

**PÓS-GRADUAÇÃO**



**ANÁLISE DE DADOS E**

**INTELIGÊNCIA ARTIFICIAL**

**DATA MINING**

**360 HORAS**

Extensão, Pós-Graduação e MBA



**lab.data**

# FIA - Fundação Instituto de Administração

A FIA dedica-se à geração de conhecimento aplicado e de novos conceitos para a excelência na administração empresarial e pública.

É uma referência de ensino desde o curso de graduação em Administração e Negócios até os cursos de Educação Executiva e em EAD, bem como nos projetos de consultoria e pesquisa aplicada.

A FIA olha para o futuro com determinação, mirando a construção de uma sociedade mais evoluída, consciente e justa, entendendo que o mundo está em evolução, em grande mudança, gerando oportunidades e desafios que construirão uma sociedade coerente com os novos tempos.



## BUSINESS SCHOOL

Graduação, Pós-graduação, MBA, Pós-MBA, Mestrado, Profissional, Cursos In-Company e EAD



## CONSULTING

Consultoria personalizada que oferece soluções baseadas em seu problema de negócio



## RESEARCH

Atualização dos conhecimentos e do material didático oferecidos nas atividades de ensino



Único curso de graduação a receber notas máximas



A primeira escola brasileira a ser finalista da maior competição de MBA do mundo



Única Business School brasileira a figurar no ranking LATAM



Signatária do pacto global da ONU



Membro fundador da ANAMBA – Associação Nacional dos MBAs



Credenciada pela AMBA – Association of MBAs



Credenciada ao Executive MBA Council



Filiada a AACSB – Association to Advance Collegiate Schools of Business



Filiada e EFMD – European Foundation for Management Development



Referência em cursos de MBA nas principais mídias de circulação



**3 vezes eleita a melhor escola de Negócios**  
do Brasil pela ÉPOCA



**Mais de 8 mil consultorias realizadas**  
em mais de 1 mil empresas



**Mais de 40 anos de Experiência com Pesquisa**  
no Brasil e no mundo

# Labdata – Laboratório de Análise de Dados

O **Laboratório de Análise de Dados – Labdata** é um Centro de Excelência que atua nas áreas de ensino, pesquisa e consultoria em análise de informação utilizando técnicas de Big Data, Analytics, Inteligência Artificial e Metaverso. É dirigido pela **Profª Dra. Alessandra de Ávila Montini**.

Veja porque escolher o Labdata:



**Centro de Excelência e pioneiro** no lançamento dos cursos de Big Data, Analytics e Inteligência Artificial no Brasil



**+ 10 anos de atuação**  
**+ 15.000 alunos formados**  
**+130 turmas de Pós e MBA**



**Lifelong Learning**  
**Atividades Gratuitas:**  
Hackathons, Palestras, Cursos de Extensão e Eventos



Número limitado de alunos por turma. **Networking profissional** qualificado e **painel de vagas** com as principais empresas



Cursos de Pós e MBA com carga horária mais completa do mercado, **plantão de dúvidas semanal** com o time de professores do Labdata.



Conteúdo programático completo com **aulas práticas e resolução de cases reais**.



Docentes com sólida formação acadêmica: **doutores e mestres** em sua maioria com larga experiência de mercado



**Certificados reconhecidos pelo MEC e mesma certificação** para os cursos online e presencial



Localização de **fácil acesso** de carro ou transporte público.



**Profª Drª Alessandra Montini**  
Diretora do Labdata FIA

# NOSSOS LABORATÓRIOS

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**, nossos laboratórios contam com equipamentos de última geração, conforto e segurança para nossos alunos e professores.



Área para Networking



Espaços para Estudos



Salas de Reunião



Áreas de Descompressão



Espaço para Refeições



Quadra de Basquete e Praia



# SOBRE O CURSO

## OBJETIVO

Apresentar, com base na resolução de casos, as principais técnicas de Analytics, Inteligência Artificial, Machine Learning e Deep Learning utilizadas para explorar e encontrar padrões escondidos nos dados, transformando dados brutos em valor para o negócio;

Preparar o aluno para trabalhar como Cientista de Dados com foco em análise de dados e modelagem de algoritmos; Resolver problemas complexos por meio de várias metodologias.

## PERFIL DO ALUNO

Profissionais de todas as áreas que desejam adquirir ou aprimorar seus conhecimentos em Modelagem de Dados, Analytics, Estatística Aplicada, Inteligência Artificial, Machine Learning e Deep Learning.

