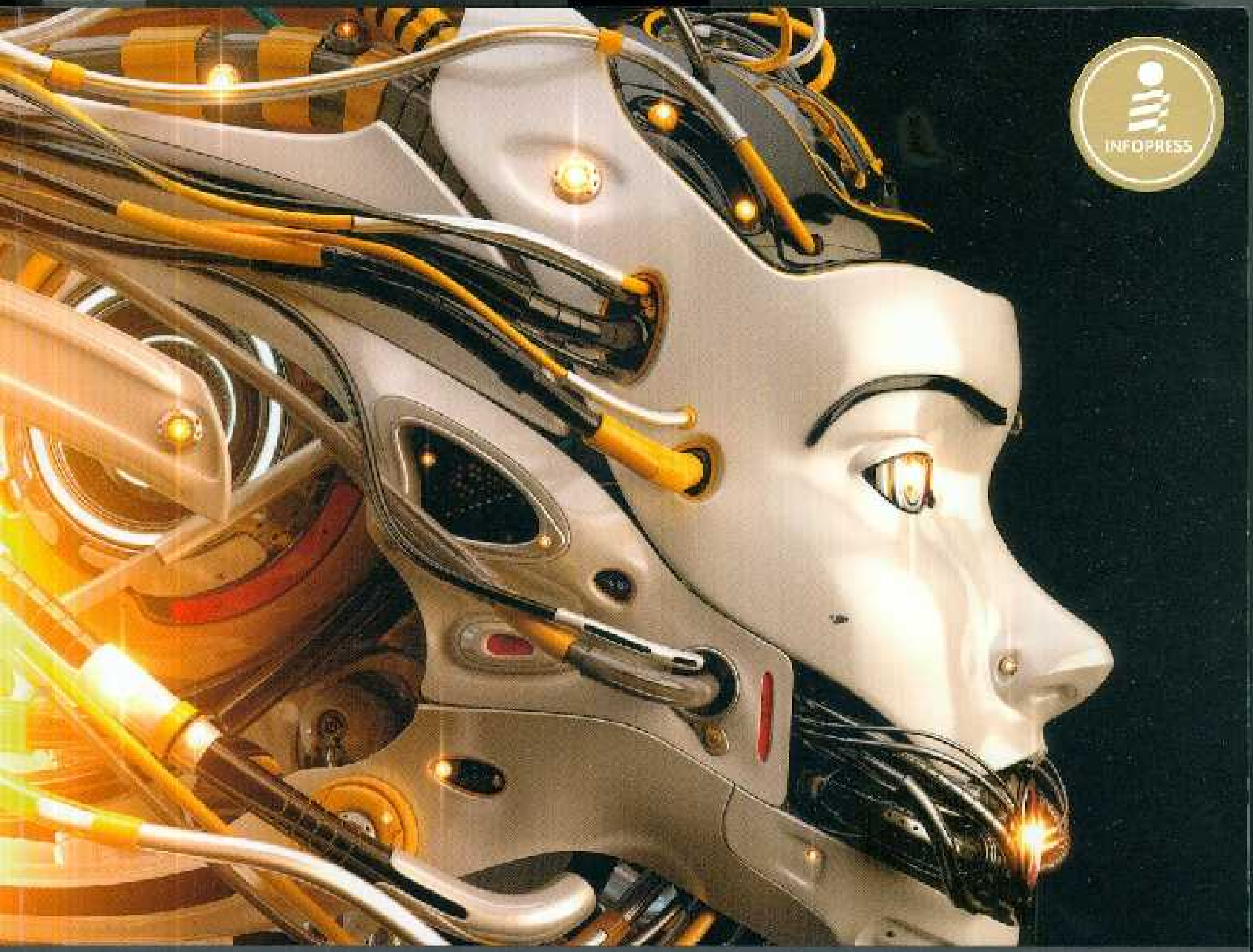

INFOPRESS

Artificial Intelligence with Machine Learning

AI สร้างได้ด้วยแมชชีนเลิร์นนิ่ง

Python Edition

เรียนอัลกอริทึมของ Machine Learning เพื่อสร้างสมองอันทรงพลังให้กับงานด้าน AI, Data Mining, Pattern Recognition, Computer Vision และงานสาขาอื่นที่เกี่ยวข้อง รวมถึงการปรับปรุงขั้นตอนวิธีเพื่อให้เหมาะสมกับการใช้งานได้อย่างมีประสิทธิภาพ

