



GUIA

DATA MASTERS

ENGENHEIRO DE MACHINE LEARNING



DATA MASTERS

ENGENHEIRO DE MACHINE LEARNING

Olá, seja muito bem-vindo ao Guia **DATA MASTERS** de Engenheiro de Dados de Machine Learning! Aqui, você encontrará uma seleção de conteúdos recheados para auxiliá-lo na sua certificação.

O Engenheiro de Dados de Machine Learning tem como principal função garantir que todas as tarefas de produção de um sistema ML/IA estão trabalhando corretamente, de modo a automatizar as tarefas relacionadas à modelagem, além de extrair a melhor performance e resultado das bibliotecas de ML.

Preparamos para você artigos, livros, cursos, vídeos e palestras que vão lhe proporcionar um aprendizado realmente incrível.

VOCÊ ENCONTRARÁ...

06 Sistemas Operacionais

- Conceitos e prática da arquitetura de sistemas operacionais.

08 Programação

- Lógica;
- Boas práticas de programação;
- Bibliotecas e APIs de Machine Learning.

11 Manipulação de Dados

- Conceitos de modelagem de dados;
- Comandos SQL e NoSQL;
- Conceitos de Big Data.

14 Big Data

- Automação da instalação e uso de ferramentas de armanejamento;
- Tableau;
- Watson Analytics;
- Sisense;
- Oracle Bi;
- Movimentação e processamento de dados .

18 Virtualização e Containers

- Conceitos de VM;
- Containers;
- Orquestração de VM vs Containers e ferramentas.

21 DevOps

- Conceitos e prática de CI/CD;
- Ferramentas de automação de infra;
- Criação de automações para ML/IA.

24 Microserviços e Gestão de Eventos

- Arquitetura e criação de micro-serviços;
- Patterns de mensageria;
- Processamento de eventos.

26 Estatística e Machine Learning

- Conceitos e prática de ML/AI;
- Uso de APIs de machine learning.

29 Cursos e Livros

- K-means;
- Algoritmos hierárquicos aglomerativos e divisivos;
- Latent Dirichlet Allocation.

SUMÁRIO

engenheiro de dados machine learning

06 Sistemas Operacionais

08 Programação

11 Manipulação de Dados

14 Big Data

18 Virtualização e Containers

21 DevOps

24 Microserviços e Gestão de
Eventos

26 Estatística e Machine
Learning

29 Cursos e Livros



ICONOGRAFIA

No conteúdo deste e-book, você encontrará diversos materiais, dentre eles opções pagas e abertas.

Abaixo você confere os significados de cada ícone e ao visualizá-los, basta clicar para acessar.



Artigo



PDF



Curso



Vídeo



Livro



Nível Básico



Nível Intermediário



Nível Avançado

SISTEMAS OPERACIONAIS


O Sistema operacional corresponde à um programa ou um conjunto de programas cuja função é gerenciar os recursos do sistema.

Iniciaremos com a conceituação e indicação da prática da arquitetura de sistemas operacionais.

CONCEITOS E PRÁTICA DA ARQUITETURA DE SISTEMAS OPERACIONAIS



How Do Operating Systems Work?

⌚ 4min 



Organização e arquitetura de computadores

⌚ 2h



Sistemas Operacionais

⌚ 8min



● ○ ○

Sistemas Operacionais - UFSM

🕒 1h30



● ○ ○

What is an operating system

🕒 20min 



● ○ ○

Structures of Operating System

🕒 1min 



● ○ ○

Sistemas Operacionais Modernos por Andrew S. Tanenbaum

Pago R\$ 267,96



● ○ ○

Guia Foca GNU/Linux

PROGRAMAÇÃO

Agora, vamos lhe apresentar os seguintes conteúdos: Conceitos básicos de Lógica; Algoritmos; Linguagem de Programação; Boas práticas e algumas bibliotecas e APIs de Machine Learning.

Fique tranquilo, pois, se você não tiver familiaridade com a linguagem de programação, nessa seção você encontrará tudo o que precisa saber sobre Python, e como está presente no cotidiano de um Engenheiro de Dados de Machine Learning.

