Artificial Intelligence with Machine Learning

AI สร้างได้ด้วยแมชชีนเลิร์นนิ่ง

Python Edition

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

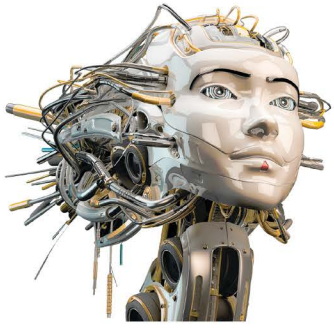
ผู้แต่ง รองศาสตราจารย์ ดร.ปริญญา สงวนสัตย์ บรรณาธิการ ภีรพล คชาเจริญ

IDC PREMIER

มีเพียง “ความรู้” เท่านั้นที่มนุษย์ใช้พลิก “โลก” และเปลี่ยน “ชีวิต”
เราจึงสร้างสรรค์ และส่งมอบ “ความรู้” ในรูปแบบที่ดีกว่า
เพื่อให้คนไทย “เรียนรู้” ได้ตลอดชีวิต

ตัวอย่าง

INFOPRESS Think Beyond Read Comics Little Heart



Artificial Intelligence with Machine Learning,

AI สร้างได้ด้วยแมชชีนเลิร์นนิ่ง

Python Edition

Writer	รศ. ดร.ปริญญา สงวนสัตย์
Editor	ภีรพล คชาเจริญ
Graphic Designer	วสันต์ พึ่งพูลผล
Page Layout	สิริลักษณ์ วาระเลิศ
Proofreader	พรรณรัตน์ ขุราษี
Publishing Coordinators	วรพล ณธิกุล, สุพัตรา อาจปรุ, นิวัช ยะหัวดง

ข้อมูลทางบรรณานุกรมของหอสมุดแห่งชาติ
ปริญญา สงวนสัตย์
Artificial Intelligence with Machine Learning, AI สร้างได้ด้วยแมชชีนเลิร์นนิ่ง.-- นนทบุรี : ไอดีซี พรีเมียร์, 2562.
356 หน้า.
1. ปัญญาประดิษฐ์. 2. การเรียนรู้ของเครื่อง. I. ชื่อเรื่อง.
006.3
ISBN 885-916-100-774-6

Microsoft Windows เป็นเครื่องหมายการค้าของบริษัท Microsoft Corporation, Python IDLE เป็นเครื่องหมายการค้าของบริษัท Python Software Foundation, PyCharm เป็นเครื่องหมายการค้าของบริษัท JetBrains และเครื่องหมายการค้าอื่น ๆ ที่อ้างถึงเป็นของบริษัทนั้น ๆ

บริษัท ไอดีซี พรีเมียร์ จำกัด จัดตั้งขึ้นเพื่อเผยแพร่ความรู้ที่มีคุณภาพสู่ผู้อ่านชาวไทย เรายินดีรับงานเขียนของนักวิชาการและนักเขียนทุกท่าน ท่านผู้สนใจกรุณาติดต่อผ่านทางอีเมลที่ infopress@idcpremier.com หรือทางโทรศัพท์หมายเลข 0-2962-1081 (อัตโนมัติ 10 คู่สาย) โทรสาร 0-2962-1084

สร้างสรรค์โดย

INFOPRESS

พิมพ์ครั้งที่ 1 พฤศจิกายน 2562

สมาชิกสัมพันธ์
โทรศัพท์ 0-2962-1081-3 ต่อ 121
โทรสาร 0-2962-1084

ร้านค้าและตัวแทนจำหน่าย
โทรศัพท์ 0-2962-1081-3 ต่อ 112-114
โทรสาร 0-2962-1084

จัดพิมพ์และจัดจำหน่ายโดย

IDC PREMIER

บริษัท ไอดีซี พรีเมียร์ จำกัด
200 หมู่ 4 ชั้น 19 ห้อง 1901
อาคารจัสมินอินเตอร์เนชั่นแนลทาวเวอร์
ถ.แจ้งวัฒนะ อ.ปากเกร็ด จ.นนทบุรี 11120
โทรศัพท์ 0-2962-1081 (อัตโนมัติ 10 คู่สาย)
โทรสาร 0-2962-1084

ราคา 495 บาท

ตัวอย่าง

EDITOR'S NOTE

ทุกวันนี้ AI ไม่ใช่เรื่องไกลตัว แต่มันเริ่มเข้ามาเป็นส่วนหนึ่งในวิถีชีวิตของเรา หากแต่เป็นช่วงเริ่มต้นที่ยังไม่ใช่ AI โดยสมบูรณ์แบบ เช่น รถยนต์ไร้คนขับ, การควบคุมการบินด้วย Auto Pilot, การซื้อขายหลักทรัพย์, การสแกนใบหน้าระบุตัวตน, ระบบสืบค้นข้อมูลของ Google, การ Tag อัตโนมัติบุคคลในภาพของ Facebook, การเรียนรู้พฤติกรรมการชมภาพยนตร์ของ Netflix, การเรียนรู้พฤติกรรมการชมคลิป VDO ของ YouTube, การรับคำสั่งด้วยเสียงของ Google Assistant เป็นต้น

"AI หรือ ปัญญาประดิษฐ์" เป็นความอัจฉริยะที่มนุษย์เราสร้างขึ้น ซึ่งที่ผ่านมาเรามีประสบการณ์กับระบบอัตโนมัติในลักษณะ Automation หมายความว่า ระบบหรือเครื่องจักรที่ทำงานโดยอัตโนมัติในแบบที่คาดการณ์ไว้แล้ว และโปรแกรมการทำงานเอาไว้ล่วงหน้าเท่านั้น แต่หากเกิดสถานการณ์อื่นนอกเหนือจากที่โปรแกรมเอาไว้ ระบบ Automation ก็จะตอบสนองไม่ได้ เนื่องจากไม่ได้ถูกโปรแกรมเอาไว้ว่าจะต้องทำอย่างไรนั่นเอง สิ่งที่พอจะทำได้ก็คือ การปรับปรุงโปรแกรมและอัพเกรดเพิ่มเติม นี่จึงเป็นข้อจำกัดของเทคโนโลยีในยุคปัจจุบัน เรียกได้ว่า ระบบ Automation ก็ยังไม่ดีพอด้วยเพราะมันเลิร์นนิ่งเองไม่ได้ ซึ่งต่างจาก AI ที่มี Machine Learning ที่ก้าวข้ามเรื่องนี้ไปได้

ระบบหรือเครื่องที่เลิร์นนิ่งได้ดีกว่าระบบเดิมอย่างไร ? ยกตัวอย่างเช่น Netflix สร้าง Machine Learning มาศึกษาพฤติกรรมลูกค้า เพื่อแนะนำภาพยนตร์และซีรีส์ให้แก่ลูกค้าได้อย่างถูกใจชนิดที่คาดไม่ถึง ซึ่งการตอบสนองความต้องการที่มนุษย์ออกแบบอาจทำได้เพียงการนำเสนอหนังใหม่ การจัด Rating การรีวิวจากผู้ชม ก็อาจตอบโจทย์ได้ดีในยุคที่การแข่งขันยังไม่สูง แต่ในปัจจุบันนี้ใคร ๆ ก็ทำได้ดีไม่แพ้กัน สมรภูมิที่รบกันด้วยสงครามราคาจึงเกิดขึ้นอย่างไม่ต้องสงสัย แต่ Netflix เลือกใช้ Machine Learning มาพิจารณาคัดเลือกภาพยนตร์หรือซีรีส์ให้แก่ลูกค้า แล้วพัฒนาโปรแกรมตอบสนองความต้องการของลูกค้าเป็นรายคน ไม่ใช่กลุ่มคนอย่างที่ผ่านมา ตรงนี้เองที่ทำให้เกิดมูลค่ามหาศาลที่ทำให้ Netflix ประสบความสำเร็จแบบไร้คู่ต่อกร

เพราะฉะนั้น หัวใจสำคัญของหนังสือเล่มนี้ คือ เรียน Machine Learning Algorithms เพื่อศึกษาว่า AI สามารถเรียนรู้ได้อย่างไร ? และอีกคำถามสำคัญก็คือ เราจะตอบสนองต่อสิ่งที่เรียนรู้มาอย่างไร ? เพราะนี่คือ คุณค่าที่แท้จริงของการพัฒนา AI ที่มี Machine Learning อยู่เบื้องหลังความสำเร็จ

แนวรบใหม่ในโลกใบเดิม "AI Change the World"

ภีรพล คชาเจริญ

Senior Editor

Digital Technology

นักวิชาการและนักเขียนทางด้าน Digital Technology และ Digital Marketting กรุณาส่งต้นฉบับของท่านมาพิจารณาที่อีเมล infopress@idcpremier.com หรือ peeidcpremier@gmail.com

IDC PREMIER

WRITOR'S NOTE

หนังสือเล่มนี้ปรับปรุงขึ้นใหม่จากฉบับเดิม โดยมีการเรียบเรียง เพิ่มเติม แก้ไขเนื้อหา และเปลี่ยนการเขียนโปรแกรมจากภาษา MATLAB เป็น Python ซึ่งกำลังเป็นที่นิยมในปัจจุบัน จุดประสงค์หลักของหนังสือเล่มนี้คือ เป็นแหล่งข้อมูลให้กับนักศึกษาและบุคคลทั่วไปที่สนใจในงานการเรียนรู้ของเครื่องหรือสาขาที่เกี่ยวข้อง สามารถนำไปใช้ประโยชน์ได้อย่างแท้จริง การเรียนรู้ของเครื่อง (Machine Learning) เป็นเครื่องมือที่ถูกนำไปประยุกต์ใช้งานอย่างกว้างขวาง อาทิเช่น งานด้านปัญญาประดิษฐ์ (Artificial Intelligence: AI) งานด้านการทำเหมืองข้อมูล (Data Mining) งานด้านการรู้จำรูปแบบ (Pattern Recognition) งานด้านคอมพิวเตอร์วิทัศน์ (Computer Vision) และงานในอีกหลายสาขาที่เกี่ยวข้อง การเข้าใจในเครื่องมือพื้นฐานของการเรียนรู้ของเครื่องมีประโยชน์ในการเลือกใช้เครื่องมือต่าง ๆ ให้เหมาะสมกับงานที่ต้องการ รวมทั้งยังสามารถปรับปรุงขั้นตอนวิธีเพื่อให้เหมาะสมกับการใช้งานในแต่ละงานได้อย่างมีประสิทธิภาพ