ผู้แต่ง รองศาสตราจารย์ ดร.ปริญญา สงวนสัตย์ บรรณาธิการ ภีรพล คชาเจริญ

CONTENTS

บทบรรณาธิการ III

คำนำ IV

สารบัญ VI

สารบัญรูป XII

สารบัญตาราง XIX

บทที่ 1 บทนำ 1

1.1 ปัญญาประดิษฐ์ (Artificial Intelligence: AI) 2

1.2 คำศัพท์ 4

1.3 สัญลักษณ์ทางคณิตศาสตร์ 6

1.4 การติดตั้งภาษา Python 6

1.5 เอกสารอ้างอิง 8

บทที่ 2 การเรียนรู้ของเครื่อง (Machine Learning) 9

2.1 จุดประสงค์ประจำบท 10

2.2 การเรียนรู้ของเครื่องและการเรียนรู้ของมนุษย์ 10

2.3 ลักษณะงานที่เหมาะสมกับการเรียนรู้ของเครื่อง 11

2.4 การออกแบบระบบเรียนรู้ 12

- 2.4.1 กำหนดปัญหา 12
- 2.4.2 เลือกรูปแบบของตัวอย่างหรือประสบการณ์ที่จะใช้ในการฝึกฝน 13
- 2.4.3 เลือกฟังก์ชันเป้าหมาย (Target Function) 14
- 2.4.4 การเรียนรู้ (Learning) 15

2.5 ตัวอย่างเกม Tic-Tac-Toe 18

- 2.5.1 ตัวอย่างโปรแกรมเกม Tic-Tac-Toe 20

2.6 เอกสารอ้างอิง 22

2.7 คำถามท้ายบท 23

บทที่ 3 การเรียนรู้แนวคิด (Concept Learning) 25

3.1 จุดประสงค์ประจำบท 26

3.2 สมมติฐาน (Hypothesis) 26

3.3 ขั้นตอนวิธี Find-S 27

3.4 ขั้นตอนวิธี List-Then-Eliminate 29

3.5 ขั้นตอนวิธี Candidate Elimination 33

3.6 เอกสารอ้างอิง 36

3.7 คำถามท้ายบท 36

บทที่ 4 การจัดแบ่งคลาสด้วยขั้นตอนวิธี k-Nearest Neighbors 37

4.1 จุดประสงค์ประจำบท 39

4.2 ขั้นตอนวิธี k-Nearest Neighbors 39

- ▸ 4.2.1 ตัวอย่างการประยุกต์ใช้งานขั้นตอนวิธี k-Nearest Neighbors กับชุดข้อมูลดอกไอริส 41

4.3 ขั้นตอนวิธี Condensed Nearest Neighbor (CNN) 45

4.4 ขั้นตอนวิธี k-dimensional Tree (k-d Tree) 49

- ▸ 4.4.1 ขั้นตอนวิธีการสร้างต้นไม้ k มิติ (k-d Tree) 49

4.5 Locality-Sensitive Hashing (LSH) 56

4.6 เอกสารอ้างอิง 61

4.7 คำถามท้ายบท 62

บทที่ 5 ต้นไม้ตัดสินใจ (Decision Tree) 63

5.1 จุดประสงค์ประจำบท 64

5.2 องค์ประกอบของต้นไม้ตัดสินใจ 64

5.3 ตัวอย่างต้นไม้ตัดสินใจ 65

- ▸ 5.3.1 ต้นไม้ตัดสินใจแบบบูลีน (Boolean Decision Tree) 65
- ▸ 5.3.2 ต้นไม้ตัดสินใจแบบหลายคลาส (Multiclass Decision Tree) 56
- ▸ 5.3.3 ต้นไม้ตัดสินใจแบบค่าจริง (Real-Value Decision Tree) 67

5.4 ขั้นตอนวิธีสร้างต้นไม้ตัดสินใจแบบอุปนัย 68

5.5 การเลือกลักษณะประจำสำหรับสร้างปม 69

5.6 ขั้นตอนวิธี Iterative Dichotomiser 3 (ID3) 70

- ▸ 5.6.1 เอนโทรปี (Entropy) 70
- ▸ 5.6.2 Information Gain (IG) 71
- ▸ 5.6.3 ขั้นตอนวิธีการสร้างต้นไม้ตัดสินใจแบบ Iterative Dichotomiser 3 (ID3) 73
- ▸ 5.6.4 การเขียนโปรแกรมสำหรับต้นไม้ตัดสินใจแบบ ID3 81

