

Pós-Graduação

Análise de Big Data – Data Engineering

360 horas



Índice

Mais Renomada Escola de Negócios do Brasil	3
Fundação Instituto de Administração	4
FIA Business School	5
Reconhecimento Comprovado: Rankings, Credenciamentos e Afiliações	6
Labdata – Laboratório de Dados	10
Direção	12
Conteúdo Premium	13
Laboratórios de Ponta	14
Perfil do Aluno	15
Formação para Empreendedores	16
Sobre o Curso: Objetivo, Corpo Docente e Metodologia	17
Requisitos para o EaD 100% ao Vivo	18
Matriz Curricular	18
Bloco Optativo Internacional	30



BUSINESS SCHOOL

A mais renomada
escola de negócios
do Brasil.

Enquanto instituições muitas vezes produzem conhecimento pouco aplicável, fórmulas prontas e conteúdos superficiais. Na FIA, unimos base teórica sólida ao conhecimento especializado que dialoga com os desafios do mercado.

Em um ambiente que estimula conexões de alto valor, transformamos teoria em prática, pensamento em impacto. Porque pra nós, saber é agir.

Quem sabe lidera.



Unidade Pinheiros: Avenida Doutora Ruth Cardoso, 7.221
Pinheiros – São Paulo/SP

FIA – Fundação Instituto de Administração

Criada em **1980 por professores do Departamento de Administração da FEA-USP**, a FIA – Fundação Instituto de Administração é uma instituição privada, sem fins lucrativos, reconhecida como uma das **maiores referências em educação executiva, pesquisa aplicada e consultoria estratégica** no Brasil.

Por meio de um Acordo de Cooperação Técnica com a Universidade de São Paulo (USP), a FIA desenvolve projetos em parceria que reúnem rigor acadêmico, visão de mercado e impacto real nos negócios e na sociedade.



Consulting

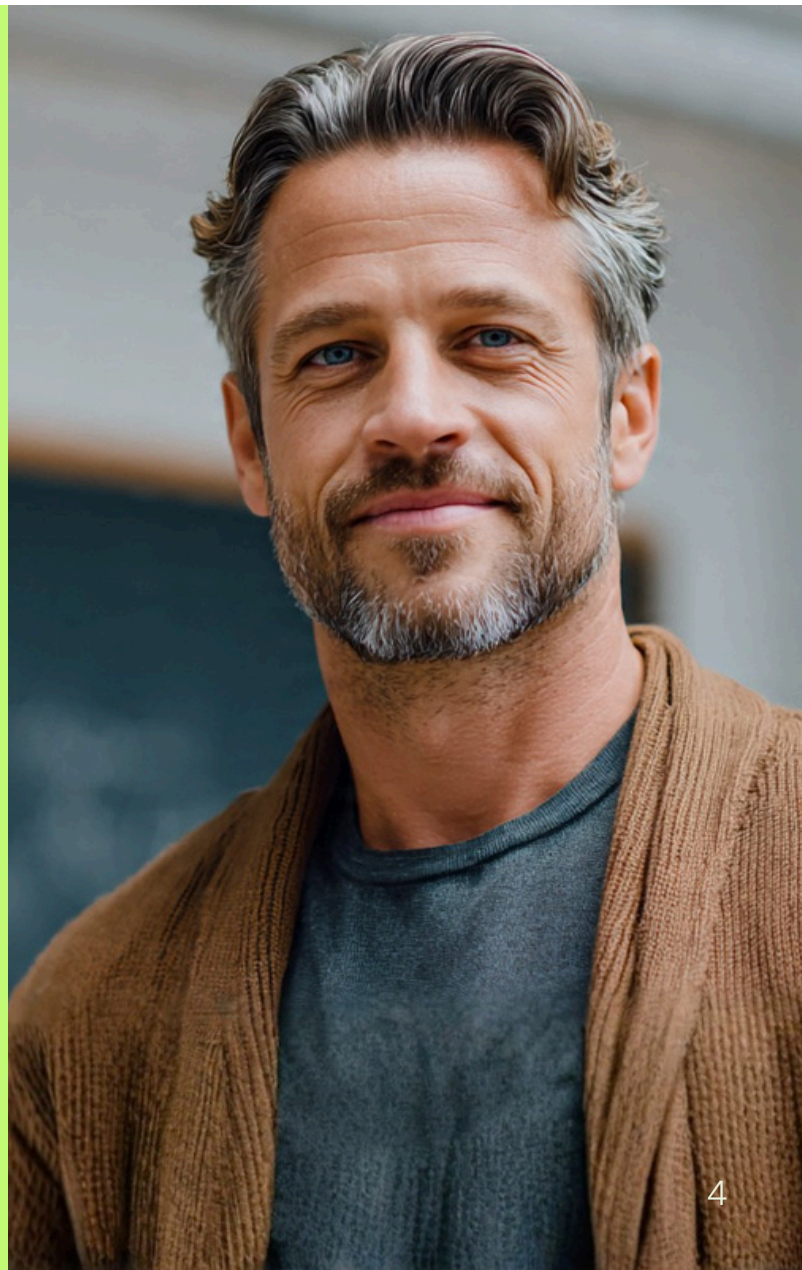
FIA conta com uma equipe de consultoria composta por profissionais com formação multidisciplinar e especializada em gestão pública e privada.

