

Caracterização da disciplina

| | | | | | | | | | |
|-----------------------------|------------------|---------------------|--------------------|---------------|---|---------------|----|------|------|
| Código da disciplina: | MCZA015-13 | Nome da disciplina: | Mineração de dados | | | | | | |
| Créditos (T-P-I): | (3 - 1 - 4) | Carga horária: | 48 horas | Aula prática: | | Câmpus: | SA | | |
| Código da turma: | TNA1MCZA015-13SA | Turma: | A1 | Turno: | N | Quadrimestre: | 1 | Ano: | 2024 |
| Docente(s) responsável(is): | Debora Medeiros | | | | | | | | |

Alocação da turma

| | Segunda | Terça | Quarta | Quinta | Sexta | Sábado |
|---------------|---------|-------|--------|--------|-------|--------|
| 8:00 - 9:00 | | | | | | |
| 9:00 - 10:00 | | | | | | |
| 10:00 - 11:00 | | | | | | |
| 11:00 - 12:00 | | | | | | |
| 12:00 - 13:00 | | | | | | |
| 13:00 - 14:00 | | | | | | |
| 14:00 - 15:00 | | | | | | |
| 15:00 - 16:00 | | | | | | |
| 16:00 - 17:00 | | | | | | |
| 17:00 - 18:00 | | | | | | |
| 18:00 - 19:00 | | | | | | |
| 19:00 - 20:00 | | | X | | | |
| 20:00 - 21:00 | | | X | | | |
| 21:00 - 22:00 | | | | | X | |
| 22:00 - 23:00 | | | | | X | |

Planejamento da disciplina
Objetivos gerais

Assimilar conceitos de mineração de dados e descoberta de conhecimento em bases dados.

Objetivos específicos

Apresentar/experimentar técnicas de mineração de dados e descoberta de conhecimento de bases de dados para a modelagem de problemas em áreas científicas e tecnológicas.

Ementa

Introdução. Seleção, Preparação e pré-processamento dos dados. Modelagem: o processo de mineração. Pós-processamento do conhecimento adquirido. Ferramentas.

Conteúdo programático

| Aula | Conteúdo | Estratégias didáticas | Avaliação |
|------------|---------------------------|--|---|
| 07/Fev (t) | Introdução; Dados | - Nas aulas teóricas (t) serão utilizadas apresentação da teoria em slides e propostas discussões sobre o tema. - Nas aulas práticas (p) serão apresentados | - A parte teórica será avaliada por meio de duas provas. - A parte prática será avaliada de duas maneiras: - - Nas aulas práticas (p) poderão ser propostas atividades complementares para serem solucionadas individualmente pelos alunos cuja solução deverá ser posteriormente entregue pelo sistema Moodle. |
| 09/Fev (p) | Classificação e regressão | | |
| 16/Fev (p) | Classificação e regressão | | |
| 21/Fev (t) | Classificação e regressão | | |
| 23/Fev (p) | Classificação e regressão | | |
| 28/Fev (t) | Agrupamento | | |
| 1o/Mar (p) | Agrupamento | | |
| 06/Mar (t) | Agrupamento | | |
| 08/Mar (p) | Agrupamento | | |
| 13/Mar (t) | Mineração de padrões | | |
| 15/Mar (p) | Mineração de padrões | | |
| 20/Mar (t) | Prova 1 | | |
| 22/Mar (p) | Detecção de outliers | | |

| | | | | |
|------------|--|------------------|---|---|
| 27/Mar (t) | Padrões sequenciais | exemplos | e | - Os alunos deverão realizar, em grupos (max. 3 integrantes), o desenvolvimento de um projeto prático que deverá ser documentado e entregue via plataforma Moodle. Os grupos também deverão apresentar seu projeto durante as 2 últimas aulas práticas. |
| 03/Abr (t) | Mineração de textos e seleção de atributos | propostos | | |
| 05/Abr (p) | Mineração de textos e seleção de atributos | problemas | a | |
| 10/Abr (t) | Deep learning | serem resolvidos | | |
| 12/Abr (p) | Deep learning | utilizando | | |
| 17/Abr (t) | Inteligência artificial generativa | ferramentas de | | |
| 19/Abr (p) | Inteligência artificial generativa | Mineração de | | |
| 24/Abr (t) | Prova 2 | dados para a | | |
| 26/Abr (p) | Projetos | linguagem | | |
| 03/Mai (t) | Prova substitutiva | Python. | | |
| 06/Mai (p) | Projetos | | | |
| 2024q2 | Recuperação | | | |

Descrição dos instrumentos e critérios de avaliação qualitativa

A média será composta por 3 componentes:

- 2 provas teóricas: 75%
- Projeto: 25%
- Atividades práticas: bônus na média (total: 0,7)

O conceito será calculado da seguinte maneira:

- A: Média final $\geq 8,5$
- B: $7,0 \leq$ Média final $< 8,5$
- C: $6,0 \leq$ Média final $< 7,0$
- D: $5,0 \leq$ Média final $< 6,0$
- F: Média final $< 5,0$

A recuperação será realizada por meio de uma prova teórica aplicada no começo do 2º quadrimestre de 2024. Após a recuperação a nota será calculada da seguinte maneira: $0,5 * \text{média_final} + 0,5 * \text{recuperação}$.

Referências bibliográficas básicas

1. HAN, J.; KAMBER, M. Data mining: concepts and techniques. 2. ed. New York, USA: Morgan Kaufmann, 2006.
2. WITTEN, I. H.; FRANK, E. Data mining: practical machine learning tools and techniques with Java implementations. New York, USA: Morgan Kaufmann, 2000.
3. TAN, P. N.; STEINBACH, M.; KUMAR, V. Introduction to data mining. Reading, USA: Addison Wesley, 2005.

Referências bibliográficas complementares

1. BISHOP, C. M. Pattern recognition and machine learning. New York, USA: Springer, 2006.
2. DUDA, R. O.; HART, P. E.; STORK, D. G. Pattern classification. 2. ed. New York, USA: John Wiley & Sons, 2000.
3. MITCHELL, T. M. Machine learning. São Paulo, SP: McGraw-Hill, 1997.
4. REZENDE, S. O. Sistemas inteligentes: fundamentos e aplicações. Barueri, SP: Manole, 2003.
5. RUSSEL, S.; NORVIG, P. Artificial intelligence: a modern approach. 2. ed. Upper Saddle River, NJ: Prentice Hall, 2003.