5.7 ขั้นตอนวิธี C4.5 93

- ▸ 5.7.1 Generalization 93
- ▸ 5.7.2 การตัดทอน (Pruning) 94
- ▸ 5.7.3 ตัวอย่างที่มีลักษณะประจำที่มีค่าแบบต่อเนื่อง 96
- ▸ 5.7.4 ตัวอย่างที่มีลักษณะประจำที่มีค่าจำนวนมาก 97
- ▸ 5.7.5 ตัวอย่างที่มีลักษณะประจำสูญหายบางส่วน 98

5.8 เอกสารอ้างอิง 99

5.9 คำถามท้ายบท 99

CONTENTS

CONTENTS

บทที่ 6 โครงข่ายประสาทเทียม (Artificial Neural Networks) 101

6.1 จุดประสงค์ประจำบท 104

6.2 แบบจำลองเซลล์ประสาทเทียม 104

6.3 เพอร์เซปตรอน (Perceptron) 105

- 6.3.1 ขั้นตอนวิธีการเรียนรู้แบบเพอร์เซปตรอน 107
- 6.3.2 การใช้เพอร์เซปตรอนแก้ปัญหาทางตรรกศาสตร์ 112

6.4 นิวรอนแบบหน่วยเชิงเส้น (Linear Unit) 115

6.5 นิวรอนแบบหน่วยซิกมอยด์ (Sigmoid Unit) 116

6.6 ขั้นตอนวิธี Gradient Descent 117

- 6.6.1 โมเมนตัม (Momentum) 124

6.7 โครงข่ายหลายชั้น (Multi-Layer Networks) 127

- 6.7.1 ตัวอย่างการสกัดคุณลักษณะของชั้นซ่อนที่สามารถแปลความหมายได้ 128
- 6.7.2 การแก้ปัญหา Exclusive OR (XOR) ด้วยโครงข่ายหลายชั้น 130

6.8 ขั้นตอนวิธี Backpropagation 132

6.9 ปัญหาแบบหลายคลาส (Multi-Class Problem) 142

6.10 เอกสารอ้างอิง 144

6.11 คำถามท้ายบท 144

บทที่ 7 ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) 145

7.1 จุดประสงค์ประจำบท 146

7.2 ระนาบเกิน (Hyperplane) 146

7.3 ระนาบเกินแบบบัญญัติ (Canonical Hyperplane) 147

7.4 ปัญหาควบคู่ (Dual Problem) 154

7.5 Soft Margin 160

7.6 ปริภูมิไม่เชิงเส้น (Nonlinear Space) 165

7.7 มิติ VC 166

7.8 กลวิธีเคอร์เนล (Kernel Trick) 167

- 7.8.1 ตัวอย่างฟังก์ชันเคอร์เนลที่นิยมใช้งาน 169
- 7.8.2 การเขียนโปรแกรมคำนวณฟังก์ชันเคอร์เนลด้วย Numpy 169

7.9 SVM แบบไม่เชิงเส้น 172

7.10 Sequential Minimal Optimization (SMO) 175

7.11 ซัพพอร์ทเวกเตอร์แมชชีนแบบหลายคลาส (Multi-Class SVM) 184

- 7.11.1 one-against-all (1-v-a) 184
- 7.11.2 one-against-one 189
- 7.11.3 Directed Acyclic Graph Support Vector Machines (DAG-SVM) 195

7.12 การประยุกต์ใช้งาน SVM บนชุดข้อมูลจริง 202

7.13 เอกสารอ้างอิง 212

7.14 คำถามท้ายบท 212

บทที่ 8 การเรียนรู้แบบเบย์ (Bayesian Learning) 213

8.1 จุดประสงค์ประจำบท 214

8.2 ทฤษฎีของเบย์ (Bayes' Theorem) 214

- 8.2.1 ความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) 214
- 8.2.2 การจำแนกประเภทด้วยทฤษฎีของเบย์ 216

