



UNIVERSIDADE FEDERAL DO PARÁ
Instituto de Ciências Exatas e Naturais
Pós-Graduação em Ciência da Computação - PPGCC

1. Nome da Disciplina

Tópicos Especiais em Computação: Aprendizado de Máquina

2. Professor Responsável

Jefferson Magalhães de Moraes

3. Carga Horária

60 (sessenta) horas

4. Créditos

4 (quatro) créditos

5. Horário: Segunda e Quarta das 13:00 às 14:40

6. Pré-requisitos

Programação, conceitos básicos de probabilidade, estatística e álgebra linear.

7. Objetivos

Proporcionar ao aluno a capacidade de compreender as principais técnicas utilizadas na área de aprendizado de máquina. Mais especificamente, ao final do curso o aluno deverá conhecer os principais métodos de aprendizado supervisionado e não supervisionado. Deverá ter o entendimento necessário para aplicar algoritmos de aprendizado de máquina para resolver problemas de complexidade moderada. Deverá ser capaz de avaliar criticamente artigos científicos atuais que abordam o assunto.

8. Ementa

Introdução ao aprendizado de máquina, análise estatística e pré-processamento de dados, aprendizado supervisionado (preditivo), avaliação de métodos de aprendizado supervisionado, aprendizado não supervisionado (descritivo) e tópicos avançados em aprendizado de máquina.

9. Conteúdo programático

1. Introdução ao Aprendizado de Máquina
 - 1.1. Motivação
 - 1.2. Contextualização
 - 1.3. Histórico

- 1.4. Aplicações
2. Análise e pré-processamento de dados
 - 2.1. Análise estatística de dados
 - 2.2. Pré-processamento de dados: Limpeza, redução de dimensionalidade (seleção e transformação de parâmetros) e normalização
3. Aprendizado Supervisionado
 - 3.1. Introdução
 - 3.2. Classificação
 - 3.2.1. *K-Nearest Neighbor*
 - 3.2.2. Árvores de Decisão e Regras de Decisão
 - 3.2.3. Classificadores probabilísticos
 - 3.2.4. Rede Neural Artificial
 - 3.2.5. *Support Vector Machine*
 - 3.3. Regressão
4. Avaliação de métodos de aprendizado supervisionado
 - 4.1. Métricas de Erro
 - 4.2. Matriz de Confusão
 - 4.3. Amostragem
 - 4.4. Validação Cruzada
 - 4.5. Curvas ROC
 - 4.6. Seleção de Modelo
5. Aprendizado não-supervisionado
 - 5.1. Introdução
 - 5.2. *K-means*
 - 5.3. Algoritmos de Cluster hierárquico
6. Tópicos Avançados
 - 6.1. Leitura de artigos com os tópicos mais atuais sobre aprendizado de máquina tais como: aprendizado por reforço, aprendizado semi-supervisionado, classificação de séries temporais, detecção de outliers, HMM e etc.

10. Metodologia

As aulas serão expositivas, guiadas pelo uso de data show, bom como quadro-branco. A aplicação dos conceitos vistos em sala de aula será reforçada através de provas, seminários e de um projeto final da disciplina.

11. Recursos

Data show, quadro-branco, Web e laboratório de computação (eventualmente).

12. Avaliação

No decorrer do semestre serão aplicadas três avaliações, constando de provas escritas e/ou trabalhos. A última avaliação consistirá de um artigo e seminário sobre a aplicação de técnicas de aprendizado de máquina as áreas de estudo do aluno ou áreas de interesse.

13. Bibliografia

FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. **Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina**. LTC Gen, 2021.

RASCHKA, SEBASTIAN. Python machine learning. Packt publishing ltd, 2015.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference and Prediction.** Springer, 2009.

MITCHEL, T. M. **Machine Learning.** New York: McGraw-Hill. Series in Computer Science, 1997.

DUDA, R. O.; HART, P. E.; and STOCK, D. G.. **Pattern Classification**, 2nd Edition, Wiley, 2001.

BISHOP, C. M. **Pattern Recognition and Machine Learning**, Springer, 2006.

WITTEN, I. H; FRANK, E. **Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations.** Morgan Kaufmann, 2005.