COM2028/12/Semr2 13/14

UNIVERSITY OF SURREY ©

Faculty of Engineering and Physical Sciences

Department of Computing

Undergraduate Programmes in Computing
Undergraduate Programmes in Mathematics

Module COM2028: 15 credits

Introduction to Artificial Intelligence

FHEQ Level 5 Examination

Time allowed: Two hours

Semester 2 2013/14

Answer **ALL** questions

Approved calculators ARE permitted

Question 1, 2, 3, 4, 5, carry 35, 18, 17, 15, and 15 marks respectively

Where appropriate the mark carried by an individual part of a question is indicated in square brackets [ ]

COM2028/12/Semr2 13/14

Solutions

1. This question is about *general understanding of the subject area.*

(a) Examining the image in Figure 1, what is the best option to remove both salt and pepper noise (i.e. those tiny white and black pixels with intensity values of 255 and 0)?

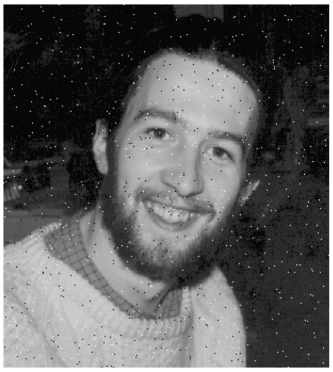

Figure 1: A picture with salt and pepper noise

A. A Gaussian filter

B. A median filter

C. Histogram equalisation

D. Edge detection

[3 marks]

B, median filter

(b) The following are descriptions of features that can be used for classification.

A. Full red, green and blue colour histograms.

B. Average red, green and blue colours in the image.

C. The height and width of the image.

D. Histogram of Oriented Gradients (HOG).

E. Frequency of a set of keywords.

QUESTION 1 CONTINUES ON NEXT PAGE

F. Sizes of image objects.

G. Centre coordinates of the objects in the images.

Examining the following scenarios, choose appropriate features for each application.

I. Finding heavy duty lorries in surveillance videos.

II. Classification of the types of fruits (e.g. oranges, kiwi fruits, bananas) scanned in images.

III. Classification of human and landscape pictures.

IV. Email spam filtering.

[8 marks]

I: D II: A. Answer of B, D, F are also acceptable. III. D IV: E

(c) The images from left to right in Figure 2 show an original picture and its two convolution results respectively. Choose from the following an appropriate convolution kernel that will transform the original image into each convolved result.

Figure 2: left: original image; middle and right: convolution results

Convolution kernels:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix} \; B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 1 \end{pmatrix} \; C = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix} \; D = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

[4 marks]

SEE NEXT PAGE

middle: C right: A

(d) Examine the following techniques, which can be categorised as supervised learning?

A. A single perceptron.

B. Back-propagation neural network.

C. Simulated annealing.

D. Hierarchical clustering.

E. K-means clustering.

F. K-nearest neighbour.

[6 marks]

A B F

(e) Given an image with dimensions 256 x 256. The parameters for its feature descriptor using Histogram of Oriented Gradients (HOG) are: orientations=8, pixels per cell = 12 x 12, cells per block = 2 x 2.

Estimate the dimension of its HOG feature descriptor.

A. 12

B. 256

C. 144

D. 400

E. 12800

F. 4000000

[3 marks]

SEE NEXT PAGE

E, 12800

(f) When building a classifier, why should we split the data into two parts, i.e. training data and testing data ? [3 marks]

We split the data set into these two parts because we wouldn't want to test the system performance based on the samples that the system has already seen before. We want to know how much the system has generalised its knowledge about a class after learning from some examples.

(g) In K-nearest neighbour, if we haven't split the data into training and testing sets, what will be the outcome if we have set up k=1, and use one of the samples for testing? [3 marks]

It will always be 100% accurate. We will retrieve same sample as its nearest point will be itself.

(h) Which of the following can describe Support Vector Machines (SVMs)?

A. The algorithm needs to find the support vectors that determine maximum-margin hyperplane.

B. Support vectors are on the hyperplane.

C. The data samples are usually transformed to another feature space so that the data of different classes are divided by a minimum-margin hyperplane.

D. The data samples are usually mapped to another feature space so that the data of different classes are divided by a clear gap that is as wide as possible.

E. SVMs are unsupervised methods.

[5 marks]

A, D

Total marks: 35

SEE NEXT PAGE

2. This question is about *K-means clustering, K-nearest mean, and K-nearest neighbour.*

   The following is a list of possible processing steps for building the above classifiers.

   A. extract features.

   B. divide the data samples into training sets and testing sets.

   C. calculate the centres of current group by averaging current members.

   D. calculate the distance between the query and individual samples.

   E. calculate the distance between the query and each cluster/group centres.

   F. calculate the distance between each sample and each current class centre.

   G. assign random centres for each group.

   H. select a testing sample.

   I. decide parameter K.

   J. choose k most similar samples and check their majority class.

   For each classifier, i.e. K-means clustering, K-nearest mean, K-nearest neighbour, choose appropriate steps from above list and put them in the order that describes the processes for training and testing.

   [18 marks](6 marks for each method)

   K-means Clustering: Training: B, A, I, G, F, C; Testing: H, E

   K-nearest mean: Training: B, A, C; Testing: H, E

   K-NN:

   B, A, I, H, D, J (I can be anywhere before J).

Total marks: 18

SEE NEXT PAGE

3. This question is about *Neural Networks*.

(a) Design an autonomous vehicle using a multilayer perceptron neural network. The input of the network is a $30 * 32$ two dimensional image which receives input from the vehicle video camera. There is one hidden layer with four hidden neurons. The output layer is a linear representation of the directions the vehicle should travel in order to keep the vehicle on the road. Assume there are 30 steering directions.

