

 UNIVERSIDADE FEDERAL DE SANTA CATARINA		Campus Araranguá Rodovia Governador Jorge Lacerda, 3201, Jardim das Avenidas Araranguá - Santa Catarina – Brasil / CEP 88900-000 www.ararangua.ufsc.br / +55 (48) 3721.6448	
Avaliação II			
Disciplina	Tópicos Especiais III (DEC7553)	08655	2022.2
Aluno(a)s			

1) Relacione os conceitos da 1ª coluna com as respectivas definições na 2ª coluna – (1,5):

[1]	Descoberta de Conhecimento em Bases de Dados	[18]	Cria árvores de decisão a partir da seleção randômica de exemplos dos dados e obtém a predição de cada árvore selecionando a melhor solução.
[2]	Pré-processamento	[19]	Conceito criado para unificar a estatística, análise de dados, aprendizado de máquina e seus métodos relacionados visando entender e analisar fenômenos reais com dados.
[3]	Aprendizado de Máquina	[1]	Processo não trivial, interativo e iterativo, visando a identificação de padrões compreensíveis, que sejam válidos, novos, potencialmente úteis a partir de grandes conjuntos de dados.
[4]	Aprendizado Supervisionado	[8]	Objetiva descobrir um relacionamento entre os atributos precursores e o atributo meta, usando registros cuja classe é conhecida, para se construir um modelo que possa ser aplicado a objetos ainda não classificados.
[5]	Aprendizado Não Supervisionado	[13]	Processo pela qual determinado conjunto de dados de teste é dividido em N subconjuntos e a partir desses N subconjuntos são realizados N testes.
[6]	Avaliação	[16]	Expressa a distância, similaridade, correlação ou associação entre dois objetos quaisquer.
[7]	Pós-processamento	[15]	Proporção de exemplos de uma classe (neste caso valores positivos) que foram corretamente classificados pelos exemplos classificados incorretamente na classe - $TP / (TP + FP)$.
[8]	Tarefa de Classificação	[14]	Estrutura que, na tarefa de classificação envolvendo classes de valores discretos, permite estimar a qualidade dos resultados obtidos.
[9]	Tarefa de Agrupamento	[4]	Processo pelo qual determinado algoritmo de aprendizado (indutor) recebe um conjunto de exemplos de treinamento para os quais os rótulos da classe associada são conhecidos e, a partir disso, constrói um modelo visando a correta determinação da classe para novos exemplos ainda não rotulados.
[10]	Tarefa de Associação	[24]	Consiste na localização de materiais (usualmente documentos) de natureza não estruturada (usualmente textos) que satisfazem determinada necessidade por informação a partir de uma grande coleção.
[11]	Árvore de Decisão	[22]	Técnica baseada na noção intuitiva de que determinados padrões de interação, representados na forma de um grafo, são importantes características capazes de descrever o comportamento de unidades de análise.
[12]	ID3	[10]	Objetiva descobrir relacionamentos importantes em um conjunto de dados, tal que, a presença de um item em uma determinada transação irá implicar na presença de outro item na mesma transação.

 UNIVERSIDADE FEDERAL DE SANTA CATARINA		Campus Araranguá Rodovia Governador Jorge Lacerda, 3201, Jardim das Avenidas Araranguá - Santa Catarina – Brasil / CEP 88900-000 www.ararangua.ufsc.br / +55 (48) 3721.6448	
Avaliação II			
Disciplina	Tópicos Especiais III (DEC7553)	08655	2022.2
Aluno(a)s			

[13]	<i>Cross-validation</i>	[5]	Processo pelo qual o algoritmo indutor analisa os exemplos fornecidos e tenta determinar quais destes podem ser agrupados de alguma maneira, formando agrupamentos ou <i>clusters</i> .
[14]	Matriz de Confusão	[23]	Medida que avalia a frequência com que determinado nodo aparece no caminho mais curto entre dois nodos quaisquer.
[15]	Precisão (<i>Precision</i>)	[2]	Etapa que visa a preparação e a transformação de dados de modo que possam ser utilizados por algoritmos de mineração de dados e, deste modo, conduzir à descoberta de conhecimento.
[16]	Medidas de Similaridade	[11]	Técnica de classificação que, através de um processo de otimização heurística e baseado na teoria da informação, seleciona iterativamente determinadas variáveis que promovem a melhor separação de classes de acordo com alguma função de custo, visando criar regiões disjuntas e estabelecendo uma fronteira de decisão.
[17]	Centralidade de Proximidade	[6]	Processo pelo qual se objetiva aferir a qualidade da aplicação/utilização de algoritmos de Mineração de Dados sobre determinado conjunto de dados.
[18]	<i>Random Forest</i>	[9]	Objetiva servir como um passo anterior à tarefa de classificação quando não se possui um conjunto de dados previamente classificado, possibilitando a reunião de itens semelhantes em determinado grupo.
[19]	Ciência de Dados	[3]	Um programa aprende a partir da experiência E, em relação a uma classe de tarefas T, com medida de desempenho P, se seu desempenho em T, medido por P, melhora com E.
[20]	Dendrograma	[7]	Etapa do processo de KDD que consiste na avaliação dos resultados obtidos, isto é, analisa se o conhecimento descoberto é relevante ou não.
[21]	Grafo	[25]	Determina a porcentagem de determinado <i>itemset</i> dentre todas as transações da Base de Dados.
[22]	Análise de Rede Social	[17]	Representa a distância natural entre todos os pares de nodos, definida pelo tamanho dos caminhos mais curtos.
[23]	Centralidade de Intermediação	[20]	Árvore que iterativamente divide o conjunto de dados em subconjuntos menores até que cada subconjunto consista de somente um objeto.
[24]	Recuperação de Informação	[12]	Algoritmo baseado na teoria da informação que constrói uma árvore de decisão onde cada vértice (nodo) corresponde a um atributo, e cada aresta da árvore a um valor possível do atributo.
[25]	Medida do Suporte	[21]	Conjunto de vértices conectados por arestas que podem ser direcionadas ou não.

 UNIVERSIDADE FEDERAL DE SANTA CATARINA		Campus Araranguá Rodovia Governador Jorge Lacerda, 3201, Jardim das Avenidas Araranguá - Santa Catarina – Brasil / CEP 88900-000 www.ararangua.ufsc.br / +55 (48) 3721.6448	
Avaliação II			
Disciplina	Tópicos Especiais III (DEC7553)	08655	2022.2
Aluno(a)s			