

# Aprendizado de Máquina

Prof. Marcos Prates

## 1 Contato

**Horário de Aula:** Segundas e Quartas, das 09:25 às 11:05hs

**Site da Disciplina:** [www.est.ufmg.br/~marcosop](http://www.est.ufmg.br/~marcosop)

**Emails:** marcosop@est.ufmg.br – Por favor adicione “EST 171” no assunto de emails relacionados ao curso. Eu preciso de ao menos 48 horas para responder emails. Emails enviados na véspera da data de entrega de listas de exercício não serão respondidos!

## 2 Informações sobre o Curso

**Objetivos:** Este curso tem o objetivo de **introduzir** o aluno a algumas ideias e **técnicas de aprendizado de máquina**. Os tópicos abordados serão: classificação, regressão, regularização, técnicas de redução de dimensionalidade, clustering e regras de associação. Se o tempo permitir, outros tópicos serão abordados.

Ao final do curso, o aluno deverá:

- Saber descrever um problema de predição matematicamente
- Conseguir descrever que suposições estão implícitas em modelos de predição
- Saber comparar a performance de diferentes algoritmos preditivos
- Conseguir implementar algoritmos básicos de predição
- Descrever a motivação de diferentes métodos de redução de dimensionalidade
- Conseguir implementar algoritmos básicos de redução de dimensionalidade
- Conseguir implementar algoritmos básicos de clustering, e interpretar seus resultados

**Bibliografia:** Os slides usados em aula estarão disponíveis no site da disciplina e foram gentilmente cedidos pelo **prof. Rafael Izbicki** da Universidade Federal de São Carlos. O livro base do curso é

James, G., Witten, D., Hastie, T. e Tibshirani, R. *An Introduction to Statistical Learning, with Applications in R*, Springer 2013. Disponível gratuitamente em <http://www-bcf.usc.edu/~gareth/ISL/>

Utilizaremos partes deste livro, e não necessariamente na ordem em que os capítulos são apresentados. Algumas referências complementares:

- Hastie, T., Tibshirani, R. e Friedman, J. *The Elements of Statistical Learning*. New York: Springer, 2009. Disponível gratuitamente em <http://statweb.stanford.edu/~tibs/ElemStatLearn/>
- Bishop, C. M. *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
- Wasserman, L. *All of Statistics*. New York: Springer, 2004.
- Wasserman, L. *All of Nonparametric Statistics*. New York: Springer, 2006.

### 3 Avaliação

- **Listas de Exercício:**

Uma a cada uma ou duas semanas de aula. As listas serão em duplas sorteadas sem repetição.

**Exercícios copiados (mesmo que não ao pé da letra) terão pontos descontados.** Discussões são encorajadas, mas cada aluno (ou dupla) deve escrever suas soluções **separadamente**.

*Exemplo de discussão que **não constitui** em cópia: João, Maria e José discutem como resolver o exercício. Após conseguir resolvê-lo, cada um vai para seu quarto e escreve sua solução separadamente. Exemplo de discussão que **constitui** em cópia: João, Maria e José discutem como resolver o exercício na lousa. Após a discussão, cada um escreve sua solução olhando para o que foi escrito no quadro.*

As listas consomem muito tempo, e frequentemente envolvem trabalho computacional. **Não deixe para fazê-las na última hora!** A proposta das listas de exercícios é praticar as ideias vistas em aula. Assim, nenhum ponto será dado para respostas corretas que não forem justificadas.

**Para cada dia de atraso, pontos serão descontados.** Fale com o professor *ao menos 3 dias antes da data de entrega* caso haja dificuldades para cumprir os prazos.

Programa (sujeito a mudanças):

## Tópicos

---

Introdução

Regressão Linear e Seleção de Modelos 1

Manipulando Imagens e Textos

Regressão Linear Moderna: lasso e penalização

Nearest Neighbors e Árvores de Regressão

Redes Neurais

Regressão Logística e Seleção de Modelos

Naive Bayes e Classificadores Plug-in

Análise Discriminante e Nearest Neighbors

Árvores e Florestas de Classificação

Bagging, Boosting

SVM

Redução de Dimensionalidade

Clustering

Sistemas de Recomendação

Modelos Gráficos

Online Learning

Tópicos Adicionais

---