## 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 presenciais serão realizadas em um ambiente propício para aprendizagem;
- As aulas na modalidade EaD ao vivo serão transmitidas por meio de uma plataforma digital;
- Todas as aulas terão a presença de um professor titular com apoio de pelo menos um professor assistente;
- Serão utilizados diversos recursos tecnológicos para o melhor entendimento dos conceitos;
- Todas as aulas serão gravadas e permanecerão disponíveis em nossa biblioteca digital até o final do curso.

## **REQUISITOS PARA EAD AO VIVO**

- As aulas serão transmitidas através de uma plataforma digital;
- Conexão com a internet - banda larga com ou sem fio com pelo menos 1MB de velocidade final, e para acompanhar as aulas ao vivo e para a realização de exercícios;
- Computador com configuração mínima: Dual Core 2Ghz ou superior (13/i5/17 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;
- Caso haja determinação legal para aplicação de avaliação presencial, ela será realizada em uma das unidades educacionais da FIA, em São Paulo/SP.

## **MATRIZ CURRICULAR**

A ampla carga horária permite a resolução de vários exercícios e a realização de estudos de casos. Nossa proposta é apresentar um curso prático oferecendo toda a base teórica necessária para a adequada tomada de decisão.

As informações estão sujeitas a alterações sem aviso prévio.

# TÉCNICAS PARA ANÁLISE DE DADOS ESTRUTURADOS

## >Introdução à Ciência de Dados e Inteligência Artificial

- Conceito de Ciência de Dados
- Conceito de Analytics
- Conceito de 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

- Trajetória dos Líderes
- Primeiras experiências de liderança;
- Principais desafios enfrentados e fatores de sucesso;
- Liderança em ambientes de mudança e incerteza;
- Liderança para gestão da diversidade e inclusão;
- O papel da liderança para ética e responsabilidade social.

## >Fundamentos de Estatística

- Conceito de Base de Dados
- Dados Estruturados vs. Não Estruturados
- Variáveis Quantitativas Discretas e Contínuas
- Variáveis Qualitativas Nominiais e Ordinais
- Estatística Descritiva vs. Estatística Inferencial
- População e Amostra
- Principais Tipos de Amostras
- Aplicações

## >Introdução ao Python

- Criação de Objetos
- Realização de Cálculos
- Módulos e Bibliotecas
- Funções
- Verificações Lógicas
- Listas e Vetores
- Métodos

## >Estruturação de Dados para Análise e Modelagem

- Leitura de Arquivos de Dados
- Tipagem de Colunas
- Seleção de Colunas
- Filtros
- Empilhamento
- Agrupamento
- Ordenação
- Cruzamento (Left Join, Right Join, Inner Join, Full Outer Join)
- Remoção de Duplicatas
- Tratamento de Valores Ausentes
- Engenharia de Variáveis
- Aplicações

## >Análise Exploratória de Dados

- Tabelas de Frequências Absolutas e Relativas
- Medidas de Posição: Média, Média Ponderada, Mediana, Mediana Ponderada, Moda, Quartis, Percentis
- Medidas de Dispersão: Desvio Padrão, Desvio Absoluto Médio, Desvio Absoluto Mediano, Amplitude
- Conceito de Outlier
- Aparamento de Variáveis
- Análise Bidimensional: Qualitativa Versus Qualitativa
- Análise Bidimensional: Qualitativa Versus Quantitativa
- Análise Bidimensional: Quantitativa Versus Quantitativa
- Aplicações

## >Visualização de Dados

- Gráfico de Barras/Colunas
- Gráfico de Barras/Colunas Empilhadas (Absoluto e Relativo)
- Gráfico de Setores
- Histograma
- Boxplot
- Gráfico de Linha
- Simetria de Distribuições para Variáveis Quantitativas
- Boas Práticas em Visualização de Dados
- Aplicações

## >Probabilidade e Inferência Estatística

- Conceito de Probabilidade
- Casos de Uso da Probabilidade no Dia a Dia
- Interpretação de Probabilidade como Medida de Frequência
- Distribuições de Variáveis
- Distribuição Normal
- Teorema do Limite Central
- Intervalos de Confiança para Médias e Proporções
- Cálculo de Tamanho Amostral para Estimação de Médias e Proporções
- Correção para Populações Finitas
- Aplicações

