



การเรียนรู้ของเครื่อง **MACHINE LEARNING**

สุพจน์ เสงพระพรหม

สาขาวิชาวิทยาการข้อมูล

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม



การเรียนรู้ของเครื่อง

ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ เสงี่ยมพรหม
วศ.ด. (วิศวกรรมคอมพิวเตอร์)

คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยราชภัฏนครปฐม

2566

การเรียนรู้ของเครื่อง MACHINE LEARNING

ผู้แต่ง ศศ.ดร.สุพจน์ เสงพะพรหม
พิมพ์ครั้งที่ 1
ปีที่พิมพ์ 2566
จำนวน 200 เล่ม
พิมพ์ที่ โรงพิมพ์มหาวิทยาลัยศิลปากร วิทยาเขตพระราชวังสนามจันทร์
เลขที่ 6 ถนนราชมรรคาใน อำเภอเมือง จังหวัดนครปฐม 73000
จัดทำโดย ศศ.ดร.สุพจน์ เสงพะพรหม
สาขาวิชาวิทยาการข้อมูล คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยราชภัฏนครปฐม ถนนมาลัยแมน ตำบลหนองปากโลง
อำเภอเมืองนครปฐม จังหวัดนครปฐม 73000
E-mail : supojn@yahoo.com
ราคา 290 บาท

ข้อมูลทางบรรณานุกรมของหอสมุดแห่งชาติ

สุพจน์ เสงพะพรหม.

การเรียนรู้ของเครื่อง MACHINE LEARNING. -- นครปฐม : สาขาวิชาวิทยาการข้อมูล
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม , 2566.

209 หน้า.

1. การเรียนรู้ของเครื่อง. |. ชื่อเรื่อง.

006.31

ISBN 978-616-608-398-9

สงวนลิขสิทธิ์ตามพระราชบัญญัติการพิมพ์
ห้ามมิให้ทำซ้ำหรือลอกเลียนแบบโดยไม่ได้รับอนุญาต

คำนำ

ตำราเรื่อง การเรียนรู้ของเครื่อง เล่มนี้ ได้ทำการพัฒนามาจากเอกสารคำสอน รายวิชา การเรียนรู้ของเครื่อง ที่ใช้ประกอบการศึกษารายวิชา การเรียนรู้ของเครื่อง รหัสวิชา 7163402 ซึ่งเป็นรายวิชาบังคับในการจัดการเรียนการสอนในระดับปริญญาตรี สาขาวิชาวิทยาการข้อมูล ของคณะวิทยาศาสตร์และเทคโนโลยีมหาวิทยาลัยราชภัฏนครปฐม โดยได้มีการปรับปรุงคำอธิบายเนื้อหาให้ละเอียด ครบคลุมและสามารถทำความเข้าใจได้ง่ายขึ้น มีการเพิ่มเติมเนื้อหาทางด้านการพัฒนาโปรแกรมด้วยภาษาไพธอนเพื่อให้สามารถนำความรู้ไปประยุกต์ใช้ได้อย่างกว้างขวางมากยิ่งขึ้น รวมถึงมีการเพิ่มสื่อวีดิทัศน์ในการอธิบายเนื้อหาภาคทฤษฎีในแต่ละบท ทำให้ผู้อ่านสามารถศึกษาเนื้อหาและทำความเข้าใจได้เหมือนกับการเข้าชั้นเรียน ซึ่งจะช่วยให้เกิดความเข้าใจในเนื้อหาได้ง่ายยิ่งขึ้น

หวังเป็นอย่างยิ่งว่าตำราเล่มนี้จะเป็นประโยชน์แก่ผู้อ่านที่สนใจเรียนรู้ในศาสตร์ด้านการเรียนรู้ของเครื่องได้ไม่มากนักน้อย หากท่านมีข้อเสนอแนะ/ข้อคำถามประการใด ผู้เขียนยินดีรับฟังจากทุกท่าน โดยสามารถส่งข้อมูลมาได้ทางจดหมายอิเล็กทรอนิกส์ได้ที่ supojn@yahoo.com และขอขอบคุณในความอนุเคราะห์มา ณ โอกาสนี้

