



**Observe que:**

Deverá realizar as partes A e B em folhas separadas, colocando em ambas nome e nº de estudante.

Exame com consulta

São permitidos meios eletrónicos (computador, calculadora, etc.), porém com todas as comunicações desativadas.

Qualquer tentativa de fraude conduzirá à anulação da prova para todos os intervenientes.

**Parte A:**

1 – Considere o seguinte conjunto de variáveis categóricas adquiridas de um conjunto de 11 doentes:

Antecedentes familiares	S	S	S	S	S	N	N	N	N	N	N
Risco Cardiovascular	A	A	A	A	B	B	B	A	A	A	A

Assumindo que S="Sim", N="Não" e que A="Alto", B="Baixo", considere que lhe é pedido que projete um preditor do risco cardiovascular usando o atributo "Antecedentes Familiares". Nesse contexto, responda às seguintes questões:

- a) Apresente o ranking de atributos usando a metodologia baseada na Informação mútua. Deverá apresentar todos os cálculos necessários.

2 – Considere o conjunto de pontos (x,y) das classes C1 e C2, respetivamente:

Tabela 1: Pontos das classes C1 e C2.

Classe C1 (x,y)	Classe C2 (x,y)
P1=(4, 5)	P26=(6, 2)
P2=(3, 6)	P27=(13, 14)
P3=(4, 8)	P28=(25, 31)
P4=(8, 10)	P29=(26, 28)
P5=(6, 1)	P30=(32, 29)
P6=(18, 5)	
P7=(8, 16)	
P8=(3, 0)	
P9=(18, 18)	
P10=(19, 14)	
P11=(8, 9)	
P12=(2, 11)	
P13=(5, 4)	
P14=(8, 9)	
P15=(11, 19)	
P16=(5, 10)	
P17=(12, 10)	
P18=(14, 4)	
P19=(4, 9)	
P20=(2, 12)	

P21=(28, 32) P22=(22, 33) P23=(22, 30) P24=(34, 16) P25=(15, 20)	
--	--

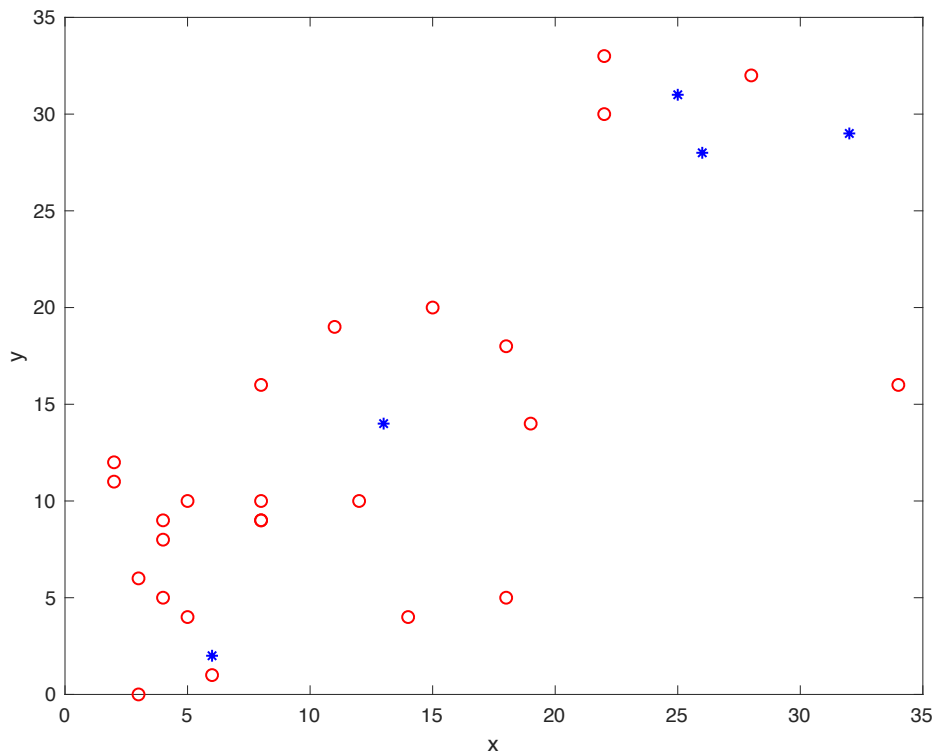


Figura 1: 'o' pontos da classe C1, '\*' pontos da classe C2.