## >Modelagem de Projeção: Regressão Linear

- Introdução à Modelagem Supervisionada
- Conceitos de Variável Resposta e Variável Explicativa
- Equação da Reta
- Regressão Linear Simples e Múltipla
- Método dos Mínimos Quadrados
- Análise de Significância de Variáveis: p-Valor
- Seleção de Variáveis (Stepwise Backward)
- Interpretação dos Coeficientes
- Colinearidade (VIF)
- Inclusão de Variáveis Qualitativas em Dummies
- Aplicações

## >Modelagem de Projeção: Árvore de Regressão

- Estrutura do Algoritmo de Árvore de Regressão
- Otimização de MAE ou MSE
- Interpretação do Modelo
- Conceito de Hiperparâmetro
- Hiperparâmetros da Árvore de Regressão
- Otimização de Hiperparâmetros: Grid Search e Random Search
- Validação Cruzada e Validação Cruzada Aninhada
- Aplicações

### >Modelagem de Projeção: Florestas Aleatórias

- Estrutura do Algoritmo de Floresta Aleatória
- Hiperparâmetros do Algoritmo
- Explicabilidade de Modelos de Ensemble: Feature Importance, SHAP, LIME
- Aplicações

### >Modelagem de Projeção: Métodos de Impulsioneamento

- Estrutura dos Algoritmos de Impulsioneamento
- Variantes de Impulsioneamento: Gradient Boosting, XGBoost, LightGBM e CatBoost
- Hiperparâmetros dos Algoritmos
- Aplicações

### >Validação e Diagnóstico de Modelos de Projeção

- Superajuste de Modelos
- Validação Hold-Out
- $R^2$  e  $R^2$  Ajustado
- MAE, MSE, MAPE
- Normalidade e Homocedasticidade dos Resíduos (para Regressão Linear)

### >Modelagem de Projeção Temporal: Modelos Lineares Estacionários

- Conceito de Série Temporal
- Estacionariedade vs. Ruído Branco
- Equações dos Modelos AR, MA e ARMA
- Interpretação dos Modelos
- Identificação Automática do Melhor Modelo
- Previsão
- Validação Temporal
- Aplicações

### >Modelagem de Projeção Temporal: Modelos Lineares Não Estacionários

- Tendência e Sazonalidade
- Diferenciação Simples e Diferenciação Sazonal
- Equações dos Modelos ARIMA e SARIMA
- Interpretação dos Modelos
- Identificação Automática do Melhor Modelo
- Previsão
- Validação Temporal
- Aplicações

### >Modelagem de Projeção Temporal: Modelos Lineares com Variáveis Exógenas

- Extensão dos Modelos ARIMA e SARIMA com Inclusão de Variáveis Explicativas
- Aplicações

### >Modelagem de Projeção Temporal: Engenharia de Features

- Criação de Features Temporais
- Extensão dos Modelos Baseados em Árvores para Projeção Temporal
- Aplicações

### >Modelagem de Classificação: Regressão Logística

- Função Logística
- Regressão Logística Simples e Múltipla
- Análise de Significância de Variáveis: p-Valor
- Seleção de Variáveis (Stepwise Backward)
- Interpretação dos Coeficientes
- Colinearidade (VIF)
- Agrupamento em Faixas de Probabilidade
- Definição de Ponto de Corte
- Aplicações

### >Modelagem de Classificação: Árvore de Decisão Binária

- Extensão da Árvore de Regressão para Variáveis Resposta Qualitativas Binárias
- Medidas de Impureza
- Interpretação do Modelo
- Aplicações

### >Modelagem de Classificação: Árvore de Decisão Multinomial

- Extensão da Árvore de Decisão Binária para Variáveis Resposta Multinomiais
- Interpretação do Modelo
- Aplicações

### >Modelagem de Classificação: Florestas Aleatórias e Impulsioneamento

- Extensão dos Algoritmos de Floresta Aleatória e Impulsioneamento para Variáveis Resposta Qualitativas (Binárias ou Multinomiais)
- Aplicações

## >Validação e Diagnóstico de Modelos de Classificação

- Matriz de Confusão
- Acurácia
- Sensibilidade (Recall)
- Especificidade
- KS
- Área Abaixo da Curva ROC (AUC)
- Balanceamento para Melhoria de Performance
- Calibração de Probabilidades
- Aplicações

## >Modelagem de Segmentação: Conceitos Gerais

- Introdução à Modelagem Não Supervisionada
- Objetivo da Tarefa de Segmentação
- Medidas de Distância
- Padronização de Variáveis Quantitativas
- Aplicações