A. Draw the architecture of the neural network. [4 marks]

B. What is the dimension of the feature vector? [2 marks]

C. Explain the learning process. [4 marks]

D. Explain the testing process, i.e. the actual autonomous driving process. [4 marks]

(b) Could the above autonomous vehicle be implemented using a self-organising feature map? Why? [3 marks]

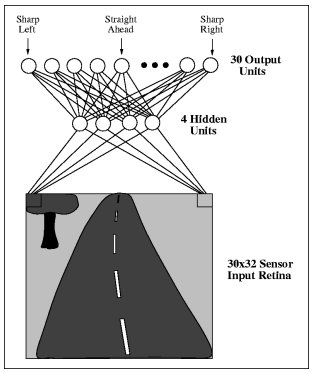

Figure 3: A neural network architecture for autonomous vehicle

A. The architecture of the system.

See Figure 3

B. The dimension of the feature vector?

30x32 = 960

C. The learning process.

Samples of road images are captured with the corresponding steering directions. Such pairs of samples are fed into the neural network for training.

SEE NEXT PAGE

The input layer has 960 node, one middle layer with 4 hidden nodes. There are 30 output nodes each representing a steering direction. Initially the weights on all connections are randomly assigned. Through feeding forward and back propagation the weights are gradually adjusted until converged.

D. The testing process, i.e. the actual autonomous driving process.

During the autonomous driving process, the system will extract features from the image captured from the camera, feed them into the neural network and check the most activated output node for corresponding steering action.

(b) If we have 30 output neurons in SOM, it is possible to train the SOM to associate those 30 neruons to 30 kinds of inputs. However this is not guaranteed to be accurate as SOM is unsupervised method, the data are grouped together based on their similarity. When data are ambiguous, the grouping may not be the same way as human would like it to be.

Total marks: 17

SEE NEXT PAGE

4. This question is about *the use of a Gaussian Classifier.*

The Gaussian probability density function for category w_i is given as below:

$$P(x \mid w_i) = [(2\pi)^d \mid \sum\nolimits_i \mid]^{-\frac{1}{2}} \exp[-\frac{1}{2}(x-u_i)^T \sum\nolimits_i^{-1}(x-u_i)] \quad (1)$$

where

u_i — the mean vector of class w_i,

$\sum_i$ — the ith class covariance matrix.

x is the feature vector.

u_i and $\sum_i$ are calculated from training samples belonging to category w_i

$u_i = \frac{1}{N_i} \sum_{j=1}^{N_i} x_j, x_j \in w_i$, where N_i is the number of training samples from class w_i.

The covariance matrix as

$$\sum\nolimits_i = \frac{1}{N_i} \sum_{j=1}^{N_i} (x_j - u_i)(x_j - u_i)^T \quad (2)$$

Design a system that is able to classify images of mountain and building scenes using a Gaussian Classifier. Describe your proposed system with regards to the following aspects:

A. What features can be used? [3 marks]

B. After the system is trained, how does the system hold the knowledge about each class? [4 marks]

C. How to test the system? [8 marks]

A, Feature: Since the shape plays an important role to distinguish mountains and buildings, so HOG descriptors will be appropriate as features.

B, During the training processing, u_i and $\sum_i$ are calculated. $P(x \mid w_1)$ for mountain class w_1 and $P(x \mid w_2)$ for building class w_2 according to equation (1) will be obtained as knowledge about each class.

C, Testing stage

SEE NEXT PAGE

Based on Bayes Theorem, we have

$$P(x, w_i) = P(x) * P(w_i \mid x) = P(w_i) * P(x \mid w_i)$$
$$\Rightarrow P(w_i \mid x) = \frac{P(w_i) * P(x \mid w_i)}{P(x)} \quad (3)$$

To simplify the problem here, we suppose $P(x)$ and $P(w_i)$ are scale factors or constants, and $P(w_1) = P(w_2)$, so

$$P(w_i \mid x) = \lambda\, P(x \mid w_i), \quad (4)$$

where λ is a constant to make sure $\sum_i P(w_i \mid x) = 1$

After the training stage, we have obtained one Gaussian function for category $w_1 : P(x \mid w_1)$, and the another Gaussian function for category $w_2 : P(x \mid w_2)$. When testing on an unknown sample x, from the above theoretical inference, we get the posterior probability: $P(w_1 \mid x)$ and $P(w_2 \mid x)$.

If $P(w_1 \mid x) > P(w_2 \mid x)$, assign x to w_1; otherwise assign x to w_2.

Total marks: 15

SEE NEXT PAGE

5. This question is about *Genetic Algorithms*. Design a method for the computer to draw your own portrait using a genetic algorithm. Answer the questions and complete the descriptions as below.

   A. How should a candidate solution be represented, i.e. how does the chromosomes look like in a candidate solution? [3 marks]

   B. Describe the objective function. [3 marks]

   C. Describe the mutation process. [3 marks]

   D. Describe the crossover process. [3 marks]

   E. How to choose the best individuals? [3 marks]

   A. representation of solution.

      The solution can be just list of 1D image intensities flattened from 2D image.

   B. Describe the objective function.

      Objective function can be the Euclidean distance between the target image (your portrait image) and the current evolving image. Essentially it is the Euclidean distance between the vector of current evolving image and the target vector of the portrait image.

   C. mutation process.

      Individual chromosome in a solution is randomly changed to another intensity.

   D. crossover process.

      Two solutions exchange some part of the chromosomes.

   E. How should the best individuals be chosen?

      The current solutions are ranked based on their Euclidean distances to the target image. The top ones are chosen.

Total marks: 15

END OF PAPER

INTERNAL EXAMINER: DR. H.L.TANG
EXTERNAL EXAMINER: PROF. DAVID MARSHALL