สุพจน์ เสงี่ยมพรหม

พฤศจิกายน 2566

สารบัญ

	หน้า
คำนำ.....	(1)
สารบัญ.....	(3)
สารบัญภาพ.....	(11)
สารบัญตาราง.....	(17)
บทที่ 1 บทนำ.....	1
1.1 ความหมายและความสำคัญของการเรียนรู้ของเครื่อง.....	1
1.1.1 ความหมายของการเรียนรู้ของเครื่อง.....	1
1.1.2 ความสำคัญของการเรียนรู้ของเครื่อง.....	2
1.2 ประวัติความเป็นมาของการเรียนรู้ของเครื่อง.....	3
1.2.1 จุดกำเนิดของการเรียนรู้ของเครื่อง.....	3
1.2.2 พัฒนาการของการเรียนรู้ของเครื่องในยุคต่าง ๆ.....	4
1.3 ประเภทของการเรียนรู้ของเครื่อง.....	8
1.3.1 การเรียนรู้แบบมีผู้สอน.....	8
1.3.2 การเรียนรู้แบบไม่มีผู้สอน.....	10
1.3.3 การเรียนรู้แบบเสริมแรง.....	12
1.4 แขนงวิชาที่นำการเรียนรู้ของเครื่องไปใช้.....	13
1.4.1 การประมวลผลภาพ.....	13
1.4.2 การประมวลผลเสียง.....	14
1.4.3 ระบบแนะนำ.....	15
1.4.4 การประมวลผลภาษาธรรมชาติ.....	15
1.4.5 วิทยาการหุ่นยนต์.....	16
1.5 ความท้าทายของการเรียนรู้ของเครื่อง.....	17
1.5.1 ความท้าทายด้านข้อมูล.....	18
1.5.2 ความท้าทายด้านการสร้างตัวแบบ.....	18
สรุป.....	19
แบบฝึกหัดท้ายบทที่ 1.....	19
เฉลยแบบฝึกหัดท้ายบทที่ 1.....	20
เอกสารอ้างอิง.....	21

สารบัญ (ต่อ)

	หน้า
บทที่ 2 ปัญหาและการค้นหาคำตอบ	25
2.1 ปัญหาและการแก้ปัญหา	25
2.1.1 นิยาม	25
2.1.2 การแก้ปัญหา	26
2.2 ปริภูมิสถานะ	27
2.2.1 สถานะ	27
2.2.2 ตัวดำเนินการ	29
2.3 เทคนิคการค้นหา	31
2.3.1 การค้นหาแบบบอด	31
2.3.2 การค้นหาแบบฮิวริสติก	35
สรุป	42
แบบฝึกหัดท้ายบทที่ 2	42
เฉลยแบบฝึกหัดท้ายบทที่ 2	43
เอกสารอ้างอิง	44
บทที่ 3 การถดถอยเชิงเส้น	45
3.1 ทฤษฎีการถดถอยเชิงเส้น	45
3.1.1 นิยามการถดถอยเชิงเส้น	45
3.1.2 การสร้างตัวแบบการถดถอยเชิงเส้นอย่างง่าย	47
3.1.3 ตัวอย่างการคำนวณและการใช้งานการถดถอยเชิงเส้น	51
3.2 การพัฒนาตัวแบบการถดถอยเชิงเส้นด้วยโปรแกรม MS-Excel	53
3.2.1 การใช้สูตรคำนวณ	53
3.2.2 การใช้เส้นแนวโน้ม	54
3.3 การพัฒนาตัวแบบการถดถอยเชิงเส้นด้วยการเขียนโปรแกรมภาษาไพธอน	57
3.3.1 การใช้สูตรคำนวณ	58
3.3.2 การใช้ไลบรารี scikit-learn	59
สรุป	60
แบบฝึกหัดท้ายบทที่ 3	61

สารบัญ (ต่อ)

	หน้า
เฉลยแบบฝึกหัดท้ายบทที่ 3.....	61
เอกสารอ้างอิง.....	64
บทที่ 4 เพื่อนบ้านใกล้ที่สุด	65
4.1 ทฤษฎีเพื่อนบ้านใกล้ที่สุด.....	65
4.1.1 นิยามเพื่อนบ้านใกล้ที่สุด.....	65
4.1.2 ตัวอย่างการคำนวณและการใช้งานเพื่อนบ้านใกล้ที่สุด.....	68
4.2 การพัฒนาระบบวิเคราะห์ข้อมูลแบบเพื่อนบ้านใกล้ที่สุดด้วยโปรแกรม MS-Excel.....	70
4.2.1 การหาระยะห่างด้วยฟังก์ชันระยะห่างยูคลิดี언.....	71
4.2.2 การหาลำดับที่ของความเหมือน.....	72
4.2.3 การหาคำตอบของตัวอย่างที่มีลักษณะเหมือนกับข้อมูลสอบถามมากที่สุด.....	72
4.2.4 การให้ค่าน้ำหนักคะแนนเสียงของคำตอบ.....	72
4.2.5 การหาผลรวมคะแนนเสียงของแต่ละคำตอบ.....	74
4.2.6 การหาคะแนนเสียงรวมมากที่สุด.....	75
4.2.7 การแสดงคำตอบเป็นผลการทำนาย.....	76
4.3 การพัฒนาระบบวิเคราะห์ข้อมูลแบบเพื่อนบ้านใกล้ที่สุดด้วยการเขียนโปรแกรมภาษาไพธอน.....	77
4.3.1 การพัฒนาตามขั้นตอนวิธีการคำนวณ.....	77
4.3.2 การใช้ไลบรารี scikit-learn.....	79
สรุป.....	81
แบบฝึกหัดท้ายบทที่ 4.....	81
เฉลยแบบฝึกหัดท้ายบทที่ 4.....	81
เอกสารอ้างอิง.....	84
บทที่ 5 เบื้อง่าย	85
5.1 การเรียนรู้แบบเบื่อง่าย.....	85
5.1.1 ทฤษฎีเบย์.....	85
5.1.2 นิยามเบื่อง่าย.....	86
5.1.3 ตัวอย่างการคำนวณและการใช้งานเบื่อง่าย.....	87

