



Aprendizado de máquina

Apresentação da disciplina

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
Prof. Jefferson Moraes
Email: jeffersonmoraes@gmail.com

Informações gerais

- Profs. Responsáveis : Jefferson Moraes (jeffersonmoraes@gmail.com) e Ronnie Alves (alvesrco@gmail.com)
- Todo o conteúdo do curso será disponibilizado na Página da disciplina no SIGAA
- Carga horária: 60 horas
- Créditos : 4 créditos
- Pré-requisitos: Programação, conceitos básicos de probabilidade, estatística e álgebra linear
- Horário: Segunda e Quarta das 16:40 às 18:20

Objetivos

- Compreender as principais técnicas utilizadas na área de Aprendizado de Máquina (AM)
- Mais especificamente
 - Conhecer os principais métodos de aprendizado supervisionado e não supervisionado
 - Ter o entendimento necessário para aplicar algoritmos de AM para resolver problemas de complexidade moderada
 - Avaliar criticamente artigos científicos atuais que abordam o assunto

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

Conteúdo Programático

1. Introdução ao Aprendizado de Máquina
 1. Motivação
 2. Contextualização
 3. Histórico
 4. Aplicações
2. Análise e pré-processamento de dados
 1. Análise estatística de dados
 2. Pré-processamento de dados
3. Aprendizado Supervisionado (Preditivo)
 1. Introdução
 2. Classificação: K-Nearest Neighbor, Árvores de Decisão e Regras de Decisão, Classificadores probabilísticos , Rede Neural Artificial e SVM
 3. Regressão

Conteúdo Programático

4. Avaliação de técnicas de aprendizado de máquina

1. Métricas de Erro
2. Matriz de Confusão
3. Amostragem
4. Validação Cruzada
5. Curvas ROC
6. Seleção de Modelo

5. Aprendizado não-supervisionado

4. Introdução
5. K-means
6. Algoritmos de Cluster hierárquico

Conteúdo Programático

- 6. Aprendizado por reforço
- 7. Deep Neural Network
- 8. Ensemble classifiers

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.

Avaliação

- Avaliação através de provas, atividades, seminários e projeto
 - 2 provas no valor de 10 pts cada
 - N Atividades: 10 pts
 - Seminário: 10 pts
 - Projeto (Competição): 10 pts
 - Media final: Prova*0.4 +
Seminario*0.2+kaggle*0.3+Atividades*0.1

Atividades

- Práticas de Programação / simulação
- Teórica
 - Leitura de artigos recomendados
 - Entrega de resumo com as respostas das perguntas a respeito do artigo
 - » Máximo duas páginas
 - » Individual

- Tópicos avançados não abordados no curso
 - Sugestões:
 - Aprendizado semi-supervisionado
 - Classificação de séries temporais
 - Detecção de *outliers*
 - Cadeias ocultas de Markov (HMM)
 - *Statistical Relational Learning*
 - Aprendizado incremental
 - Aprendizado Ativo (Active Learning)
 - Classificação multi rótulo
 - ...

Projeto

- Utilizar AM para resolver problema real
 - Preferencialmente relacionado a dissertação ou tese
 - Artigo científico individual no formato SBC
 - http://www.sbc.org.br/index.php?option=com_jdownloads&Itemid=195&task=viewcategory&catid=32
 - 8 a 10 páginas, coluna simples
 - Apresentação no máximo 15 minutos

Cronograma de atividades

- Prova 1: 27/05/2019
- Prova 2: 19/06/2019
- Atividades: datas determinadas em cada aula
- Seminários
 - Entrega do tema do Seminário: até 17/06/2019
 - Apresentação do seminários: 24, 26 e 28/06/2019
- Projeto (competição)
 - Entrega do artigo: até 19/06/2019
 - Apresentação dos projetos: 01/07/2019 e 03/07/2019

Bibliografia

- FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. **Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina**, LTC, 2011
- Han, Jiawei, Jian Pei, and Micheline Kamber. **Data mining: concepts and techniques**. Elsevier, 2011.
- Tan, Pang-Ning. **Introduction to data mining**. Pearson Education India, 2006.
- *Flach, Peter. **Machine learning: the art and science of algorithms that make sense of data**. Cambridge University Press, 2012.
- 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.
- Hall, Mark, Ian Witten, and Eibe Frank. **Data mining: Practical machine learning tools and techniques**. Kaufmann, Burlington (2011).