## >Modelagem de Segmentação: Algoritmo Hierárquico

- Estrutura do Algoritmo Hierárquico
- Análise de Dendrograma
- Definição da Quantidade de Clusters
- Métodos de Ligação: Complete, Single, Average, Ward
- Análise Exploratória dos Clusters
- Aplicações

## >Modelagem de Segmentação: Algoritmos de Partição

- Estrutura dos Algoritmos de Partição
- Algoritmo K-médias
- Algoritmo K-medoides
- Representação Gráfica dos Clusters
- Análise Exploratória dos Clusters
- Aplicações

## >Modelagem de Segmentação: DBSCAN

- Estrutura do Algoritmo DBSCAN
- Definição de Ruído
- Teste de Hiperparâmetros
- Aplicações

## >Aprendizado de Máquina Automatizado (Auto ML)

- Conceitos de Auto ML
- Principais Pacotes de Auto ML em Python
- Comparação e Escolha do Melhor Modelo
- Aplicações

## >MLOps

- Conceito de Sistemas de ML
- Conceito de MLOPs
- Perfil do profissional de MLOPs
- Ciclo de vida de um modelo de ML
- Implementação de um modelo de ML
- Monitoramento de modelos de ML
- Casos de Uso

# TÉCNICAS PARA ANÁLISE DE DADOS NÃO ESTRUTURADOS

## >Introdução ao Deep Learning e as Redes Neurais

- Entender a evolução da Inteligência Artificial
- Entender os conceitos básicos de Redes Neurais e as principais redes mais utilizadas (Redes Densas, Redes Convolucionais, Redes Recorrentes)
- Entender os cálculos que são realizados pelo algoritmo Perceptron
- Desenvolver os códigos para rodar o algoritmo Perceptron
- Apresentação do Projeto Deep Learning

## >Redes Neurais Densas

- Entender os conceitos de Redes Neurais Densas (algoritmo MLP - Multi Layer Perceptron)
- Entender as funções de ativação
- Entender como o método do Gradiente Descendente reduz o erro
- Entender o algoritmo de Backpropagation (regra da cadeia e cálculo das derivadas parciais)
- Entender o funcionamento da biblioteca TensorFlow (infraestrutura de manipulação de tensores)
- Entender o 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

- Entender como as imagens digitais são representadas e como podem ser processadas utilizando diferentes bibliotecas do Python
- Entender as operações básicas como negação, adição, subtração, multiplicação, divisão, exponenciação e rotação
- Entender a operação de convolução para remover ruído, realçar detalhes, suavizar imagens, detectar bordas, adicionar borda (padding) e extrair características
- Entender e desenvolver códigos para aplicar as principais técnicas de pré-processamento de imagens
- Entender as principais camadas dentro das redes convolucionais (convolucional, pooling, flatten, densa), e como elas são adequadas para tarefas de processamento de imagens devido à sua capacidade de aprendizado de características.
- Entender as técnicas de regularização como dropouts, regularização L1/L2, early stopping
- Entender como data augmentation pode ajudar no processo de treinamento
- Entender como podemos utilizar a técnica de transfer learning para melhorar e agilizar o treinamento dos modelos (VGG 16, ResNet)
- Entender como podemos interpretar os resultados das redes convolucionais (GradCAM)
- Desenvolver os códigos para aplicar e comparar todos os conceitos estudados neste tópico

## >Redes Neurais Recorrentes

- Entender a estrutura básica das redes neurais recorrentes (RNNs) "simples" e entender suas limitações
- Entender como as dependências temporais com feedback loop podem auxiliar no treinamento
- Entender como problemas de vanishing e exploding gradients impactam no aprendizado
- Entender os princípios das Long Short-Term Memory (LSTMs) com Cell State, Hidden State, Input Gate, Forget Gate e Output Gate
- Entender as 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

## >Projeto Deep Learning

- Integração dos conhecimentos adquiridos ao longo do curso, desenvolvendo um modelo de classificação e um modelo generativo utilizando Keras.
- Trabalhar com dados no formato de imagens ou texto, desde o pré-processamento até a implementação final de modelos.
- Aplicar técnicas para adaptar os dados de classificação para um modelo generativo, explorando sua versatilidade e uso prático.

