

Aprendizado de Máquina A

EMENTA: Aprendizado de Máquina com ênfase em técnicas de aprendizado clássico. Análise matemática e estatística de diferentes modelos de aprendizado de máquina. Regressões Regularizadas Ridge e Lasso, Análise Linear de Discriminantes, Máquinas Vetores de Suporte, Florestas Aleatórias, Truques de Kernel, Ensembles de Métodos, Métodos de Amostragem. Técnicas de aprendizado não supervisionado.

A – OBJETIVOS

Os principais objetivos da disciplina são possibilitar ao aluno:

- Entender a matemática por trás dos diferentes modelos
- Saber implementar modelos que já existem em ambiente científico computacional
- Saber implementar novos modelos de aprendizado de máquina
- Entender as garantias teóricas de algoritmos de aprendizado de máquina

B – PROGRAMA

1. O problema de aprendizado supervisionado e não-supervisionado
Introdução. Classificação. Treinamento, validação e teste. Overfitting e underfitting
2. Modelos Lineares de Regressão
 - a. Regressão Ridge
 - b. Regressão Lasso
 - c. Implementação de Ridge e Lasso
 - d. Modelo do Perceptron
3. Modelos Lineares de Classificação
 - a. Regressão logística regularizada
 - b. Modelo Linear Discriminante
4. Teoria do Aprendizado: A dimensão VC (Vapnik-Chervonenkis)
5. Validação e Avaliação de Métodos
 - a. Bayesian Information Criterion
 - b. Akaike information Criterion
 - c. Ligação de K-Fold com o Bayesian Information Criterion
6. Teoria da decisão (minimização de risco):
 - a. Risco vs. Risco empírico
 - b. Regra de decisão de Bayes e
 - c. Risco de Bayes.
 - d. Garantias teóricas de máxima verossimilhança
7. Modelos aditivos (ensembles) e de árvores
 - a. Florestas Aleatórias
 - b. Bagging
 - c. Boosting
 - d. Florestas para regressão e classificação

- e. Ligação com métodos com base em funções de perda
- 8. Métodos de Amostragem
 - a. Amostragem por Rejeição
 - b. Algoritmos MCMC
 - c. Garantias teóricas do Bootstrap
- 9. Máquinas de Vetor Suporte
 - a. Kernels
 - b. Truque do Kernel
 - c. Regularização e Espaços de Hilbert
 - d. SVM para Classificação e Regressão
- 10. Visão geral de: agrupamento
 - a. Redução de dimensionalidade
 - b. Estimativa de densidade
 - c. Descoberta de estrutura intrínseca e organização de dados
 - d. Problemas clássicos em aprendizado não supervisionado
- 11. Espaços métricos e coberturas, agrupamento em espaços métricos, problema do k-medóides, problema do k-means
 - a. Resultados de dureza
 - b. Garantias de aproximação
 - c. NP-Completeness do problema
 - d. K-Means Bayesiano
 - e. Variações de K-Means com divergências de Bregman
- 12. Agrupamentos Espectrais
 - a. Grafo epsilon
 - b. Matrizes de Afinidade (Grafo de vizinhos próximos)
 - c. Grafo completamente conectado
 - d. Laplaciana e Laplaciana Normalizada
 - e. Diferentes variações do Spectral Clustering
- 13. Assuntos avançados

E - BIBLIOGRAFIA

- *Pattern recognition and machine learning*. Bishop, C. M. 2006. Springer.
- *An Introduction to Statistical Learning: with Applications in R*. JAMES, Gareth; WITTEN, Daniela; HASTIE, Trevor; TIBSHIRANI, Robert. New York: Springer, 2013.
- *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Second Edition. Trevor Hastie. Robert Tibshirani. Jerome Friedman 2009