



UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA E CIÊNCIAS ATUARIAIS

Componente Curricular:	ESTAT0016 – TOPICOS ESPECIAIS EM ESTATISTICA (Introdução à aprendizagem de máquina)		
Carga horária:	60 horas (4 créditos)	Horário:	Segundas - 17h00 às 18h50 Quintas - 17h00 às 18h50
Unidade Responsável:	Departamento de Estatística e Ciências Atuariais		
Docente:	Sadraque E. F. Lucena		
Tipo de Componente:	Disciplina	Quantidade de Avaliações:	2
Período Letivo:	2023-2		

Ementa: Pré-processamento de dados. Aprendizado supervisionado. Regressão. Classificação. k -Nearest Neighbors. Naïve Bayes. Árvores de Decisão. Florestas aleatórias. Avaliação de desempenho. Validação Cruzada. Métricas de desempenho de Classificação. Ajuste de parâmetros. Bagging, Boosting e Stacking. Aprendizado não supervisionado. Regras de Associação. k -means.

Objetivos: Introduzir os conceitos fundamentais e técnicas da aprendizagem de máquina.

Metodologia: Serão ministradas aulas teóricas expositivas; utilizados recursos visuais e computacionais; aplicação dos métodos em sala de aula; solicitação de atividades extraclasse. O cronograma de aulas encontra-se no Quadro 1.

Habilidades e Competências: No final deste curso, os(as) alunos(as) deverão estar equipados(as) com as habilidades e o entendimento necessários para aplicar a aprendizagem de máquina em diversos contextos, preparando-os(as) para carreiras em ciência de dados, inteligência artificial e análise de dados.

Avaliação: Será realizada uma avaliação contínua, com solicitação de atividades extraclasse e apresentações em sala de aula. Duas avaliações também serão realizadas ao longo do semestre.

Conteúdo:

1. Pré-processamento de Dados
 - 1.1. Exploração
 - 1.2. Limpeza
 - 1.3. Transformação

1.4. Redução

2. Aprendizado Supervisionado

2.1. Regressão

2.1.1. Regressão linear

2.1.2. Regressão logística

2.2. k -Nearest Neighbors

2.3. Naïve Bayes

2.4. Árvores de Decisão

2.5. Florestas aleatórias

2.6. Avaliação de desempenho

2.6.1. Validação Cruzada

2.6.2. Amostragem bootstrap

2.6.3. Acurácia

2.6.4. Kappa

2.6.5. Precisão e revocação

2.6.6. Sensibilidade e especificidade

2.7. Ajuste de parâmetros

2.8. Métodos de conjunto (ensemble methods)

2.8.1. Bagging

2.8.2. Boosting

2.8.3. Stacking

3. Aprendizado Não Supervisionado

3.1. Regras de Associação

3.2. k -means

Bibliografia:

Básica:

- FACELI, Katti. **Inteligência artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2011. 378 p.
- IZBICKI, Rafael; DOS SANTOS, Tiago Mendonça. **Aprendizado de máquina: uma abordagem estatística** [livro eletrônico]. São Carlos: Rafael Izbicki, 2020.
- NWANGANGA, Fred; CHAPPLE, Mike. **Practical machine learning in R**. Indianapolis: John Wiley & Sons, 2020.
- ABU-MOSTAFA, Yaser S.; MAGDON-ISMAIL, Malik; LIN, Hsuan-Tien. **Learning from data**. New York: AMLBook, 2012.
- HASTIE, Trevor et al. **The elements of statistical learning: data mining, inference, and prediction**. New York: springer, 2009.

Complementar:

- RAMASUBRAMANIAN, Karthik; SINGH, Abhishek. **Machine learning using R**. New Delhi, India: Apress, 2017.
- LANTZ, Brett. **Machine learning with R: expert techniques for predictive modeling**. Packt publishing ltd, 2019.

Quadro 1: Cronograma de aulas previstas para a componente curricular Tópicos Especiais em Estatística para o período 2023-2.

Data	Dia da semana	Aula	Assunto Previsto
13/11/23	Segunda	1	Apresentação da disciplina.
16/11/23	Quinta	2	Introdução ao Quarto.
20/11/23	Segunda	3	Tratando os dados.
23/11/23	Quinta	4	Tratando os dados.
27/11/23	Segunda	5	IX SEMAC - Semana Acadêmico-Cultural da UFS
30/11/23	Quinta	6	IX SEMAC - Semana Acadêmico-Cultural da UFS
04/12/23	Segunda	7	Regressão linear.
07/12/23	Quinta	8	Regressão linear.
11/12/23	Segunda	9	Regressão logística.
14/12/23	Quinta	10	Regressão logística.
18/12/23	Segunda	11	k -Nearest Neighbors.
21/12/23	Quinta	12	k -Nearest Neighbors.
25/12/23	Segunda	–	Recesso natalino
28/12/23	Quinta	–	Recesso natalino
01/01/24	Segunda	–	Férias coletivas
04/01/24	Quinta	–	Férias coletivas
08/01/24	Segunda	–	Férias coletivas
11/01/24	Quinta	–	Férias coletivas
15/01/24	Segunda	13	Naive Bayes.
18/01/24	Quinta	14	Naive Bayes.
22/01/24	Segunda	15	Avaliação 1.
25/01/24	Quinta	16	Árvores de decisão.
29/01/24	Segunda	17	Árvores de decisão.
01/02/24	Quinta	18	Floresta aleatória.
05/02/24	Segunda	19	Floresta aleatória.
08/02/24	Quinta	20	Avaliando a performance do modelo.
12/02/24	Segunda	–	Carnaval (ponto facultativo)
15/02/24	Quinta	21	Atividade.
19/02/24	Segunda	22	Avaliando a performance do modelo.
22/02/24	Quinta	23	Melhorando a performance dos modelos.
26/02/24	Segunda	24	Melhorando a performance dos modelos.
29/02/24	Quinta	25	Descoberta de padrões por regras de associação.
04/03/24	Segunda	26	Descoberta de padrões por regras de associação.
07/03/24	Quinta	27	Agrupamento de dados com k -Means.
11/03/24	Segunda	28	Agrupamento de dados com k -Means.
14/03/24	Quinta	29	Aplicações diversas.
18/03/24	Segunda	30	Projeto.
21/03/24	Quinta	31	Aplicações diversas.
25/03/24	Segunda	32	Aplicações diversas.
28/03/24	Quinta	–	Quinta-feira Santa (ponto facultativo)
01/04/24	Segunda	33	Primeira apresentação do projeto.
04/04/24	Quinta	34	Aplicações diversas.
08/04/24	Segunda	35	Apresentação do projeto.
11/04/24	Quinta	36	Encerramento.