Resumo – Aprendizado Supervisionado

Machine Learning: permite que sistemas aprendam e melhorem com a experiência, sem serem explicitamente programados.

Dados: Informações coletadas que servem como base para o aprendizado. Podem ser de todo tipo: tabelas, imagens, sons, vídeos, etc.

Algoritmos: Conjuntos de instruções que processam os dados para identificar padrões. São os algoritmos que produzem os modelos de machine learning.

Em outras palavras, o modelo de machine learning é um produto da execução do lgoritmo sobre uma base de dados.

Métricas de Avaliação: Critérios utilizados para medir a performance e precisão dos modelos. São importantes para validar a qualidade do modelo, como sua performance e seu valor de negócio.

**O que é Aprendizado Supervisionado?**

Modelo treinado para **determinar saídas com base nos dados de entrada.**

**Objetivo:** aprender função que mapeia entrada para saída, permitindo previsões precisas.

Utiliza rótulos (dados anotados) durante o treinamento.

Foco na capacidade de generalização para novos dados.

**Por que o Aprendizado é Chamado de "Supervisionado"?**

**Presença de Dados Rotulados:** O modelo é treinado com conjuntos de dados que incluem entradas e suas respectivas saídas corretas, permitindo que aprenda a mapear entradas para saídas desejadas.

**Processo de Correção:** Durante o treinamento, o modelo faz previsões e as compara com as saídas reais, ajustando-se com base nos erros para melhorar sua precisão.

**Analogia com Ensino Tradicional:** Assim como um aluno aprende com a orientação de um professor, o modelo aprende com os dados rotulados que atuam como supervisores.

**Objetivo de Generalização:** Após o treinamento, o modelo deve ser capaz de aplicar o

conhecimento adquirido para prever corretamente saídas de novos dados não vistos

anteriormente.

**O que é Aprendizado Supervisionado?**

**Objetivo**: atribuir categoria/classe (classificação) ou valor (regressão) a um conjunto

de dados com base em suas características.

**Classificação**: Prever categorias ou classes.

**Regressão:** Prever valores numéricos.

**Features vs Target**

**Features**

* Também chamadas de variáveis preditoras ou variáveis independentes.
* Dados de entrada do modelo.
* Categóricas e Numéricas.

**Target:**

* Também chamadas de variável de saída, variável alvo, variável resposta, e outros.
* Dados de saída do modelo.
* y: variável de saída alvo (target).
* y\_hat: variável de saída predita (prediction)

**PROCESSO**

1. Preparação das Features

2. Treinamento do Modelo

3. Validação

4. Utilização (predição)

**Regressão Linear**

**O modelo é a linha que melhor se ‘ajusta’ aos pontos.**

Texto, Carta

O conteúdo gerado por IA pode estar incorreto.

**Treinar o modelo, é encontrar todos os valores de w (parâmetros), que encontre a curva que melhor se ajusta ao conjunto de dados.**

Tabela

O conteúdo gerado por IA pode estar incorreto.

Interface gráfica do usuário, Aplicativo

O conteúdo gerado por IA pode estar incorreto.

Tabela

O conteúdo gerado por IA pode estar incorreto.

Gráfico, Gráfico de dispersão

O conteúdo gerado por IA pode estar incorreto.

Texto

O conteúdo gerado por IA pode estar incorreto.

Uma imagem contendo Diagrama

O conteúdo gerado por IA pode estar incorreto.

Texto

O conteúdo gerado por IA pode estar incorreto.

Gráfico, Diagrama

O conteúdo gerado por IA pode estar incorreto.

Texto, Carta

O conteúdo gerado por IA pode estar incorreto.

Gráfico, Linha do tempo, Gráfico de dispersão

O conteúdo gerado por IA pode estar incorreto.

Interface gráfica do usuário

O conteúdo gerado por IA pode estar incorreto.

Interface gráfica do usuário, Texto, Aplicativo, Email

O conteúdo gerado por IA pode estar incorreto.

Interface gráfica do usuário, Texto, Aplicativo, Email

O conteúdo gerado por IA pode estar incorreto.

Importar pacotes

Importar dados

Identificar variável y

Exploração de dados (distribuição, analise descritiva, outlier, correlação)

Preparar dados para modelagem

Treino e teste (80/20)

Regressão linear simples

Análise das métricas

Visualização gráfica

Validações

Regressão múltipla