futuro.

Auxilia gestores a responder com confiança às mudanças, controlar as operações e tomar decisões estratégicas que impulsionam o crescimento futuro.

MELHORIAS DE PROCESSOS

Auxilia na identificação de oportunidade de melhorias em processo e operações

OTIMIZAÇÃO DE RECURSOS

Ajuda a otimizar a alocação de recursos, como estoque, produção e pessoal

TOMADA DE DECISÃO

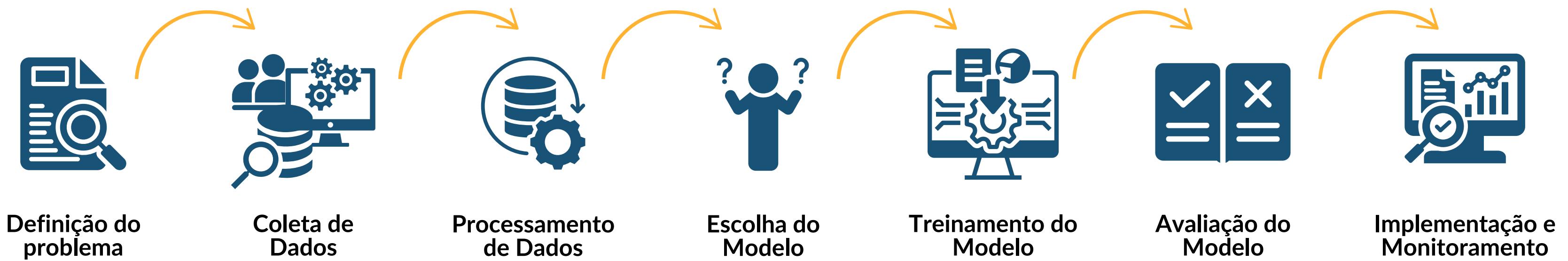
Permite antecipar tendências e tomar decisões estratégicas com mais confiança

IDENTIFICAÇÃO DE ANOMALIAS

Permite detectar eventos inesperados que podem exigir uma ação imediata

Ciclo de vida de Desenvolvimento

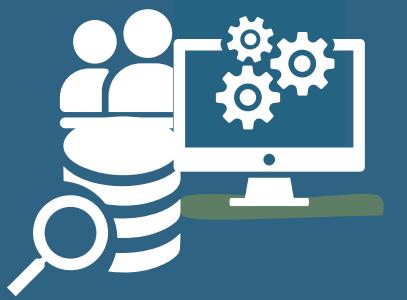
Desenvolver um modelo de previsão eficaz envolve seguir uma série de etapas.





Na **definição do problema** é importante se atentar aos seguintes pontos:

- **Objetivo:** Qual a pergunta que o modelo deve responder? O que se deseja prever?
- **Precisão:** Como medir a precisão do modelo? Quais métricas serão utilizadas para avaliar o desempenho?
- **Tempo de previsão:** É necessário analisar o período de tempo que se deseja prever, quanto mais tempo mais incertezas.



Na **coleta de dados** é necessário:

- **Identificar as fontes:** verificar quais dados serão utilizados e sua origem (Banco de dados, APIs, Planilhas, etc.)
- **Definir as variáveis:** determinar quais variáveis são relevantes para a previsão (ex.: região, temperatura, quantidade de atendimento, etc.)
- **Realizar a coleta:** extrair os dados das fontes identificadas e armazenar em um ambiente para os testes.



No **processamento de dados** é realizado:

- **Transformação:** conversão dos dados em um formato adequado para análise, normalizando e padronizando os dados.
- **Engenharia de features:** criação de novas *features* (variáveis) relevantes a partir dos dados existentes (ex: médias móveis, sazonalidade).
- **Seleção de features:** escolha das *features* mais relevantes para o modelo.



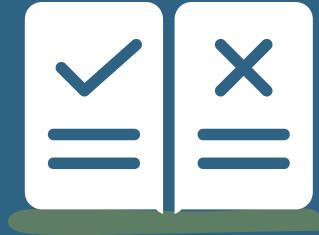
A **escolha do modelo** adequado é uma das etapas críticas do ciclo de desenvolvimento.

No próximo capítulo deste ebook abordaremos as principais categorias de modelos que irão te auxiliar em sua escolha.