Consultoria personalizada que oferece soluções baseadas em seu problema de negócio.



+8 mil consultorias

realizadas em mais de mil empresas privadas e órgãos públicos.



Aqui é onde líderes se formam para transformar o futuro.



FIA Business School

A faculdade FIA Business School é mantida pela fundação FIA – Fundação Instituto de Administração. **Reconhecida como uma das mais tradicionais e respeitadas escolas de negócios do Brasil**, consolidou-se pela oferta de programas executivos premium de Graduação, Pós-Graduação, MBA, Mestrado Profissional, Doutorado Profissional, Extensões, Cursos In-Company, além de soluções educacionais customizadas para organizações dos setores privado e público, no Brasil e em diversos países.

Destaca-se como um dos mais importantes centros de educação executiva do país, faculdade responsável pela **formação de líderes, executivos e tomadores de decisão** preparados para atuar em um mercado cada vez mais complexo, tecnológico e global.

Está entre as escolas de negócios mais reconhecidas globalmente, com programas entre os **Top 100 do mundo** e creditações internacionais que posicionam a instituição no seletivo grupo de business school de padrão global.

Estudar na FIA Business School é preparar-se para liderar decisões estratégicas, transformar organizações e deixar um legado real no mundo dos negócios.



Reconhecimento Comprovado

A FIA Business School possui **reconhecimento nacional e internacional** por sua excelência em ensino executivo, reconhecido por rankings globais, credenciamentos acadêmicos e afiliações estratégicas que poucas instituições conseguem alcançar.

Enquanto a maioria das instituições alcança apenas requisitos regulatórios, como o MEC, a FIA é reconhecida internacionalmente, prestigiada com selos globais e excelência acadêmica orientada ao mercado corporativo. Nossos selos não são apenas “troféus acadêmicos”, eles atestam a qualidade do corpo docente, a consistência acadêmica, a metodologia premium, a relevância prática dos programas e a forte conexão com o mercado.

Rankings

QS Stars

A 1ª instituição brasileira a ser classificada, destacando-se com nota máxima pela excelência em ensino, empregabilidade, educação em gestão e educação online.

Global Business Education Awards

Importante título de “Escola mais inovadora da América do Sul” pela European CEO. Uma publicação comandada por um time editorial composto por jornalistas e correspondentes dos veículos mais influentes do mundo.

QS MBA Executivo

O MBA Executivo Internacional obteve 5 estrelas e está entre os TOP MBAs, sendo o único do Brasil entre os 100 melhores do mundo. Destacando-se com uma das melhores pontuações globais.



Mais que um diploma,
uma marca
global no seu
currículo.

Credenciamentos

Ministério da Educação (MEC)

Integra as melhores escolas de negócios do Brasil, com **Nota 5 no MEC**.

Business Graduates Association (BGA)

2ª universidade brasileira a conquistar a acreditação internacional, **um dos selos mais rigorosos e respeitados do mundo**. Reforça nosso compromisso com excelência acadêmica, impacto real e elevados padrões de qualidade.

Association of MBAs (AMBA)

1ª **escola** brasileira a adquirir o credenciamento. Significando que a FIA atende aos mais altos padrões de qualidade internacional na educação em gestão.

Associação Nacional de MBA (ANAMBA)

Reconhecimento global que assegura padrões de excelência para MBAs, alinhados às melhores práticas internacionais.

European Foundation for Management Development (EFMD)

Os MBAs Internacionais são acreditados, garantindo padrões de excelência e posicionamento entre os **melhores do mundo**.

Business School Impact System (BSIS Impact)

Integramos um seletivo grupo de menos de 100 escolas no mundo reconhecidas pelo BSIS Impact, avaliados pelo impacto gerado na economia, na sociedade e ecossistema.



Da sala de aula
pra sala de
decisão.



Afiliações

Association to Advance Collegiate Schools of Business (AACSB)

Atesta a conformidade com **elevados padrões de excelência** em ensino, currículo, recursos e resultados dos alunos.

Associação Brasileira de Educação a Distância (ABED)

Associação científica sem fins lucrativos que promove o desenvolvimento e a excelência da educação a distância.

Principles for Responsible Management Education (PRME)

Com base no Pacto Global da ONU, promove práticas de **responsabilidade social na educação executiva**.



Principles for Responsible Management Education

3x eleita a melhor escola de negócios do Brasil pela ÉPOCA.



Labdata

O Laboratório de Dados – Labdata é um centro de conhecimento que atua nas áreas de ensino, pesquisa e consultoria em análise de informação utilizando big data, engenharia de dados, arquitetura de dados, análise de dados e inteligência artificial.

Seja você um profissional mais técnico ou de negócios, esta especialização irá aprofundar seus conhecimentos e desenvolver as habilidades necessárias para ser um líder orientado a dados de alto desempenho.

Laboratório de Dados da FIA Business School



Centro de ensino pioneiro

no lançamento dos cursos de engenharia de dados, arquitetura de dados, análise de dados e inteligência artificial no Brasil.

Lifelong learning e atividades gratuitas:

Hackathons, palestras, cursos de curta duração, pós-graduação, advanced MBA e eventos.

Número limitado de alunos por turma,

networking profissional de elite e divulgação de oportunidades nas principais empresas.