LÓGICA



● ○ ○

Introdução à programação com Python

por Nilo Ney Coutinho Menezes - Novatec

Pago R\$ 73,00



● ○ ○

Learning Python

por Mark Lutz - Editora O'Reilly

Pago R\$ 270,01



● ○ ○

Lógica de Programação e Algoritmos

🕒 21min Pago R\$ 270,01



● ○ ○

Lógica de programação

🕒 1h30

BOAS PRÁTICAS DE PROGRAMAÇÃO



Código limpo com Clean Code

🕒 20min



Design Patterns com Python

🕒 1h



Design Patterns in Python by Peter Ullrich

🕒 28min

BIBLIOTECAS E APIS DE MACHINE LEARNING



TOP 20 bibliotecas Python para Data Science

🕒 30min



Melhores bibliotecas de machine e deep learning

🕒 30min



11 ferramentas open source para machine learning

🕒 15min



Machine learning APIs

🕒 3min 



Scikit-Learn Machine Learning APIs



Tensorflow Machine Learning APIs



MANIPULAÇÃO DE DADOS

Em Manipulação de Dados vamos analisar requisitos, definir o modelo conceitual e modelo lógico. Por fim, entender como aplicar o modelo físico. Ficou confuso? Não se preocupe! Nessa seção vamos abordar os conceitos de NoSQL e modelagem de dados e como aplicar os comandos da Linguagem SQL.

CONCEITOS DE MODELAGEM DE DADOS



● ○ ○

Entenda o que é a modelagem de banco de dados

🕒 15min



● ○ ○

Modelagem de Dados -
Conceitos de Bancos de Dados

🕒 9h



● ○ ○

Introdução a Sistemas de Bancos de Dados por Date, C. J. - Elsevier

Pago R\$ 215,90



● ○ ○

Sistemas de Banco de Dados por Elmasri, Ramez - Navathe, Shamkant B.

Pago R\$ 263,84

COMANDOS SQL E NoSQL



● ○ ○

Comandos básicos em SQL insert
update delete e select

🕒 20min



● ○ ○

Guia completo de SQL



● ○ ○

Introdução a SQL: Consulta e
gerenciamento de dados

🕒 1h30



● ○ ○

SQL Tutorial

🕒 2h 🇬🇧



● ○ ○

O que é NoSQL

🕒 2h



● ○ ○

Top 6 NoSQL data bases

🕒 20min

CONCEITOS DE BIG DATA



● ○ ○

Big Data

🕒 7min



● ○ ○

Big data - O que é e qual sua importância?

🕒 30min



● ○ ○

O que é Big Data - Conceitos básicos

🕒 18min

BIG DATA

Você sabe o que é Big Data? Este é um termo que trata de grandes conjuntos de dados que precisam ser processados e armazenados.

Neste momento, nos deteremos no processo de instalação e uso de ferramentas de armazenamento, além da movimentação e do processamento de dados.

AUTOMAÇÃO DE INSTALAÇÃO E USO DE FERRAMENTAS DE ARMAZENAMENTO



Best Practices to Help You Leverage the Power of Storage Automation

🕒 15min 🇬🇧



Storage resource management (SRM)

🕒 10min 🇬🇧



Red Hat Ansible



Chef



TABLEAU



● ● ● ○

Complete tutorial TABLEAU

🕒 9h 

HADOOP



● ● ● ○

Hadoop: The Definitive Guide
Editora O'Reilly

Pago R\$ 157,73 



● ● ● ○

Hadoop Application Architectures:
Designing Real-World Big Data
Applications Editora O'Reilly

Pago R\$ 74,89 



● ● ● ○

Plataforma Hadoop e framework
de aplicação

🕒 20h



● ● ● ○

Big Data Essentials: HDFS,
MapReduce and Spark RDD

🕒 44h



● ● ● ○

Modern Big Data Processing
with Hadoop Packt Publishing





● ● ● ○

Business Intelligence



Pago R\$ 23,99

MOVIMENTAÇÃO E PROCESSAMENTO DE DADOS



Kafka Hadoop Integration

🕒 20min 🇬🇧



Talend and Apache Spark:
A Technical Primer

🕒 30min 🇬🇧



Kafka Tutorial | Apache Kafka
Tutorial For Beginners

🕒 17min 🇬🇧



Etapas de Big Data

🕒 1h



Integrating Hadoop with
SQL Server

🕒 40min 🇬🇧



Data processing

🕒 10min 🇬🇧



Impala



5 different types of data processing

🕒 50min 🇬🇧



Introduction to Data and Data Processing

🕒 12h



VIRTUALIZAÇÃO E CONTAINERS

Muito bem! Agora que você chegou até aqui, apresentaremos conceitos de VM (Virtual Machine), Containers e Organização de VM em contraste com Containers e Ferramentas.

