
UNIVERSITI SAINS MALAYSIA

First Semester Examination
2010/2011 Academic Session

November 2010

CCS513 – Computer Vision and Image Analysis
[Penglihatan Komputer dan Analisis Imej]

Duration : 2 hours
[Masa : 2 jam]

INSTRUCTIONS TO CANDIDATE:
[ARAHAN KEPADA CALON:]

- Please ensure that this examination paper contains **FOUR** questions in **SEVEN** printed pages before you begin the examination.

*[Sila pastikan bahawa kertas peperiksaan ini mengandungi **EMPAT** soalan di dalam **TUJUH** muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]*

- Answer **ALL** questions.

*[Jawab **SEMUA** soalan.]*

- You may answer the questions either in English or in Bahasa Malaysia.

[Anda dibenarkan menjawab soalan sama ada dalam bahasa Inggeris atau bahasa Malaysia.]

- In the event of any discrepancies, the English version shall be used.

[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi bahasa Inggeris hendaklah diguna pakai.]

1. (a) Given three binary bit strings 010111, 101001 and 111100, compute the dissimilarity metrics (distance metrics) for these three points as given below, using the Jacquard and Cosine distance metrics.

JACQUARD	PT1	PT2	PT3
----------	-----	-----	-----

PT 1			
------	--	--	--

PT 2			
------	--	--	--

PT3			
-----	--	--	--

COSINE	PT1	PT2	PT3
--------	-----	-----	-----

PT 1			
------	--	--	--

PT 2			
------	--	--	--

PT3			
-----	--	--	--

(40/100)

- (b) Given the 5 data points below, show all steps for a 2-means clustering algorithm ($k=2$) where the initial mean estimates are (1.67, 0.67) and (3.25, 1).

$$\langle 2.1, 3.4 \rangle, \langle 1.8, 3.4 \rangle, \langle 0.7, 2.0 \rangle, \langle 3.4, 2.0 \rangle, \langle 3.6, 2.0 \rangle$$

Run the algorithm to convergence, that is until the clusters do not change.

(30/100)

- (c) Complete this problem without the use of a computer to make sure that you understand the details of the clustering algorithms. Consider the following "data" to be clustered as described below:

$$10 \ 20 \ 40 \ 80 \ 85 \ 121 \ 160 \ 168 \ 195$$

Assume that Euclidean distance will be used to measure the distance between the data points.

Use hierarchical agglomerative clustering with single linkage to cluster the data. Draw a dendrogram to illustrate your clustering.

(30/100)

2. (a) Give an example of a decision tree with two internal nodes (including the root), and explain how it classifies an example.

(30/100)

- (b) What kinds of techniques can be used to counter the problem of over-fitting in decision trees?

(30/100)

- (c) Consider the following data set with three Boolean predictive attributes, W, X, Y and Boolean classification C.

W	X	Y	C
T	T	T	T
T	F	T	F
T	F	F	F
F	T	T	F
F	F	F	T

We now encounter a new example: W=F, X=T, Y=F. If we apply the Naive Bayes method, what probability is assigned to the two values of C?

(40/100)

3. (a) Using image examples, explain the following concepts:

- (i) Filtering
- (ii) Low pass filtering
- (iii) High pass filtering

(60/100)

- (b) For the following image in Figure 1(a), compute the Sobel edge magnitude and direction images using the filters in Figures 1(b).

100	80	200	180	160
90	100	220	150	80
120	90	160	50	90
110	130	140	90	120
70	70	90	90	80

Figure 1(a): Image

-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1

Figure 1(b): Sobel Filters

(40/100)

4. (a) Briefly describe each of the following image segmentation methods and compare them:

(i) Segmentation by clustering

(ii) Segmentation by thresholding

(60/100)

(b) For the image in Figure 2, compute the Gray Level Co-Occurrence Matrix (GLCM).