Docentes com sólida formação acadêmica

Doutores e mestres com larga experiência de mercado.

Certificação de alto valor

com reconhecimento internacional nos Advanced MBAs e Pós-Graduações, tanto nas modalidades EaD 100% ao vivo quanto presencial.

Localização privilegiada

nos principais centros de negócios de São Paulo, com fácil acesso por transporte público e privado: Unidade Paulista e Unidade Pinheiros.



9 em cada 10
alunos são
promovidos
ou
conquistam
novas
oportunidades.



Direção

Prof.ª Dr.ª Alessandra Montini

É uma das principais referências brasileiras em dados, estatística e inteligência artificial (IA) aplicada aos negócios. Em razão de sua **reconhecida autoridade** nos temas, é frequentemente convidada para painéis, eventos nacionais e grandes veículos de comunicação, como a [Globo News](#) e a [Record](#), para falar sobre análise de dados, IA e tomada de decisão orientada por dados.

Professora e pesquisadora com carreira acadêmica construída na Universidade de São Paulo ([USP](#)), é graduada e **mestre em Estatística Aplicada** pelo Instituto de Matemática e Estatística da USP ([IME-USP](#)) e **doutora em Administração de Empresas** pela Faculdade de Economia, Administração e Contabilidade da USP ([FEA-USP](#)).

Sua trajetória é marcada pela excelência acadêmica e pela habilidade de conectar teoria, tecnologia e aplicação prática em negócios. Ao longo de **mais de 20 anos de atuação**, recebeu **mais de 30 prêmios** de excelência acadêmica na USP e acumula mais de 30 prêmios de excelência como docente dos cursos de MBA da FIA Business School, consolidando seu reconhecimento tanto no meio acadêmico quanto executivo.

No Labdata, lidera pesquisas e programas avançados de educação executiva nas áreas de estatística, ciência de dados, engenharia de dados e IA. Criou laboratórios dedicados ao ensino prático em São Paulo, além de conduzir a expansão desse modelo de excelência para o ensino remoto, ampliando o acesso à formação de alto nível.

Atua como palestrante, orientadora de mestrado e doutorado na FEA-USP; é membro do Conselho Curador da FIA Business School e é colunista em relevantes portais de tecnologia, como [Olhar Digital](#), [TI Inside](#) e [Febraban Tech](#).

Conteúdo Premium

No mercado brasileiro, é comum que programas de MBA e Pós-Graduação anunciem cargas horárias elevadas, porém a **carga real de aulas** com professor é significativamente inferior, o restante é composto por atividades assíncronas.

No Labdata, **carga horária não é discurso, é compromisso com a excelência.** A Pós-Graduação tem uma imersão profunda, intelectualmente exigente e orientada a práticas reais do mercado. Por isso, o programa conta com 360 horas efetivas de aulas em tempo real com professores, oferecendo **profundidade técnica, por meio de aulas práticas e resolução de cases reais.**

Não formamos simples operadores de ferramentas, nem colecionadores de certificados. Formamos profissionais preparados para conduzir equipes e projetos orientados a dados.

360 horas

de conhecimento profundo das mais atualizadas técnicas e ferramentas utilizadas no mercado.

Formação de elite se mede em horas reais de profundidade.



Laboratórios de Ponta

A parceria Labdata e GoWork traz para o ecossistema do coworking, um núcleo de educação executiva, onde pesquisadores, empreendedores e estudantes contribuem para a maximização dos resultados criando um ambiente de busca por conhecimento, sintetizado em muitos eventos, hackathons, aulas de alto nível e amplo espaço para networking.

Localizado na Avenida Paulista e na Avenida das Nações Unidas, nossos laboratórios contam com **equipamentos individuais, segurança, áreas de lazer e networking** para nossos alunos e professores.



Espaços de Estudos



Quadra e Praia



Salas de Reunião



Espaços de Coffee Break



Laboratórios Bem Equipados

“ É um background que eu não tinha até o momento. Eu consegui me desenvolver muito nessa frente, tive bastante conteúdo relevante que consegui fazer com que eu conversasse com pessoas na minha empresa no mesmo nível. ”

Thiago Cordon Rodrigues

Pós-Graduação Análise de Big Data -
Data Engineering



Perfil do Aluno

Este programa é direcionado a profissionais de todas as áreas que desejam acelerar sua trajetória na área de engenharia e arquitetura de dados, desenvolvendo competências avançadas para projetar e construir ecossistemas analíticos. O curso atende aqueles que buscam fortalecer o pensamento crítico, dominar as tecnologias mais atuais e desenvolver habilidades de empreendedorismo.

Acreditamos que o domínio avançado de engenharia de dados, arquitetura de dados, estatística aplicada e inteligência artificial é fundamental para profissionais que desejam extrair insights significativos a partir de grandes volumes de dados e transformar informações em ativos estratégicos, traduzindo desafios complexos de negócio em soluções escaláveis e de alto valor para as organizações.

Após concluir a Pós-Graduação, o aluno está preparado para ocupar cargos de maior relevância. **A formação amplia o acesso a projetos estratégicos, melhores condições de trabalho e maior protagonismo nas decisões organizacionais, posicionando o profissional como referência.**

Formação para Empreendedores

O curso foi concebido para formar líderes capazes de transformar conhecimento técnico em **negócios escaláveis, inovadores e orientados por dados**. Ao longo do programa, a FIA Business School – Labdata oferece uma formação que vai além da técnica: desenvolve **mentalidade empreendedora, visão estratégica e capacidade de execução**, alinhadas aos padrões das principais escolas de negócios do mundo.

