

---

## Aprendizagem Automática

FICHA LABORATORIAL N. 3

ENUNCIADO

Semestre de Inverno 2017/2018

Nome: Ana Vanessa Dos Santos Nunes

Número: A41499

---

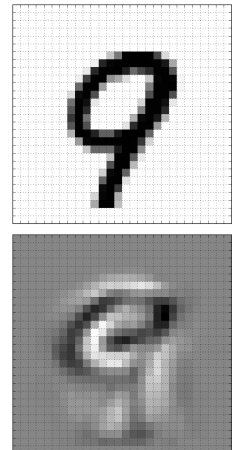
ATENÇÃO: Fixa de respostas múltiplas. Só uma única resposta em cada alínea está correta. Cada alínea vale 2 valores. Respostas erradas descontam 0.5 valores.

---

1. Considere o conjunto de dígitos manuscritos do ficheiro `mnist_small.p`. Considere ainda as 363 primeiras imagens do conjunto de treino do dígito “9”.

- (a) 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 1659.
  - ii. O desvio padrão da dimensão 206 dos dados é 38.
  - iii. Todas as respostas anteriores.
  - iv. Nenhuma das respostas anteriores.
- (b) . 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 214º dígito do conjunto (coluna 213 da matriz de dados).
- ii. A figura de baixo é a imagem do 1º vetor próprio da matriz de covariância dos dados.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.



- (c) Para esta alínea, arredonde os valores pedidos a 0 casas decimais. Considere que os dígitos das colunas 24, 62, 347, e 354 da matriz de dados (exemplos nº 25, 63, 348 e 355) 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 9036.
  - ii. O produto interno entre os vetores  $\mathbf{x}_3$  e  $\mathbf{x}_4$  é 8692.
  - iii. Todas as respostas anteriores.
  - iv. Nenhuma das respostas anteriores.
- (d) 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 337 dimensões dos dados que não variam.
  - ii. O coeficiente de correlação entre a dimensão 719 e 629 é -0.02.

- iii. Todas as respostas anteriores.  
iv. Nenhuma das respostas anteriores.
- (e) O número de pixeis com valor igual a zero para todos os exemplos deste conjunto é:
- i. 302                      ii. 447                      iii. 219                      iv. 233
- (f) Pretende-se projetar os dígitos nas suas componentes principais. O número mínimo de componentes principais para que os dados projetados tenham acima de 90 da variância total dos dados originais, é:
- i. 50                      ii. 54                      iii. 32                      iv. 49
- (g) O número de valores próprios superiores a  $10^{-10}$  da matriz de covariância dos dados é:
- i. 384                      ii. 362                      iii. 219                      iv. 435
- (h) 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.
- (i) Considere o dígito da coluna 312 da matriz de dados (exemplo nº 313). Considere ainda que se pretende projetar e reconstruir este dígito usando as 14 primeiras componentes principais. O erro quadrático médio entre o dígito original e o reconstruído é (arredondando a 0 casas decimais):
- i. 625                      ii. 788                      iii. 889                      iv. 949
- (j) Para esta alínea arredonde os valores pedidos a 0 casas decimais. Considere o dígito da coluna 362 da matriz de dados (exemplo nº 363). Pretende-se projetar e reconstruir este dígito usando um número mínimo de componentes principais tal que os dados projetados tenham 68% do total da variância dos dados originais. O erro absoluto médio entre o dígito original e o reconstruído é:
- i. 15                      ii. 29                      iii. 11                      iv. 20