

UNIVERSIDADE DE PERNAMBUCO		
UNIDADE – ESCOLA POLITÉCNICA DA UNIVERSIDADE DE PERNAMBUCO		
DISCIPLINA – CIÊNCIA ORIENTADA A DADOS		
CÓDIGO DA DISCIPLINA – FIS50		
CARGA HORÁRIA TOTAL – 60 HORAS TEÓRICAS		
EMENTA <i>Introdução ciência de dados suas aplicações em física, biologia, tecnologia, ciências sociais e engenharia, decomposição em valores singulares, transformada de Fourier, escassez e compressão de dados, regressão e seleção de modelos, aglomeração e classificação, redes neurais e aprendizagem de máquina, sistemas dinâmicos orientados a dados, teoria de controle linear, modelos balanceados, controle orientado a dados, modelos de ordem reduzida, interpolação paramétrica para modelos de ordem reduzida.</i>		
ÁREA/EIXO/NÚCLEO <i>CIÊNCIAS EXATAS FÍSICA DE MATERIAIS NÚCLEO PROFISSIONALIZANTE</i>	COMPETÊNCIA(S) 1. <i>Compreender as técnicas e transformações matemáticas para o tratamento de dados.</i> 2. <i>Compreender os métodos de tratamento e aproximação de dados.</i> 3. <i>Converter conjuntos de dados em modelos por meio de análise preditiva.</i> 4. <i>Implementar algoritmos de aprendizado de máquina para proposição de modelos físicos e matemáticos.</i>	HABILIDADES COMPETÊNCIA 1 <ul style="list-style-type: none"> • Compreender o método de aproximação de matrizes, suas propriedades matemáticas e manipulações. • Compreender o método de mínimos quadrados e regressão, a análise de componentes principais (ACP) e o truncamento e alinhamento. • Realizar a decomposição aleatória em valores singulares, decomposições de tensores e matrizes de dados. • Entender e aplicar o método de decomposição em valores singulares (DVS). • Compreender e aplicar as transformadas de Fourier e Wavelets: transformação discreta de Fourier (TDF) e transformação rápida de Fourier (TRF), transformação de equações diferenciais parciais, transformação de Gabor e espectrograma, wavelets e análise de alta resolução, transformações 2d e processamento de imagem. COMPETÊNCIA 2 <ul style="list-style-type: none"> • Compreender o conceito de escassez e detecção compactada: geometria da compressão, regressão esparsa, representação esparsa, análise robusta de componentes principais (ARCP), sensor disperso. • Compreender o mecanismo de regressão e seleção de modelo: ajuste de curva clássico, regressão não-linear e redução de gradiente, regressão e $ax = b$, sistemas super e sub-



5. Domine as melhores práticas para o controle de modelos e interpolação.

determinados, otimização como pedra angular da regressão, abordagem de Pareto e Lex Parsimoniae, seleção de modelo: validação cruzada, seleção de modelo e critérios de informação.

COMPETÊNCIA 3

- Compreender os fundamentos do armazenamento em aglomerados e classificação de dados: seleção de recursos e mineração de dados, aprendizado supervisionado versus aprendizado não supervisionado
- Entender o funcionamento do aprendizado não-supervisionado: agrupamento de k-means, agrupamento hierárquico não-supervisionado, dendrograma, modelos de mistura e o algoritmo de expectativa e maximização, aprendizado supervisionado e discriminadores lineares, máquinas de vetores de suporte (MVS), árvores de classificação e floresta aleatória.
- Entender e aplicar os principais algoritmos da mineração de dados.

COMPETÊNCIA 4

- Compreender o funcionamento de redes Neurais e Aprendizado Profundo: redes de uma camada, redes multicamadas e funções de ativação, algoritmo de retropropagação, algoritmo estocástico de descida de gradiente, redes neurais convolucionais profundas, redes neurais para sistemas dinâmicos, diversidade de redes neurais.
- Compreender os conceitos de sistemas dinâmicos orientados a dados, visão geral, motivações e desafios, decomposição no modo dinâmico (DMD), identificação esparsa de dinâmica não-linear (SINDy), teoria do operador de Koopman, análise Koopman e dados.
- Entender os fundamentos da teoria do Controle Linear: controle de feedback em malha fechada, sistemas lineares invariantes no tempo, controlabilidade e observabilidade, controle de estado completo ideal, regulador quadrático linear (RQL), estimativa do estado completo, o filtro de Kalman, controle ideal baseado em sensor gaussiano quadrático linear (GQL), estudo de caso, pêndulo invertido em um carrinho, técnicas robustas de controle e domínio de frequência.

