

Lista #3

Curso: Ciência da Computação

Disciplina: Inteligência Artificial

Prof^a. Cristiane Neri Nobre

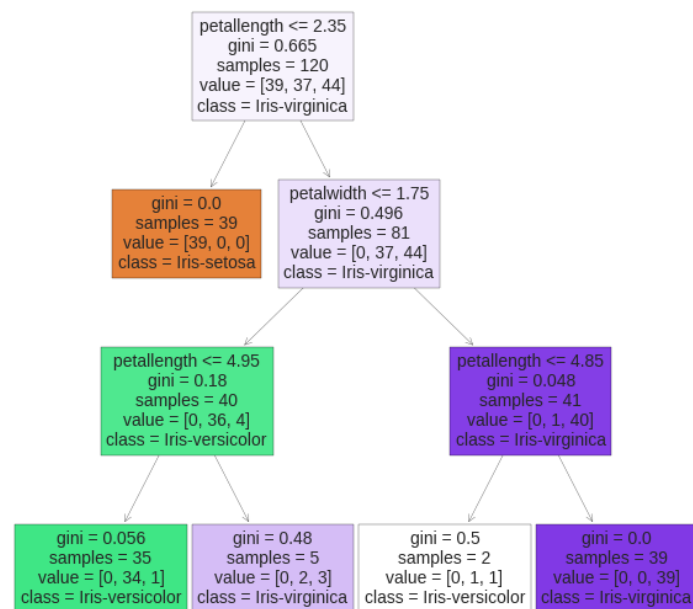
Data de entrega: 10/09

Valor: 1,5 pontos

Questão 01

A figura abaixo mostra uma árvore de decisão construída por um algoritmo de aprendizado indutivo a partir de um conjunto de dados em que as instâncias são descritas por quatro atributos: **Tamanho** da Pétala, **largura** da Pétala, **Tamanho** da Sépala e **Largura** da Sépala.

Dado um objeto de classe desconhecida, essa árvore classifica o objeto nas classes: **Iris_Setosa**, **Iris_Virgínica** e **Iris_Versicolor**. Esta árvore foi gerada com os hiperparâmetros (DecisionTreeClassifier(criterion='gini', max_depth=3)), usando a linguagem Python.



Com base nestas informações, qual as saídas da árvore para os seguintes **registros de teste**, respectivamente?

Registros de teste	Tamanho da Pétala	Largura da Pétala	Tamanho da Sépala	Largura da Sépala
Instância 1	3.46	0.87	2.45	1.78
Instância 2	1.67	1.89	0.78	1.32
Instância 3	2.56	2.34	2.45	1.78
Instância 4	6.67	2.34	2.45	1.78

- a) Iris_Virgínica, Íris_Setosa, Iris_Versicolor, Iris_Virgínica
- b) Iris_Virgínica, Íris_Setosa, Iris_Virgínica , Iris_Versicolor
- c) Iris_Versicolor, Íris_Setosa, Iris_Versicolor, Iris_Virgínica
- d) Íris_Setosa, Iris_Virgínica, Iris_Virgínica , Iris_Versicolor
- e) Iris_Versicolor, Íris_Setosa, Iris_Versicolor, Iris_Virgínica

Questão 02

Considerando a árvore da questão anterior, e as seguintes afirmações:

- I. Esta árvore possui 5 regras de classificação
- II. Das regras geradas, há apenas uma com cobertura por classe de 100%
- III. A menor cobertura por classe é de 6.8% e corresponde à classe Iris_Virgínica

É **correto** o que se afirma em:

- a) I, apenas.
- b) III, apenas.
- c) I e II, apenas.
- d) I e III, apenas.
- e) I, II e III.

Questão 03

Considere a seguinte matriz de confusão obtida por meio do classificador, **Árvore de decisão**, para um problema de quatro classes:

		Foi classificado como			
		A	B	C	D
Era da classe	A	10	4	2	1
	B	1	15	2	0
	C	2	3	20	5
	D	4	1	2	50

Quais os valores para as métricas abaixo para cada uma das classes A, B, C e D?

	Precisão	Recall	F1Score	TVP	TFN	TFP	TVN
A							
B							
C							
D							

Questão 04

Considerando a base de dados abaixo. Mostre o primeiro e o segundo nível da árvore gerada.

História do crédito	Dívida	Garantias	Renda Anual	Risco
Ruim	Alta	Nenhuma	<15000	Alto
Desconhecida	Alta	Nenhuma	>=15000 a <=35000	Alto
Desconhecida	Baixa	Nenhuma	>=15000 a <=35000	Moderado
Desconhecida	Baixa	Nenhuma	>35000	Alto
Desconhecida	Baixa	Nenhuma	>35000	Baixo
Desconhecida	Baixa	Adequada	>35000	Baixo
Ruim	Baixa	Nenhuma	<15000	Alto
Ruim	Baixa	Adequada	>35000	Moderado
Boa	Baixa	Nenhuma	>35000	Baixo
Boa	Alta	Adequada	>35000	Baixo
Boa	Alta	Nenhuma	<15000	Alto
Boa	Alta	Nenhuma	>=15000 a <=35000	Moderado
Boa	Alta	Nenhuma	>35000	Baixo
Ruim	Alta	Nenhuma	>=15000 a <=35000	Alto

Questão 05

Investigue a biblioteca <https://pypi.org/project/chefboost/>

E implemente os algoritmos **ID3** e **C45** para gerar a árvore da base de dados da questão anterior.

Questão 06

Investigue como é o funcionamento do algoritmo CART. Mostre todos os cálculos necessários para a geração da árvore.

Questão 07

Utilizando a biblioteca acima, rode o CART e veja a árvore gerada. Compare as 3 árvores geradas a partir da mesma base de dados.

Importante:

- | |
|---|
| <ol style="list-style-type: none">1. Você deve apresentar todas as discussões necessárias para uma completa compreensão do que foi feito, em todas as questões.2. Adicionar também o link para o código desenvolvido |
|---|