หนังสือเล่มนี้ผู้เขียนได้อาศัยประสบการณ์ในการสอนและการวิจัยในสาขาวิชานี้ตั้งแต่ระดับชั้นปริญญาตรีถึงปริญญาเอก โดยพยายามสอดแทรกตัวอย่างที่ง่ายต่อการเข้าใจที่ได้เคยใช้ในการอธิบายนักศึกษาในชั้นเรียน ตลอดจนประสบการณ์ในการทำวิจัย ในแต่ละบทประกอบด้วยเนื้อหา ขั้นตอนวิธี และโปรแกรมตัวอย่างที่เขียนขึ้นด้วยภาษา Python ซึ่งเป็นภาษาที่เข้าใจได้ง่ายและได้รับความนิยมอย่างมากในปัจจุบัน ทั้งนี้รหัสต้นฉบับ (Source code) เกือบทั้งหมดในหนังสือเล่มนี้ผู้เขียนมีความตั้งใจเขียนขึ้นมาจากพื้นฐาน โดยไม่ได้ใช้ไลบารีระดับสูง เพื่อให้เข้าใจหลักการทำงานเบื้องหลังอย่างแท้จริง ทำให้ผู้อ่านสามารถนำไปประยุกต์ใช้ด้วยภาษาที่ตนเองถนัดได้ง่าย

สุดท้ายนี้ผู้เขียนหวังเป็นอย่างยิ่งว่าหนังสือเล่มนี้จะเป็นประโยชน์สำหรับผู้อ่านที่สนใจเพื่อใช้ศึกษาหรือนำไปใช้ประโยชน์ในงานต่าง ๆ ได้อย่างเหมาะสม หากผู้อ่านมีปัญหาหรือพบข้อผิดพลาดประการใดในหนังสือเล่มนี้ทางผู้เขียนขออภัยไว้ ณ ที่นี้ และสามารถติดต่อผู้เขียนเพื่อแจ้งปัญหาหรือข้อผิดพลาดที่พบได้ รวมทั้งคำถามต่าง ๆ ในการใช้งานได้ทาง Facebook Page หรือ YouTube Channel AI บ้าน บ้าน

รองศาสตราจารย์ ดร.ปริญญา สงวนสัตย์
คณะวิศวกรรมศาสตร์และเทคโนโลยี
สถาบันการจัดการปัญญาภิวัฒน์

ตัวอย่าง

ARTIFICIAL
INTELLIGENCE

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

CDU

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) 66
- 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

ตัวอย่าง

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 ทริกเคอร์เนล (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

ตัวอย่าง

7.11 ซัพพอร์ตเวกเตอร์แมชชีนแบบหลายคลาส (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 (Naïve 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 การไขว้เปลี่ยนแบบจุดเดี่ยว (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 แบบจำลองเกาส์แบบผสม (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

ตัวอย่าง

รูปที่ 1.1 แผนภูมิแสดงความสัมพันธ์ของปัญญาประดิษฐ์และการเรียนรู้ของเครื่อง 3
รูปที่ 2.1 ตัวอย่างลายจุดไข่ปลาตัวอักษร ก.ไก่ สำหรับคัดลายมือ 10
รูปที่ 2.2 ระบบ Black Box 11
รูปที่ 2.3 ตัวอย่างการเก็บข้อมูลทางตรง (Direct) 13
รูปที่ 2.4 ตัวอย่างการเก็บข้อมูลทางอ้อม (Indirect) 13
รูปที่ 2.5 ตัวอย่างการเลือกฟังก์ชันเป้าหมายแบบการให้คะแนน 14
รูปที่ 2.6 ตัวอย่างกรณีฝ่ายน้ำเงินชนะ V(b) = -100 14
รูปที่ 2.7 ตัวอย่างกรณีเสมอกัน $V(b)$ = 0 14
รูปที่ 2.8 ตัวอย่างกรณีฝ่ายน้ำเงินแพ้ $V(b)$ = 100 15
รูปที่ 2.9 ตัวอย่างการคำนวณคะแนนในทางเดินหมากแบบต่าง ๆ 16
รูปที่ 2.10 ตัวอย่างการเดินหมากตาที่ 0 ถึงตาที่ 1 16
รูปที่ 2.11 ตัวอย่างการเดินหมากถึงตาที่ 0 ถึงตาที่ 3 17
รูปที่ 2.12 ตัวอย่างการเดินหมากตาที่ 3 ถึงตาที่ 4 17
รูปที่ 2.13 ตัวอย่างการเดินหมากตาที่ 4 ถึงตาที่ 6 17
รูปที่ 2.14 ตัวอย่างเกม Tic-Tac-Toe ขั้นตอนที่ 2 19
รูปที่ 2.15 ตัวอย่างเกม Tic-Tac-Toe ขั้นตอนที่ 3 19
รูปที่ 2.16 ตัวอย่างเกม Tic-Tac-Toe ขั้นตอนที่ 4 20
รูปที่ 2.17 ตัวเลขแสดงตำแหน่งในการกำหนดการวางหมากในเกม Tic-Tac-Toe 20
รูปที่ 3.1 ปริภูมิแบบ (Version Space) 29
รูปที่ 3.2 ตัวอย่างขั้นตอนการสร้างปริภูมิแบบด้วยขั้นตอนวิธี Candidate Elimination 35
รูปที่ 3.3 ตัวอย่างปริภูมิแบบที่ได้จากขั้นตอนวิธี Candidate Elimination 35
รูปที่ 4.1 ปริภูมิตัวอย่าง 39
รูปที่ 4.2 ระยะทางจากตัวอย่างที่ค่า k เท่ากับ 1, 3 และ 6 40
รูปที่ 4.3 ผลการจำแนกประเภทชุดข้อมูลดอกไอริสด้วยขั้นตอนวิธี 17-NN 44
รูปที่ 4.4 อัตราความถูกต้องของขั้นตอนวิธี k-NN กรณีค่า k ต่าง ๆ สำหรับชุดข้อมูลดอกไอริส 44
รูปที่ 4.5 อัตราส่วนขอบเขต 46
รูปที่ 4.6 เซตต้นแบบเทียบกับชุดตัวอย่างฝึกฝนทั้งหมดของชุดข้อมูลดอกไอริส 48
รูปที่ 4.7 อัตราความถูกต้องของขั้นตอนวิธี CNN กรณีค่า k ต่าง ๆ สำหรับชุดข้อมูลดอกไอริส 48
รูปที่ 4.8 ผลการจำแนกประเภทชุดข้อมูลดอกไอริสด้วยขั้นตอนวิธี CNN กรณีค่า k = 1 49
รูปที่ 4.9 ตัวอย่างการแบ่งปริภูมิ 2 มิติด้วย k-d Tree 51
รูปที่ 4.10 ตัวอย่าง k-d Tree 51

ตัวอย่าง

สารบัญรูป

รูปที่ 4.11 ต้นไม้ k มิติ (k-d Tree) ของชุดข้อมูลดอกไอริส 55
รูปที่ 4.12 กราฟแสดงเวลาในการประมวลผลการค้นหาตัวอย่างที่ใกล้ที่สุดด้วย k-d Tree 56
รูปที่ 4.13 ตัวอย่างการแฮชในปริภูมิตัวอย่างสำหรับวิธี LSH 57
รูปที่ 4.14 อัตราความถูกต้องในการจำแนกประเภทด้วย k-NN แบบปกติเทียบกับแบบ LSH 60
รูปที่ 4.15 เวลาประมวลผลในการจำแนกประเภทด้วย k-NN แบบปกติเทียบกับแบบ LSH 61
รูปที่ 5.1 ตัวอย่างต้นไม้ตัดสินใจ 64
รูปที่ 5.2 ตัวอย่างต้นไม้ตัดสินใจแบบบูลีน 65
รูปที่ 5.3 ตัวอย่างปัญหาหลายคลาส 66
รูปที่ 5.4 ตัวอย่างต้นไม้ตัดสินใจแบบปัญหาหลายคลาส 66
รูปที่ 5.5 ตัวอย่างปัญหาแบบค่าจริง 67
รูปที่ 5.6 ตัวอย่างต้นไม้ตัดสินใจแบบค่าจริง 67
รูปที่ 5.7 ต้นไม้ตัดสินใจที่สร้างจากการเลือกปมที่ต่างกันจากชุดข้อมูลในตารางที่ 5.1 69
รูปที่ 5.8 เอนโทรปีของเหตุการณ์การโยนเหรียญ 1 เหรียญ 1 ครั้ง ด้วยการถ่วงน้ำหนักเหรียญต่าง ๆ กัน 71
รูปที่ 5.9 ตัวอย่างต้นไม้ตัดสินใจที่สร้างจากการเลือกปมที่ต่างกันจากตารางที่ 5.2 72
รูปที่ 5.10 แผนภาพแจกแจงค่าของลักษณะประจำ Outlook 75
รูปที่ 5.11 แผนภาพแจกแจงค่าของลักษณะประจำ Temp 76
รูปที่ 5.12 แผนภาพแจกแจงค่าของลักษณะประจำ Humidity 76
รูปที่ 5.13 แผนภาพแจกแจงค่าของลักษณะประจำ Wind 77
รูปที่ 5.14 ต้นไม้ตัดสินใจที่สร้างโดยใช้ Outlook เป็นปมแรก 78
รูปที่ 5.15 ต้นไม้ตัดสินใจที่สร้างโดยใช้ Outlook เป็นปมแรกและใช้ Humidity เป็นปมที่สอง 79
รูปที่ 5.16 ต้นไม้ตัดสินใจสมบูรณ์ที่สร้างจากชุดตัวอย่างในตารางที่ 5.3 81
รูปที่ 5.17 ตัวอย่างการแทนค่าต้นไม้แบบ Parent/Child 81
รูปที่ 5.18 ตัวอย่างต้นไม้ตัดสินใจจากคลาส ID3 ของชุดข้อมูล Play Tennis 89
รูปที่ 5.19 ตัวอย่างต้นไม้ตัดสินใจจากคลาส ID3 ของชุดข้อมูลการประเมินรถยนต์ 92
รูปที่ 5.20 ปัญหา Overfitting 93
รูปที่ 5.21 การแบ่งข้อมูลอยู่เป็น K กลุ่มย่อย 94
รูปที่ 5.22 K-fold Cross-Validation 94
รูปที่ 5.23 ต้นไม้ตัดสินใจที่ได้จากลักษณะประจำที่มีค่าจำนวนมาก 97
รูปที่ 5.24 การใช้งานต้นไม้ตัดสินใจในกรณีมีข้อมูลสูญหายบางส่วน 98
รูปที่ 6.1 เซลล์ประสาท 102
รูปที่ 6.2 ศักยะงาน (Action Potential) 103

