#### JORNADA DE APRENDIZAGEM





**Python para Análise de Dados** 

Lógica de Programação

Criação de Script e IDE

Coleta de Dados

Tratamento de Dados

Preparação de Dados

Visualização de **Dados** 



# Pré-Modelagem

Pré-Processamento de Dados

Análise Exploratória dos Dados

Machine Learning 2: Tipos de Aprendizados

Correlação, Balanceamento, Variáveis Categóricas



FASE 5

**VAMOS** COMEÇAR!

FASE 1

FASE 2

da **Dados** 

FASE 3

FASE 4





Regressão Linear

Aplicada

Baysiana

Decisão



EW

Estatística

Aprendizagem

Árvore de

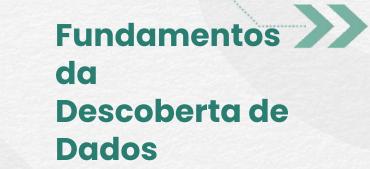


# **Fundamentos** da Ciência de **Dados** e Machine Learning

Introdução à Ciência de Dados;

Machine Learning 1: Conceitos Primários

Programação para Ciência de Dados



Analytics 1: Storytelling

**Analytics 2: Elementos Visuais** 

Estatística básica

Git

JORNADA DE APRENDIZAGEM





# Aprofundamento de Analytics

Analytics 3: Looker Studio

Analytics 4: Introdução ao PowerBl

Analytics 5: Bancos de Dados + SQL

Analytics 6: Operações Básicas com SQL

Analytics 7: Operações Avançadas com SQL

# Modelos avançados

Regressão Logística

Duelo entre Modelos: Árvore x Regressão

Na prática do dia a dia: o Pós Modelo

Agrupamento com K-means

Técnicas de avaliação de modelos

Random Forest

Duelo entre Modelos: Árvore x Random Forest







#### FASE 3

Conceitos avançados em Modelagem Estatística

Regressão

Melhoramento de Modelos

**Cross Validation** 

**Duelo entre Modelos: Cross Validation** 

PCA - Redução da Dimensionalidade

# Técnicas avançadas em modelos de Machine Learning

Métodos Essemble

Xgboost

Duelo entre modelos 4: Xgbost X Árvore de Decisão

Modelos de Linguagem









#### JORNADA DE APRENDIZAGEM



Integrar conhecimentos teóricos e práticos da Ciência de Dados, Machine Learning e Programação para abordar desafios complexos na área.



## Habilidade 1

Utilizar ambientes como Google Colab e Jupyter Notebook para programação em Python e compreender seu papel no contexto de Ciência de Dados.

# Habilidade 2

Aplicar métodos preditivos e descritivos, utilizando técnicas de Machine Learning para analisar e extrair informações.

## Habilidade 3

Gerenciar versões de código usando o GitHub, incluindo a criação de repositórios, controle de alterações, colaboração com outros usuários, e a criação de um portfólio profissional para exibir projetos.





#### JORNADA DE APRENDIZAGEM



# Competência

Desenvolver scripts em Python para coletar, tratar, preparar e visualizar dados de forma eficiente, utilizando boas práticas de programação e ferramentas apropriadas, como IDEs e bibliotecas específicas.

#### Habilidade 1

Escrever e executar scripts em Python, utilizando operadores, variáveis, estruturas condicionais e de repetição para solucionar problemas.

## Habilidade 2

Coletar dados de fontes como web e APIs e saber como extrair informações relevantes e formatá-las para análise.

## Habilidade 3

Limpar e tratar dados, normalizar e padronizar informações e lidar com outliers.

## Habilidade 4

Criar visualizações informativas utilizando bibliotecas as Matplotlib e Seaborn comunicar insights dos dados.



## JORNADA DE APRENDIZAGEM



# Competência

Comunicar insights de dados de maneira impactante através de storytelling e elementos visuais e aplicar conceitos estatísticos fundamentais na interpretação e análise de dados.

## Habilidade 1

Criar narrativas a partir de dados com conceitos de storytelling para a comunicação eficaz com stakeholders.

# Habilidade 2

Criar gráficos interativos e contar histórias usando elementos visuais, tornando as informações acessíveis e compreensíveis.

#### Habilidade 3

Compreender conceitos como média, mediana, moda, desvio padrão, variância e distribuição dos dados, para uma interpretação mais informada e análises estatísticas mais avançadas.



## JORNADA DE APRENDIZAGEM



# Competência

Conduzir todas as etapas na manipulação e análise de dados aplicando algoritmos de Machine Learning, lidar com os desafios associados à preparação de dados para projetos de Machine Learning, garantindo que os dados estejam em condições ideais para a construção de modelos robustos e eficientes.

#### Habilidade 1

Realizar uma análise exploratória de dados para compreender a distribuição e comportamento dos dados.