O **treinamento do modelo** é necessário realizar ações de:

- **Dividir os dados:** Separar os dados em conjuntos de treinamento e teste.
- **Ajustar os parâmetros:** a alteração dos parâmetros do modelo auxilia a minimizar o erro das previsões aumentando sua taxa de acerto.



Já na **avaliação do modelo** é preciso:

- **Métricas:** Utilizar métricas apropriadas para avaliar o desempenho do modelo (ex.: RMSE, MAE, precisão).
- **Análise de resíduos:** Analisar a diferença entre os valores previstos e os valores reais para identificar padrões e *outliers*.
- **Interpretação:** analisar os resultados da previsão para atender os padrões e tendências do mercado



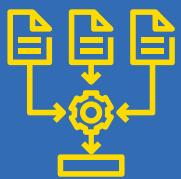
Na **implementação e monitoramento** são feitas:

- **Visualização:** apresentação dos resultados de gráfica.
- **Implantação:** integração do modelo em um sistema de produção para previsões automatizadas.
- **Monitoramento:** Acompanhar o desempenho do modelo ao longo do tempo e retrainá-lo periodicamente para garantir a sua precisão.

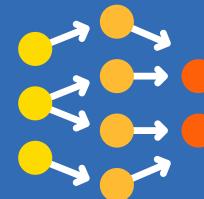
Modelos de Previsão

Escolher o modelo adequado é uma das etapas mais críticas para garantir previsões precisas e úteis.

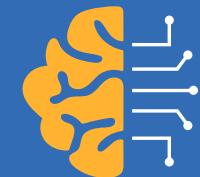
Os modelos de previsão de dados podem ser classificados em **três grandes categorias**:



Modelos clássicos



Modelos de Machine Learning



Modelos de Deep Learning



Escolher o modelo adequado é um passo crucial para garantir que as **previsões sejam precisas e úteis**.

Fazer essa escolha envolve considerar vários fatores e responder a uma série de perguntas essenciais.

Confira ao lado os principais aspectos a serem considerados e as perguntas que vão te apoiar a compreendendo o **objetivo** da previsão, a **natureza dos dados** e os padrões presentes, o **nível de precisão necessário** para a utilidade da previsão, os **recursos necessários** e o **custo envolvido**, além de definir o **horizonte temporal** da previsão, quais os **usuários finais** e suas necessidades.

Perguntas essenciais na escolha de um modelo de previsão:

- »»» Por que a previsão é necessária?
- »»» Quais as características dos dados existentes?
- »»» Qual a precisão desejada?
- »»» Qual o custo da previsão?
- »»» Que período deverá ser previsto?
- »»» Quem fará uso da previsão?

