

---

## Aprendizagem Automática

FICHA LABORATORIAL N. 3

ENUNCIADO

Semestre de Inverno 2017/2018

Nome: Rui Filipe Guimarães Dos Santos

Número: A39286

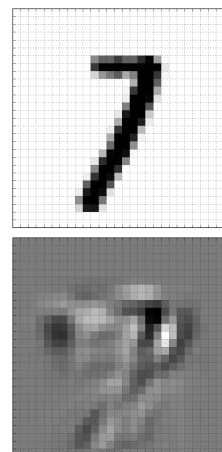
---

ATENÇÃO: Fixa de respostas múltiplas. Só uma única resposta em cada alínea está correta. Cada alínea vale 2 valores. Respostas erradas descontam 0.5 valores.

---

1. Considere o conjunto de dígitos manuscritos do ficheiro `mnist_small.p`. Considere ainda as 393 primeiras imagens do conjunto de treino do dígito “7”.
  - (a) Para esta alínea, arredonde os valores pedidos a 0 casas decimais. Considere que os dígitos das colunas 3, 81, 212, e 378 da matriz de dados (exemplos nº 4, 82, 213 e 379) são representados pelos vetores  $\mathbf{x}_1$ ,  $\mathbf{x}_2$ ,  $\mathbf{x}_3$ , e  $\mathbf{x}_4$ .
    - i. O produto interno entre os vetores  $\mathbf{x}_1$  e  $\mathbf{x}_2$  é 1725005.
    - ii. O produto interno entre os vetores  $\mathbf{x}_3$  e  $\mathbf{x}_4$  é 2770255.
    - iii. Todas as respostas anteriores.
    - iv. Nenhuma das respostas anteriores.
  - (b) Considere o dígito da coluna 152 da matriz de dados (exemplo nº 153). Considere ainda que se pretende projetar e reconstruir este dígito usando as 41 primeiras componentes principais. O erro quadrático médio entre o dígito original e o reconstruído é (arredondando a 0 casas decimais):
    - i. 766
    - ii. 696
    - iii. 1115
    - iv. 774
  - (c) O número de pixeis com valor igual a zero para todos os exemplos deste conjunto é:
    - i. 279
    - ii. 241
    - iii. 280
    - iv. 278
  - (d) Considere que se estima a matriz de covariância deste conjunto.
    - i. A dimensão da matriz é  $784 \times 784$
    - ii. O determinante da matriz de covariância é igual a 0
    - iii. Todas as respostas anteriores.
    - iv. Nenhuma das respostas anteriores.
  - (e) . Pretende-se visualizar os dígitos deste conjunto e os vetores próprios da matriz de covariância dos dados.

- i. A figura de cima é a imagem invertida do 208º dígito do conjunto (coluna 207 da matriz de dados).
- ii. A figura de baixo é a imagem do 5º vetor próprio da matriz de covariância dos dados.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.



- (f) Para esta alínea arredonde os valores pedidos a 0 casas decimais. Considere o dígito da coluna 184 da matriz de dados (exemplo nº 185). Pretende-se projetar e reconstruir este dígito usando um número mínimo de componentes principais tal que os dados projetados tenham 70% do total da variância dos dados originais. O erro absoluto médio entre o dígito original e o reconstruído é:

- i. 11
- ii. 29
- iii. 20
- iv. 27

- (g) Para esta alínea, arredonde os valores pedidos a 0 casas decimais. Note ainda, que tendo os dados  $d$  dimensões, estas serão indexadas com números de 1 a  $d$ .

- i. A norma do vetor de média dos dados é 810.
- ii. O desvio padrão da dimensão 321 dos dados é 92.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.

- (h) O número de valores próprios superiores a  $10^{-10}$  da matriz de covariância dos dados é:

- i. 385
- ii. 392
- iii. 391
- iv. 333

- (i) Para esta alínea, arredonde os valores pedidos a 2 casas decimais. Note ainda, que tendo os dados  $d$  dimensões, estas serão indexadas com números de 1 a  $d$ .

- i. Existem 308 dimensões dos dados que não variam.
- ii. O coeficiente de correlação entre a dimensão 185 e 249 é -0.04.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.

- (j) Pretende-se projetar os dígitos nas suas componentes principais. O número mínimo de componentes principais para que os dados projetados tenham acima de 95% da variância total dos dados originais, é:

- i. 43
- ii. 91
- iii. 86
- iv. 113