O aluno conclui o programa com **domínio técnico profundo, visão estratégica de alto nível e mentalidade de negócio**, plenamente capacitado para conceber, estruturar e escalar iniciativas data-driven — seja fundando sua própria startup, liderando projetos e equipes ou atuando em empreendimentos de alto crescimento orientados por dados e inteligência artificial.



“ Eu tinha “um projeto técnico”: um catálogo de dados com IA. As aulas me fizeram dar um passo atrás e responder o que eu evitava.

Fiz pequenas POCs, conversei com as áreas de negócio – esses retornos moldaram o produto melhor do que qualquer diagrama. Também aprendi a falar a língua do negócio: antes de arquitetura, resultado; antes de feature, hipótese validada.

O impacto foi muito prático: **o catálogo deixou de ser “mais um dev entregando um script” e virou ferramenta que economiza horas de trabalho, traz governança (LGPD) e dá autonomia** para quem usa dado no dia a dia.

No fim, a disciplina não me ensinou só a empreender um produto, mas a empreender dentro da empresa: ouvir, priorizar, simplificar, testar rápido e contar a história com números. Foi o impulso que faltava para transformar uma boa ideia em valor em produção. ”

Victor Hugo Ferreira da Silva

3º Bloco: Transforme Ideias em Negócios – Empreender com Inovação e Propósito

Sobre o curso

Objetivo

- Desenvolvimento de habilidades de resolução de problemas e pensamento analítico para aplicar as tecnologias de big data e inteligência artificial de forma eficaz para a resolução de problemas complexos;
- Apresentar componentes de infraestrutura visando a criação de ambientes de desenvolvimento e execução dos pipelines de dados, auxiliando em tomadas de decisões e na avaliação de soluções escaláveis, eficientes com melhor custo benefício;
- Apresentar os principais componentes da engenharia de software envolvidos em projetos de big data;
- Apresentar algumas abordagens de arquiteturas escaláveis, como computação em nuvem e arquiteturas distribuídas, para lidar com grandes volumes de dados eficientemente;
- Apresentação de estratégias para dimensionamento e otimização de infraestruturas de big data para garantir desempenho e escalabilidade;
- Aplicação de governança de dados em camadas de armazenamento e distribuição de dados;
- Definição de arquitetura completa de plataforma de dados utilizando softwares open source.
- Capacitar o aluno com conhecimento técnico e estratégico para estruturar, desenvolver e escalar uma startup orientada por dados e tecnologia.

Corpo Docente

O corpo docente conta com professores altamente capacitados com experiência no mundo corporativo. Nos critérios de seleção do corpo docente, serão priorizadas sua qualificação e experiências profissionais nas distintas matérias, de maneira que o curso permita não somente a transmissão de conhecimentos, mas também experiências enriquecedoras para os alunos.

Metodologia

- As aulas na modalidade presencial são realizadas em um ambiente que

propicia à aprendizagem;

- As aulas na modalidade EaD ao vivo são transmitidas por meio de uma plataforma digital que permite contato em tempo real entre alunos e professor;
- Todas as aulas terão a presença de um professor titular com apoio de pelo menos um monitor;
- São utilizados diversos recursos tecnológicos para o melhor entendimento dos conceitos;
- Todas as aulas são gravadas e permanecem disponíveis em nossa biblioteca digital até o final do curso.

Requisitos para EaD 100% ao vivo

Requisitos MEC

Caso haja determinação legal para aplicação de avaliação presencial, ela será realizada em uma das unidades educacionais da FIA, em São Paulo – SP.

Requisitos para Acompanhar o Curso

- Conexão com a internet – banda larga com ou sem fio com pelo menos 1MB de velocidade final para acompanhar as aulas ao vivo e para a realização de exercícios;
- Computador com configuração mínima: Dual Core 2Ghz ou superior (i3/i5/i7 ou AMD equivalente) com no mínimo 8GB de RAM;
- Os navegadores recomendados são Internet Explorer 11+, Edge 12+, Firefox 27+, Chrome 30+ ou Safari +7;
- Permissão de administrador da máquina.

Matriz Curricular

A especialização reúne as principais metodologias e tecnologias exigidas pelo mercado para formar especialistas capazes de projetar, estruturar e escalar ecossistemas de dados organizados em 3 blocos, proporcionando uma visão técnica, inovadora e empreendedora.

Empresas orientadas por dados dependem de arquiteturas robustas, governança eficiente e profissionais capazes de conectar tecnologia e negócio. Ao dominar as competências de engenharia de dados, o profissional

passa a:

- Projetar e implementar pipelines e arquiteturas de dados escaláveis e seguras;
- Integrar engenharia de dados com soluções de IA avançadas;
- Garantir qualidade, governança e confiabilidade das informações;
- Traduzir necessidades de negócio em soluções técnicas sustentáveis;
- Liderar iniciativas data-driven com visão estratégica e foco em resultados.

As informações a seguir estão sujeitas a alterações sem aviso prévio.