ตัวอย่าง

รูปที่ 6.3 แบบจำลองเซลล์ประสาทเทียม 104
รูปที่ 6.4 ฟังก์ชันถ่ายโอนฮาร์ดลิมิต 105
รูปที่ 6.5 ฟังก์ชันถ่ายโอนฮาร์ดลิมิตแบบสมมาตร 105
รูปที่ 6.6 โครงสร้างของ Perceptron 106
รูปที่ 6.7 เพอร์เซปตรอนขั้นตอนที่ 1 108
รูปที่ 6.8 เพอร์เซปตรอนขั้นตอนที่ 2 108
รูปที่ 6.9 เพอร์เซปตรอนขั้นตอนที่ 3 109
รูปที่ 6.10 เพอร์เซปตรอนขั้นตอนที่ 4 110
รูปที่ 6.11 เพอร์เซปตรอนขั้นตอนที่ 5 110
รูปที่ 6.12 ตัวอย่างระนาบเกินที่ได้จากเพอร์เซปตรอน 112
รูปที่ 6.13 เพอร์เซปตรอนกับปัญหา AND 112
รูปที่ 6.14 เพอร์เซปตรอนกับปัญหา OR 113
รูปที่ 6.15 ปัญหา XOR 114
รูปที่ 6.16 ฟังก์ชันถ่ายโอนแบบเชิงเส้น 115
รูปที่ 6.17 นิวรอนแบบหน่วยเชิงเส้น 115
รูปที่ 6.18 ฟังก์ชันถ่ายโอนแบบซิกมอยด์ 116
รูปที่ 6.19 นิวรอนแบบหน่วยซิกมอยด์ 116
รูปที่ 6.20 ตัวอย่างการหาค่าต่ำสุดสัมพัทธ์ด้วยขั้นตอนวิธี Gradient Descent โดยกำหนดให้ ε = 0.01, η = 0.01 และ x_0 = 2 119
รูปที่ 6.21 ตัวอย่างการหาค่าต่ำสุดสัมพัทธ์ด้วยขั้นตอนวิธี Gradient Descent โดยกำหนดให้ ε = 0.01, η = 0.001 และ x_0 = 2 120
รูปที่ 6.22 ตัวอย่างการหาค่าต่ำสุดสัมพัทธ์ด้วยขั้นตอนวิธี Gradient Descent โดยกำหนดให้ ε = 0.01, η = 0.1 และ x_0 = 2 121
รูปที่ 6.23 ปัญหา Local Trap 124
รูปที่ 6.24 การแก้ปัญหา Local Trap ด้วยการเพิ่มโมเมนตัม 125
รูปที่ 6.25 ตัวอย่างการหาค่าต่ำสุดสัมพัทธ์ด้วยขั้นตอนวิธี Gradient Descent แบบไม่มีโมเมนตัม 126
รูปที่ 6.26 ตัวอย่างการหาค่าต่ำสุดสัมพัทธ์ด้วยขั้นตอนวิธี Gradient Descent แบบมีโมเมนตัม 127
รูปที่ 6.27 โครงข่ายหลายชั้น 127
รูปที่ 6.28 ตัวอย่างโครงข่ายหลายชั้นที่แปลความหมายได้ 128
รูปที่ 6.29 การแก้ปัญหา Exclusive OR (XOR) ด้วยโครงข่ายหลายชั้น (ขั้นตอนที่ 1) 130
รูปที่ 6.30 การแก้ปัญหา Exclusive OR (XOR) ด้วยโครงข่ายหลายชั้น (ขั้นตอนที่ 2) 131
รูปที่ 6.31 การแก้ปัญหา Exclusive OR (XOR) ด้วยโครงข่ายหลายชั้น (ขั้นตอนที่ 3) 131
รูปที่ 6.32 ค่าความผิดพลาดกำลังสองเฉลี่ยของการแก้ปัญหา XOR ด้วยโครงข่ายประสาทเทียมแบบหลายชั้น โดยใช้ฟังก์ชันกระตุ้นแบบ Sigmoid 139

ตัวอย่าง

สารบัญรูป

รูปที่ 6.33 ระนาบเกินที่เกิดจากการแก้ปัญหา XOR ด้วยโครงข่ายประสาทเทียมแบบหลายชั้น โดยใช้ฟังก์ชันกระตุ้นแบบ Sigmoid...139

รูปที่ 6.34 ค่าความผิดพลาดกำลังสองเฉลี่ยของการแก้ปัญหา XOR ด้วยโครงข่ายประสาทเทียมแบบหลายชั้น โดยใช้ฟังก์ชันกระตุ้นแบบ ReLU 141

รูปที่ 6.35 ระนาบเกินที่เกิดจากการแก้ปัญหา XOR ด้วยโครงข่ายประสาทเทียมแบบหลายชั้น โดยใช้ฟังก์ชันกระตุ้นแบบ ReLU.... 141

รูปที่ 6.36 ค่าความผิดพลาดกำลังสองเฉลี่ยของชุดข้อมูลดอกไอริสด้วยโครงข่ายประสาทเทียมแบบหลายชั้น 143

รูปที่ 7.1 ตัวอย่างระนาบเกินในปริภูมิ 1 มิติ..... 146

รูปที่ 7.2 ตัวอย่างระนาบเกินในปริภูมิ 2 มิติ..... 146

รูปที่ 7.3 ตัวอย่างระนาบเกินในปริภูมิ 3 มิติ..... 147

รูปที่ 7.4 ระนาบเกินที่สามารถแบ่งข้อมูล 2 คลาสได้ 147

รูปที่ 7.5 ระนาบเกินที่เกิดปัญหา Overfitting 148

รูปที่ 7.6 เวกเตอร์ซัพพอร์ต..... 148

รูปที่ 7.7 ความสัมพันธ์ระหว่างระยะขอบกับเวกเตอร์ปกติของระนาบเกิน 149

รูปที่ 7.8 พื้นที่แสดงอาณาเขตของแต่ละคลาสในกรณีไบนารี..... 150

รูปที่ 7.9 SVM แบบเชิงเส้นโดยใช้โมดูล cvxopt 154

รูปที่ 7.10 ผลเฉลย SVM แบบเชิงเส้นที่ผิดพลาด..... 154

รูปที่ 7.11 ผลเฉลย SVM แบบเชิงเส้นด้วยการแก้ปัญหาควบคู่..... 158

รูปที่ 7.12 ค่า α_i ที่ได้จากการแก้ปัญหาควบคู่ 158

รูปที่ 7.13 เวกเตอร์ซัพพอร์ต..... 159

รูปที่ 7.14 ตัวแปร Slack สำหรับ SVM แบบเชิงเส้นกรณี Soft Margin..... 160

รูปที่ 7.15 ระนาบเกินที่ได้จากการแก้ปัญหา SVM แบบเชิงเส้นกรณี Soft Margin เมื่อค่า C = 10..... 163

รูปที่ 7.16 ระนาบเกินที่ได้จากการแก้ปัญหา SVM แบบเชิงเส้นกรณี Soft Margin เมื่อค่า C = 1..... 164

รูปที่ 7.17 ระนาบเกินที่ได้จากการแก้ปัญหา SVM แบบเชิงเส้นกรณี Soft Margin เมื่อค่า C = 100..... 164

รูปที่ 7.18 ปริภูมิไม่เชิงเส้น 165

รูปที่ 7.19 ตัวอย่างการแปลงปริภูมิแบบไม่เชิงเส้นเป็นปริภูมิแบบเชิงเส้น..... 165

รูปที่ 7.20 การกำหนดฉลากแบบไบนารีที่เป็นไปได้ทั้งหมดกรณี 3 จุด 166

รูปที่ 7.21 การกำหนดฉลากแบบไบนารีในกรณี 2 มิติที่เส้นตรงไม่สามารถแบ่งได้..... 167

รูปที่ 7.22 ระนาบเกินที่ได้จากการแก้ปัญหา SVM แบบไม่เชิงเส้นด้วยเคอร์เนล RBF..... 174

รูปที่ 7.23 ค่า α_i ที่ได้จากการแก้ปัญหา SVM แบบไม่เชิงเส้นด้วยเคอร์เนล RBF 174

รูปที่ 7.24 กราฟเส้นตรงแสดงค่า α_1 และ α_2 ในเงื่อนไข $0 \le \alpha_i \le C$ ในกรณี $y_1 \neq y_2$ ที่ $k > 0$ (ซ้าย) และที่ $k < 0$ (ขวา) 175

ตัวอย่าง

รูปที่ 7.25 กราฟเส้นตรงแสดงค่า α_1 และ α_2 ในเงื่อนไข $0 \le \alpha_i \le C$ ในกรณี $y_1 = y_2$ ที่ $k < C$ (ซ้าย) และที่ $k > C$ (ขวา) 175

รูปที่ 7.26 ระนาบเกินที่ได้จากการแก้ปัญหา SVM แบบไม่เชิงเส้นด้วยขั้นตอนวิธี SMO 183

รูปที่ 7.27 ค่า α_i ที่ได้จากการแก้ปัญหา SVM แบบไม่เชิงเส้นด้วยขั้นตอนวิธี SMO 183

รูปที่ 7.28 ระนาบเกินของคลาสที่ 1 (● สีแดง) กับคลาสที่เหลือแบบ 1-v-a 187

รูปที่ 7.29 ระนาบเกินของคลาสที่ 2 (★ สีเขียว) กับคลาสที่เหลือ แบบ 1-v-a 187

รูปที่ 7.30 ระนาบเกินของคลาสที่ 3 (● สีน้ำเงิน) กับคลาสที่เหลือแบบ 1-v-a 188

รูปที่ 7.31 ระนาบเกินทั้งหมดแบบ 1-v-a 188

รูปที่ 7.32 ตัวอย่างขอบเขตแบบแข็ง (Hard Margin) กรณีหลายคลาสแบบ 1-v-a 188

รูปที่ 7.33 ตัวอย่างขอบเขตแบบอ่อน (Soft Margin) กรณีหลายคลาสแบบ 1-v-a 189

รูปที่ 7.34 ระนาบเกินของคลาสที่ 1 (● สีแดง) และคลาสที่ 2 (★ สีเขียว) แบบ 1-v-1 193

รูปที่ 7.35 ระนาบเกินของคลาสที่ 1 (● สีแดง) และคลาสที่ 3 (● สีน้ำเงิน) แบบ 1-v-1 193

รูปที่ 7.36 ระนาบเกินของคลาสที่ 2 (★ สีเขียว) และคลาสที่ 3 (● สีน้ำเงิน) แบบ 1-v-1 194

รูปที่ 7.37 ระนาบเกินของคลาสทั้งสามแบบ 1-v-1 194

รูปที่ 7.38 ตัวอย่างขอบเขตแบบแข็ง (Hard Margin) ด้วยวิธีการออกเสียงข้างมากกรณีหลายคลาสแบบ 1-v-1 194

รูปที่ 7.39 ตัวอย่างขอบเขตแบบอ่อน (Soft Margin) กรณีหลายคลาสแบบ 1-v-1 195

รูปที่ 7.40 ตัวอย่างขอบเขตแบบแข็ง (Hard Margin) ด้วยวิธีการสะสมค่ากรณีหลายคลาสแบบ 1-v-1 195

รูปที่ 7.41 ตัวอย่างกราฟสำหรับ DAG-SVM แบบ 3 คลาส 196

รูปที่ 7.42 ดัชนีอ้างอิงปมสำหรับกราฟ DAG-SVM 196

รูปที่ 7.43 ระนาบเกินของคลาสที่ 1 (● สีแดง) และคลาสที่ 3 (● สีน้ำเงิน) แบบ DAG-SVM 200