สารบัญ (ต่อ)

	หน้า
5.2 การพัฒนาระบบวิเคราะห์ข้อมูลแบบเบย์อย่างง่ายด้วยโปรแกรม MS-Excel.....	89
5.2.1 การหาจำนวนเหตุการณ์ตามเงื่อนไขที่กำหนด.....	90
5.2.2 การหาค่าความน่าจะเป็นของเหตุการณ์.....	90
5.2.3 การหาค่าความน่าจะเป็นของการเกิดคลาสด่าง ๆ เมื่อรู้ค่าคุณลักษณะของข้อมูล.....	93
5.2.4 การหาคำตอบที่ได้จากการทำนาย.....	94
5.3 การพัฒนาระบบวิเคราะห์ข้อมูลแบบเบย์อย่างง่ายด้วยการเขียนโปรแกรมภาษาไพธอน.....	95
5.3.1 การพัฒนาตามขั้นตอนวิธีการคำนวณ.....	95
5.3.2 การใช้ไลบรารี scikit-learn.....	98
สรุป.....	101
แบบฝึกหัดท้ายบทที่ 5.....	101
เฉลยแบบฝึกหัดท้ายบทที่ 5.....	101
เอกสารอ้างอิง.....	103
บทที่ 6 ต้นไม้ตัดสินใจ.....	105
6.1 ทฤษฎีต้นไม้ตัดสินใจ.....	105
6.1.1 โครงสร้างต้นไม้ตัดสินใจ.....	106
6.1.2 การสร้างต้นไม้ตัดสินใจด้วยทฤษฎีสารสนเทศ.....	107
6.1.3 ตัวอย่างการคำนวณเพื่อสร้างต้นไม้ตัดสินใจ.....	109
6.1.4 การสร้างกฎจากต้นไม้ตัดสินใจ.....	113
6.2 การสร้างตัวแบบต้นไม้ตัดสินใจด้วยโปรแกรม WEKA.....	114
6.2.1 การจัดเตรียมเพิ่มข้อมูลซีเอสวี.....	114
6.2.2 การวิเคราะห์ข้อมูลด้วยโปรแกรม WEKA.....	116
6.3 การนำตัวแบบมาสร้างระบบการจำแนกประเภทข้อมูลด้วยโปรแกรม MS-Excel.....	122
6.4 การพัฒนาตัวแบบต้นไม้ตัดสินใจด้วยการเขียนโปรแกรมภาษาไพธอน.....	124
6.4.1 การพัฒนาจาก “กฎถ้าแล้ว”.....	124
6.4.2 การใช้ไลบรารี scikit-learn.....	125
สรุป.....	128
แบบฝึกหัดท้ายบทที่ 6.....	129

สารบัญ (ต่อ)