Arquitetura e Engenharia de Dados

Este bloco tem como objetivo apresentar, de forma prática, as principais tecnologias utilizadas nas organizações por engenheiros e arquitetos de dados, para a implementação de projetos de big data e de inteligência artificial.

Introdução

- Conceito de big data, inteligência artificial e machine learning;
- Panorama para o surgimento do big data;
- Necessidades e vantagens da utilização das técnicas de big data e das metodologias de inteligência artificial;
- 7 Vs do big data: variedade, veracidade, valor, volume, velocidade, visualização e vulnerabilidade;
- O profissional adequado para trabalhar com o big data: data scientist, data engineer e data architect;
- Cases nacionais e internacionais de big data.

Cloud Computing

- Fundamentos Cloud Computing;
- Cloud Adoption Framework;
- Conceitos e casos de uso de aplicações utilizando SaaS (Software-as-a-Service), Paas (Platform-as-a-Service) e IaaS (Infrastructure-as-a-Service);
- Finops;
- Ingestão de dados.

Programação

- Introdução a linguagem Python e suas principais bibliotecas;
- Análise de dados utilizando Python;
- Microservices em Python;

- Ingestão de dados em Python;
- Introdução a linguagem SQL;
- Conceitos de DML, DDL e DQL;
- Acesso a dados com linguagem SQL;
- Padrões de Microservices;
- CI/CD;
- Observabilidade.

Plataformas de Dados

- Criação de plataforma de dados utilizando tecnologias de código aberto;
- Hadoop: Arquitetura e conceitos do framework pioneiro dos ecossistemas big data;
- Conceitos e utilização de conjuntos de práticas de automatização de processos DevOps;
- Introdução a Docker e containers e Kubernetes;
- Abordagem de arquitetura em micros serviços independentes;
- EDA – Conceitos da Arquitetura Orientada a Eventos e sua utilização com Streams de Eventos;
- MLOPS: Automatização de projetos de machine learning e suas práticas.

Armazenamento de Dados

- Utilização de dados estruturados, semiestruturados e não estruturados;
- Tipos de arquivos (CSV, JSON, Avro, Parquet, Delta e outros);
- Arquitetura e armazenamento de dados destruído com HDFS;
- LakeHouse: Transações ACID, Time Travel, Schema Evolution, Vacuum e Particionamento;
- Conceito, arquitetura e utilização de padrão S3 em cloud e on premise com Minio;
- Conceito, arquitetura, modelagem e utilização de banco de dados NoSQL das famílias Column Store (Cassandra), Document (MongoDB e Elastic), Key/Value (Redis), Neo4j (Grafo) e GEN IA Database (Vector);
- Arquitetura de dados relacional, internal engine e utilização de bancos de dados relacionais e Data Warehouse com MySQL e PostgreSQL.

Ingestão de Dados

- Fontes de dados e coleta de dados;
- Streaming de dados com Kafka Platform (Kafka, Kafka Connect , KSqlDb);
- Eventos em banco de dados utilizando CDC (Change Data Capture);
- Ingestão em tempo real e fluxo de dados com Kafka Connect, Spark, Nifi e AirByte;
- Ingestão batch, ETL e ELT utilizando Spark e AirFlow;
- Orquestração e deploy de pipeline de ingestão de dados.

Processamento e Análise de Dados

- Conceitos e arquitetura Spark;
- Introdução ao processamento distribuído;
- Análise de dados com Python (PySpark) e SQL;
- Análise exploratória de dados com Spark;
- Processamento de dados em tempo real;
- Análise e cruzamento de dados em modo batch;
- Criação de aplicações distribuídas com Spark;
- Processamento ETL com DBT.

Acesso e Distribuição de Dados

- Conceito de acesso a arquivos com engine SQL, ClickHouse, Duckdb, Athena, Synapse, Trino, Presto e Dremio;
- Arquitetura e acesso a dados com Hive;
- Federação de dados e acesso In Memory com Trino/Presto;
- Entregando dados com API.

Governança

- Catálogo de metadados;
- Qualidade de dados;
- Linhagem;
- Gestão de acesso.

1º Projeto Prático: Engenharia de Dados

Será realizado um projeto de big data com as principais tecnologias utilizadas por engenheiros e arquitetos de dados.

O objetivo deste projeto é aplicar os conhecimentos adquiridos no curso para criar uma solução ou produto baseado em dados que seja resiliente, escalável e confiável utilizando as principais tecnologias de dados destacadas em aula. Para isso, é proposto desenhar e implementar uma arquitetura de dados com os seguintes pilares:

Coleta de Dados

- Implementar um sistema de ingestão de dados capaz de lidar com diferentes fontes, como bancos de dados, logs e streaming de dados em tempo real;
- Utilizar ferramentas como Plataforma Kafka, Spark Streaming, AirByte, Nifi e Airflow para garantir a escalabilidade e a confiabilidade na ingestão de dados.

Armazenamento

- Projetar e implementar um sistema de armazenamento utilizando tecnologias como HDFS e S3;
- Integrar bases de dados NoSQL, como Elastic ou MongoDB, para acesso de dados em baixa latência.

Processamento de Dados

- Aplicar técnicas de processamentos distribuídos utilizando Apache Spark e Python para realizar análises relevantes ao projeto;
- Implementar pipelines de processamento de dados eficientes para transformação e limpeza dos dados.

