



Pós-Graduação em Processamento Distribuído com Plataformas Modernas de Dados (Apache Spark e Databricks)

366 H/A



O QUE É A DATA SCIENCE ACADEMY?



A Data Science Academy (DSA) é uma instituição brasileira criada em 2016, especializada em oferecer cursos e treinamentos na área de Ciência de Dados, Inteligência Artificial, Engenharia de Dados e áreas correlatas. Fundada com o objetivo de capacitar profissionais e estudantes com conhecimentos e habilidades necessárias para atuar no mercado de trabalho, a DSA é reconhecida por sua abordagem prática e voltada para o mercado, com cursos de alta qualidade.

Este Programa de Pós-Graduação Lato Sensu, com certificação reconhecida pelo MEC no Brasil, amplia ainda mais as oportunidades de qualificação para você que busca alavancar sua carreira na área de tecnologia e Ciência de Dados.



O QUE É A

PÓS-GRADUAÇÃO EM PROCESSAMENTO DISTRIBUÍDO COM PLATAFORMAS MODERNAS DE DADOS?

A PÓS-GRADUAÇÃO EM PROCESSAMENTO DISTRIBUÍDO COM PLATAFORMAS MODERNAS DE DADOS da DSA é um programa de extensão Lato Sensu para quem deseja buscar especialização no processamento distribuído de dados com o Apache Spark e Databricks. O programa é reconhecido pelo MEC e oferece título de especialista.

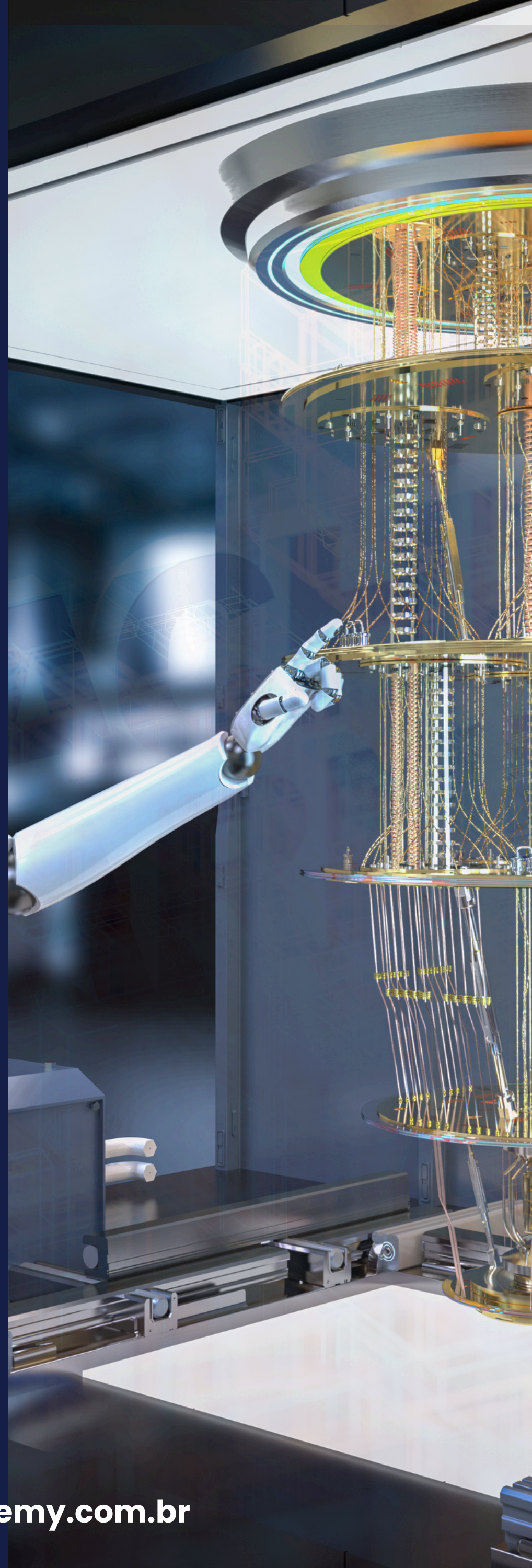
Confira abaixo a descrição das disciplinas.



PIPELINES DE ETL E MACHINE LEARNING COM APACHE SPARK

Esta disciplina é totalmente prática e focada na construção e otimização de pipelines de dados escaláveis. Ela aborda exploração de dados com Spark SQL, desenvolvimento de pipelines ETL robustos para clusters Spark e processamento de dados em tempo real. Inclui um projeto integrando LLMs com LangChain e Apache Spark para análise textual avançada.

84 H/A





PYSPARK E APACHE KAFKA PARA PROCESSAMENTO DE DADOS EM BATCH E STREAMING

Esta é uma disciplina completa que cobre desde a preparação do ambiente com PySpark até ETL avançado com Python e integração de dados em tempo real com Apache Kafka. Foca no processamento distribuído e em técnicas de limpeza e análise de dados em lote e em tempo real, atendendo à crescente demanda por habilidades essenciais para Engenheiros de Dados modernos.

90 H/A



MODELAGEM DE SÉRIES TEMPORAIS E REAL-TIME ANALYTICS COM APACHE SPARK E DATABRICKS

Esta Disciplina irá ampliar suas habilidades em análise e modelagem de séries temporais para interpretar e prever tendências, padrões e anomalias em dados. Isso pode ser um diferencial decisivo em setores como finanças, energia, comércio eletrônico e agronegócio. A capacidade de extrair insights acionáveis e criar modelos preditivos robustos, especialmente em tempo real, é uma competência valiosa para você como profissional de dados.

