

## Programa de Disciplina Especialização em Ciência dos Dados

### **Módulo: III**

**Disciplina:** Classificadores Lineares

**Carga Horária:** 15 horas

**Ofertante:** Departamento de Estatística – DEEST/UFOP

### **Objetivo:**

Capacitar o estudante a usar classificadores lineares. Em particular, serão apresentadas duas técnicas que estão entre as mais utilizadas, Regressão Logística, SVM (Support Vector Machine) e Naive Bayes. Ao final do curso o aluno será capaz de ajustar, calibrar e testar esses classificadores.

### **Ementa:**

Introdução aos tipos de classificadores. Classificadores Lineares. Regressão Logística. Support Vector Machine. Naive Bayes

### **Conteúdo Programático:**

1. Introdução aos tipos de classificadores:
  - a) Classificadores supervisionados.
  - b) Classificadores não-supervisionados.
  - c) Classificadores Lineares.
  - d) Função de perda.
2. Regressão Logística:
  - a) Definições básicas do modelo.
  - b) Ajuste e validação do modelo.
  - c) Modelo binário.
  - d) Modelo Multinomial.
  - e) Aplicações.
3. Support Vector Machine:
  - a) Definições básicas.
  - b) Ajuste e validação do modelo.

c) Aplicações.

4. Naive Bayes:

- a) Definições básicas.
- b) Suposições e ajuste do modelo.
- c) Aplicações.

### **Metodologia:**

A disciplina será composta por aulas expositivas e aulas práticas. Nessas aulas práticas os alunos irão resolver problemas utilizando as técnicas aprendidas nas aulas teóricas.

### **Atividade Prática Proposta:**

Análise de um conjunto de dados para aplicação dos métodos de classificação e comparação entre eles.

### **Softwares:**

1. Todas as técnicas e modelos serão implementados utilizando o software R. Esse é um software que funciona através de linha de comandos e é voltado para análise de dados. Ele é um software livre e de código aberto.

### **Bibliografia:**

FARAWAY, J. J. *Extending the Linear Model with R: Generalized linear, mixed effects and nonparametric regression models*. 2. ed. United Kingdom: Chapman & Hall/CRC, 2016. 413 p. ISBN 978-149872-096-0.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)*, Wiley for the Royal Statistical Society, v. 135, n. 3, p. 370–384, 1972. Disponível em: <<http://www.jstor.org/stable/2344614>>. Acesso em: 16 abr. 2020.

PAMPEL, F. C. *Logistic regression: A primer*. 1. ed. Nova York: SAGE Publications Inc, 2000. 96 p. ISBN 978-076192-010-6.

PENG, C.-Y. J.; LEE, K. L.; INGERSOLL, G. M. An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, Routledge, v. 96, n. 1, p. 3–14, 2002. Disponível em: <<https://doi.org/10.1080/00220670209598786>>. Acesso em: 16 abr. 2020.

RUSSELL, S.; NORVIG, P. *Artificial intelligence: A modern approach*. 3. ed. Malaysia: Pearson, 2016. 1152 p. eBook Kindle.