- Usando o DBScan com  $\epsilon = 10$  e  $\text{minPts} = 4$  indique, fundamentando, o resultado da classificação do ponto P24.
- Considere o ponto  $P31 = (10, 10)$  cuja classe não é conhecida. Sugira uma estratégia de imputação para este ponto e apresente o respectivo resultado.
- Usando a metodologia *Adaptive Synthetic Sampling Approach* com  $K=5$ , determine o conjunto de dados decorrentes do processamento dos pontos P24 e P28. Considere que  $\beta = 1$ . Apresente todos os cálculos necessários.
- Determine a projeção que maximiza a variância explicada dos dados da tabela 1. Apresente todas as matrizes e indique todos os cálculos necessários.

### Parte B:

3 – Considere um sinal de tempo discreto  $x[n]$ , periódico de período  $N = 40$ , que resultou da amostragem de um sinal  $x(t)$  à frequência de amostragem  $f_s = 200$  Hz.

3.1 Calcule a frequência angular  $\omega_0$  do sinal  $x(t)$ .

3.2 Considerando que os coeficientes não nulos da Série de Fourier trigonométrica do sinal são dados pela seguinte tabela, indique as frequências angulares do sinal  $x(t)$ , em rad/s. Nota: se não conseguiu responder à alínea 3.1, considere  $\omega_0 = 5\pi$ .

<b>N</b>	1	5	7
<b>C</b>	2	3	4
<b><math>\theta</math> (rad)</b>	$\pi$	0	$\pi$

3.3 Com base nos coeficientes da alínea anterior, calcule o valor de  $x(10)$ .

4 – Considere o sinal  $y(t)$  representado na figura 2. Sobre ele foi realizado um processo de “Empirical Mode Decomposition”, do qual resultaram duas Intrinsic Mode Functions (IMF-A e IMF-B).

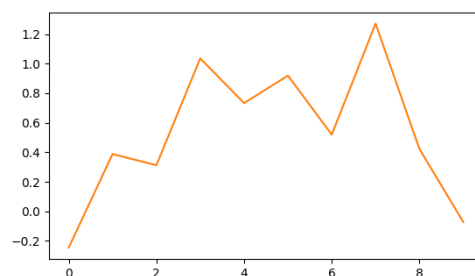


Figura 2 – Sinal  $y(t)$

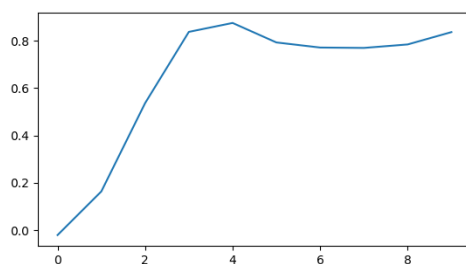


Figura 4 – IMF-A.

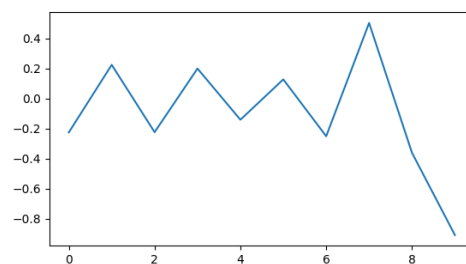


Figura 5 – IMF-B.

4.1 – Descreva os diferentes passos da decomposição EMD

4.2 – Indique duas vantagens da decomposição EMD face à análise tempo-frequência por wavelets

4.3 – Qual das IMFs (A ou B) corresponde à primeira IMF gerada pelo processo EMD? Justifique a sua resposta.