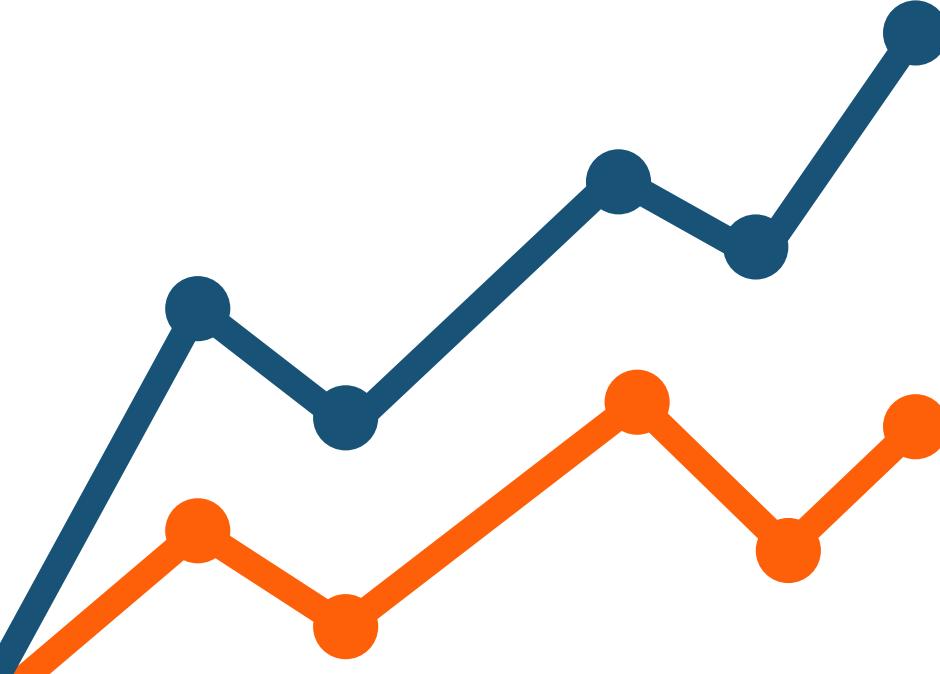




Exemplo de Modelos Clássicos

Os modelos clássicos referem-se a métodos estatísticos tradicionais que foram amplamente utilizados antes do advento do *Machine Learning* e outros métodos mais moderno. Esses modelos geralmente assumem relações lineares ou específicas entre variáveis.

Esse tipo de Modelo é utilizado principalmente em problemas com conjuntos de dados menores e menos complexos, onde as suposições subjacentes são mais prováveis de serem atendidas.



Modelos ARIMA: Usados para modelar e prever séries temporais, assumindo que os dados são estacionários após algumas transformações.

Modelos de Suavização Exponencial: Como o Holt-Winters, que são usados para previsões em séries temporais, considerando tendências e sazonalidade.

Regressão Linear Clássica: Um método simples, mas poderoso, para modelar a relação entre uma variável dependente contínua e uma ou mais variáveis independentes.

Análise de Componentes Principais: Usado para redução de dimensionalidade, preservando o máximo possível da variabilidade dos dados.

[Saiba mais sobre Modelos Clássicos](#)



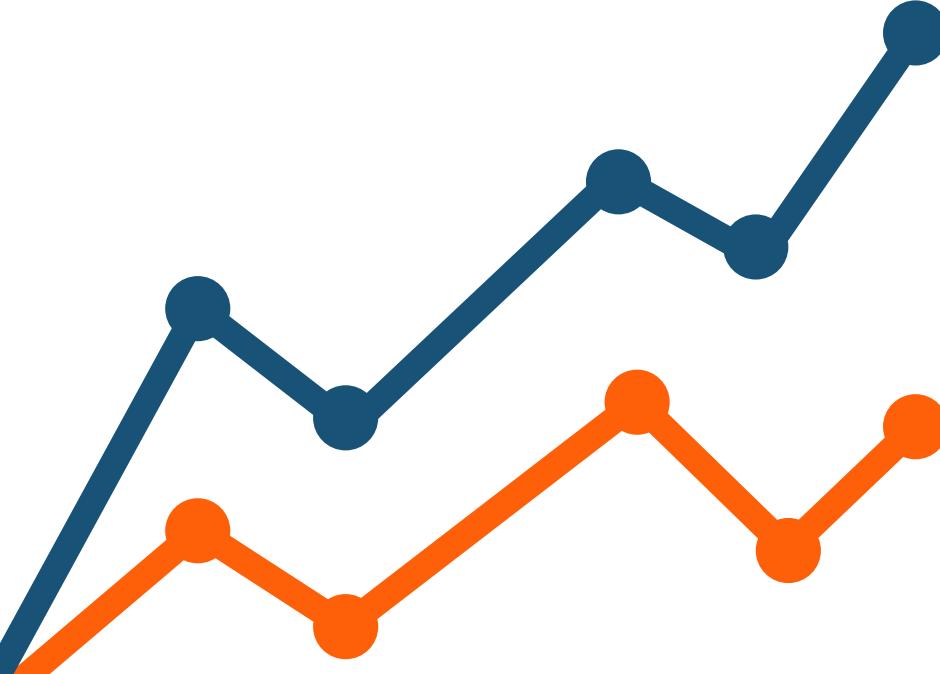


Exemplo de Modelos de Machine Learning



Trata-se da construção de algoritmos que aprendem a partir dos dados, em que a estatística tem um papel fundamental. O aprendizado de máquina envolve a programação de computadores para que, com base em experiências passadas (dados de entrada), possam aprender a resolver problemas.

Esse tipo de Modelo é utilizado principalmente em problemas com o objetivo de aprendem por si mesmos por meio de experiências e comportamentos passados.