รูปที่ 7.44 ระนาบเกินของคลาสที่ 1 (● สีแดง) และคลาสที่ 2 (★ สีเขียว) แบบ DAG-SVM 200

รูปที่ 7.45 ระนาบเกินของคลาสที่ 2 (★ สีเขียว) และคลาสที่ 3 (● สีน้ำเงิน) แบบ DAG-SVM 201

รูปที่ 7.46 ระนาบเกินของคลาสทั้งสามแบบ DAG-SVM 201

รูปที่ 7.47 ตัวอย่างขอบเขตแบบแข็ง (Hard margin) กรณีหลายคลาสแบบ DAG-SVM 201

รูปที่ 7.48 อัตราความถูกต้องบนชุดข้อมูลฝึกฝนของ SVM เปรียบเทียบวิธีการ 1-v-a 1-v-1 และ DAG-SVM 209

รูปที่ 7.49 เวลาที่ใช้ในการฝึกฝนของ SVM เปรียบเทียบวิธีการ 1-v-a 1-v-1 และ DAG-SVM 209

รูปที่ 7.50 เวลาที่ใช้ในการทดสอบของ SVM เปรียบเทียบวิธีการ 1-v-a 1-v-1 และ DAG-SVM 210

รูปที่ 8.1 ตลับยาซึ่งประกอบด้วยยาเม็ดและยาแคปซูล 214

รูปที่ 8.2 การเปลี่ยนแปลงของความน่าจะเป็นภายหลังเมื่อเพิ่มจำนวนตัวอย่างฝึกฝน 217

รูปที่ 9.1 การสุ่มคำศัพท์ด้วยขั้นตอนวิธี Brute Force 232

รูปที่ 9.2 การวิวัฒนาการของมนุษย์ 235

ตัวอย่าง

สารบัญรูป

รูปที่ 9.3 การไขว้เปลี่ยนของโครโมโซม 236
รูปที่ 9.4 การไขว้เปลี่ยนแบบจุดเดี่ยว 236
รูปที่ 9.5 การไขว้เปลี่ยนแบบสองจุด 237
รูปที่ 9.6 การไขว้เปลี่ยนแบบเอกรูป 237
รูปที่ 9.7 การกลายพันธุ์แบบจุด 237
รูปที่ 9.8 การกลายพันธุ์แบบสลับ 238
รูปที่ 9.9 การเก็บอันดับต้น 239
รูปที่ 9.10 วงล้อรูเล็ต 239
รูปที่ 9.11 ผังงานแสดงขั้นตอนวิธีเชิงพันธุกรรม 240
รูปที่ 9.12 การสุ่มคำศัพท์ด้วยขั้นตอนวิธีเชิงพันธุกรรม 244
รูปที่ 9.13 ค่า Fitness ของขั้นตอนวิธีเชิงพันธุกรรมที่ใช้แก้ปัญหาการทายคำศัพท์ปริศนา 245
รูปที่ 9.14 ค่า Fitness ของขั้นตอนวิธีเชิงพันธุกรรมที่ใช้แก้ปัญหาการปรับเส้นโค้ง 249
รูปที่ 9.15 ค่าพิกัดข้อมูลจริง (วงกลม) เทียบกับฟังก์ชันพหุนามที่ประมาณขึ้นค่าขั้นตอนวิธีเชิงพันธุกรรม (เส้น) 249
รูปที่ 9.16 การไขว้เปลี่ยนแบบวน 251
รูปที่ 9.17 ค่า Fitness และเส้นทางที่ได้จากการแก้ปัญหาการเดินทางของพนักงานขายด้วยขั้นตอนวิธีเชิงพันธุกรรม 255
รูปที่ 10.1 ชุดข้อมูลดอกไอริสในปริภูมิเดิม 263
รูปที่ 10.2 ชุดข้อมูลดอกไอริสในปริภูมิใหม่ที่ผ่านการแปลงด้วยวิธี PCA 264
รูปที่ 10.3 ค่าลักษณะเฉพาะจากวิธี PCA ของชุดข้อมูลดอกไอริส 265
รูปที่ 10.4 ค่าลักษณะเฉพาะจากวิธี PCA ของชุดข้อมูล ORL 269
รูปที่ 10.5 ความถูกต้องในการระบุภาพใบหน้ามนุษย์ชุดข้อมูล ORL ด้วย PCA 270
รูปที่ 10.6 ค่าลักษณะเฉพาะจากวิธี LDA ของชุดข้อมูลดอกไอริส 276
รูปที่ 10.7 ชุดข้อมูลดอกไอริสในปริภูมิใหม่ที่ผ่านการแปลงด้วยวิธี LDA 276
รูปที่ 10.8 ความถูกต้องในการระบุภาพใบหน้ามนุษย์ชุดข้อมูล ORL ด้วย PCA และ LDA 279
รูปที่ 10.9 ความถูกต้องในการระบุภาพใบหน้ามนุษย์ชุดข้อมูล ORL ด้วยวิธี GRP และ SGP 283
รูปที่ 11.1 ผังงานแสดงขั้นตอนวิธี K-Means 287
รูปที่ 11.2 ตัวอย่างการแบ่งกลุ่มข้อมูล 2 มิติด้วยขั้นตอนวิธี K-Means 290
รูปที่ 11.3 ตัวอย่างการแบ่งกลุ่มด้วยขั้นตอนวิธี K-Means ที่ผิดพลาด 290
รูปที่ 11.4 ตัวอย่างการแบ่งกลุ่มข้อมูล 1 มิติด้วยขั้นตอนวิธี K-Means 291
รูปที่ 11.5 ตัวอย่างการแบ่งกลุ่มข้อมูล 3 มิติด้วยขั้นตอนวิธี K-Means 291
รูปที่ 11.6 ตัวอย่างการแบ่งกลุ่มข้อมูล 2 มิติด้วยขั้นตอนวิธี Fuzzy C-Means 295
รูปที่ 11.7 ตัวอย่างการแบ่งกลุ่มด้วยขั้นตอนวิธี Fuzzy C-Means ที่ผิดพลาด 296

ตัวอย่าง

รูปที่ 11.8 ตัวอย่างการแบ่งกลุ่มข้อมูล 1 มิติด้วยขั้นตอนวิธี Fuzzy C-Means 296
รูปที่ 11.9 ตัวอย่างการแบ่งกลุ่มข้อมูล 3 มิติด้วยขั้นตอนวิธี Fuzzy C-Means 297
รูปที่ 11.10 โครงข่าย Kohonen แบบ 1 มิติ 298
รูปที่ 11.11 โครงข่าย Kohonen แบบ 2 มิติ 298
รูปที่ 11.12 การเปลี่ยนแปลงค่าของเวกเตอร์น้ำหนักจากตัวอย่าง 299
รูปที่ 11.13 เวกเตอร์ระบุพิกัดในชั้น Kohonen 300
รูปที่ 11.14 การหดตัวของรัศมีด้วยฟังก์ชัน $\sigma(t)$ 301
รูปที่ 11.15 ตัวอย่างการแบ่งกลุ่มข้อมูล 2 มิติด้วย SOM 304
รูปที่ 11.16 ตัวอย่างการแบ่งกลุ่มข้อมูล 1 มิติด้วย SOM 305
รูปที่ 11.17 ตัวอย่างการแบ่งกลุ่มข้อมูล 3 มิติด้วย SOM 305
รูปที่ 11.18 ตัวอย่างการกำหนดค่าเริ่มต้นของเวกเตอร์น้ำหนักของการสร้างทัศนภาพ 3 มิติของชุดข้อมูลดอกไอริสด้วย SOM...306
รูปที่ 11.19 ตัวอย่างการสร้างทัศนภาพ 3 มิติของชุดข้อมูลดอกไอริสด้วย SOM 307
รูปที่ 11.20 ตัวอย่างฟังก์ชันการแจกแจงแบบปกติหลายตัวแปร 309
รูปที่ 11.21 ตัวอย่างแบบจำลองเกาส์แบบผสมของปริภูมิขนาด 2 มิติ 312
รูปที่ 11.22 ตัวอย่างแบบจำลองเกาส์แบบผสมจำนวน 3 ส่วนผสมของปริภูมิขนาด 2 มิติ โดยแยกวาดแต่ละส่วนผสม 312
รูปที่ 11.23 ตัวอย่างแบบจำลองเกาส์แบบผสมของปริภูมิขนาด 1 มิติ 313
รูปที่ 11.24 ตัวอย่างแบบจำลองเกาส์แบบผสมจำนวน 3 ส่วนผสมของปริภูมิขนาด 1 มิติ โดยแยกวาดแต่ละส่วนผสม 313
รูปที่ 11.25 ตัวอย่างการแบ่งกลุ่มข้อมูล 2 มิติด้วยขั้นตอนวิธี EM 320
รูปที่ 11.26 ตัวอย่างแบบจำลองเกาส์แบบผสมจากการแบ่งกลุ่มข้อมูล 2 มิติด้วยขั้นตอนวิธี EM 320
รูปที่ 11.27 ตัวอย่างการแบ่งกลุ่มข้อมูล 1 มิติด้วยขั้นตอนวิธี EM 321
รูปที่ 11.28 ตัวอย่างแบบจำลองเกาส์แบบผสมจากการแบ่งกลุ่มข้อมูล 1 มิติด้วยขั้นตอนวิธี EM 321
รูปที่ 11.29 ตัวอย่างการแบ่งกลุ่มข้อมูล 3 มิติด้วยขั้นตอนวิธี EM 322
รูปที่ 11.30 ตัวอย่างการแบ่งกลุ่มด้วยขั้นตอนวิธี EM ที่ผิดพลาด 322

ตัวอย่าง

สารบัญตาราง

ตารางที่ 3.1 ตัวอย่างข้อมูลจากชุดข้อมูล Play Tennis 26

ตารางที่ 3.2 การจำแนกตัวอย่างด้วยปริภูมิแบบ 36

ตารางที่ 4.1 ตารางค่าลอการิทึมฐานสิบ 38

ตารางที่ 4.2 ผลการจำแนกประเภทด้วยค่า k ต่าง ๆ 40

ตารางที่ 4.3 ตัวอย่างการเข้ารหัสสำหรับวิธี LSH 58

ตารางที่ 5.1 ตัวอย่างข้อมูลแบบไบนารี 69

ตารางที่ 5.2 ตัวอย่างข้อมูลเพิ่มเติมจากตารางที่ 5.1 72

ตารางที่ 5.3 ชุดข้อมูล Play Tennis 74

ตารางที่ 5.4 ชุดตัวอย่างย่อยของตัวอย่างในตารางที่ 5.3 ในกรณี Outlook = Sunny 78

ตารางที่ 5.5 ชุดตัวอย่างย่อยของตัวอย่างในตารางที่ 5.3 ในกรณี Outlook = Rain 80

ตารางที่ 5.6 ตัวอย่างการเก็บข้อมูลต้นไม้ตัดสินใจด้วยอาร์เรย์ 82

ตารางที่ 5.7 ตัวอย่างข้อมูลที่มีค่าแบบต่อเนื่อง 96

ตารางที่ 6.1 ตารางค่าความจริงแสดงคุณสมบัติของ AND, OR และ XOR 112

ตารางที่ 6.2 ตัวอย่างข้อมูลนำเข้าและเป้าหมายของโครงข่ายหลายชั้น 129

ตารางที่ 6.3 ตัวอย่างข้อมูลในชั้นซ่อนของโครงข่ายหลายชั้น 129

ตารางที่ 6.4 ตัวอย่างค่าประมาณของข้อมูลในชั้นซ่อนของโครงข่ายหลายชั้น 130

ตารางที่ 7.1 จำนวน SVM แบบไบนารีที่ใช้แก้ปัญหาหลายคลาสวิธีการต่าง ๆ 196

ตารางที่ 7.2 เปรียบเทียบเวลาประมวลผลสำหรับการแก้ปัญหาหลายคลาสในกรณีต่าง ๆ (วินาที) 211