Análise e Visualização

- Disponibilizar acesso aos dados armazenados e processados com ferramentas de acesso como Presto, Hive, API, entre outros;
- Desenvolver dashboards interativos utilizando frameworks Metabase ou Apache Superset para fornecer insights acionáveis aos usuários finais.

Segurança e Gerenciamento de Dados:

- Implementar políticas de segurança para proteger dados sensíveis;
- Utilizar ferramentas de gerenciamento de metadados para rastrear a proveniência dos dados e garantir a conformidade com regulamentações.

Entregáveis

- Documentação técnica detalhada do sistema, abrangendo arquitetura, processos de ETL, segurança e manutenção;
- Código-fonte do sistema, hospedado em um repositório versionado (por exemplo, GitHub);
- Apresentação executiva destacando os principais resultados e benefícios obtidos com a implementação da solução.

Inteligência Artificial

Este bloco tem como objetivo apresentar os fundamentos e a importância da inteligência artificial para a tomada de decisão.

Introdução à Ciência de Dados e Inteligência Artificial

- Áreas da inteligência artificial;
- Aplicações de aprendizado de máquina (machine learning);
- Aplicações de aprendizado profundo (deep learning);
- Aplicações de inteligência artificial generativa (GenAI);
- Casos de uso no mundo corporativo e acadêmico.

Transformação Digital

- Contexto e dimensões da transformação digital;
- Tecnologias disruptivas e exponenciais e seu impacto nos negócios;
- Transformação digital e o alinhamento entre as estratégias tecnológica e organizacional;
- Jornada da transformação digital e sua implantação adequada nas organizações;
- Tecnologias emergentes.

Introdução ao Python

- Introdução ao Google Colab;
- Comentários em código;
- Criação e remoção de objetos;
- Realização de cálculos;
- Módulos e bibliotecas;
- Funções;
- Verificações lógicas;
- Listas e vetores;
- Atributos e métodos de classes;
- Criação de uma base de dados;
- Erros de execução.

Modelos Lineares: Regressão Linear

- Introdução à modelagem supervisionada;
- Conceitos de variável resposta e variável explicativa;
- Equação da reta;
- Regressão linear simples e múltipla;

Métodos de Validação e Regularização

- Superajuste de modelos;
- Conceito de hiperparâmetro.

Modelos Baseados em Árvores: Árvore de Regressão

- Estrutura do algoritmo;
- Medidas de impureza;
- Interpretação do modelo;
- Hiperparâmetros.

Modelos Baseados em Árvores: Random Forest

- Estrutura do algoritmo;
- Otimização de hiperparâmetros;
- Aplicações e estudos de caso.

Modelagem de Projeção: Boosting

- Estrutura dos algoritmos;
- Variantes: Gradient Boosting, AdaBoost, XGBoost, LightGBM e CatBoost;
- Otimização de hiperparâmetros;
- Aplicações e estudos de caso.

Modelos Baseados em Árvores: Árvore de Decisão Binária

- Extensão da árvore de regressão para respostas qualitativas binárias;
- Medidas de impureza;
- Interpretação do modelo;
- Balanceamento de variável resposta;
- Calibração de probabilidades;
- Aplicações e estudos de caso.

Modelos Baseados em Árvores: Random Forest e Boosting

- Extensão dos algoritmos de floresta aleatória e impulsionamento para respostas qualitativas (binárias ou multinomiais);
- Aplicações e estudos de caso.

Introdução à Modelagem Não Supervisionada

- Principais tarefas não supervisionadas;
- Objetivo da tarefa de segmentação;
- Medidas de distância;
- Padronização de variáveis quantitativas.

Modelos de Segmentação: Hierárquico

- Estrutura do algoritmo;
- Métodos de ligação: complete, single, average e ward;
- Análise de dendrograma;
- Soma de quadrados intracluster (WSS);
- Gráficos de silhueta e cotovelo;
- Definição da quantidade de clusters.
- Análise exploratória dos clusters;
- Aplicações e estudos de caso.

Modelagem Automatizada: AutoML

- Conceito de Auto ML em modelagem supervisionada;
- Principais pacotes de Auto ML em Python;
- Comparação e escolha do melhor modelo;
- Aplicações e estudos de caso.

Modelos em Produção: MLOps

- Conceito de sistemas de ML;
- Conceito de MLOPs;

- Ciclo de vida de modelos de ML;
- Exercício prático de implementação de um modelo de ML;
- Monitoramento de modelos de ML;
- Casos de uso.

Introdução ao Deep Learning e as Redes Neurais

- Evolução da inteligência artificial;
- Conceitos básicos de redes neurais e as principais redes mais utilizadas (redes densas, redes convolucionais e redes recorrentes);
- Cálculos realizados pelo algoritmo Perceptron.

Redes Neurais Densas

- Conceitos de redes neurais densas (algoritmo MLP – Multi Layer Perceptron);
- Funções de ativação;
- Como o método do gradiente descendente reduz o erro;
- Algoritmo de Backpropagation (regra da cadeia e cálculo das derivadas parciais);
- Funcionamento da biblioteca TensorFlow (infraestrutura de manipulação de tensores);
- Funcionamento da biblioteca Keras (camadas, modelos, otimizadores, perdas, métricas, etc);
- Desenvolver os códigos para rodar o algoritmo MLP para problemas de regressão e classificação;
- Desenvolver os códigos para rodar o algoritmo MLP para problemas de classificação de imagem.