Modelos de aprendizado supervisionado: é treinado com um conjunto de dados rotulados, aprende com exemplos de entradas e as saídas corretas conhecidas. A depender do seu problema, pode ser utilizada uma das abordagens a seguir:

- **Classificação:** classifica dados em segmentos específicos. Entre os modelos estão: regressão logística, K-Nearest Neighbors (KNN), floresta aleatória e Naïve Bayes.
- **Regressão:** classifica utilizando método estatístico para avaliar a relação entre uma variável dependente e, um ou diversos fatores isolados. Entre os modelos estão: regressão ridge, lasso e regressão logística.

Modelos de aprendizado não supervisionado: é treinado com dados que não possuem rótulos ou saídas conhecidas, buscando encontrar padrões ou agrupamentos nos dados. Os modelos não supervisionados mais conhecidos são:

- K-Means Clustering
- Hierarchical Clustering
- GMM

[Saiba mais sobre Machine Learning](#)

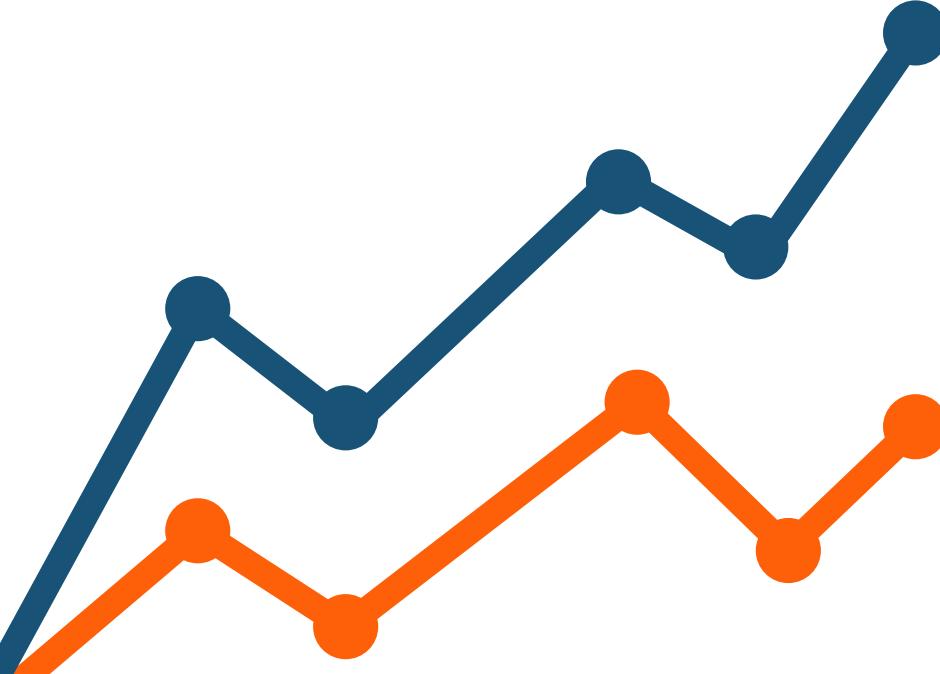




Exemplo de Modelos de Deep Learning

Os modelos de Aprendizado Profundo (*Deep Learning*) são uma subcategoria do *Machine Learning* baseada em redes neurais artificiais com várias camadas (daí o termo "profundo"). Esses modelos se destacam por sua complexidade e capacidade de lidar com grandes volumes de dados, estruturas complexas, ou que precisem capturar padrões sutis e complexos nos dados.

Esse modelo é utilizado principalmente em problemas complexos como: reconhecimento de fala, tradução automática, reconhecimento de imagem, diagnósticos médicos, etc.



Redes Neurais Convolucionais (CNNs): Usadas principalmente para reconhecimento de imagens e processamento de dados visuais.

Redes Neurais Recorrentes (RNNs): Comumente aplicadas em séries temporais e processamento de linguagem natural, onde a sequência de dados é importante.