	หน้า
เฉลยแบบฝึกหัดท้ายบทที่ 6.....	130
เอกสารอ้างอิง.....	132
บทที่ 7 โครงข่ายประสาทเทียม.....	133
7.1 ทฤษฎีโครงข่ายประสาทเทียม.....	133
7.1.1 แนวคิดเบื้องต้นเกี่ยวกับโครงข่ายประสาทเทียม.....	133
7.1.2 กฎการเรียนรู้เพอร์เซปตรอน.....	134
7.1.3 ตัวอย่างการคำนวณและการใช้งานเพอร์เซปตรอน.....	136
7.2 การพัฒนาตัวแบบเพอร์เซปตรอนด้วยโปรแกรม MS-Excel.....	139
7.2.1 การเตรียมแผนงานสำหรับการเรียนรู้เพอร์เซปตรอน.....	139
7.2.2 การเขียนชุดคำสั่งด้วย VBA สำหรับการเรียนรู้เพอร์เซปตรอน.....	141
7.3 การพัฒนาตัวแบบเพอร์เซปตรอนด้วยการเขียนโปรแกรมภาษาไพธอน.....	144
7.3.1 การพัฒนาตามขั้นตอนวิธีการเรียนรู้เพอร์เซปตรอน.....	145
7.3.2 การสร้างตัวแบบเพอร์เซปตรอนด้วยไลบรารี scikit-learn.....	146
7.3.3 การสร้างตัวแบบเพอร์เซปตรอนหลายชั้น.....	148
สรุป.....	149
แบบฝึกหัดท้ายบทที่ 7.....	150
เฉลยแบบฝึกหัดท้ายบทที่ 7.....	151
เอกสารอ้างอิง.....	155
บทที่ 8 การจัดกลุ่มแบบค่าเฉลี่ย.....	157
8.1 ทฤษฎีการจัดกลุ่มแบบค่าเฉลี่ย.....	157
8.1.1 ขั้นตอนวิธีการจัดกลุ่มแบบค่าเฉลี่ย.....	157
8.1.2 ตัวอย่างการคำนวณและการใช้งานการจัดกลุ่มแบบค่าเฉลี่ย.....	158
8.2 การพัฒนาระบบการจัดกลุ่มแบบค่าเฉลี่ยด้วยโปรแกรม MS-Excel.....	162
8.2.1 การเตรียมแผนงานสำหรับการจัดกลุ่มแบบค่าเฉลี่ย.....	162
8.2.2 การเขียนชุดคำสั่งด้วย VBA สำหรับการจัดกลุ่มแบบค่าเฉลี่ย.....	165
8.3 การพัฒนาระบบการจัดกลุ่มแบบค่าเฉลี่ยด้วยการเขียนโปรแกรมภาษาไพธอน.....	167
8.3.1 การพัฒนาตามขั้นตอนวิธีการจัดกลุ่มแบบค่าเฉลี่ย.....	167

สารบัญ (ต่อ)

	หน้า
8.3.2 การสร้างตัวแบบการจัดกลุ่มข้อมูลแบบค่าเฉลี่ยเคด้วยไลบรารี scikit-learn.....	170
สรุป.....	172
แบบฝึกหัดท้ายบทที่ 8.....	172
เฉลยแบบฝึกหัดท้ายบทที่ 8.....	172
เอกสารอ้างอิง.....	177
บทที่ 9 การวัดประสิทธิภาพของตัวแบบ	179
9.1 การวัดประสิทธิภาพตัวแบบการถดถอย.....	179
9.1.1 ค่าเฉลี่ยความผิดพลาดกำลังสอง.....	179
9.1.2 รากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง.....	180
9.1.3 ค่าเฉลี่ยความผิดพลาดสัมบูรณ์.....	181
9.1.4 การเขียนชุดคำสั่งภาษาไพธอนสำหรับการวัดประสิทธิภาพของตัวแบบการถดถอย.....	181
9.2 การวัดประสิทธิภาพตัวแบบการจำแนกประเภท.....	182
9.2.1 เมตริกซ์ความสับสน.....	182
9.2.2 ค่าความถูกต้อง.....	184
9.2.3 ค่าความแม่นยำ.....	184
9.2.4 ค่าระลอก/ค่าความไว.....	185
9.2.5 ค่าความจำเพาะ.....	185
9.2.6 ค่าคะแนนเอฟหนึ่ง.....	185
9.2.7 การเขียนชุดคำสั่งภาษาไพธอนสำหรับการวัดประสิทธิภาพของตัวแบบการจำแนก ประเภท.....	186
9.3 การวัดประสิทธิภาพตัวแบบการจัดกลุ่ม.....	189
9.3.1 ผลรวมกำลังสองของค่าเบี่ยงเบนภายในกลุ่ม.....	187
9.3.2 ผลรวมกำลังสองของค่าเบี่ยงเบนระหว่างกลุ่ม.....	188
9.3.3 การเขียนชุดคำสั่งภาษาไพธอนสำหรับการวัดประสิทธิภาพของตัวแบบการจัดกลุ่ม.....	189
9.4 การแบ่งชุดข้อมูลสำหรับการทดสอบตัวแบบ.....	190
9.4.1 การแบ่งแยกชุดข้อมูลสอนและข้อมูลทดสอบ.....	190
9.4.2 การตรวจสอบความถูกต้องแบบไขว้เคกลุ่ม.....	191

