

Plano de Ensino – 2025/2

Código	DCC831/ICE163
Disciplina	Mineração de Dados Avançada: Interpretabilidade
Turma	(a definir) 3a/5a às 17:00
Professor	Wagner Meira Jr.

Ementa: Interpretabilidade. Modelos interpretáveis. Modelos agnósticos. Modelos baseados em exemplos. Interpretabilidade em aprendizado profundo. Avaliação. Perspectivas. Estudos de caso.

Programa

#	Data	Conteúdo previsto	Ref.	Obs
1		Interpretabilidade e seus objetivos	1-3	
2		Visão Geral dos Métodos	4	
3		Dados e Modelos	5	
4		Regressão Linear e Regressão Logística	6-7	
5		GLM, GAM e Árvores de Decisão	8-9	
6		Regras de Decisão e Ajuste de Regras	10-11	
7		Ceteris Paribus e ICE	12-13	
8		Lime e Contrafactuais	13-14	
9		Âncoras	15	
10		Valores Shapley e Shap	17-18	
11		PDP e ALE	19-20	
12		Interação de Atributos e Decomposição Funcional	21-22	
13		Importâncias PF e LOFO	23-24	
14		Surrogates e Prototypes	25-26	
15		Redes Neurais - Atributos aprendidos	27	
16		Redes Neurais - Mapas de Saliência e Conceitos	28-29	
17		Redes Neurais - Adversarial	30	
18		Redes Neurais - Influentes	31	
19		Avaliação de Modelos de Interpretabilidade	32	
20		Estórias	33	
21		Perspectivas Futuras	34	
22		Seminários		
23		Seminários		
24		Seminários		
25		Seminários		
26		Seminários		
27		Seminários		
28		Seminários		
29		Projetos		
30		Projetos		

Material de apoio

Livro: Interpretable Machine Learning

(<https://christophm.github.io/interpretable-ml-book>)

Moodle (slides e outros materiais)

Artigos e projetos

Avaliações

Haverá atividades a serem realizadas em grupo e/ou individualmente. Cada aluno vai desenvolver um projeto.