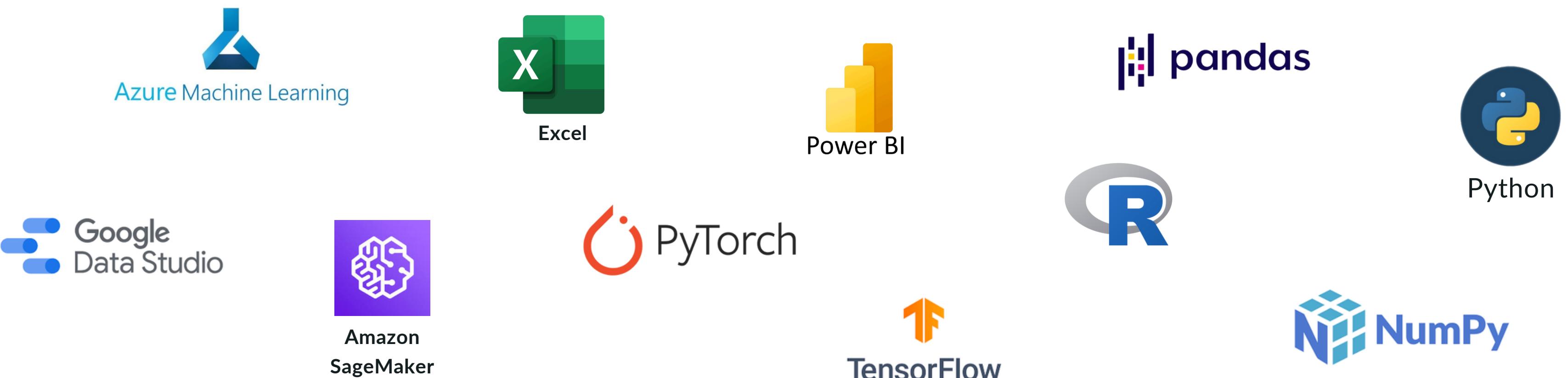
Autoencoders: Utilizados para redução de dimensionalidade e aprendizado de características.

[Saiba mais sobre Deep Learning](#)



Ferramentas e Bibliotecas

Analise o seu problema e técnica a ser empregada antes da escolha da melhor ferramenta para cada fase do desenvolvimento.



Apresentamos a seguir, algumas ferramentas e bibliotecas disponíveis atualmente. Alguma delas pode exigir custos com licenciamento ou consumo, consulte caso a caso.

Dicas finais

A criação de um **modelo de previsão** de dados é um **processo iterativo** que requer conhecimento básico em **estatística, programação e domínio do negócio.**

A **escolha do modelo e a interpretação dos resultados** devem ser feitas com cuidado, considerando o contexto do problema e os objetivos da análise.

Lembre-se de monitorar continuamente seus modelos e retreiná-los periodicamente para garantir que continuem precisos.



Referências

ALBRECHT, Denise Lange; PANZENHAGEN, Ilana Renata Lizi; GUIMARÃES, Iochane Garcia; FERREIRA, Mateus Freitas; THOMASI, Virginia. Modelo de previsão por séries temporais: uma aplicação para serviços comerciais em uma concessionária de energia elétrica. Iberoamerican Journal of Industrial Engineering, Florianópolis, SC, Brasil, v. 11, n. 21

CASTRO, Daniel Gomes Ferrari Leandro Nunes de. Introdução à Mineração de Dados: Conceitos Básicos, Algoritmos e Aplicações. São Paulo: SRV Editora LTDA, 2016. E-book. ISBN 978-85-472-0100-5.

FILHO, Oscar Gabriel. Inteligência artificial e aprendizagem de máquina: aspectos teóricos e aplicações. 1. ed. São Paulo, SP: Blucher, 2023.

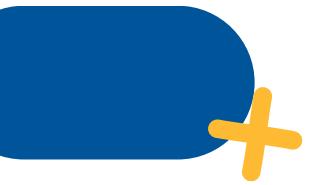
HANKE, J.E & Reitsch A.G. (1998) Business Forecasting. 6th Edition, Prentice Hall, Upper Sadle River, NJ
Modelagem e Previsão em 6 Passos: Fluxo de trabalho com séries temporais.

Práticas recomendadas para criar dados de treinamento tabulares.

SILVA, Fabrício M.; LENZ, Maikon L.; FREITAS, Pedro H C.; et al. Inteligência artificial. Porto Alegre: Grupo A, 2018. E-book. ISBN 9788595029392.

SHARDA, Ramesh; DELEN, Dursun; TURBAN, Efraim. Business intelligence e análise de dados para gestão do negócio. Porto Alegre: Grupo A, 2019. E-book. ISBN 9788582605202.

Slides da disciplina de Processos Estocásticos - Prof. Dr. João Ferreira Netto.



Nós confiamos nos dados
In data we trust