Redes Neurais Convolucionais

- Como as imagens digitais são representadas e como podem ser processadas utilizando diferentes bibliotecas do Python;
- Operações básicas como negação, adição, subtração, multiplicação, divisão, exponenciação e rotação;
- Operação de convolução para remover ruído, realçar detalhes, suavizar imagens, detectar bordas, adicionar borda (padding) e extrair características;
- Desenvolvimento de códigos para aplicar as principais técnicas de pré-processamento de imagens;
- Técnicas de regularização como dropouts, regularização L1/L2, early stopping;
- Como Data Augmentation pode ajudar no processo de treinamento;
- Como podemos utilizar a técnica de transfer learning para melhorar e agilizar o treinamento dos modelos (VGG 16, ResNet);

Redes Neurais Recorrentes

- Estrutura básica das redes neurais recorrentes (RNNs) "simples" e entender suas limitações;
- Como as dependências temporais com feedback loop podem auxiliar no treinamento;
- Como problemas de vanishing e exploding gradients impactam no aprendizado;
- Princípios das Long Short-Term Memory (LSTMs) com Cell State, Hidden State, Input Gate, Forget Gate e Output Gate;
- Principais etapas de processamento de linguagem natural – Tokenização e embeddings (Word2Vec);
- Desenvolver os códigos para aplicar e comparar todos os conceitos estudados neste tópico;
- Desenvolver aplicações do mundo real, mostrando como essas redes podem ser utilizadas em tarefas como processamento de linguagem natural (NLP) para análise de sentimento e previsão de séries temporais.

Inteligência Artificial Generativa

- Diferença entre modelos discriminativos e generativos;
- Diferentes tipos de aplicações (entrada e saída do modelo);
- Casos de uso da IA Generativa (texto, imagem, som e multimodal);
- Evolução da IA Generativa ao longo do tempo.

Arquiteturas para Geração de Imagens

- Fundamentos das arquiteturas generativas voltadas para geração de imagens;
- Funcionamento da arquitetura GAN, incluindo o papel do gerador e do discriminador;
- Explorar a arquitetura DCGAN, com foco em convolução transposta, funções de ativação como Leaky ReLU e mode collapse;
- Funcionamento da Arquitetura Stable Diffusion e sua relevância no mercado atual;
- Criação de dados artificiais para treinamento de modelos de machine learning.

Arquitetura Transformer

- Compreender os fundamentos da arquitetura Transformer, base dos modelos de linguagem modernos;
- Funcionamento dos mecanismos de atenção e autoatenção;
- Analisar como modelos Transformer processam grandes volumes de informação de forma contextual;
- Avaliar as principais capacidades e limitações dessa arquitetura em aplicações reais.

Modelos GPT

- O que são os modelos GPT e como eles se diferenciam de abordagens tradicionais;
- Compreender como os modelos GPT funcionam internamente e como geram respostas;
- Analisar as capacidades e limitações dos modelos GPT em cenários reais;
- Explorar casos de uso práticos em diferentes áreas de negócio;
- Aplicar modelos GPT para geração de textos, respostas e conteúdos automatizados;
- Integrar modelos GPT a aplicações de negócio de forma prática.

Representação de Linguagem Natural

- Entender como o texto é convertido em dados que os modelos conseguem processar;
- Compreender o papel da tokenização no funcionamento dos modelos de linguagem;
- Explorar técnicas de Word Embeddings e representação vetorial de palavras;
- Uso de Positional Encoding para preservação de contexto e ordem;
- Compreender como o GPT interpreta, relaciona e gera linguagem natural.

Prompt Engineering

- Compreender o que é prompt engineering e por que ele é essencial para o uso eficiente de LLMs;
- Como a qualidade do prompt impacta diretamente a qualidade das respostas;
- Explorar os benefícios do prompt engineering para aumento de eficiência, controle e previsibilidade;
- Principais conceitos e terminologias utilizados em prompt engineering;
- Aplicar Explanations para orientar o comportamento dos modelos;
- Utilizar Completion Prompts e System Prompts de forma estratégica;
- Compreender o impacto de parâmetros como Temperature, Max Tokens, Top-K e Top-P no comportamento do modelo.

Técnicas de Prompt Engineering

- Aplicar técnicas de Zero-Shot, One-Shot e Few-Shot para diferentes cenários;
- Estruturar prompts utilizando listas e formatos organizados;
- Utilizar Chain of Thought para respostas mais explicáveis e consistentes;
- Criar Meta-Prompts para padronização e reutilização;
- Aplicar técnicas de prompts defensivos para reduzir erros e alucinações;
- Gerar saídas estruturadas utilizando JSON, Schemas e Constraints.

Avaliação de Prompts

- Como avaliar a qualidade das respostas geradas por modelos de linguagem;
- Aplicar métricas tradicionais como BLEU, ROUGE e METEOR;
- Realizar testes A/B de prompts em ambientes de produção;
- Comparar desempenho entre diferentes prompts;
- Comparar desempenho entre diferentes modelos.

