

การทำเหมือง ข้อมูล

รองศาสตราจารย์ ดร.สุรพงศ์ เอื้อวัฒนมงคล
สำนักพิมพ์
สถาบันบัณฑิตพัฒนบริหารศาสตร์ (NIDA)

สารบัญ

	หน้า
บทที่ 1 การทำเหมืองข้อมูล	1
1.1 ขบวนการทำเหมืองข้อมูล	1
1.2 รูปแบบการทำเหมืองข้อมูล	2
1.3 ขั้นตอนการทำเหมืองข้อมูล	4
แบบฝึกหัด	6
บทที่ 2 ข้อมูล	7
2.1 ประเภทของข้อมูล	7
2.2 ชุดข้อมูล	8
2.3 การวัดความเหมือนของข้อมูล	9
แบบฝึกหัด	18
บทที่ 3 การเตรียมข้อมูล (Data Preprocessing)	19
3.1 คุณภาพข้อมูลและการเตรียมข้อมูล	19
3.2 การทำ Normalization	22
3.3 การทำ Standardization	23
3.4 การสุ่มข้อมูล (Sampling)	23
3.5 การลดจำนวนตัวแปร (Dimension Reduction)	24
แบบฝึกหัด	28
บทที่ 4 การสำรวจข้อมูล	29
4.1 Summary Statistics	29
4.2 Data Visualization	31
4.3 Online Analytical Processing (OLAP)	35
แบบฝึกหัด	37
บทที่ 5 การจำแนกประเภทของข้อมูล (Classification)	39
5.1 ตัวแบบ K-nearest Neighbors	42
5.2 ต้นไม้การตัดสินใจ (Decision Tree)	44
5.3 ตัวแบบการจำแนกประเภทข้อมูลแบบเบย์ส (Bayes Classifier)	53
5.4 ตัวแบบจำแนกประเภทข้อมูลแบบเครือข่ายประสาท (Neural Network)	55

	หน้า
5.5 ตัวแบบ Support Vector Machine (SVM)	65
5.6 Ensemble Method	75
5.7 การประเมินตัวแบบการจำแนกประเภทข้อมูล	77
5.8 Imbalance Issue	82
แบบฝึกหัด	84
บทที่ 6 การจับกลุ่มข้อมูล (Data Clustering)	87
6.1 วิธีจับกลุ่มข้อมูลแบบ K-Means	90
6.2 Hierarchical Clustering	95
6.3 การจับกลุ่มข้อมูลด้วยวิธี DBSCAN	99
6.4 การประเมินผลลัทธิการจับกลุ่มข้อมูล (Cluster Validity)	101
แบบฝึกหัด	106
บทที่ 7 การหากฎความสัมพันธ์ (Association Rule Analysis)	109
7.1 การหากฎความสัมพันธ์โดยวิธี Apriori	111
7.2 FP-Tree	115
7.3 การสร้างกฎความสัมพันธ์จาก Frequent Itemsets	118
7.4 Multiple Minimum Support	120
แบบฝึกหัด	123
บทที่ 8 การตรวจจับข้อมูลผิดปกติ (Anomaly Detection)	125
8.1 ขบวนการตรวจจับข้อมูลผิดปกติ	125
8.2 วิธีการตรวจจับข้อมูลผิดปกติ	126
แบบฝึกหัด	130
บรรณานุกรม	131