ตารางที่ 9.1 จำนวนครั้งการทายกรณีมากที่สุดเทียบกับคำศัพท์ที่มีจำนวนตัวอักษรจำนวนต่าง ๆ 234

ตารางที่ 10.1 ผลทดลองเปรียบเทียบการลดมิติวิธีต่าง ๆ สำหรับการระบุภาพใบหน้ามนุษย์บนฐานข้อมูล ORL 283

ตารางที่ 11.1 ผลทดลองเปรียบเทียบค่าคลาดเคลื่อนเฉลี่ยของ Centroid 327

ตารางที่ 11.2 ผลทดลองเปรียบเทียบร้อยละอัตราความถูกต้อง 327

ตารางที่ 11.3 ผลทดลองเปรียบเทียบเวลาที่ใช้ในการประมวลผล (วินาที) 327

ตารางที่ 11.4 สรุปอันดับประสิทธิภาพด้านต่าง ๆ ของขั้นตอนวิธีการจัดกลุ่มบนชุดข้อมูลดอกไอริส 328

ตัวอย่าง

ARTIFICIAL
INTELLIGENCE

CHAPTER

01

บทนำ

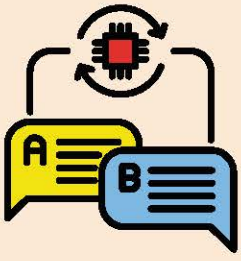
เริ่มต้นบทแรกด้วยการทำความเข้าใจความหมาย และคำจำกัดความของปัญญาประดิษฐ์ การเรียนรู้ของเครื่อง และการเรียนรู้เชิงลึก เพื่อให้ผู้อ่านแยกแยะได้ว่าสิ่งที่พูดถึงนั้นหมายถึงอะไร ตลอดจนการใช้คำศัพท์ภาษาอังกฤษ ภาษาไทย และการใช้สัญลักษณ์ทางคณิตศาสตร์ เรื่องสุดท้ายจะเป็นคำแนะนำในการติดตั้งภาษา Python และโมดูลที่เกี่ยวข้อง

ตัวอย่าง

1.1 ปัญญาประดิษฐ์ (Artificial Intelligence: AI)

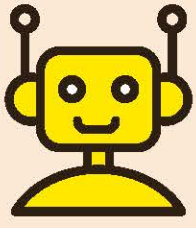
คำว่า **ปัญญาประดิษฐ์ (Artificial Intelligence: AI)** ถูกใช้ครั้งแรกเมื่อประมาณ ค.ศ. 1956 (Russell, Norvig, & Davis, 2010) โดย John McCarthy เพื่อจัดงานสัมมนาเชิงปฏิบัติการสำหรับนักวิจัยที่สนใจในหัวข้อที่เกี่ยวข้องกัน เช่น ทฤษฎีออโตมาตา โครงข่ายประสาทเทียม และการศึกษาเกี่ยวกับความชาญฉลาด

ความหมายของปัญญาประดิษฐ์มีการให้ความหมาย และคำจำกัดความไว้มากมาย แต่มักเกี่ยวข้องกับกระบวนการคิดและการให้เหตุผลโดยมีมนุษย์เป็นต้นแบบ ซึ่งแตกแขนงออกเป็นหลายสาขา เช่น



การประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) ซึ่งเกี่ยวข้องกับการสื่อสารด้วยภาษาที่มนุษย์ใช้งาน เช่น ข้อความ ตัวอักษร เสียงพูด โดยตัวอย่าง NLP ที่พบเจอได้ง่ายในปัจจุบัน เช่น Chatbot ต่าง ๆ ไม่ว่าจะเป็น Siri ของบริษัท Apple, Alexa ของบริษัท Amazon และ Google Assistant ของบริษัท Google เป็นต้น

คอมพิวเตอร์วิทัศน์ (Computer Vision) ซึ่งเทียบได้กับการสร้างตาให้กับเครื่องจักร ที่สามารถรับรู้สิ่งแวดล้อมต่าง ๆ จากภาพ เช่น การรู้จำใบหน้ามนุษย์ การจำแนกประเภทของสิ่งของในภาพ หรือการตรวจหาตำแหน่งสิ่งของที่ต้องการจากภาพ เป็นต้น

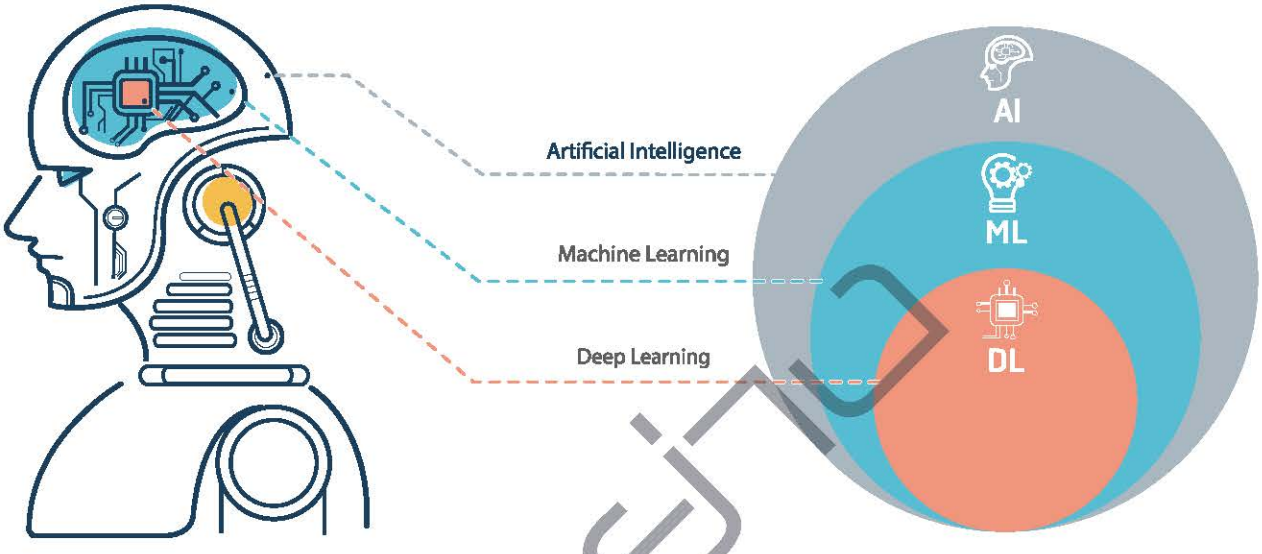


หุ่นยนต์ (Robotics) ซึ่งเกี่ยวกับการสร้างเครื่องกลที่สามารถทำงานในลักษณะเฉพาะต่าง ๆ ได้อย่างแม่นยำและชาญฉลาด คล้ายกับการสร้างร่างกายในกับระบบปัญญาประดิษฐ์

ระบบผู้เชี่ยวชาญ (Expert System) เป็นการจำลองผู้เชี่ยวชาญในงานนั้น ๆ เป็นชุดโปรแกรมคอมพิวเตอร์เพื่อให้สามารถทำงานนั้น ๆ ได้แทนผู้เชี่ยวชาญ หรือช่วยประกอบการตัดสินใจ เช่น ระบบซื้อขายหุ้น ระบบวินิจฉัยโรค เป็นต้น

ตัวอย่าง

อย่างไรก็ตามเครื่องมือที่สาขาต่าง ๆ เหล่านี้ใช้ในการแก้ปัญหาโดยส่วนใหญ่ในปัจจุบันก็คือ **การเรียนรู้ของเครื่อง (Machine Learning: ML)** ซึ่งประสบความสำเร็จอย่างรวดเร็วในยุคปัจจุบัน โดยเฉพาะอย่างยิ่ง **การเรียนรู้เชิงลึก (Deep Learning: DL)** ซึ่งก็เป็นเพียงวิธีการหนึ่งในหลาย ๆ วิธีของการเรียนรู้ด้วยเครื่องเช่นกัน ดังรูปที่ 1.1 แสดงความสัมพันธ์ระหว่าง AI, ML และ DL



▲ รูปที่ 1.1 แผนภูมิแสดงความสัมพันธ์ของปัญญาประดิษฐ์และการเรียนรู้ของเครื่อง

การเรียนรู้ของเครื่อง (Machine Learning: ML) เป็นเครื่องมือหนึ่งของปัญญาประดิษฐ์ที่มุ่งเน้นในการใช้ตัวอย่างหรือประสบการณ์เพื่อการเรียนรู้งาน โดยมนุษย์มีส่วนร่วมเพียงการออกแบบระบบเท่านั้น หลังจากนั้นระบบจะสกัดสาระสำคัญจากตัวอย่างเหล่านี้เอง หลังจากการเรียนรู้เสร็จสิ้นด้วยตัวอย่างจำนวนหนึ่งอย่างเพียงพอ เครื่องหรือระบบที่เรียนรู้แล้วนี้สามารถนำไปใช้ในการประมวลผลของตัวอย่างใหม่ที่ไม่เคยพบมาก่อนได้อย่างมีประสิทธิภาพ

ปัจจุบันระบบปัญญาประดิษฐ์ส่วนใหญ่แทบจะประกอบด้วยการเรียนรู้ของเครื่องเป็นส่วนประกอบหลักเกือบทั้งหมด ดังนั้นการศึกษาการเรียนรู้ของเครื่องจึงเป็นศาสตร์ที่สำคัญที่สุดในการสร้างระบบปัญญาประดิษฐ์ที่สามารถเอาไปประยุกต์ใช้งานได้จริง