CONCEITOS DE VM



O que é virtualização?

🕒 5min



O que é virtualização?

🕒 20min



Virtualização de Servidores:
conceito e vantagens

🕒 20min



VMS vs Containers quais
diferenças e usos?

🕒 20min

CONTAINERS




O que é container?

🕒 20min



Docker e containers: uma introdução rápida

🕒 41min 



Container, o novo passo para a virtualização

🕒 10min

ORQUESTRAÇÃO DE VM VS CONTAINERS E FERRAMENTAS



VMS vs Containers: quais são as diferenças e os usos?

🕒 10min



Simplifique a gestão de cloud através da orquestração

🕒 10min



Ferramentas de orquestração de containers: como escolher?

🕒 20min



O que é orquestração de container?

🕒 5min



● ● ○
O que é Docker?

🕒 15min

DEVOPS

Devops é um termo que foi criado para descrever um conjunto de práticas para integração entre as equipes de desenvolvimento de softwares, operações de apoio envolvidas e a adoção de processos automatizados para produção rápida e segura de aplicações e serviços.

Nesse momento, você entrará em contato com conceitos e prática de CI/CD, com as Ferramentas de Automação de Infra e com a criação de automações para ML/IA.

CONCEITOS E PRÁTICA DE CI/CD



DevOps

🕒 10min



O que é DevOps?

🕒 10min



O que é CI/CD?

🕒 20min 

FERRAMENTAS DE AUTOMAÇÃO DE INFRA



O que é o AWS Systems Manager?



SaltStack Get Started



Learn about secrets management and data protection with HashiCorp Vault



Chef



Get Started with Docker



CRIAÇÃO DE AUTOMAÇÕES PARA ML/IA




A Beginner's Guide to Automated Machine Learning & AI

🕒 30min 



Why AutoML Is Ready To Be The Future Of Artificial Intelligence

🕒 20min 



Tudo o que você precisa saber sobre AutoML e Neural Architecture Search

🕒 25min



Jenkins



Sonarqube



Fortify static code



Automação e provisionamento ágil com Ansible

🕒 40min 



Terraform



MICROSSERVIÇOS E GESTÃO DE EVENTOS

Agora, entram em cena a arquitetura e criação de microsserviços, patterns de mensageria e as etapas que compõem o processamento de eventos.

ARQUITETURA E CRIAÇÃO DE MICROSSERVIÇOS



Microsserviços em poucas
palavras

🕒 25min



Introdução à Arquitetura de
microsserviços

🕒 40min



Como implantar uma verdadeira
Arquitetura de Microsserviços na
sua empresa

🕒 25min 

PATTERNS DE MENSAGERIA



Mensageria: o pombo-correio
de seus sistemas

🕒 25min



Enterprise Integration Patterns
Editora Addison Wesley

Pago R\$ 129,36

PROCESSAMENTO DE EVENTOS



Eventos: denifinição e
processamento

🕒 25min



What is event processing?

🕒 15min 



Complex Event Processing

🕒 1h30

ESTATÍSTICA E MACHINE LEARNING

Há quem diga que os primeiros estudos utilizando estatística foram iniciados em 1662, e suas principais aplicações estavam voltadas para formulação de políticas públicas e fornecimento de dados demográficos e econômicos.

Apesar de ser uma ciência relativamente antiga, com o passar do tempo ela vem sendo aplicada de diversas maneiras, é comum encontramos estatística atrelado aos conceitos e prática de ML/AI e o uso de APIs de Machine Learning.

CONCEITOS E PRÁTICA DE ML/AI



● ○ ○

O que é Machine Learning
(Inteligência artificial)?