สารบัญ (ต่อ)

	หน้า
9.4.3 การตรวจสอบความถูกต้องแบบไขว้ออกหนึ่ง.....	193
สรุป.....	194
แบบฝึกหัดท้ายบทที่ 9.....	194
เฉลยแบบฝึกหัดท้ายบทที่ 9.....	196
เอกสารอ้างอิง.....	199
บรรณานุกรม.....	201
บรรณานุกรม.....	207

สารบัญภาพ

ภาพที่	หน้า
1.1 แฟรงก์ โรเซนเบลตต์ ผู้พัฒนาสถาปัตยกรรม เพอร์เซปตรอน.....	3
1.2 อาเธอร์ ซามูเอล ผู้พัฒนาโปรแกรมการเรียนรู้การเล่นหมากฮอสด้วยการค้นหาแบบฮิวริสติก.....	4
1.3 อลัน ทัวริง ผู้เสนอวิธี “การทดสอบของทัวริง” (turing test).....	5
1.4 การทดสอบของทัวริง (turing test).....	5
1.5 จอห์น โรเจอร์ส เซียร์ล ผู้เสนอการทดลองทางความคิดที่ชื่อว่า “ห้องอักษรจีน” (Chinese room).....	6
1.6 การทดลองทางความคิดที่ชื่อว่า “ห้องอักษรจีน” (Chinese room).....	6
1.7 ตัวอย่างโปรแกรมแชทจีพีที (ChatGPT).....	7
1.8 ตัวอย่างการพยากรณ์ด้วยวิธีการถดถอยเชิงเส้น.....	9
1.9 ตัวอย่างการจำแนกประเภทข้อมูลด้วยวิธีต้นไม้ตัดสินใจ.....	10
1.10 ตัวอย่างการจัดกลุ่มข้อมูล.....	11
1.11 ตัวอย่างการหาความสัมพันธ์.....	12
1.12 การแข่งขันโกะ ระหว่างอัลฟาโกะ กับ ลี เซดอล แชมป์โกะโลกชาวเกาหลีใต้.....	13
1.13 การติดตามยานพาหนะในการจราจร.....	14
1.14 ตัวอย่างการตัดคำในการประมวลผลภาษาธรรมชาติ.....	16
1.15 หุ่นยนต์ขุดสวนของ จอห์น เดียร์ (John Deere's robot planter) ที่สามารถใส่ปุ๋ยและเมล็ดพันธุ์ได้อย่างแม่นยำ.....	17
2.1 นิยามปัญหาของเคปเนอร์และเทอร์โก.....	25
2.2 ตัวอย่างสถานะเริ่มต้น และสถานะสุดท้ายที่ต้องการของปัญหาปริศนา 8 ตัว.....	27
2.3 ปริภูมิสถานะของปัญหาปริศนา 8 ตัว.....	28
2.4 ปัญหาโลกของกล่อง.....	29
2.5 ปริภูมิสถานะของปัญหาโลกของกล่อง.....	29
2.6 ตัวอย่างการใช้ตัวดำเนินการสำหรับปัญหาปริศนา 8 ตัว.....	30
2.7 ตัวอย่างการใช้ตัวดำเนินการสำหรับปัญหาโลกของกล่อง.....	30
2.8 ภาพรวมของเทคนิคการค้นหาในศาสตร์ด้านปัญญาประดิษฐ์.....	31
2.9 รูปแบบที่เป็นไปได้ทั้งหมดของปัญหาโลกของกล่อง.....	32
2.10 ตัวอย่างการค้นหาตามแนวกว้างก่อนในปัญหาโลกของกล่อง.....	33
2.11 ตัวอย่างการค้นหาตามแนวลึกก่อนในปัญหาโลกของกล่อง.....	34
2.12 การหาค่าฟังก์ชันแมนฮัตตันสำหรับปัญหาปริศนา 8 ตัว.....	36

สารบัญภาพ (ต่อ)