1.2 คำศัพท์

หนังสือเล่มนี้พยายามใช้ศัพท์บัญญัติราชบัณฑิตยสถานให้ครอบคลุมเนื้อหา แต่เนื่องจากคำศัพท์ที่ใช้ในหนังสือเล่มนี้บางคำไม่มีศัพท์บัญญัติราชบัณฑิตยสถานออกมาใช้งาน หรือบางคำมีการกำหนดศัพท์บัญญัติราชบัณฑิตยสถานไว้หลายคำ ดังนั้นเพื่อความเข้าใจที่ตรงกันจึงขอชี้แจ้งคำศัพท์ที่สำคัญที่ใช้ในหนังสือเล่มนี้ดังนี้

คำศัพท์ภาษาอังกฤษ	คำศัพท์ภาษาไทย	หมายเหตุ
Accuracy	ความแม่น	ศัพท์บัญญัติราชบัณฑิตยสถาน (คอมพิวเตอร์ 2544, คณิตศาสตร์ 2547)
Acyclic Graph	กราฟไม่มีวง	ศัพท์บัญญัติราชบัณฑิตยสถาน (คอมพิวเตอร์ 2544)
Attribute	ลักษณะประจำ	ศัพท์บัญญัติราชบัณฑิตยสถาน (คอมพิวเตอร์ 2544, คณิตศาสตร์ 2547)
Axon	แกนประสาทนำออก	ศัพท์บัญญัติราชบัณฑิตยสถาน (แพทยศาสตร์ 2544)
Bias	ไบแอส	ขอทับศัพท์เพื่อความเข้าใจ
Canonical	แบบบัญญัติ	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Centralize	เซ็นทรัลไลซ์	ขอทับศัพท์เพื่อความเข้าใจ
Classifier	ตัวจำแนกประเภท	
Conjunction	ประพจน์เชื่อม	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Crossover	การไขว้เปลี่ยน	
Cross-Validation	การตรวจสอบไขว้	
Dendrite	ใยประสาทนำเข้า	ศัพท์บัญญัติราชบัณฑิตยสถาน (แพทยศาสตร์ 2544)
Dimension	มิติ	
Directed Acyclic Graph	กราฟไม่มีวงมีทิศทาง	
Discriminant	ดิสคริมิแนนต์	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Disjunction	ประพจน์เลือก	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Dual Problem	ปัญหาควบคู่	
Edge	เส้นเชื่อม	ศัพท์บัญญัติราชบัณฑิตยสถาน (คอมพิวเตอร์ 2544)
Eigenvalue	ค่าลักษณะเฉพาะ	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Eigenvector	เวกเตอร์ลักษณะเฉพาะ	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Excitation	การเร้า	ศัพท์บัญญัติราชบัณฑิตยสถาน (แพทยศาสตร์ 2544)
Expectation Expected Value	ค่าคาดหมาย	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Exponential	เลขชี้กำลัง	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)

ตัวอย่าง

คำศัพท์ภาษาอังกฤษ	คำศัพท์ภาษาไทย	หมายเหตุ
Feature	คุณลักษณะ	
Fuzzy	คลุมเครือ	ศัพท์บัญญัติราชบัณฑิตยสถาน (เทคโนโลยีสารสนเทศ 2545)
Generalize	ทำให้ใช้ทั่วไปได้	ศัพท์บัญญัติราชบัณฑิตยสถาน (คอมพิวเตอร์ 2544)
Hyperplane	ระนาบเกิน	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Hypothesis	สมมุติฐาน	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Kernel	เคอร์เนล	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Label	ป้ายกำกับ	
Map	การส่ง	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Mode	โหมด	ใช้ในบริบทเทนเซอร์
Mutation	การกลายพันธุ์	
Node	ปม	ศัพท์บัญญัติราชบัณฑิตยสถาน (คอมพิวเตอร์ 2544)
Normalize	นอร์มอลไลซ์	ขอทับศัพท์เพื่อความเข้าใจ
Order	อันดับ	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Positive Definite	บวกแน่นอน	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Positive Semi-Definite	กึ่งบวกแน่นอน	
Precision	ความเที่ยง	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Pruning	การตัดทอน	
Rank	ค่าลำดับชั้น	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Recall	การระลึกได้	ศัพท์บัญญัติราชบัณฑิตยสถาน (แพทย์ศาสตร์ 2544)
Sensitivity	ความไว	ศัพท์บัญญัติราชบัณฑิตยสถาน (แพทยศาสตร์ 2544)
Space	ปริภูมิ	ศัพท์บัญญัติราชบัณฑิตยสถาน (คณิตศาสตร์ 2547)
Specificity	ความจำเพาะ	ศัพท์บัญญัติราชบัณฑิตยสถาน (แพทยศาสตร์ 2544)
Supervised	แบบมีผู้สอน	
Target	เป้าหมาย	
Unsupervised	แบบไม่มีผู้สอน	

ทั้งนี้ในหนังสือเล่มนี้จะใช้คำภาษาอังกฤษในรูปแบบตัวอักษรภาษาอังกฤษ ในกรณีที่การทับศัพท์ไม่นิยมอย่างแพร่หลายหรือยังไม่มีศัพท์บัญญัติ เพื่อให้ง่ายในการสื่อสาร

ตัวอย่าง

1.3 สัญลักษณ์ทางคณิตศาสตร์

ขอกำหนดความหมายของสัญลักษณ์ทางคณิตศาสตร์ในหนังสือเล่มนี้ดังนี้

ตัวแปร	คำอธิบาย
x	สเกลาร์ (Scalar)
X	เซต (Set)
$\mathbf{x}$	เวกเตอร์ (Vector)
$\mathbf{X}$	เมทริกซ์ (Matrix)
$E\mathbf{x}, E[\mathbf{x}]$	ค่าคาดหมาย (Expected value) ของ $\mathbf{x}$
$\bar{x}$	ค่าเฉลี่ยเลขคณิตของ x
$\Re$	เซตจำนวนจริง

1.4 การติดตั้งภาษา Python

ภาษา Python ที่นิยมใช้งานในปัจจุบันมี 2 รุ่น ได้แก่ 2.X และ 3.X ซึ่งหนังสือเล่มนี้จะใช้รุ่น 3.6 และใช้ติดตั้งผ่าน Conda โดยสามารถดาวน์โหลด Miniconda ได้จาก https://docs.conda.io/en/latest/miniconda.html

ให้เลือกตามระบบปฏิบัติการที่ใช้ โดยรองรับการใช้งานกับ Windows, Mac OS หรือ Linux เท่านั้น และควรเลือกแบบ 64 บิต และ Python รุ่น 3 จากนั้นทำการติดตั้ง เมื่อติดตั้งเรียบร้อยจะทำการสร้าง Virtual Environment ด้วยคำสั่ง

```
conda create -n AIML python=3.6
```

ใน Anaconda Promt สำหรับระบบปฏิบัติการ Windows และใน Terminal สำหรับระบบปฏิบัติการ Mac OS และระบบปฏิบัติการ Linux

การสร้าง Virtual Environment มีประโยชน์ในการทำงานหลายงานบนเครื่องเดียวกัน เพราะโมดูลในภาษา Python มีการขึ้นต่อกัน ทำให้บางโมดูลต้องการใช้ต่างรุ่นกัน การใช้เพียง Environment เดียวกันจะทำให้มีปัญหาความขัดแย้งนี้ และจะเกิดความผิดพลาดจากความต้องการรุ่นต่าง ๆ ของแต่ละโมดูลที่แตกต่างกันจนไม่สามารถใช้งานร่วมกันได้

สำหรับระบบปฏิบัติการ Windows ที่กำหนดภาษาของเครื่องเป็นภาษาไทย อาจมีการแสดงความผิดพลาดเกี่ยวกับการเข้ารหัสภาษา cp874 ให้เข้าไปแก้ไขที่ไฟล์ aliases.py โดยจะอยู่ในแฟ้มที่ติดตั้ง เช่น

```
C:\Users\<Username>\Miniconda3\Lib\encodings
```

ตัวอย่าง

แล้วทำการเพิ่มรหัส cp874 ดังตัวอย่าง

```
# cp869 codec
'869'               : 'cp869',
'cp_gr'             : 'cp869',
'csibm869'          : 'cp869',
'ibm869'            : 'cp869',

# cp874 codec
'874': 'cp874',

# cp932 codec
'932'               : 'cp932',
'ms932'             : 'cp932',
'mskanji'           : 'cp932',
'ms_kanji'          : 'cp932',
```

สำหรับระบบปฏิบัติการอื่น ๆ จะไม่พบปัญหานี้

เมื่อได้ Virtual Environment ชื่อ AIML แล้วจะทำการ activate ด้วยคำสั่ง

สำหรับระบบปฏิบัติการ Windows (ใน Anaconda Prompt) ระบบปฏิบัติการ Mac OS หรือระบบปฏิบัติการ Linux (ใน Terminal)

```
conda activate AIML
```

หากสำเร็จจะปรากฏ (AIML) อยู่ด้านหน้า เช่น

```
(AIML) C:\Users\<Username>>
```

จากนั้นจะทำการติดตั้งโมดูลที่จำเป็นสำหรับหนังสือเล่มนี้ ด้วยคำสั่งดังนี้

```
conda install numpy scipy matplotlib pandas mayavi cvxopt
```

ในความเป็นจริงแล้วชุดคำสั่งหลักที่เกี่ยวข้องกับขั้นตอนวิธีต่าง ๆ ที่ใช้ทั้งหมดในหนังสือเล่มนี้ จะทำการเขียนขึ้นมาได้เองทั้งหมดจาก numpy เพียงโมดูลเดียว ส่วนโมดูล cvxopt เอาไว้แสดงตัวอย่างการแก้ปัญหา Quadratic Programming ของ Support Vector Machine (SVM) เท่านั้น แต่จะใช้จริงด้วยขั้นตอนวิธี Sequential Minimal Optimization (SMO) ที่มีประสิทธิภาพเหมาะสมกับ SVM มากกว่า และสามารถสร้างขึ้นเองด้วย numpy ได้ ทั้งนี้โมดูลอื่น ๆ มีไว้เพื่อแสดงผลและนำเข้าข้อมูลเท่านั้น

1.5 เอกสารอ้างอิง

Russell, S. J., Norvig, P., & Davis, E. (2010). *Artificial Intelligence:* A Modern Approach: Prentice Hall.

