



Machine Learning usando o R

Vítor Wilher, Mestre em Economia e Cientista de Dados

1 Ementa do Curso

O Curso se divide em 17 seções e 15 laboratórios onde o aluno será apresentado aos principais algoritmos de aprendizado supervisionado e não supervisionado, bem como fará aplicações dos mesmos a datasets previamente escolhidos pelo professor.¹

Nos laboratórios, o aluno será encorajado a passar por todo o processo de coleta, tratamento, análise e apresentação dos dados, como se estivesse já aplicando seus modelos de ML na prática.

2 Programa Detalhado das Aulas

1. Introdução à Machine Learning;
2. Entendendo os dados;
3. Regressão Linear;
4. Classificação;
5. Classificação usando algoritmos k-NN;
6. Classificação usando Naive Bayes;
7. Classificação com Árvores de Decisão;
8. Previsão numérica usando métodos de regressão;
9. Redes Neurais;
10. Support Vector Machine (SVM);
11. Detecção de Padrões com regras de associação;
12. Clustering;
13. Métodos de Reamostragem;
14. Seleção e regularização linear de modelos;
15. Bagging, Random Forests e Boosting;
16. Avaliando o desempenho dos modelos;
17. Aumentando o desempenho dos modelos.

¹O nosso curso será baseado, principalmente, em Lantz (2013), Kuhn and Johnson (2016) e James et al. (2017).

3 Programa Detalhado dos Laboratórios

1. Usando o R para gerar relatórios;
2. Análise Exploratória de Dados no R;
3. Rodando modelos de regressão no R;
4. Previsão de Churn com Regressão Logística;
5. Diagnosticando câncer de mama com algoritmos k-NN;
6. Filtrando ligações spam de celular com Naive Bayes;
7. Identificando empréstimos bancários de risco;
8. Prevendo gastos médicos com regressão linear;
9. Modelando resistência de concreto com ANNs;
10. Executando OCR com SVM;
11. Identificação de mantimentos comprados com frequência;
12. Segmentando mercados com k-means clustering;
13. Cross-Validation e Bootstrap na prática;
14. Rodando regressões Ridge, Lasso e PLS no R;
15. Aplicação de algoritmos de meta aprendizagem no R.

Referências

- James, G., D. Witten, T. Hastie, and R. Tibshirani. 2017. *An Introduction to Statistical Learning with Applications in R*. Springer.
- Kuhn, M., and K. Johnson. 2016. *Applied Predictive Modeling*. Springer.
- Lantz, Brett. 2013. *Machine Learning with R*. Packt Publishing.