# Habilidade 2

Detectar e lidar com outliers usando Box Plot e Gráfico de Dispersão e aplicar técnicas como SMOTE e Tomek Links.

#### Habilidade 3

Identificar e compreender os tipos de aprendizado em Machine Learning (supervisionado, não supervisionado e por reforço) para selecionar abordagens adequadas em projetos práticos.



## JORNADA DE APRENDIZAGEM



# Competência

Analisar, avaliar e interpretar modelos estatísticos e probabilísticos para tomar decisões fundamentadas em evidências estatísticas, contribuindo para a resolução de problemas e a formulação de estratégias eficientes em diversas áreas.

#### Habilidade 1

Interpretar coeficientes de regressão linear para inferir relações entre variáveis.

# Habilidade 2

Calcular métricas de desempenho como R<sup>2</sup> para avaliar modelos de regressão linear.

# Habilidade 3

Aplicar o conceito de probabilidade para fundamentar análises estatísticas.

#### Habilidade 4

Realizar testes de significância para validar hipóteses estatísticas.



# JORNADA DE APRENDIZAGEM



# Competência

Utilizar SQL e ferramentas de visualização para obter insights, permitindo tomada de decisão informada e estratégica. Explorar o potencial dos dados disponíveis, transformando-os em vantagens competitivas e oportunidades de negócios.

## Habilidade 1

Criar visualizações de dados utilizando ferramentas como Looker Studio e PowerBI.

# Habilidade 2

Integrar dados de diferentes fontes para análise e visualização.

# Habilidade 3

Aplicar filtros e criar campos personalizados para aprimorar análises visuais.

## Habilidade 4

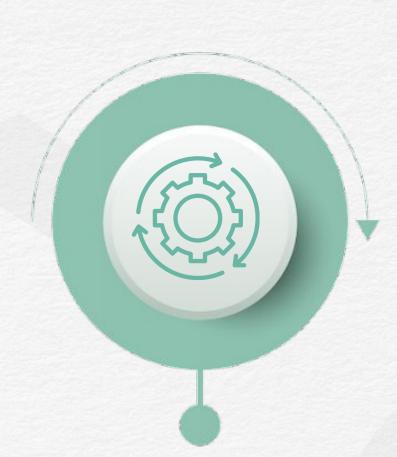
Executar consultas
SQL para filtrar,
agrupar e juntar
dados de bancos de
dados.



## JORNADA DE APRENDIZAGEM

# Competência

Analisar, implementar, avaliar e otimizar modelos preditivos, utilizando técnicas estatísticas e algoritmos de machine learning, para tomada de decisão embasada.



## Habilidade 1

Interpretar coeficientes e probabilidades na regressão logística para entender a influência de variáveis aplicando nas probabilidades de ocorrência.

## Habilidade 2

Utilizar bibliotecas como scikit-learn para codificar e treinar modelos de regressão logística, conceitos estatísticos e de machine learning.

#### Habilidade 3

Utilizar métricas como AUC-ROC e cross-validation para avaliar e ajustar o modelo visando melhorar sua precisão e generalização.

#### Habilidade 4

Utilizar gráficos como curvas ROC e matrizes de confusão para visualizar e interpretar o desempenho do modelo, facilitando a tomada de decisões.



## JORNADA DE APRENDIZAGEM



# Competência

Otimizar modelos avançados, aplicando regularização, seleção de recursos, validação cruzada e redução de dimensionalidade para aprimorar o desempenho e a generalização dos modelos de machine learning.

#### Habilidade 1

Aplicar técnicas de regularização (L1, L2) na regressão logística para lidar com overfitting.

## Habilidade 2

Utilizar métodos como Grid Search e Random Search para otimização de hiperparâmetros de modelos.

## Habilidade 3

Implementar a técnica de K-Fold Cross-Validatio n para avaliar modelos de forma robusta.

## Habilidade 4

Reduzir a
dimensionalidade
dos dados utilizando
Análise de
Componentes
Principais (PCA) para
simplificar modelos.



## JORNADA DE APRENDIZAGEM



# Competência

Integrar e aplicar métodos ensemble, bem como modelagem de linguagem, para desenvolver e aprimorar modelos preditivos, contribuindo para a tomada de decisão informada e ética.

#### Habilidade 1

Explorar técnicas de Bagging e Boosting para combinar modelos de forma eficaz.

## Habilidade 2

Implementar o algoritmo XGBoost em Python para análise de dados.

#### Habilidade 3

Avaliar e otimizar o desempenho do modelo XGBoost utilizando métricas como AUC-ROC.

#### Habilidade 4

Utilizar modelos de linguagem para auxiliar análises, insights e codificação, seguindo melhores práticas e considerações éticas.