🕒 6min



● ○ ○


Machine Learning

🕒 15min



● ○ ○

Desmistificando Machine Learning

🕒 20min 



● ○ ○

A Diferença Entre Inteligência
Artificial, Machine Learning e Deep
Learning

🕒 15min

USO DE APIS DE MACHINE LEARNING



● ● ●

Uma introdução prática ao
Machine Learning

🕒 57min



● ● ●

Applied Machine Learning
in Python

🕒 24h 🇬🇧



● ● ●

Introduction to Machine
Learning

🕒 15h 🇬🇧



● ● ●

What is the Azure Face APIs?

🇬🇧



● ● ●

Azure Cognitive Services
Documentation

🇬🇧



● ● ●

Machine box

🇬🇧



● ● ●

Amazon Connect



APIs



Build on the AI platform
for business



CURSOS E LIVROS

Para você que aprecia um estudo de forma linear, como uma trilha de aprendizado, separamos alguns Cursos e Livros para complementar no seu processo de aprendizagem.

Os cursos abordam de forma transversal os assuntos do Guia **DATA MASTERS**, mas indicamos que você também estude todos os conteúdos, pois assim você vai garantir a total compreensão dos temas.

Bons estudos!



Machine learning

🕒 55h 🇬🇧

- Introduction;
- Linear Regression with One Variable;
- Linear Algebra Review;
- Linear Regression with Multiple Variables;
- Octave/Matlab Tutorial;
- Logistic Regression;
- Regularization;
- Neural Networks: Representation;
- Neural Networks: Learning;
- Advice for Applying Machine Learning;
- Machine Learning System Design;
- Support Vector Machines;
- Unsupervised Learning;
- Dimensionality Reduction;
- Anomaly Detection;
- Recommender Systems;
- Large Scale Machine Learning;
- Application Example: Photo OCR.



Programa de cursos integrados aprendizagem automática

🕒 7h 🇬🇧

- Fundações do aprendizado de máquina: uma abordagem por estudo de caso;
- Machine Learning: Regression;
- Machine Learning: Classification;
- Machine Learning: Clustering & Retrieval.



Python para Data Science e Machine

🕒 18h Pago R\$ 129,99

- Uso de Python para Data Science e Machine Learning;
- Visualização profissional de dados com Matplotlib;
- Aprender sobre visualização de dados intuitiva com o Plotly;
- Aprender a usar o Seaborn para análise estatística dos dados;
- Uso do Pandas para análise de dados;

- Uso do Numpy para operações vetoriais e matriciais;
- Regressão Logística;
- K-Nearest-Neighbors (KNNs);
- K-means-clustering;
- Support-Vector-Machines (SVM);
- Neural-Networks (Redes Neurais);
- Uso do Spark para Big-Data;
- Processamento natural de linguagem e sistemas de filtro de spam;
- Sistemas de recomendação (usados no NetFlix, Spotify, etc)...;
- Projetos com dados reais;
- Spark e AWS.



My Mooc

Pesquisa pela competência que quer desenvolver



● ○ ○

Machine Learning e Data Science com Python de A à Z

Pago R\$ 144,99

- Tenha uma base teórica sólida sobre os principais algoritmos de Machine Learning;
- Utilize as bibliotecas numpy, scikit-learn e pandas aplicado em Data Science e Machine Learning;
- Aprenda na teoria e na prática sobre os algoritmos de Machine Learning para classificação, regressão, regras de associação e agrupamento;
- Aprenda a realizar o pré-processamento em bases de dados com pandas e scikit-learn;
- Entenda como funcionam as técnicas para redução de dimensionalidade PCA, KernelPCA e LDA;
- Aprenda a avaliar os algoritmos de Machine Learning usando estatística não paramétrica;
- Aprenda a detectar outliers em bases de dados.



Machine learning

Casa do código

Pago R\$ 44,90



Machine learning

Springer Verlag Usa-id

Pago R\$ 878,60



Python para Data Science e Machine Learning

Wiley

Pago R\$ 271,30



Python Data Science Handbook: Essential Tools for Working with Data

Editora O'Reilly

Pago R\$ 155,82



