
Aprendizagem Automática

FICHA LABORATORIAL N. 3

ENUNCIADO

Nome: Luis Carlos Semedo Da Fonseca

Número: A45125

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

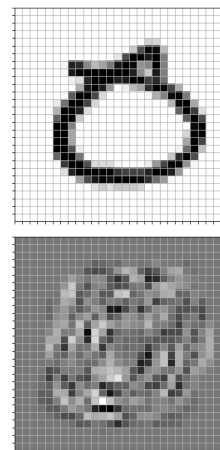
ATENÇÃO: Considere para esta ficha o conjunto de dígitos manuscritos do ficheiro `MNISTsmall.p`. Não modifique os dados (ex: não remova a média ao conjunto de dados).

Os exemplos do conjunto (colunas da matriz X) bem como as dimensões dos dígitos (linhas da matriz X) estão indexadas a começar por um e não zero.

1. Considere os 321 primeiros exemplos do conjunto de teste do dígito “0”.
 - (a) Considere que se estima a matriz de covariância deste conjunto.
 - i. A dimensão da matriz é 784×784 .
 - ii. O determinante da matriz de covariância é igual a 0.
 - iii. Todas as respostas anteriores.
 - iv. Nenhuma das respostas anteriores.
 - (b) Para esta alínea arredonde os valores pedidos a 0 casas decimais. Considere o 283º dígito da matriz de dados. Pretende-se projetar e reconstruir este dígito usando um número mínimo de componentes principais tal que os dados projetados tenham 75% do total da variância dos dados originais. O erro absoluto médio entre o dígito original e o reconstruído é:

i. 36	ii. 23	iii. 18	iv. 19
-------	--------	---------	--------
 - (c) Para esta alínea, arredonde os valores pedidos a 0 casas decimais.
 - i. A norma do vetor de média dos dados é 959.
 - ii. O desvio padrão da dimensão 153 dos dados é 107.
 - iii. Todas as respostas anteriores.
 - iv. Nenhuma das respostas anteriores.
 - (d) . 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 239º dígito do conjunto.
- ii. A figura de baixo é a imagem (ou a imagem invertida) do 17º vetor próprio da matriz de covariância dos dados.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.



(e) Considere o 262º dígito da matriz de dados. Considere ainda que se pretende projetar e reconstruir este dígito usando as 43 primeiras componentes principais. O erro quadrático médio entre o dígito original e o reconstruído é (arredondando a 0 casas decimais):

- i. 691
- ii. 711
- iii. 681
- iv. 561

(f) Pretende-se projetar os dígitos nas suas componentes principais. O número mínimo de componentes principais para que os dados projetados tenham acima de 60% da variância total dos dados originais, é:

- i. 9
- ii. 14
- iii. 4
- iv. 7

(g) O número de pixels com valor igual a zero para todos os exemplos deste conjunto é:

- i. 333
- ii. 322
- iii. 449
- iv. 410

(h) Para esta alínea, arredonde os valores pedidos a 2 casas decimais.

- i. Existem 303 dimensões dos dados que não variam.
- ii. O coeficiente de correlação entre a dimensão 245 e 635 é -0.02.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.

(i) Para esta alínea, arredonde os valores pedidos a 0 casas decimais. Considere que os 45º, 149º, 267º e 278º dígitos 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 5311517.
- ii. O produto interno entre os vetores \mathbf{x}_3 e \mathbf{x}_4 é 5567856.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.

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

- i. 267
- ii. 282
- iii. 274
- iv. 320