DIRETORIA ACADÊMICA

PLANO DE ENSINO

Curso: Análise e Desenvolvimento de Sistemas		
Unidade Curricular: Ciência de dados e BIG DATA		**Carga Horária:** 80h
Modalidade: Presencial (X) Semipresencial ( ) EAD ( )		
Período: 4º	**Ano/ Semestre Letivo:** 2026.2	
Professor (es): Fernando Marlon Soares Figueiredo		

1. EMENTA

Estudo prático dos fundamentos de Ciência de Dados e Big Data. Coleta, limpeza, transformação e análise de dados. Mineração de dados, aprendizado de máquina e visualização. Ferramentas práticas como Python (Pandas, Scikit-learn), Power BI, Tableau, Hadoop e Spark.

2. OBJETIVO (S)

2.1. Geral:

Capacitar os alunos a aplicar técnicas modernas de Ciência de Dados para solucionar problemas reais por meio de análise de dados e ferramentas de suporte à decisão.

2.2. Específicos:

- Compreender fundamentos de Ciência de Dados, Big Data e mineração de dados.
- Aplicar métodos de análise exploratória de dados, limpeza e transformação de dados.
- Implementar modelos de aprendizado de máquina com bibliotecas Python.
- Utilizar ferramentas de visualização como Tableau e Power BI para criar dashboards interativos

3. CONTEÚDO PROGRAMÁTICO									
U	**Conteúdos**	**Conhecimentos**	**Habilidades**	**Atitudes**	**Nº Horas /Aulas**				
					T	P	L	SP	EAD
I	Introdução à Ciência de Dados e seu impacto nas organizações. Ciclo de Vida dos Dados e aplicações práticas. Fundamentos de Big Data: conceitos dos 4Vs (Volume, Velocidade, Variedade e Valor) e desafios tecnológicos. Arquitetura de Big Data: Introdução a sistemas distribuídos como Hadoop e Spark	Compreender o papel da Ciência de Dados na tomada de decisão organizacional. Reconhecer as etapas do ciclo de vida dos dados e sua importância na análise de dados. Identificar as características do Big Data e os desafios associados ao processamento de grandes volumes de dados. Arquitetura básica de sistemas de Big Data (Hadoop e Spark)	Aplicar conceitos de Ciência de Dados para resolver problemas baseados em dados reais. Explorar exemplos práticos dos 4Vs para demonstrar desafios e oportunidades. Explorar ambientes e ferramentas de processamento distribuído em atividades orientadas.	Demonstrar iniciativa na exploração de ferramentas e conceitos de Ciência de Dados. Adotar uma postura colaborativa no desenvolvimento de soluções baseadas em dados. Ser crítico e atento aos impactos sociais e éticos do uso de grandes volumes de dados.	15	10			
II	Técnicas de coleta e pré-processamento de dados (limpeza, normalização, remoção de ruídos). Ferramentas de armazenamento e análise em sistemas distribuídos como Hadoop e Spark. Fundamentos da mineração de dados e suas aplicações práticas. Algoritmos de mineração de dados: Regras de Associação (Apriori) e Clusterização.	Etapas de coleta, limpeza e preparação de dados. Reconhecer a importância do pré-processamento na qualidade dos resultados de análise. Identificar técnicas e ferramentas de mineração de dados para extrair padrões significativos. Compreender o uso de algoritmos como Apriori e K-means na análise de grandes conjuntos de dados.	Desenvolver pipelines de coleta e pré-processamento usando bibliotecas Python (Pandas e NumPy). Implementar algoritmos de mineração de dados para identificar padrões e associações. Utilizar bibliotecas Python para manipulação e análise de conjuntos de dados.	Manter uma postura de persistência frente a problemas técnicos e desafios complexos. Colaborar com colegas para realizar análises e interpretar resultados. Demonstrar senso crítico na aplicação de técnicas de mineração, respeitando questões éticas e de privacidade.	15	15			

