

PLANO DE ENSINO

Identificação			Ementa
Curso: Estatística			Conceitos básicos sobre Machine Learning. Aprendizado Supervisionado versus Aprendizado não Supervisionado. Erro Médio de Teste. Validação Cruzada. Métodos de Regressão. Métodos de Classificação.
Disciplina: Optativa I - Tópicos em Estatística - Introdução ao Aprendizado de Máquina		Código: E43	
Professor: Luana Lúcia Alves de Azevêdo			
Coordenador: Lenilson Sergio Candido			
Período: 5º	Semestre: 01	Ano: 2019	
Carga Horária: 80		Créditos: 04	
Teórica: 40	Prática: 40	Total: 80	

Objetivo da Disciplina no Curso

O objetivo desta disciplina é apresentar uma introdução aos conceitos, técnicas e aplicações do Aprendizado de Máquina, que podem ser utilizados em diferentes etapas do processo de extração de conhecimento e tomada de decisões a partir da bases de dados.

Metodologia de Trabalho do Professor na Disciplina

A explanação do conteúdo se dará por meio do uso de slides, anotações no quadro e demonstrações em computador. Será enfatizado com seminários, listas de exercícios e trabalhos como forma de consolidar os conteúdos.

Conteúdo Programático

Unidade I - Conceitos básicos sobre Machine Learning, Aprendizagem Estatística, Aprendizado Supervisionado versus Aprendizado não Supervisionado, problemas de Regressão e problemas de Classificação. Big Data.

Unidade II – Erro Médio de Teste (*Average Test Error*). Erro de Treinamento e Erro de Teste. Bias-Variance Trade-Off.

Unidade III - Introdução à Validação Cruzada. Tipos de Validação Cruzada: holdout, k -Fold e Validação Cruzada Leave-One-Out (LOOCV).

Unidade IV – Aprendizado Supervisionado: Regressão e Classificação. Regressão Linear. Regressão Logística. Análise Discriminante Regularizada. Máquina de Vetor de Suporte (SVM). Árvores de Decisão. Bootstrap. Florestas Aleatórias.

Unidade V – Utilização de software para demonstrações, análises e cálculos.

Avaliação e Critérios de Avaliação da Disciplina no Curso

Será procedida por meio de avaliações (provas e trabalhos).

Provas: Serão realizadas duas provas em sala de aula. A Média das Provas (MP) é dada por:

$$MP = \frac{N1+N2}{2},$$

Em que **N1** é a nota da primeira prova e **N2** é a nota da segunda prova.

Trabalhos: Poderão ser realizados trabalhos propostos pela professora, nos prazos pré-estabelecidos, com um valor total de 100 pontos. Neste, estão inclusos seminário e exercícios avaliativos. A nota final será dada por:

$$MT = \text{Soma dos Trabalhos.}$$

A Média Final (MF) será obtida por meio da expressão:

$$MF = \frac{MP+MT}{2}.$$

Se $MF \geq 60$ o aluno estará aprovado.

Se $MF < 60$ o aluno fará a avaliação repositiva.

Bibliografia da Disciplina no Curso

Básica

Complementar

<p>JAMES, G., WITTEN, D., HASTIE, T., TIBSHIRANI, R., (2013). An Introduction to Statistical Learning: with Applications in R. New York: Springer Velarg.</p> <p>HASTIE, T., TIBSHIRANI, R., FRIEDMAN, J., (2001). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer Velarg.</p> <p>CLARKE, B., FOKOUÉ, E., ZHANG, H. H., (2009). Principles and Theory for Data Mining and Machine Learning. New York: Springer Velarg.</p> <p>AZEVÊDO, L. A. (2018). Métodos Estatísticos em Aprendizado de Máquinas para problemas de Classificação. Dissertação (Mestrado em Estatística).</p>	<p>TIBSHIRANI, R., HASTIE, T., NARASIMHAN, B. AND CHU, G., (2003). Class Prediction by Nearest Shrunken Centroids, with Applications to DNA Microarrays. <i>Statistical Science</i> 18 (1), 104-107.</p> <p>CHOLLET FRANÇOIS, ALLAIRE J.J. Deep Learning with R. Manning Publications Company, 2018.</p> <p>WICKHAM H., GROLEMUND G. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly Media, California, 2017.</p> <p>GÉRON, A. Hands-On Machine Learning with Scikit-Learn and TensorFlow. O'Reilly Media, 2017.</p> <p>SEBASTIAN, R. (2015). Python Mahine Learning. Packt Publishing.</p> <p>KUHN, M., JOHNSON, K. Applied Predictive Modeling. New York, Springer. 2013</p>
---	---

Ji-Paraná, Dezembro de 2018.