ARTIFICIAL
INTELLIGENCE

CHAPTER

02

การเรียนรู้ของเครื่อง

(Machine Learning)

การเรียนรู้ของเครื่อง (Machine Learning) หมายถึง การให้เครื่อง (ในที่นี้หมายถึงเครื่องคำนวณที่สามารถทำงานตามชุดคำสั่งได้ เช่น คอมพิวเตอร์) เรียนรู้งานใดงานหนึ่ง (Task) จากตัวอย่าง (Sample) หรือประสบการณ์ (Experience) จำนวนหนึ่งเพื่อให้ทำงานนั้น ๆ ได้อย่างมีประสิทธิภาพ (Performance) อีกทั้งยังสามารถเพิ่มประสิทธิภาพได้จากการเรียนรู้จากตัวอย่างหรือประสบการณ์ที่เพิ่มขึ้นได้

2.1 จุดประสงค์ประจำบท

หลังจากผู้อ่านทำการศึกษาบทนี้จบแล้ว ผู้อ่านควรจะสามารถ

- เข้าใจความหมายของการเรียนรู้ของเครื่อง
- เข้าใจว่างานรูปแบบใดเหมาะสมกับการประยุกต์ใช้การเรียนรู้ของเครื่อง
- ออกแบบระบบเรียนรู้ได้อย่างเหมาะสม
- เขียนโปรแกรมเพื่อแก้ไขปัญหาอย่างง่ายโดยอาศัยการเรียนรู้ของเครื่อง

2.2 การเรียนรู้ของเครื่องและการเรียนรู้ของมนุษย์

ประสบการณ์การเรียนรู้ของมนุษย์ที่เกิดขึ้นกับตัวเราโดยตรง สามารถเป็นตัวอย่างที่ดีให้สามารถเข้าใจการเรียนรู้ของเครื่องได้ง่ายขึ้น เช่น ตอนวัยเด็กเราเรียนรู้การเขียนตัวอักษร ก.ไก่ ด้วยวิธีการคัดลายมือ กล่าวคือ สมุดคัดลายมือเล่มแรกจะมีจุดประหรือจุดไข่ปลาให้เขียนตามได้ และเมื่อคล่องแล้วจะเริ่มคัดลายมือโดยไม่มีจุดไข่ปลาอีกต่อไป จนในที่สุดเราสามารถจดจำและทำการเขียนตัวอักษร ก.ไก่ ได้โดยไม่จำเป็นต้องทำการคัดลายมืออีก ซึ่งเหตุการณ์นี้สามารถเปรียบเทียบกับการเรียนรู้ของเครื่องได้ดังนี้

- **งาน (Task)** คือ การเขียนตัวอักษร ก.ไก่ ได้
- **ประสบการณ์ (Experience)** คือ การฝึกคัดลายมือ
- **ประสิทธิภาพ (Performance)** คือ ความถูกต้องของลายมือที่เขียนอักษร ก.ไก่ หลังจากที่ได้ทำการฝึกคัดลายมือแล้ว

▲ รูปที่ 2.1 ตัวอย่างลายจุดไข่ปลาตัวอักษร ก.ไก่ สำหรับคัดลายมือ

นอกจากนี้ยังสามารถเปรียบเทียบโดยละเอียดเพื่อความชัดเจนยิ่งขึ้นได้ดังนี้ ลายมือที่ถูกเขียนบนจุดไข่ปลาดังรูปที่ 2.1 สามารถเทียบได้กับ **ชุดตัวอย่างฝึกฝน (Training Set)** ซึ่งควรมีจำนวนมากพอ เพราะถ้าคัดลายมือน้อยเกินไปจะทำให้เกิด **ปัญหาขนาดตัวอย่างน้อย (Small Sample Size (SSS) Problem)** ซึ่งส่งผลให้ผลลัพธ์หรือประสิทธิภาพในการเขียนไม่ดีตามมา การคัดลายมือลักษณะนี้เป็น **การเรียนรู้แบบมีผู้สอน (Supervised Learning)** เนื่องจากมีจุดไข่ปลาซึ่งเทียบได้กับ **เป้าหมาย (Target)** หรือเป็นตัวบอกถึง **ป้ายกำกับ (Label)** ของตัวอักษรว่าเป็นตัวอักษรอะไร **ประสิทธิภาพ (Performance)** อาจวัดได้จาก **ความผิดพลาด (Error)** ที่เกิดขึ้นจากการคัดลายมือ เช่น จำนวนจุดไข่ปลาที่ไม่ถูกลากผ่าน หรือจากจำนวนบุคคลอื่นที่สามารถอ่านอักษรที่เขียนได้ถูกต้อง เป็นต้น หลังจากนั้นการคัดลายมือบนกระดาษเปล่าที่ไม่มีจุดไข่ปลา แต่ยังทราบว่ากำลังเขียนตัวอักษร ก.ไก่ นั้นเทียบได้กับ **ชุดประเมินผล (Validation Set)** ซึ่งคล้ายกับเป็นการทดลองเขียนก่อนจะนำไปใช้จริง สุดท้ายการเขียนตัว ก.ไก่ เพื่อใช้จริงในการประกอบกับตัวอักษรต่าง ๆ ให้เกิดคำภายหลัง ซึ่งไม่ทราบล่วงหน้าว่าจะเขียนตัวอักษรใดเทียบได้กับ **ชุดทดสอบ (Test Set)**

ตัวอย่าง

นอกจากตัวอย่างการคัดลายมือที่กล่าวมายังมีตัวอย่างอีกมากมายในชีวิตประจำวัน เช่น การฝึกฝีมือต่าง ๆ การจำแนกหรือคัดแยกสิ่งของ การประมาณค่าต่าง ๆ การทำนายหรือการพยากรณ์ การเล่นเกมต่าง ๆ เป็นต้น โดยทั่วไปศาสตร์ด้านการเรียนรู้ของเครื่องเน้นการศึกษา และพัฒนาเครื่องมือเบื้องหลังให้ระบบอัจริยะต่าง ๆ เช่น การทำเหมืองข้อมูล (Data Mining) ระบบสนับสนุนการตัดสินใจ (Decision Support System: DSS) ระบบธุรกิจอัจฉริยะ (Business Intelligence: BI) การวินิจฉัยโรค (Medical Diagnosis) ชีวมาตร (Biometric) การประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) การรู้จำอักขระด้วยแสง (Optical Character Recognition: OCR) การรู้จำเสียงพูด (Speech Recognition) การรู้จำใบหน้า (Face Recognition) ฯลฯ

2.3 ลักษณะงานที่เหมาะสมกับการเรียนรู้ของเครื่อง

ทำไมเครื่องจักรต้องทำการเรียนรู้ เหตุใดจึงไม่สร้างเครื่องจักรที่เหมาะสมต่องานที่ต้องการใช้ไปเลยตั้งแต่แรก เช่น หากต้องการปรับอากาศก็สร้างเครื่องปรับอากาศ หากต้องการถ่ายรูปก็สร้างกล้องถ่ายรูป เป็นต้น ซึ่งหากวิเคราะห์ต่อไปจะพบว่าเครื่องจักรเหล่านี้ก็ยังคงไม่ตอบสนองความต้องการเราได้เต็มที่หากไม่มีระบบการเรียนรู้เพิ่มเติมไปด้วย เช่น เครื่องปรับอากาศควรมีการปรับอุณหภูมิให้เหมาะสมตลอดเวลาโดยไม่ให้ผู้ใช้รู้สึกร้อนหรือหนาวเกินไป หรือกล้องถ่ายรูปก็ควรมีระบบปรับจุดโฟกัสอัตโนมัติให้ภาพชัดเจนในจุดที่ต้องการ ซึ่งสิ่งเหล่านี้เราไม่สามารถสร้างเตรียมไว้ก่อนได้ เพราะมีปัจจัยและสิ่งแวดล้อมที่เปลี่ยนแปลงได้เสมอ กล่าวคือ เราไม่สามารถทราบได้ล่วงหน้าว่าผู้ใช้จะใช้เครื่องปรับอากาศในสภาพอากาศแบบใด หรือเราไม่สามารถทราบได้ล่วงหน้าว่าผู้ใช้จะนำกล้องถ่ายรูปไปใช้ถ่ายรูปอะไร และจะต้องการให้ชัดตำแหน่งใด หลักการเรียนรู้ด้วยเครื่องสามารถช่วยจัดการปัญหาเหล่านี้ได้

การเรียนรู้ของเครื่องมีความเหมาะสมกับงานมีที่ลักษณะต่าง ๆ ดังนี้

- งานที่มีความซับซ้อนมากหรือมีข้อมูลมากจนหาวิธีการวิเคราะห์หรือจัดการได้ยากด้วยมนุษย์
- งานที่มีปัจจัยและสภาพแวดล้อมมีการเปลี่ยนแปลงตลอดเวลา และต้องการเครื่องจักรที่สามารถปรับตัวให้เหมาะสมกับสภาพแวดล้อมนั้น ๆ ได้
- งานที่ไม่สามารถคาดการณ์ล่วงหน้าหรือไม่สามารถออกแบบวิธีการไว้ก่อนล่วงหน้าได้ เพราะต้องมีการตอบสนองที่แตกต่างกันไปตามสถานการณ์
- งานที่ไม่มีหรือไม่สามารถรู้รายละเอียดภายในระบบได้ ทราบเพียงแต่ข้อมูลนำเข้า (Input) และผลลัพธ์ (Output) เท่านั้น หรือบางครั้งเรียกว่าระบบ Black Box ดังรูปที่ 2.2

▲ รูปที่ 2.2 ระบบ Black Box

ตัวอย่าง

- งานที่ต้องการวิเคราะห์ความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลจำนวนมาก ซึ่งไม่อาจทราบล่วงหน้าได้ถึงความสัมพันธ์นั้นว่าคือความสัมพันธ์ของสิ่งใดกับสิ่งใด
- งานที่ต้องอาศัยความรู้ที่เกิดขึ้นใหม่หรือเปลี่ยนแปลงอยู่ตลอดเวลา
- งานที่มีข้อมูลเข้าเดียวกัน แต่มีหลายรูปแบบหรือไม่คงที่
- งานที่ไม่สามารถแก้ปัญหาด้วยการค้นหาให้ครบถ้วนหรือครบทุกกรณีได้ (Brute-Force) หรือทำได้แต่ใช้เวลานานเกินกว่าจะยอมรับได้ในงานนั้น ๆ

2.4 การออกแบบระบบเรียนรู้

งานแต่ละประเภทอาจไม่เหมือนกัน แต่มีขั้นตอนวิธีการออกแบบการเรียนรู้ที่สามารถใช้หลักการเดียวกันในการออกแบบระบบได้ดังนี้

2.4.1 กำหนดปัญหา

กำหนดปัญหาของงานที่ต้องการดังนี้

- งาน (Task) ที่ต้องการคืออะไร
- การวัดประสิทธิภาพ (Perfermance) ทำได้อย่างไร
- ตัวอย่างหรือประสบการณ์ (Experience) ที่จะนำมาฝึกฝนระบบคืออะไร

ตัวอย่างเช่น

การเล่นหมากรุก

- งาน (Task) ที่ต้องการคือ การเล่นหมากรุกให้ชนะ
- การวัดประสิทธิภาพ (Perfermance) ทำได้จากการคำนวณค่าอัตราส่วนของจำนวนครั้งที่เล่นแล้วชนะต่อจำนวนครั้งที่เล่นทั้งหมด
- ตัวอย่างหรือประสบการณ์ (Experience) ที่จะนำมาฝึกฝนระบบทำได้โดยการใช้บันทึกหมากของมนุษย์หรือได้จากการที่ให้ระบบเล่นแข่งขันกันเอง

