

PROGRAMA DE ELETIVAS ESPM

Plano de Ensino e Aprendizado

Disciplina:	Machine Learning		
Período: noturno		Carga Horária: 72 h/aula	Revisão: Junho 2018
Área:	Eletivas		
Coordenador:	Prof. Rodrigo Tafner		
Professor:	Prof. Fabrício Barth		

Objetivos Gerais

Esta disciplina tem como objetivo capacitar o estudante no uso de aprendizagem de máquina para o desenvolvimento de sistemas complexos ou para a análise de dados.

Objetivos da Aprendizagem

O aluno deverá, ao final da disciplina, ser capaz de:

- Definir o que é Aprendizagem de Máquina (AM);
- Saber onde e como aplicar os conceitos relacionados a Aprendizagem de Máquina, e;
- Compreender o processo de desenvolvimento de soluções baseadas em AM.

Ementa

Esta disciplina irá tratar os seguintes tópicos: (i) o que é Aprendizagem de Máquina; (ii) hierarquia de aprendizagem de máquina; (iii) processo de descoberta de conhecimento e CRISP-DM; (iv) aprendizagem de máquina supervisionada, e; (v) aprendizagem de máquina não supervisionada.

Conteúdo Programático

O Conteúdo programático da disciplina é composto pelas seguintes partes:

- Introdução sobre Aprendizagem de Máquina;
- Processo de descoberta de conhecimento (KDD) e CRISP-DM;
- Aprendizagem de máquina supervisionada;
- Aprendizagem de máquina não-supervisionada;
- Métodos de avaliação de modelos de aprendizagem de máquina, e;

- Como implantar, integrar e manter modelos de AM em produção.

Estratégia de Ensino e Aprendizagem

Serão utilizadas técnicas de aprendizado ativo. O processo será composto por:

- Apresentação de problemas, com os seus respectivos dados;
- Utilização do ambiente RStudio e linguagem de programação R para tratar os dados e implementar os algoritmos de aprendizagem de máquina;
- Discussão sobre quais são as melhores abordagens para aquisição, pré-processamento, construção do modelo, validação e implantação, e;
- Apresentação dos conceitos teóricos que por ventura não foram tratados na discussão dos problemas, mas fazem parte do escopo da disciplina.

Processo de Avaliação

Nota	Descrição	Peso
1	Exercícios diários – questionários que serão realizados no início de cada aula	20%
2	Avaliação intermediária	20%
3	Avaliação final	20%
4	Projeto de Aprendizagem de Máquina Supervisionado	20%
5	Projeto de Aprendizagem de Máquina Não Supervisionado	20%

Plano de Aulas

Módulo	Objetivo	Quantidade de Horas/Aulas	Método de ensino
1	Introdução a Aprendizagem de Máquina. Hierarquia de Aprendizagem de Máquina Indutiva. Processo de descoberta de conhecimento (KDD) e CRISP-DM.	10	Exercícios em R para simular todo o processo de CRISP-DM e KDD
2	Pré-processamento no R e análise exploratória.	10	Exercícios em R sobre Análise Exploratória
3	Regressão linear. Avaliação de modelos de regressão.	10	Exercícios em R sobre regressão.
4	Algoritmos indutores de árvore de decisão. Aprendizado baseado em instâncias – knn.	12	Exercícios em R sobre

	Avaliação de classificadores.		classificação.
5	Ensemble learning: Random Forest	10	Exercícios em R sobre ensemble learning.
6	Algoritmos de agrupamento e redução de dimensionalidade	10	Exercícios em R sobre aprendizagem de máquina não supervisionada: clustering
7	Regras de associação	10	Exercícios em R sobre aprendizagem de máquina não supervisionada: regras de associação.

Como esta disciplina faz uso de uma metodologia de aprendizado ativo baseado em projetos, a ordem do conteúdo não necessariamente será a apresentada acima.

Bibliografia Básica

LÉVY, Pierre. A esfera semântica: Computação, cognição, economia da informação. Tomo I. São Paulo: AnnaBlume, 2010.

RUSSEL, Stuart. NORVIG, Peter. Inteligência Artificial. Tradução da 3ª ed. Rio de Janeiro: Elsevier, 2013.

FACELI, Katti, LORENA, Ana Carolina, GAMA, João e CARVALHO, André C. P. L. F. de. Inteligência Artificial. Uma Abordagem de Aprendizado de Máquina. 1ª. ed. Rio de Janeiro: LTC, 2012. [e-book]

LUGER, George F. Inteligência Artificial. 6ª.ed. São Paulo: Pearson, 2013. [e-book].

Bibliografia Complementar

MITCHELL, Tom. The Discipline of Machine Learning. Acessado em <http://www.cs.cmu.edu/~tom/pubs/MachineLearning.pdf>.

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An Introduction to Statistical Learning with Applications in R. Springer, 4th edition, 2014.

Hadley Wickham. Tidy data. Journal of Statistical Software, 59(10):??–??, 9 2014.

Roger Peng. R Programming for Data Science. <http://leanpub.com/rprogramming>. 2015.

Hastie, Tibshirani, Friedman. The Elements of Statistical Learning: Data Mining, Inference and Prediction. Second Edition. Springer. 2008.

Demais referências serão entregues durante o desenvolvimento da disciplina.