1	2	4	8	6	5	4	5
5	4	8	6	6	4	2	1
1	1	5	6	8	4	4	1
2	3	5	7	1	8	9	7
4	5	7	1	2	6	4	2
8	5	1	2	5	9	1	2
4	2	3	3	6	4	2	7

Figure 2: Image

(40/100)

KERTAS SOALAN DALAM VERSI BAHASA MALAYSIA

[CCS513]

- 5 -

1. (a) Diberi tiga rentetan bit perduaan 010111, 101001 dan 111100, kirakan metrik ketidaksamaan (metrik jarak) untuk ketiga-tiga titik ini dengan menggunakan metrik jarak Jacquard dan Kosain (COSINE).

JACQUARD	PT1	PT2	PT3
----------	-----	-----	-----

PT 1

PT 2

PT3

KOSAIN	PT1	PT2	PT3
--------	-----	-----	-----

PT 1

PT 2

PT3

(40/100)

- (b) Diberi 5 data seperti di bawah, tunjukkan kesemua langkah untuk algoritma pengklusteran *2-means* ($k=2$) yang mana anggaran purata awal ialah (1.67, 0.67) dan (3.25, 1).

$\langle 2.1, 3.4 \rangle, \langle 1.8, 3.4 \rangle, \langle 0.7, 2.0 \rangle, \langle 3.4, 2.0 \rangle, \langle 3.6, 2.0 \rangle$

Laksanakan algoritma tersebut sehingga mencapai penumpuan, iaitu sehingga kluster-kluster tersebut tidak berubah.

(30/100)

- (c) Selesaikan masalah ini tanpa penggunaan komputer untuk memastikan bahawa anda memahami algoritma-algoritma pengklusteran dengan terperinci. Titik-titik data yang perlu diklusterkan dinyatakan di bawah:

10 20 40 80 85 121 160 168 195

Andaikan jarak Euclidean akan digunakan untuk mengukur jarak antara titik-titik data.

Gunakan pengklusteran aglomeratif berhierarki dengan kaitan tunggal untuk pengklusteran data tersebut. Lukiskan *dendrogram* untuk menggambarkan pengklusteran anda.

(30/100)

2. (a) Berikan satu contoh pepohon keputusan dengan dua nod dalaman (termasuk akar), dan jelaskan bagaimana ia mengelaskan suatu contoh.

(30/100)

- (b) Apakah jenis-jenis teknik yang boleh digunakan untuk menyelesaikan masalah terlebih-suai dalam pepohon keputusan?

(30/100)

- (c) Pertimbangkan set data berikut dengan tiga sifat ramalan Boolean, W, X, Y dan pengelasan Boolean C.

W	X	Y	C

T	T	T	T
T	F	T	F
T	F	F	F
F	T	T	F
F	F	F	T

Sekarang, kita menemui suatu contoh baru: W=F, X=T, Y=F. Jika kita menggunakan kaedah *Naive Bayes*, apakah kebarangkalian yang akan diberi kepada kedua-dua nilai C berkenaan?

(40/100)

3. (a) Dengan menggunakan contoh-contoh imej, jelaskan konsep-konsep berikut:

- (i) Penapisan
- (ii) Penapisan laluan rendah
- (iii) Penapisan laluan tinggi

(60/100)

- (b) Untuk imej berikut dalam Gambar Rajah 1(a), kirakan magnitud tepi Sobel dan imej-imej arah menggunakan penapis-penapis seperti dalam Gambar Rajah 1(b).

100	80	200	180	160
90	100	220	150	80
120	90	160	50	90
110	130	140	90	120
70	70	90	90	80

Gambar Rajah 1(a): Imej

-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1

Gambar Rajah 1(b): Penapis-penapis Sobel

(40/100)

4. (a) Huraikan secara ringkas setiap kaedah segmentasi imej berikut dan bandingkan antara mereka:

- (i) Segmentasi menggunakan pengklusteran
(ii) Segmentasi menggunakan *thresholding*

(60/100)

- (b) Untuk imej dalam Gambar Rajah 2, kirakan *Gray Level Co-Occurrence Matrix (GLCM)*.

1	2	4	8	6	5	4	5
5	4	8	6	6	4	2	1
1	1	5	6	8	4	4	1
2	3	5	7	1	8	9	7
4	5	7	1	2	6	4	2
8	5	1	2	5	9	1	2
4	2	3	3	6	4	2	7

Gambar Rajah 2: Imej

(40/100)