

Trabalho Prático do Módulo 2

Entrega 5 jun em 17:00 **Pontos** 25 **Perguntas** 15
Disponível até 5 jun em 17:00 **Limite de tempo** Nenhum

Instruções

O arquivo abaixo contém as informações para realização do Trabalho Prático.

[Enunciado Trabalho Prático - Módulo 2 \(MPT\).pdf](#) 

Histórico de tentativas

	Tentativa	Tempo	Pontuação
MAIS RECENTE	Tentativa 1	216 minutos	25 de 25

⚠ As respostas corretas estarão disponíveis em 5 jun em 21:00.

Pontuação deste teste: **25** de 25

Enviado 2 jun em 17:22

Esta tentativa levou 216 minutos.

Pergunta 1	1,66 / 1,66 pts
Qual é a idade média dos pacientes?	
<input type="radio"/> 18.	
<input type="radio"/> 23.	
<input type="radio"/> 55.	
<input checked="" type="radio"/> 37.	

Pergunta 2	1,66 / 1,66 pts

Qual é a proporção aproximada de pacientes que compareceram à consulta?

☐ 60%.

☐ 70%.

☒ 80%.

☐ 50%.

Pergunta 3

1,66 / 1,66 pts

Qual é a menor e maior idade, respectivamente, presente no dataset?

☐ 1, 113.

☐ 6, 113.

☐ 0, 115.

☒ -1, 115.

Pergunta 4

1,66 / 1,66 pts

Qual região possui o maior número de consultas marcadas?

☐ 64.

☒ 38.

☐ 16.

☐ 1.

Pergunta 5

1,66 / 1,66 pts

Qual foi o tipo de transformação dos dados categóricos realizado?

☐ Regression encoding.

☐ Linear encoding.

☐ One Hot encoding.

☒ Label encoding.

Pergunta 6

1,66 / 1,66 pts

Quantos valores diferentes encontramos para a variável dos SMS recebidos (SMS_received)?

☐ 12.

☒ 2.

☐ 4.

☐ 8.

Pergunta 7

1,66 / 1,66 pts

Em quantas localidades diferentes (Neighbourhood) os pacientes residem?

☐ 34.☐ 103.☒ 81.☐ 27.**Pergunta 8****1,66 / 1,66 pts**

Qual é a proporção de pacientes que receberam o sms e NÃO compareceram?

☐ 50%.☐ 16%.☒ 27%.☐ 72%.**Pergunta 9****1,66 / 1,66 pts**

Sobre a matriz de confusão para essa classificação, é CORRETO afirmar:

☒

Para os dados de teste existem mais instâncias da saída zero (0 - compareceram à consulta).

☐

A matriz de confusão apresenta a correlação entre as variáveis entrada x saída.



Para os dados de teste existem mais instâncias da saída um (1 - não compareceram à consulta).



Para os dados de teste, o melhor desempenho ocorreu para o a entrada idade ('Age')

Pergunta 10

1,66 / 1,66 pts

Qual das colunas do dataset possuem valores nulos?



Age.



Nenhuma.



Neighbourhood.



Handcap.

Pergunta 11

1,66 / 1,66 pts

Comparando apenas a identificação do paciente (PatientId) com o número dos agendamentos das consultas (AppointmentID) o que podemos inferir?



A idade dos pacientes é superior a 35 anos.



Existem pacientes que realizam mais de 1 atendimento.



O número de consultas realizadas em dias de semana é maior.



Pacientes com maior idade comparecem menos às consultas.

Pergunta 12**1,66 / 1,66 pts**

Quantas variáveis binárias (apenas dois valores) existem no dataset?

☒ 7.☐ 14.☐ 2.☐ 10.**Pergunta 13****1,66 / 1,66 pts**

Qual foi a acurácia obtida com a aplicação do modelo?

☒ 73%.☐ 50%.☐ 90%.☐ 30%.**Pergunta 14****1,66 / 1,66 pts**

Quantas instâncias e atributos existem no dataset, respectivamente?

☐ (14,110527).☐ (12,100456).

☐ (100456, 12).

☒ (110527, 14).

Pergunta 15

1,76 / 1,76 pts

Sobre a matriz de confusão para os dados de teste, é correto afirmar, EXCETO:



O desempenho do algoritmo para a previsão da classe 1 (não compareceu) é superior ao da classe 0 (compareceu).



Mais de 6300 instâncias foram classificadas (previstas) como pertencentes à classe 1.



Mais de 26000 instâncias de teste pertencem à classe 0 (compareceram à consulta).



Aproximadamente, 80% dos dados são da classe 0 (compareceu).

Pontuação do teste: **25** de 25