8.3 การเลือกสมมติฐานที่เหมาะสม 217

8.4 ตัวจำแนกประเภทที่เหมาะสมที่สุดแบบเบย์ (Bayes Optimal Classifier) 218

8.5 ตัวจำแนกประเภทแบบเบย์อย่างง่าย (Naive Bayes Classifier) 219

- 8.5.1 ขั้นตอนวิธีการเรียนรู้แบบเบย์อย่างง่าย 221

8.6 การจำแนกประเภทข้อความ (Text Classification) 225

8.7 เอกสารอ้างอิง 229

8.8 คำถามท้ายบท 229

บทที่ 9 ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm) 231

9.1 จุดประสงค์ประจำบท 234

9.2 การวิวัฒนาการ (Evolution) 234

9.3 การแทนพันธุกรรม (Genetic Representation) 235

9.4 การไขว้เปลี่ยน (Crossover) 236

- 9.4.1 การไขว้เปลี่ยนแบบจุดเดียว (Single point) 236
- 9.4.2 การไขว้เปลี่ยนแบบสองจุด (Two-point) 237
- 9.4.3. การไขว้เปลี่ยนแบบเอกรูป (Uniform) 237

9.5 การกลายพันธุ์ (Mutation) 237

- 9.5.1 การกลายพันธุ์แบบจุด (Point) 237
- 9.5.2 การกลายพันธุ์แบบสลับ (Swap) 238

9.6 การคัดเลือก (Selection) 238

- 9.6.1 การเก็บเฉพาะอันดับต้น 239
- 9.6.2 วงล้อรูเล็ต (Roulette Wheel) 239

CONTENTS

9.7 ขั้นตอนวิธีเชิงพันธุกรรมและการประยุกต์....................240
- 9.7.1 ตัวอย่างการแก้ปัญหาการทายคำศัพท์ปริศนาด้วยขั้นตอนวิธีเชิงพันธุกรรม....................242
- 9.7.2 ตัวอย่างการปรับเส้นโค้ง (Curve Fitting) ด้วยขั้นตอนวิธีเชิงพันธุกรรม....................246
- 9.7.3 ตัวอย่างการแก้ปัญหาการเดินทางของพนักงานขายด้วยขั้นตอนวิธีเชิงพันธุกรรม 250

9.8 เอกสารอ้างอิง....................256

9.9 คำถามท้ายบท....................256

บทที่ 10 การลดมิติ (Dimensionality Reduction) 257

10.1 จุดประสงค์ประจำบท....................258

10.2 การเลือกคุณลักษณะ (Feature Selection)....................258

10.3 Principal Component Analysis (PCA)....................258
- 10.3.1 ตัวอย่างการระบุภาพใบหน้ามนุษย์ (Face Identification) ด้วย PCA....................268

10.4 Linear Discriminant Analysis (LDA)....................271
- 10.4.1 ปัญหาขนาดตัวอย่างน้อย (Small Sample Size Problem)....................277
- 10.4.2 การลดมิติด้วย PCA และ LDA....................277

10.5 Random Projection (RP)....................279
- 10.5.1 Gaussian Random Projection (GRP)....................280
- 10.5.2 Sparse Random Projection (SRP)....................280
- 10.5.3 การลดมิติด้วย Random Projection ในการระบุภาพใบหน้ามนุษย์....................281

10.6 เอกสารอ้างอิง....................284

10.7 คำถามท้ายบท....................284

บทที่ 11 การจัดกลุ่ม (Clustering) 285

11.1 จุดประสงค์ประจำบท....................286

11.2 การจัดกลุ่มด้วยขั้นตอนวิธี K-Means....................286

11.3 การจัดกลุ่มด้วยขั้นตอนวิธี Fuzzy C-Means....................292

11.4 การจัดกลุ่มด้วยขั้นตอนวิธี Self-Organizing Map (SOM)....................297

11.5 การจัดกลุ่มด้วยขั้นตอนวิธี Expectation-Maximization (EM)....................307
- 11.5.1 การแจกแจงแบบปกติหลายตัวแปร (Multivariate Normal Distribution)....................308
- 11.5.2 แบบจำลองเกาส์เซียนผสม (Gaussian Mixture Model: GMM)....................309
- 11.5.3 ขั้นตอนวิธี Expectation-Maximization (EM)....................314

11.6 การเปรียบเทียบประสิทธิภาพขั้นตอนวิธีในการจัดกลุ่ม....................323
- 11.6.1 การวัดค่าคลาดเคลื่อนเฉลี่ยของ Centroid....................323
- 11.6.2 การวัดอัตราความถูกต้อง....................324
- 11.6.3 ผลการเปรียบเทียบประสิทธิภาพชุดข้อมูลดอกไอริส....................325

11.7 เอกสารอ้างอิง 328

11.8 คำถามท้ายบท 328

บรรณานุกรม 329

ดัชนี 332