การรู้จำใบหน้า

- งาน (Task) ที่ต้องการคือ การรู้จำบุคคลจากภาพใบหน้า
- การวัดประสิทธิภาพ (Perfermance) ทำได้จากการคำนวณค่าร้อยละความถูกต้อง
- ตัวอย่างหรือประสบการณ์ (Experience) ที่จะนำมาฝึกฝนระบบทำได้โดยการใช้ตัวอย่างภาพของกลุ่มบุคคลที่ต้องการรู้จำทั้งหมด

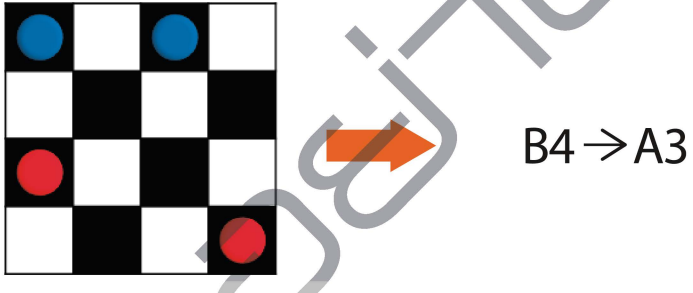
ตัวอย่าง

การยืนยันตัวตนด้วยเสียง

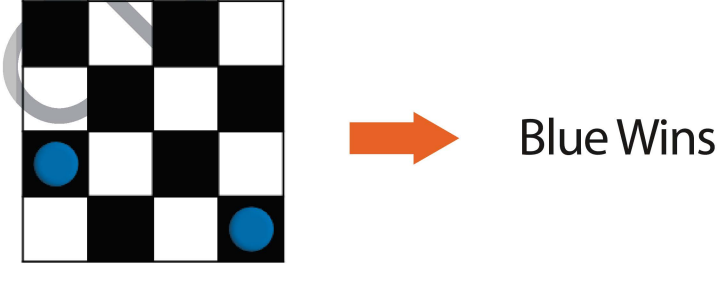
- งาน (Task) ที่ต้องการคือ การยืนยันตัวตนจากเสียงพูด
- การวัดประสิทธิภาพ (Perfermance) ทำได้จากการคำนวณค่าร้อยละความถูกต้อง ค่าความไว (Sensitivity) และความจำเพาะ (Specificity)
- ตัวอย่างหรือประสบการณ์ (Experience) ที่จะนำมาฝึกฝนระบบทำได้โดยการใช้ตัวอย่างเสียงพูดของบุคคลนั้น ๆ

2.4.2 เลือกรูปแบบของตัวอย่างหรือประสบการณ์ที่จะใช้ในการฝึกฝน

- **การเก็บข้อมูลทางตรง (Direct) หรือทางอ้อม (Indirect)** ตัวอย่างเช่น การเก็บข้อมูลการเล่นหมากฮอส (Mitchell, 1997) เราอาจจะเก็บบันทึกการเดินหมากทีละตาทั้งหมด ซึ่งแบบนี้เป็นการเก็บทางตรงดังรูปที่ 2.3 หรือเราอาจจะเก็บเฉพาะจำนวนหมากที่เหลือในตอนที่ทราบผลแพ้ชนะแล้วดังรูปที่ 2.4 ก็ได้ ซึ่งแบบหลังนี้เป็นการเก็บทางอ้อม จะเห็นว่าการเก็บข้อมูลทางตรงจะละเอียดกว่า ซึ่งต้องใช้ทรัพยากรในการเก็บข้อมูลมากกว่า แต่การเก็บข้อมูลทางอ้อมจะง่ายและใช้ทรัพยากรน้อยกว่า ซึ่งข้อมูลที่ได้จะมีรายละเอียดน้อยกว่า

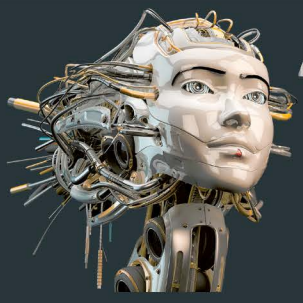


▲ รูปที่ 2.3 ตัวอย่างการเก็บข้อมูลทางตรง (Direct)



▲ รูปที่ 2.4 ตัวอย่างการเก็บข้อมูลทางอ้อม (Indirect)

- **แบบมีผู้สอนหรือแบบไม่มีผู้สอน** การใช้ข้อมูลแบบมีผู้สอนนั้นหมายถึงการที่มีความรู้จากผู้สอน หรือจากภายนอกระบบมาช่วยในการเลือกหรือจัดการข้อมูล เช่น การเดินหมากในแต่ละตานั้นควรเดินอย่างไรให้ได้ผลดี เป็นต้น ส่วนแบบไม่มีผู้สอนนั้น เครื่องหรือผู้เรียนจะเป็นผู้ตัดสินใจเองว่าจะเดินหมากแต่ละตาอย่างไร
- **วิธีการเรียนรู้** การเรียนรู้ทำได้หลายรูปแบบ ตัวอย่างเช่น การเรียนรู้ในการเดินหมาก เราอาจให้เครื่องเดินแข่งกับตัวเองหรือแข่งกับผู้เล่นจริงก็ได้ ซึ่งการเลือกแบบแรกนั้นเครื่องสามารถเรียนรู้ได้ด้วยตนเองโดยไม่ต้องการปัจจัยภายนอก แต่อาจต้องใช้เวลาในการเรียนรู้มากกว่าวิธีหลัง ซึ่งต้องอาศัยผู้เล่นอื่นมาช่วยในการฝึกฝน



Artificial Intelligence with Machine Learning

AI สร้างได้ด้วยแมชชีนเลิร์นนิ่ง

Python Edition

Supervised Learning

Unsupervised Learning

Machine Learning คืออะไร

แมชชีนเลิร์นนิ่งหรือการเรียนรู้ของเครื่อง เป็นองค์ประกอบหนึ่งในระบบ Artificial Intelligence หรือ AI มุ่งเน้นการสอนให้ระบบคอมพิวเตอร์สามารถเรียนรู้ได้ด้วยตนเอง โดยอาศัยชุดข้อมูลตัวอย่าง หรือประสบการณ์เพื่อการเรียนรู้งาน มนุษย์มีส่วนร่วมเพียงการออกแบบระบบเท่านั้น หลังจากนั้นระบบจะสกัดสาระสำคัญจากตัวอย่างเหล่านี้เอง หลังจากการเรียนรู้เสร็จสิ้นด้วยตัวอย่างจำนวนหนึ่งอย่างเพียงพอ เครื่องหรือระบบที่เรียนรู้แล้วนี้สามารถนำไปใช้ในการประมวลผลของตัวอย่างใหม่ที่ไม่เคยพบมาก่อนได้อย่างมีประสิทธิภาพ

Machine Learning มีความสำคัญอย่างไร

ในปัจจุบันระบบ AI ส่วนใหญ่แทบจะประกอบด้วย Machine Learning เป็นส่วนประกอบหลักเกือบทั้งหมด ดังนั้น การศึกษา Machine Learning จึงเป็นศาสตร์ที่สำคัญที่สุดในการสร้างระบบ AI ที่สามารถเอาไปประยุกต์ใช้งานได้จริง อีกทั้งการเข้าใจในเครื่องมือพื้นฐานของ Machine Learning มีประโยชน์ในการเลือกใช้เครื่องมือต่าง ๆ ให้เหมาะสมกับงานที่ต้องการ รวมถึงการปรับปรุงขั้นตอนวิธีเพื่อให้เหมาะสมกับการใช้งานในแต่ละงานได้อย่างมีประสิทธิภาพ

ทำไมหนังสือเล่มนี้จึงเหมาะกับคุณ

หนังสือเล่มนี้ ผู้เขียนได้อาศัยประสบการณ์ในการสอน และการวิจัยในสาขาวิชานี้ตั้งแต่ระดับชั้นปริญญาตรีถึงปริญญาเอก โดยในแต่ละบทจะประกอบด้วยเนื้อหาและตัวอย่างที่เข้าใจง่าย แสดงขั้นตอนวิธี ตลอดจนการเขียนโปรแกรมตัวอย่างด้วยภาษา Python ซึ่งเป็นภาษาที่เข้าใจง่าย และได้รับความนิยมอย่างสูงในปัจจุบัน

ทั้งนี้ Source Code เกือบทั้งหมด ผู้เขียนตั้งใจเขียนขึ้นมาจากพื้นฐานโดยไม่ใช้ Library ระดับสูง เพื่อให้ผู้อ่านได้เข้าใจในหลักการทำงานเบื้องหลังอย่างแท้จริง และสามารถนำไปประยุกต์ใช้ด้วยภาษาที่ตนเองถนัดได้ง่าย

ผู้อ่านควรมีความรู้พื้นฐานอะไรบ้าง

- มีพื้นฐานการคำนวณทางด้านคณิตศาสตร์ในระดับมัธยมปลายเป็นอย่างน้อย
- มีพื้นฐานทางด้านการเขียนโปรแกรมด้วยภาษา Python หรือภาษาอื่นๆ

การเรียนรู้ของเครื่องสร้างได้

- เห็นภาพความสัมพันธ์ของ AI, ML และ DL
- เปรียบเทียบการเรียนรู้ของเครื่องกับการเรียนรู้ของมนุษย์
- Machine Learning ดีกว่าเครื่องหรือระบบปัจจุบันอย่างไร
- Machine Learning เหมาะกับลักษณะงานแบบไหน
- การออกแบบระบบเรียนรู้ทำได้อย่างไร
- เรียน Machine Learning Algorithms ผ่านตัวอย่างง่าย ๆ
- เลือก Algorithms ให้เหมาะกับงาน
- เทคนิคการลดมิติเพื่อการวิเคราะห์ที่มีประสิทธิภาพ
- Concept Learning เพื่อเรียนรู้แนวคิด
- k-Nearest Neighbors เพื่อการจำแนกข้อมูล
- Decision Tree ช่วยตัดสินใจผ่านกราฟต้นไม้
- Artificial Neural Networks พื้นฐานของ Deep Learning
- Support Vector Machine เพื่อการจำแนกข้อมูล
- Bayesian Learning เพื่อหาข้อสรุปจากความน่าจะเป็น
- Genetic Algorithm เพื่อหาผลเฉลยที่เหมาะสม
- Clustering ด้วยการเรียนรู้แบบ Unsupervised Learning

ผู้แต่ง
รศ. ดร.ปริญญา สงวนสัตย์
บรรณาธิการ
ภีรพล คชาเจริญ

โปรดติดตามสาระเพิ่มเติมได้ทาง
Facebook Page และ YouTube
Channel AI บ้าน บ้าน

SERAZU ร้านหนังสือออนไลน์
สนใจสั่งซื้อได้ที่
www.serazu.com

จัดจำหน่ายโดย IDC PREMIER
ISBN 885-916-100-774-6
8 859161 007746
IDC 495 IIM