COMPETÊNCIA 5



- Compreender o funcionamento de modelos balanceados para controle: redução de modelo e identificação do sistema, redução equilibrada do modelo, identificação do sistema.
- Aplicar os fundamentos do controle orientado a dados: identificação não-linear do sistema para controle, controle de aprendizado de máquina, controle de procura extrema adaptável.
- Resolver problemas utilizando modelos de pedidos reduzidos (MPR): decomposição ortogonal adequada (DOA), equações diferenciais parciais, elementos de base ideais, expansão e dinâmica de soliton, formulação contínua, simetrias, rotações e traduções.
- Compreender os fundamentos da interpolação para MPRs paramétricas: gappy DOA, erro e convergência do gappy DOA, medições de gappy: minimize o número da condição, medições de gappy, variação máxima, DOA e o método de interpolação empírica discreta, implementação do algoritmo de interpolação empírica discreta, MPRs da aprendizagem de máquina.

CONTEÚDO PROGRAMÁTICO

1. *Decomposição em Valores Singulares (DVS): visão geral, aproximação de matrizes, propriedades matemáticas e manipulações, pseudo-inverso, mínimos quadrados e regressão, análise de componentes principais (ACP), truncamento e alinhamento, decomposição aleatória em valores singulares, decomposições de tensores e matrizes de dados.*
2. *Transformadas de Fourier e Wavelets: série de Fourier e transformadas de Fourier, transformação discreta de Fourier (TDF) e transformação rápida de Fourier (TRF), transformando equações diferenciais parciais, transformação de Gabor e espectrograma, wavelets e análise de alta resolução, transformações 2d e processamento de imagem.*
3. *Escassez e Detecção Compactada: introdução, detecção compactada, exemplos de detecção compactada, geometria da compressão, regressão esparsa, representação esparsa, análise robusta de componentes principais (ARCP), sensor disperso.*
4. *Regressão e Seleção de Modelo: ajuste de curva clássico, regressão não-linear e redução de gradiente, regressão e $ax = b$, sistemas super e subdeterminados, otimização como pedra angular da regressão, abordagem de Pareto e Lex Parsimoniae, seleção de modelo, validação cruzada, seleção de modelo: critérios de informação.*
5. *Armazenamento em Aglomerados e Classificação: seleção de recursos e mineração de dados, aprendizado supervisionado versus aprendizado não-supervisionado, aprendizado não-supervisionado, aglomerado de k-means.*
6. *Aglomerção Hierárquica Não-supervisionada: dendrograma, modelos de mistura e o algoritmo de expectativa e maximização, aprendizado supervisionado e discriminadores lineares, máquinas de vetores de suporte (MVS), árvores de classificação e floresta aleatória, principais algoritmos da mineração de dados.*



7. *Redes Neurais e Aprendizado Profundo: redes de uma camada, redes multicamadas e funções de ativação, algoritmo de retropropagação, algoritmo estocástico de descida de gradiente, redes neurais convolucionais profundas, redes neurais para sistemas dinâmicos, diversidade de redes neurais.*
8. *Sistemas Dinâmicos Orientados a Dados: visão geral, motivações e desafios, decomposição no modo dinâmico (DMD), identificação esparsa de dinâmica não-linear (SINDy), teoria do operador de Koopman, análise Koopman e dados.*
9. *Teoria do Controle Linear: controle de feedback em malha fechada, sistemas lineares invariantes no tempo, controlabilidade e observabilidade, controle de estado completo ideal, regulador quadrático linear (RQL), estimativa do estado completo, o filtro de Kalman, controle ideal baseado em sensor gaussiano quadrático linear (GQL), estudo de caso, pêndulo invertido em um carrinho, técnicas robustas de controle e domínio de frequência.*
10. *Modelos Balanceados para Controle: redução de modelo e identificação do sistema, redução equilibrada do modelo, identificação do sistema.*
11. *Controle Orientado a Dados: identificação não-linear do sistema para controle, controle de aprendizado de máquina, controle de procura extrema adaptável.*
12. *Modelos de Pedidos Reduzidos (MPR): decomposição ortogonal adequada (DOA), equações diferenciais parciais, elementos de base ideais, expansão e dinâmica de soliton, formulação contínua, simetrias, rotações e traduções.*
13. *Interpolação para MPRs Paramétricas: gappy DOA, erro e convergência do gappy DOA, medições de gappy, minimize o número da condição, medições de gappy, variação máxima, DOA e o método de interpolação empírica discreta, implementação do algoritmo de interpolação empírica discreta, MPRs da aprendizagem de máquina.*

BIBLIOGRAFIA

1. BRUNTON, S. L., and KUTZ, J. N. **Data-Driven Science and Engineering - Machine Learning, Dynamical Systems, and Control.** 1st Ed., Cambridge University Press, 2019.
2. Grus, J. **Data Science from Scratch: First Principles with Python.** 2nd Ed., O'Reilly Media, 2019.
3. EMC Education Services. **Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data.** 1st Ed., John Wiley & Sons, 2015.
4. Bruce, P., and Bruce, A. **Practical Statistics for Data Scientists.** 1st Ed., O'Reilly, 2017.
5. VanderPlas, J. **Python Data Science Handbook: Essential Tools for Working with Data.** 1st Ed., O'Reilly Media, 2015.