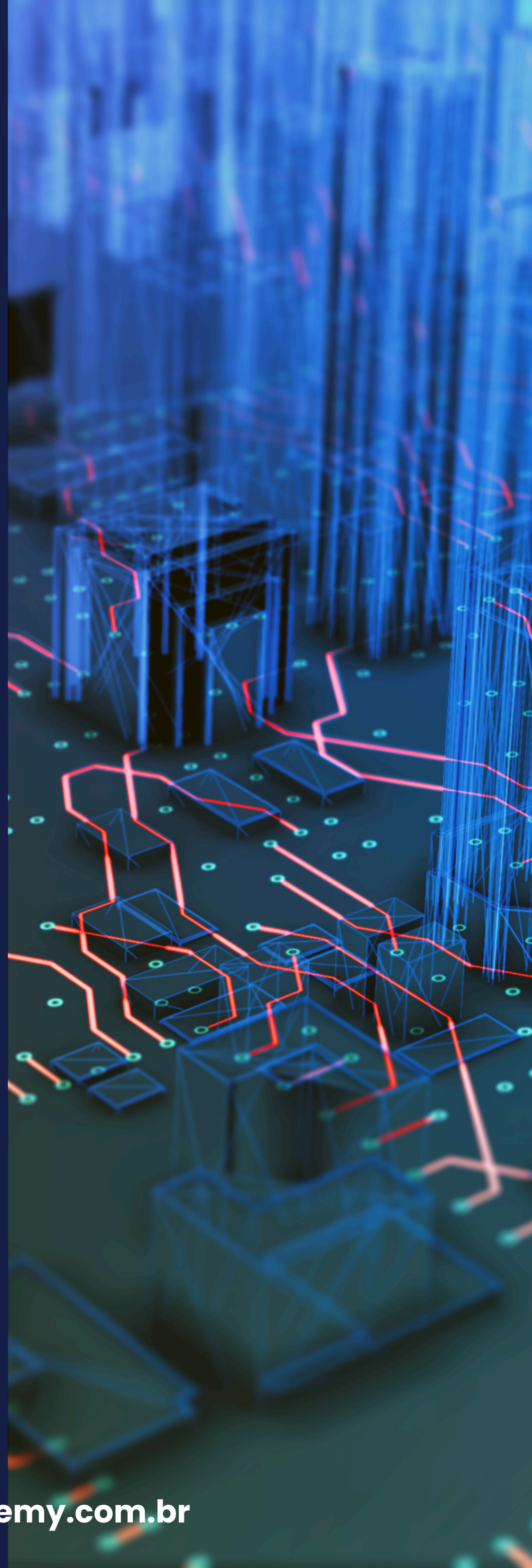
96 H/A



INFRAESTRUTURA DE DADOS, IA E AGENTES DE IA COM DATABRICKS

Esta disciplina é focada no domínio da Plataforma Databricks para construir pipelines de dados eficientes, treinar modelos de IA e implementar agentes inteligentes. Com projetos práticos e cenários reais, irá capacitar você a criar soluções escaláveis que gerem valor para as empresas, unindo infraestrutura de dados e Inteligência Artificial.

96 H/A





CURSOS BÔNUS DE APERFEIÇOAMENTO PROFISSIONAL

Além das disciplinas você terá acesso a cursos de Aperfeiçoamento Profissional incluindo:

- Planejando sua Carreira para as Profissões do Futuro
- Introdução à Lógica de Programação
- Sistema Operacional Linux, Docker e Kubernetes
- Governança de Dados
- Soft Skills – Desenvolvendo Suas Habilidades Comportamentais
- E-Gov Analytics
- Machine Learning com JavaScript e Go
- Data Science e Machine Learning com Linguagem Julia

**Com Certificado Extra Para
Cada Curso Bônus**



MÓDULO DE PÓS

Este módulo especial irá conduzir você por todo o curso de Pós-Graduação, além de ofertar mais 20 horas de atividades extra-curriculares, incluindo atividades de orientação de carreira e organização de estudo, atividades complementares, apoio de carreira e projeto final.

20 H/A



www.datascienceacademy.com.br



SOBRE A PÓS GRADUAÇÃO

PGPDPM - PÓS-GRADUAÇÃO EM PROCESSAMENTO DISTRIBUÍDO COM PLATAFORMAS MODERNAS DE DADOS

Prazo de Acesso: 12 meses

Modalidade: 100% online

Quem Pode Participar?

Qualquer pessoa pode participar do processo de seleção desde que atenda 2 requisitos mandatórios: Ter nacionalidade brasileira e ter concluído um curso de graduação reconhecido pelo MEC em qualquer área. Será necessário apresentar documento válido de identificação e o certificado de conclusão do curso de graduação durante o processo de seleção.

Qual o prazo para obter o certificado da Pós?

O prazo mínimo é de 4 meses e o máximo é de 12 meses. O aluno terá que concluir as atividades e o módulo da Pós entre 4 e 12 meses além de entregar o projeto final. Entretanto, todo o material de estudo das disciplinas, ficará disponível no perfil do aluno por 24 meses.

Para os alunos matriculados na Formação Apache Sparke e Databricks 4.0 descontaremos 100% do valor que o aluno pagou na Formação e o aluno paga somente a diferença se quiser migrar para o Programa de Pós.



Inscreva-se Agora!



Junte-se a nós.

São mais de 9 anos e mais de 1 milhão de alunos capacitados em todos os cantos do Brasil e do Mundo!

www.datascienceacademy.com.br