## >Inteligência Artificial Generativa

- Entender a diferença entre modelos discriminativos e generativos
- Entender os diferentes tipos de aplicações (entrada e saída do modelo)
- Casos de uso IA Generativa (texto, imagem, som, multimodal)
- Evolução da IA Generativa ao longo do tempo
- Arquitetura de GAN: gerador e discriminador
- Arquitetura DCGAN: convolução transposta, leaky relu, mode collapse
- Arquitetura Stable Diffusion
- Desenvolver aplicações que comparem as arquiteturas apresentadas para geração de imagem e criação de dados sintéticos para treinamento
  
- Arquitetura Transformer:
  1. Introdução ao GPT (capacidade, limitações e casos de uso)
  2. Mecanismo de atenção e autoatenção
  3. Entendendo como o GPT Compreende a Linguagem
  4. Entender a representação Vetorial de Palavras (técnicas de Word Embedding, Positional Encoding, Tokenization)
  5. Desenvolver aplicação utilizando GPT para geração de textos criativos
  
- Prompt Engineering:
  1. Introdução ao Prompt Engineering
    - Definição de Prompt Engineering
    - Importância na Interação com Modelos de Linguagem
  2. Por que estudar o Prompt Engineering?
    - Vantagens e Benefícios
    - Impacto na Eficiência e Eficácia da Comunicação com GPT
  3. Termos, Conceitos e Jargões da Engenharia de Prompt
    - Explanations (Explicações)
    - Completion Prompts (Prompts de Completamento)
    - System Prompts (Prompts do Sistema)
    - Role of Temperature and Max Tokens

#### 4. Engenharia de Prompts:

- Tarefas mais comuns: Resumo, Tradução, Código, Extração de Informação
- Estado da Arte em Engenharia de Prompt: direção, Exemplos, Formatos, Meta Prompt
- Técnicas de Engenharia de Prompt: Zero, One, Few-shot, Listas, Chain of Thought

#### 5. Métricas de Avaliação de Prompt:

- Métricas para Avaliar a Qualidade das Respostas
- BLEU, ROUGE, METEOR, etc.

#### 6. RAGs (Retrieval-Augmented Generation) e Agentes de AI:

- Introdução ao conceito de RAGs
- Funcionamento e casos de uso
- Papel de Agentes de AI na interação com modelos generativos
- Exemplos práticos de uso

#### 7. Aplicações e exemplos:

- Chatbots, assistência na escrita, perguntas e respostas

## BLOCO OPTATIVO - INTERNACIONAL

Explore novos horizontes!

Adquira habilidades em instituições internacionais por meio do nosso bloco optativo internacional.

Desperte para oportunidades inéditas e destaque-se como líder no mercado, se imergindo em experiências enriquecedoras em faculdades de renome internacional. Conheça nossas instituições conveniadas [aqui](#).

Amplie suas possibilidades e trilhe o caminho do sucesso global!

## DIREÇÃO



**PROF.ª DR.ª**  
**ALESSANDRA DE ÁVILA MONTINI**

Diretora do LABDATA-FIA, apaixonada por dados e pela arte de lecionar. Têm muito orgulho de ter criado na FIA cinco laboratórios para as aulas de Big Data e inteligência Artificial. Possui mais de 20 anos de trajetória nas áreas de Data Mining, Big Data, Inteligência Artificial e Analytics. Cientista de dados com carreira realizada na Universidade de São Paulo. Graduada e mestra em estatística aplicada pelo IME-USP e doutora pela FEA-USP. Com muita dedicação chegou ao cargo de professora e pesquisadora na FEA-USP, ganhou mais de 30 prêmios de excelência acadêmica pela FEA-USP e mais de 30 prêmios de excelência acadêmica como professora dos cursos de MBA da FIA. Orienta alunos de mestrado e de doutorado na FEA-USP. Membro do Conselho Curador da FIA, Coordenadora de Grupos de Pesquisa no CNPQ, Parecerista da FAPESP e Colunista de grandes Portais de Tecnologia.



[linkedin.com/in/alessandramontini](https://www.linkedin.com/in/alessandramontini)

# Informações

11 97132-3535

labdata@fia.com.br

labdata.fia.com.br

**UNIDADE PAULISTA - ESTAÇÃO BRIGADEIRO**

Avenida Paulista, 302, 5º andar - CEP 01310-000 - Bela vista - São Paulo/SP

**UNIDADE NAÇÕES UNIDAS - ESTAÇÃO PINHEIROS**

Avenida Doutora Ruth Cardoso, 7.221 - CEP 05425-070 - Pinheiros - São Paulo/SP