O Novo Papel do Engenheiro de Dados na Era do AI-Augmented Coding

- Do manual coding ao AI-assisted engineering;
- Onde GenAI realmente gera produtividade (e onde não gera);
- Trade-offs: velocidade × qualidade × governança;
- Ferramentas de coding com GenAI;
- Cursor;
- Contexto de código completo;
- Refatoração orientada a intenção;
- Diferença entre cursor e IDEs tradicionais;
- GitHub Copilot;
- Copilot Chat vs inline suggestions;
- Uso responsável em bases corporativas;
- Claude Code;
- Análise de bases grandes;
- Geração de código orientada a raciocínio.

Padrões Corretos de Uso em Engenharia de Dados

- Prompting orientado a:
 - o Transformação de dados;
 - o Modelagem dimensional;
 - o Pipelines ELT.
- Anti-padrões:
 - o “Geração cega de SQL”;
 - o Lógica de negócio implícita no prompt.
- Versionamento, revisão e auditoria de código gerado por IA.

O Problema Clássico dos Dados Não Estruturados

- PDFs, textos, e-mails, contratos e descrições livres;
- Por que ETL tradicional falha nesse cenário;
- Onde LLMs entram como semantic processors.

Casos de Uso Reais em Engenharia de Dados

- Parsing inteligente de documentos;
- Entity Extraction: pessoas, empresas, datas, valores e eventos;
- Classificação semântica;

- Normalização de texto;
- Data quality assistida por IA.

Arquitetura de enriquecimento com GenAI

- Bronze / Silver / Gold com LLMs;
- Separação entre: extração semântica, validação e persistência;
- Custos, latência e estratégias de batch vs streaming.

O que São Sistemas Multi-Agentes

- Agente \neq Prompt;
- Agente como unidade: objetivo, memória e ferramentas;
- Diferença entre: pipelines determinísticos vs pipelines agênticos.

Casos de Uso em Engenharia de Dados

- Agente de ingestão;
- Agente de validação;
- Agente de enriquecimento;
- Agente de observabilidade;
- Agente de correção automática de erros.

Arquitetura de Referência

- Orquestração entre agentes;
- Controle de custos;
- Logs, rastreabilidade e governança;
- Quando não usar agentes.

Transforme Ideias em Negócios: Empreender com Inovação e Propósito

Empreender para Impactar: Construindo Negócios Sustentáveis

- Empreendedorismo e intraempreendedorismo;
- Introdução a governança e ética corporativa;
- Comportamento do consumidor;
- Cálculo financeiro e análise de viabilidade de investimentos.

Criando o Futuro: Empreendedorismo, Ética e Transformação

- Aprenda a vender seu projeto dentro e fora da empresa;
- Inovação que move: da ideia ao plano de negócio;
- Plano de negócio – Model CANVAS;
- Conceitos básicos de estratégia;
- Ética.

2º Projeto Prático:

Inovando com Análise de Dados Não Estruturados usando Deep Learning e Inteligência Artificial Generativa

Integração dos conhecimentos adquiridos ao longo do curso para desenvolver uma aplicação baseada em dados não estruturados (texto, imagem e áudio), utilizando técnicas de Deep Learning e IA Generativa. O projeto tem como objetivo criar uma solução multimodal capaz de interpretar diferentes tipos de entrada e gerar respostas inteligentes para diferentes tipos de negócio.

Bloco Optativo Internacional

Explore novos horizontes por meio do nosso bloco optativo internacional que combina teoria avançada, aplicações práticas e networking internacional.

Em parceria com universidades de referência, como o ISEG – Universidade de Lisboa e a La Salle em Barcelona, você aprofunda conhecimentos em tendências globais, negociações e liderança. São vivências imersivas, com visitas técnicas e metodologias de ponta para enriquecer sua visão estratégica e ampliar suas oportunidades de atuação global.

Solicite os valores e condições no botão abaixo.

Consulte

La Salle

Curso: Habilidades Gerenciais & Negociação Internacional

O que você vai vivenciar:

- Liderança e negociação em ambientes multiculturais;
- Aulas práticas, estudos de caso e visitas empresariais;
- Desenvolvimento de competências gerenciais globais;
- Certificação internacional pela La Salle.

Campus:

- Barcelona, Espanha.



ISEG – Universidade de Lisboa

Curso: Leading to the Future

O que você vai vivenciar:

- Futurismo aplicado, cenários e estratégia;
- Design estratégico e inovação;
- Visitas técnicas e metodologia Scanning, Sensing & Acting;
- Certificação internacional do ISEG Executive Education.

Curso: Liderando com Impacto

O que você vai vivenciar:

- Liderança estratégica em ambientes voláteis;
- Simulador Harvard Everest Challenge;
- Laboratório de inovação e visitas empresariais;
- Competências de autoliderança, storytelling e gestão de equipes;
- Certificação internacional do ISEG;
- Uma jornada que transforma carreiras e abre portas para o futuro.



Da sala
de aula
pra sala
de decisão.

Contatos

labdata@fia.com.br

labdata.fia.com.br

Unidades

Unidade Paulista

Estação de metrô Brigadeiro

Avenida Paulista, 302, 5º andar – Bela Vista – São Paulo/SP

CEP 01310-000

Unidade Nações Unidas

Estação de metrô Pinheiros

Avenida Doutora Ruth Cardoso, 7.221 – Pinheiros – São Paulo/SP

CEP 05425-070

