
Aprendizagem Automática

FICHA LABORATORIAL N. 3

ENUNCIADO

Nome: Pedro Miguel Pereira Henriques

Número: A45415

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 293 primeiros exemplos do conjunto de teste do dígito “7”.

(a) Para esta alínea, arredonde os valores pedidos a 0 casas decimais. Considere que os 50º, 151º, 208º e 209º dígitos são representados pelos vetores x_1 , x_2 , x_3 , e x_4 .

- i. O produto interno entre os vetores x_1 e x_2 é 2057026.
- ii. O produto interno entre os vetores x_3 e x_4 é 1884538.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.

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

- i. 292
- ii. 315
- iii. 155
- iv. 313

(c) Considere o 162º dígito da matriz de dados. Considere ainda que se pretende projetar e reconstruir este dígito usando as 59 primeiras componentes principais. O erro quadrático médio entre o dígito original e o reconstruído é (arredondando a 0 casas decimais):

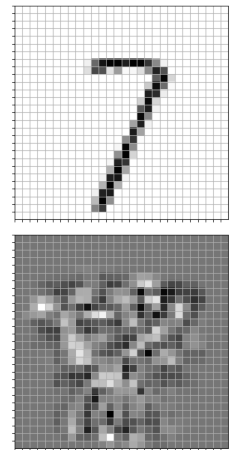
- i. 231
- ii. 181
- iii. 271
- iv. 261

(d) 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. 12
- ii. 14
- iii. 10
- iv. 7

(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 121º dígito do conjunto.
- ii. A figura de baixo é a imagem (ou a imagem invertida) do 10º vetor próprio da matriz de covariância dos dados.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.



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

- i. 457
- ii. 253
- iii. 242
- iv. 306

(g) Para esta alínea, arredonde os valores pedidos a 0 casas decimais.

- i. A norma do vetor de média dos dados é 1461.
- ii. O desvio padrão da dimensão 322 dos dados é 110.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.

(h) Considere que se estima a matriz de covariância deste conjunto.

- i. A dimensão da matriz é 293×293 .
- ii. O determinante da matriz de covariância é igual a 0.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.

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

- i. Existem 306 dimensões dos dados que não variam.
- ii. O coeficiente de correlação entre a dimensão 230 e 601 é -0.11.
- iii. Todas as respostas anteriores.
- iv. Nenhuma das respostas anteriores.

(j) Para esta alínea arredonde os valores pedidos a 0 casas decimais. Considere o 177º 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 53% do total da variância dos dados originais. O erro absoluto médio entre o dígito original e o reconstruído é:

- i. 29
- ii. 25
- iii. 23
- iv. 14