



มหาวิทยาลัยสงขลานครินทร์
คณะวิศวกรรมศาสตร์
ภาควิชาวิศวกรรมคอมพิวเตอร์

สอบกลางภาค: ภาคการศึกษาที่ 2

ปีการศึกษา: 2552

วันที่สอบ: 22 ธันวาคม 2552

เวลาสอบ: 14.30 - 16.30 น.

รหัสวิชา: 241- 588

ห้องสอบ: R 200

ชื่อวิชา: SP(INTRO TO MACHINE LEARNING)

คำสั่ง: อ่านรายละเอียดของข้อสอบ และคำแนะนำให้เข้าใจก่อนเริ่มทำข้อสอบ

อนุญาต: เอกสารการสอน, เครื่องคิดเลขและเครื่องเขียนต่าง ๆ นำเข้าห้องสอบ

ไม่อนุญาต: อุปกรณ์สื่อสาร นำเข้าห้องสอบ

เวลา: 2 ชั่วโมง (120 นาที)

คำแนะนำ:

- ข้อสอบมี 7 หน้า (รวมใบປะหน้า) มี 5 ข้อ คิดเป็นคะแนนเก็บ 20 %
- คำตอบทั้งหมดจะต้องเขียนลงในข้อสอบ
- เขียนชื่อ รหัสนักศึกษา ในทุกหน้าของข้อสอบให้ชัดเจน

ทุจริตในการสอบ โทษขั้นต่ำคือ
ปรับตกในรายวิชาที่ทุจริต และพักการเรียน 1 ภาคการศึกษา

ชื่อ _____ รหัสนักศึกษา _____

1. Learning Algorithm (10 นาที)

จงอธิบายพร้อมยกตัวอย่างประกอบความหมายของ Machine learning ในรูปแบบของ Tom Mitchell

๕๔ รหัสนักศึกษา _____

2. จงอธิบายความหมายพร้อมยกตัวอย่างของคำเหล่านี้ (20 นาที)

2.1 Learning input - output functions

2.2 Random Variables

2.3 Batch, Incremental (Stochastic) and online method

3. Naïve Bayes (30 นาที)

3.1 กำหนดให้ learning function คือ $X \rightarrow Y$ โดยที่ Y เป็นตัวแปรตระกูล, $X = \langle X_1, X_2 \rangle$ และให้ X_1 เป็นตัวแปรชนิดตรรกะ และ X_2 เป็นตัวแปรชนิดจำนวนจริง แสดงสมการในการหาค่า $P(Y|X)$ ในเทอมของตัวแปรดังกล่าว

3.2 กำหนดให้ joint distribution ของ random variables A, B, และ C เป็นดังนี้

A	B	C	P(A,B,C)
0	0	0	1/8
0	1	0	1/8
0	0	1	1/8
0	1	1	1/8
1	0	0	1/8
1	1	0	1/8
1	0	1	1/8
1	1	1	1/8

- a) จริงหรือเท็จ ที่ A เป็น conditionally independent ของ $B|C$ (B given C) จงอธิบายสั้น ๆ

- b) ถ้าคำตอบในข้อ (a) จริงให้ทำการเปลี่ยนแปลงค่าในตาราง 2 แล้วแรกเพื่อให้คำตอบเป็นเท็จ
ถ้าคำตอบในข้อ (a) เท็จให้ทำการเปลี่ยนแปลงค่าในตาราง 2 แล้วแรกเพื่อให้คำตอบเป็นจริง

4. Linear Regression (30 นาที)

กำหนดให้ข้อมูลต่อไปนี้

x	-1	0	2
y	1	-1	1

- a) กำหนดให้ $Y_i = \beta_0 + \epsilon_i$ เป็น linear regression model ของ β_0

- b) กำหนดให้ $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ เป็น linear regression model จงหาค่า β_0 และ β_1

5. Logistic Regression (30 นาที)

หากต้องสร้างโมเดลในการทำนายเพศ (ตัวแปรชนิดตระกาก, G) ของบุคคลทั่วไปโดยพิจารณาจากน้ำหนักของบุคคลนั้น (ตัวแปรชนิดจำนวนจริง, W) และเป็นนักศึกษาระดับบัณฑิตศึกษา (ตัวแปรชนิดตระกาก, S) โดยกำหนดให้ $P(W|G=\text{หญิง})$ เท่ากับ $P(W|G=\text{ชาย})$

5.1 เรายสามารถใช้ Naïve Bayes ในการ classifier ได้หรือไม่ เพราะเหตุใด

5.2 จงแสดง logistic function เพื่อใช้ในการทำนาย