ภาพที่	หน้า
2.13 ค่าฟังก์ชันแมนฮัตตันของสถานะต่าง ๆ ตามเส้นทางการค้นหาในปริภูมิสถานะของปัญหาปริศนา 8 ตัว.....	36
2.14 ค่าฟังก์ชัน h_1 ของสถานะต่าง ๆ ตามเส้นทางการค้นหาในปริภูมิสถานะของปัญหาโลกของกล่อง.....	37
2.15 ค่าฟังก์ชัน h_2 ของสถานะต่าง ๆ ตามเส้นทางการค้นหาในปริภูมิสถานะของปัญหาโลกของกล่อง.....	37
2.16 การค้นหาแบบป็นเขา.....	38
2.17 การค้นหาแบบจำลองการหลอม.....	39
2.18 ตัวอย่างการค้นหาแบบดีที่สุดก่อน.....	41
3.1 รูปแบบข้อมูลที่ใช้ในการวิเคราะห์การถดถอยเชิงเส้น.....	45
3.2 ความลาดชันของเส้นตามค่า a	46
3.3 ตัวอย่างการสร้างตัวแบบการถดถอยเชิงเส้นอย่างง่าย.....	47
3.4 กราฟเปรียบเทียบค่าความคลาดเคลื่อนเทียบกับค่า a	48
3.5 ตำแหน่งค่า a และ b ที่ทำให้ค่าความคลาดเคลื่อนน้อยที่สุด.....	48
3.6 คิวอาร์โค้ดวิดีโอที่แสดงการอธิบายเนื้อหาเรื่องการถดถอยเชิงเส้น.....	53
3.7 การออกแบบหน้าจอ MS-Excel สำหรับเตรียมสร้างตัวแบบการถดถอยเชิงเส้น.....	53
3.8 เลือกช่วงข้อมูลที่จะนำมาวิเคราะห์สร้างตัวแบบ.....	55
3.9 เมนูแทรกแผนภูมิกระจาย.....	55
3.10 แผนภูมิกระจาย.....	55
3.11 เมนูลดการเพิ่มเส้นแนวโน้ม.....	56
3.12 หน้าต่างการจัดรูปแบบเส้นแนวโน้ม.....	56
3.13 การแสดงตัวแบบการถดถอยเชิงเส้น.....	57
3.14 คิวอาร์โค้ดวิดีโอที่แสดงขั้นตอนการพัฒนาตัวแบบการถดถอยเชิงเส้นด้วยโปรแกรม MS-Excel.....	57
4.1 แนวคิดเบื้องต้นของวิธีเพื่อนบ้านใกล้ที่สุดเค.....	65
4.2 คิวอาร์โค้ดวิดีโอที่แสดงการอธิบายเนื้อหาเรื่องเพื่อนบ้านใกล้ที่สุดเค.....	70
4.3 การจัดเตรียมข้อมูลลงไฟล์ Excel สำหรับการวิเคราะห์ด้วยวิธีเพื่อนบ้านใกล้ที่สุดเค.....	71
4.4 การใส่สูตรคำนวณระยะห่างยูคลิเดียน.....	71
4.5 การใส่สูตรคำนวณเพื่อหาลำดับที่ของความคล้ายคลึง.....	72
4.6 การใส่สูตรคำนวณเพื่อหาคำตอบของตัวอย่างที่มีลักษณะคล้ายคลึงกับข้อมูลสอบถามมากที่สุดเคตัว.....	73

สารบัญภาพ (ต่อ)