III	Visualização de Dados: uso de ferramentas como Power BI, Tableau e bibliotecas Python (Matplotlib, Seaborn). Criação de dashboards interativos para apoiar a tomada de decisões. Introdução ao aprendizado de máquina: regressão, classificação e clusterização. Aplicações práticas de aprendizado de máquina em projetos reais.	Reconhecer a importância da visualização na comunicação de insights e tomada de decisões. Entender os fundamentos de aprendizado de máquina e como aplicá-los em problemas reais. Fundamentos de regressão, classificação e clusterização.	Desenvolver gráficos interativos e dashboards que sintetizem informações complexas. Implementar modelos básicos de aprendizado de máquina com Scikit-learn. Usar ferramentas modernas para explorar dados e comunicar resultados.	Valorizar a clareza e a organização ao comunicar resultados por meio de visualizações. Demonstrar criatividade e inovação na aplicação de modelos e técnicas. Ser proativo na busca por soluções práticas e replicáveis para problemas de negócio de dados	10	15			
Subtotal da Carga Horária					**40**	**40**			
Total da Carga Horária					**80**				

Legenda: U - unidade **T –** quantidade de aula(s) teórica(s) por unidade(s); **P-** quantidade de aula(s) práticas(s) por unidade(s); **L - quantidade de aula(s) por unidade(s) de Prática Pedagógica (exclusivo para o Curso de Educação Física - Licenciatura**); **SP** - quantidade de aulas(s) semipresenciais; **EAD** - quantidade de aula(s) à distância.

4. METODOLOGIA

Aulas expositivas e dialogadas, estudos de caso, atividades práticas em laboratório, utilização de software específico para manipulação e análise de dados (Python, R, Hadoop, Spark), projetos em grupo e atividades no Ambiente Virtual de Aprendizagem (AVA). Serão desenvolvidos mini-projetos práticos com bases de dados reais, promovendo aprendizagem baseada em problemas e análise de cenários do mercado.

5. RECURSOS DIDÁTICOS

Para a disciplina de Ciência de Dados e Big Data, serão utilizados diversos recursos didáticos para facilitar o aprendizado. As aulas serão complementadas pelo uso de um projetor multimídia para exibição de slides e vídeos, e um quadro branco com pincéis para diagramas e explicações. O laboratório de informática estará disponível com acesso a softwares específicos como Python, R, Hadoop e Spark, proporcionando um ambiente prático para o desenvolvimento de habilidades técnicas. Além disso, os alunos terão acesso a uma ampla bibliografia digital e artigos científicos que suportam o aprofundamento teórico dos temas abordados. O Ambiente Virtual de Aprendizagem (AVA) será utilizado para disponibilizar materiais de estudo, exercícios e fóruns de discussão, promovendo uma aprendizagem contínua e interativa.

6. AVALIAÇÃO (ÕES)

A avaliação será contínua e processual, mediante somatório de diferentes atividades, como trabalhos, resumos, relatórios, projetos integradores e provas escritas. Serão realizadas três principais verificações de aprendizagem, uma por unidade. As avaliações incluirão atividades práticas de análise de dados, construção de dashboards e implementação de modelos simples de aprendizado de máquina.

7. BIBLIOGRAFIA

7.1. Básica:

BASSO, D. G. Big Data. 1ª Ed. Curitiba: Intersaberes, 2020. [acervo digital]
ELMASRI, Ramez; NAVATHE, Shamkant B. Sistemas de banco de dados. 7ª Ed. São Paulo: Pearson, 2018. [acervo digital]
VALDATI, A. B. Inteligência Artificial – IA. 1ª Ed. Curitiba: Intersaberes, 2020. [acervo digital]

7.2. Complementar:

ALVES, W. P. Banco de dados: Teoria e desenvolvimento. 2ª Ed. São Paulo: Érica, 2020. [acervo físico]
ARAÚJO, R. C. A. Urban data analytics, urban big data e IOT. Curitiba: Intersaberes, 2020. [acervo digital]
FORTA, Ben. SQL em 10 minutos por dia. 1ª Ed. São Paulo: Novatec, 2021. [acervo físico]
NIELD, Thomas. Introdução à Linguagem SQL: Abordagem prática para iniciantes. 1ª Ed. São Paulo: Novatec, 2016. [acervo físico]
REGO, B. L. Simplificando a governança de dados: Governe os dados de forma objetiva e inovadora. 1ª Ed. Rio de Janeiro: Brasport, 2020. [acervo digital]