



CENTRO UNIVERSITÁRIO RITTER DOS REIS

PÓS-GRADUAÇÃO EM DATA SCIENCE E BIG DATA

DISCIPLINA: *Tópicos especiais I*

CARGA HORÁRIA: 12h

PROFESSOR: Wilson Pires Gavião Neto

ANO/SEM: 2018 / 1

Horário das Aulas: *Quartas-feiras* (19h - 22h40min)

PLANO DE ENSINO

EMENTA:

Processo de descoberta de conhecimento em bases de dados textuais. Categorização de Texto. Análise de sentimento. Seleção de termos importantes (feature selection).

OBJETIVOS:

- Estudar e aplicar técnicas usuais no contexto de mineração de textos
- Capacitar o aluno a compreender as etapas do processo KDD sobre bases textuais
- Utilizar ferramentas de software na prototipação de modelos

PROGRAMA:

- Pré-processamento
- Transformação
- Seleção de *features* relevantes
- Classificação de textos e análise de sentimento
- Análise de resultados de classificação supervisionada

METODOLOGIA:

- Aulas expositivas e práticas usando os computadores do laboratório de informática, para o uso dos ambientes e ferramentas a serem utilizados na disciplina.
- Desenvolvimento de exercícios envolvendo programação e aplicação das técnicas.

ATIVIDADES DISCENTES:

- Os alunos desenvolverão os exercícios propostos pelo professor utilizando os ambientes a serem estudados na disciplina, além de desenvolverem trabalhos práticos (individuais e/ou em grupo), para compor as notas de cada um dos graus.

PROCEDIMENTOS, INSTRUMENTOS E CRITÉRIOS DE AVALIAÇÃO:

A avaliação de desempenho do aluno será contínua, considerando um processo gradativo e cumulativo de construção do conhecimento através de trabalhos práticos. A nota final será dada pelo trabalho prático.

BIBLIOGRAFIA RECOMENDADA:

Bibliografia Básica

Christopher D. MANNING, Prabhakar Raghavan, and Hinrich Schütze. 2008. Introduction to Information Retrieval. Cambridge University Press, New York, NY, USA.
<https://nlp.stanford.edu/IR-book/>

FACELI, Katti et al. Inteligência artificial: uma abordagem de aprendizado de máquina. Rio

de Janeiro: LTC, c2011. 378 p. ISBN 978-85-216.-1880-5.

Liu, Bing. "Sentiment analysis and opinion mining." Synthesis lectures on human language technologies 5.1 (2012): 1-167. <https://www.cs.uic.edu/~liub/FBS/liub-SA-and-OM-book.pdf>

Bibliografia Complementar e Sites de Apoio

North, Matthew. Data mining for the masses. Global Text Project, 2012. http://www.academia.edu/download/39657706/North_-_Data_Mining_for_the_Masses_-_2012.pdf