ภาพที่	หน้า
4.7 การใส่สูตรคำนวณเพื่อหาน้ำหนักของคะแนนเสียงตามความคล้ายคลึง.....	74
4.8 การใส่สูตรคำนวณเพื่อหาค่าผลรวมคะแนนเสียงของแต่ละคำตอบ.....	75
4.9 การใส่สูตรคำนวณเพื่อหาค่าคะแนนเสียงรวมมากที่สุด.....	75
4.10 การใส่สูตรคำนวณเพื่อแสดงคำตอบเป็นผลการทำนาย.....	76
4.11 คิวอาร์โค้ดวีดิทัศน์แสดงขั้นตอนการพัฒนารีวิวเพื่อนบ้านใกล้ที่สุดได้ด้วยโปรแกรม MS-Excel.....	76
5.1 คิวอาร์โค้ดวีดิทัศน์แสดงการอธิบายเนื้อหาเรื่องเบย์อย่างง่าย.....	89
5.2 การจัดเตรียมข้อมูลลงไฟล์ Excel สำหรับการวิเคราะห์ด้วยวิธีเบย์อย่างง่าย.....	90
5.3 ผลลัพธ์จากการคำนวณหาเหตุการณ์ตามเงื่อนไขที่กำหนด.....	91
5.4 ผลลัพธ์จากการคำนวณหาค่าความน่าจะเป็นของเหตุการณ์.....	92
5.5 ผลลัพธ์จากการคำนวณหาค่าความน่าจะเป็นของการเกิดคลาสต่าง ๆ เมื่อรู้ค่าคุณลักษณะของข้อมูล.....	93
5.6 ผลลัพธ์จากการคำนวณหาคำตอบที่ได้จากการทำนาย.....	94
5.7 คิวอาร์โค้ดวีดิทัศน์แสดงขั้นตอนการพัฒนาระบบการวิเคราะห์ข้อมูลแบบเบย์อย่างง่ายด้วยโปรแกรม MS-Excel.....	94
6.1 โครงสร้างต้นไม้ตัดสินใจโดยใช้คุณลักษณะ Hair เป็นบัพแรก.....	106
6.2 ตัวอย่างต้นไม้ตัดสินใจสำหรับข้อมูลศึกษาปัจจัยที่มีผลต่อการอาบแดด.....	107
6.3 ตัวอย่างต้นไม้ตัดสินใจสำหรับข้อมูลศึกษาปัจจัยที่มีผลต่อการอาบแดดแบบอื่นที่เป็นไปได้.....	107
6.4 ชุดข้อมูล T ที่จะนำมาคำนวณหาค่าเอนโทรปี.....	108
6.5 การนำคุณลักษณะทั้งหมดมาทดสอบสร้างเป็นบัพแรก.....	110
6.6 การสร้างบัพทดสอบสำหรับกิ่งที่ยังไม่สามารถจำแนกข้อมูลได้อย่างสมบูรณ์ของต้นไม้ Hair.....	112
6.7 คิวอาร์โค้ดวีดิทัศน์แสดงการอธิบายเนื้อหาเรื่องต้นไม้ตัดสินใจ.....	113
6.8 การจัดเตรียมข้อมูลลงไฟล์ Excel.....	114
6.9 การเรียกใช้คำสั่งบันทึกเป็นในโปรแกรม MS-Excel.....	115
6.10 โครงสร้างแฟ้มข้อมูลซีเอสวี.....	115
6.11 หน้าต่างโปรแกรม WEKA.....	116
6.12 การนำเข้าแฟ้มข้อมูลซีเอสวีเข้าสู่โปรแกรม WEKA.....	117
6.13 รายละเอียดข้อมูลที่นำเข้าไปในโปรแกรม WEKA.....	117
6.14 การนำคุณลักษณะที่ไม่ต้องการใช้ในการประมวลผลออก.....	118

สารบัญภาพ (ต่อ)

ภาพที่	หน้า
6.15 หน้าต่าง Classify ในโปรแกรม WEKA.....	119
6.16 การเลือกขั้นตอนวิธีสำหรับการสร้างตัวแบบในโปรแกรม WEKA.....	119
6.17 การกำหนดตัวเลือกสำหรับการสร้างตัวแบบในโปรแกรม WEKA.....	120
6.18 กฎถ้าแล้ว ที่ได้จากการวิเคราะห์ด้วยโปรแกรม WEKA.....	121
6.19 วิธีการเรียกดูโครงสร้างต้นไม้ตัดสินใจในโปรแกรม WEKA.....	121
6.20 โครงสร้างต้นไม้ตัดสินใจที่ได้ในโปรแกรม WEKA.....	122
6.21 การนำตัวแบบต้นไม้ตัดสินใจมาสร้างระบบการจำแนกประเภทข้อมูลด้วยโปรแกรม MS-Excel.....	123
6.22 คิวอาร์โค้ดวิทัศน์แสดงขั้นตอนการพัฒนากระบวนการวิเคราะห์ข้อมูลด้วยวิธีต้นไม้ตัดสินใจโดยใช้โปรแกรม WEKA และ Excel.....	124
7.1 โครงสร้างเซลล์ประสาท.....	134
7.2 โครงสร้างของเพอร์เซปตรอน.....	134
7.3 ตัวอย่างโครงสร้างเพอร์เซปตรอนหลายชั้น.....	134
7.4 ฟังก์ชันกระตุ้นแบบต่าง ๆ.....	135
7.5 โครงสร้างเพอร์เซปตรอนสำหรับการเรียนรู้แนวคิดของตรรก “และ”.....	137
7.6 คิวอาร์โค้ดวิทัศน์แสดงการอธิบายเนื้อหาเรื่องโครงข่ายประสาทเทียม.....	139
7.7 การจัดเตรียมข้อมูลลงไฟล์ Excel สำหรับการวิเคราะห์ด้วยเพอร์เซปตรอน.....	140
7.8 การเรียกเมนู “Customize the Ribbon...”.....	140
7.9 หน้าต่าง Excel Options.....	140
7.10 เมนู “Developer” ที่ปรากฏบนแถบเมนู.....	141
7.11 ขั้นตอนการสร้างปุ่มคำสั่ง.....	141
7.12 หน้าต่างการเขียนชุดคำสั่ง VBA.....	142
7.13 การยกเลิก Design Mode.....	144
7.14 ผลลัพธ์ที่ได้หลังจากเรียกใช้ชุดคำสั่ง VBA.....	144
7.15 คิวอาร์โค้ดวิทัศน์แสดงขั้นตอนการพัฒนาตัวแบบเพอร์เซปตรอนด้วยโปรแกรม MS-Excel.....	144
7.16 โครงสร้างเพอร์เซปตรอนหลายชั้นที่ได้จากชุดคำสั่งที่กำหนดค่า hidden_layer_sizes=(5,5).....	148
8.1 ผังงานขั้นตอนวิธีการจัดกลุ่มแบบค่าเฉลี่ย.....	158
8.2 คิวอาร์โค้ดวิทัศน์แสดงการอธิบายเนื้อหาเรื่องการจัดกลุ่มแบบค่าเฉลี่ย.....	162

สารบัญภาพ (ต่อ)

ภาพที่	หน้า
8.3 การเตรียมแผ่นงานสำหรับการจัดกลุ่มแบบค่าเฉลี่ยเค.....	163
8.4 ผลลัพธ์ของกระบวนการเรียนรู้จากแผ่นงานที่ได้จากการใส่สูตรคำนวณ.....	165
8.5 ชุดคำสั่ง VBA สำหรับทำการคัดลอกข้อมูลสำหรับการจัดกลุ่มแบบค่าเฉลี่ยเค.....	165
8.6 ผลการจัดกลุ่มแบบค่าเฉลี่ยเคเมื่อกดปุ่ม CommandButton1.....	167
8.7 คิวอาร์โค้ดวิดิทัศน์แสดงขั้นตอนการพัฒนาระบบการจัดกลุ่มแบบค่าเฉลี่ยเคด้วยโปรแกรม MS-Excel.....	167
9.1 กราฟเปรียบเทียบค่าคำตอบที่แท้จริงกับผลการทำนายของตัวแบบการถดถอยเชิงเส้น.....	180
9.2 ตัวอย่างตารางเมทริกซ์ความสับสนขนาด 2x2.....	183
9.3 การแบ่งแยกชุดข้อมูลสอนและข้อมูลทดสอบ.....	190
9.4 การแบ่งชุดข้อมูลสำหรับการตรวจสอบความถูกต้องแบบไขว้เคกลุ่ม.....	192
9.5 ขั้นตอนวิธีการตรวจสอบความถูกต้องแบบไขว้เคกลุ่ม.....	192

สารบัญตาราง

ตารางที่	หน้า
3.1 ตัวอย่างข้อมูลขนาดและน้ำหนักลูกหมู.....	47
3.2 สูตรคำนวณที่ใส่ในเซลล์ต่าง ๆ.....	54
4.1 ตัวอย่างชุดข้อมูลดอกไอริส.....	68
4.2 การคำนวณความคล้ายคลึงระหว่างข้อมูลทำนายกับข้อมูลสอนด้วยฟังก์ชันระยะห่างยูคลิดีเนียน.....	69
5.1 ชุดข้อมูลการเล่นกอล์ฟของโปรกอล์ฟคนหนึ่ง.....	88
5.2 สูตรคำนวณจำนวนเหตุการณ์ตามเงื่อนไขที่กำหนด.....	91
5.3 สูตรคำนวณหาค่าความน่าจะเป็นของเหตุการณ์.....	92
6.1 ชุดข้อมูลศึกษาปัจจัยที่มีผลต่อการอาบแดด.....	106
7.1 ชุดข้อมูลสอนแนวคิดของตรรก “และ”	136
7.2 การเรียนรู้ปัญหาแนวคิดของตรรก “และ”	138
8.1 ตัวอย่างชุดข้อมูลดอกไอริสสำหรับการจัดกลุ่มแบบค่าเฉลี่ยเค.....	159
8.2 การกำหนดค่าเซนทรอยด์เริ่มต้น.....	159
8.3 ผลการคำนวณหาค่าความเป็นสมาชิกของกลุ่มย่อยรอบที่ 1.....	160
8.4 การคำนวณหาค่าเซนทรอยด์ 1 ใหม่ ในรอบที่ 1.....	160
8.5 การคำนวณหาค่าเซนทรอยด์ 2 ใหม่ ในรอบที่ 1.....	161
8.6 ผลการคำนวณหาค่าความเป็นสมาชิกของกลุ่มย่อยรอบที่ 2.....	161
8.7 ค่าเซนทรอยด์ใหม่ที่คำนวณได้ในรอบที่ 2.....	162
9.1 ตัวอย่างข้อมูลเปรียบเทียบค่าคำตอบที่แท้จริงกับผลการทำนายของตัวแบบการถดถอยเชิงเส้น.....	180
9.2 ตัวอย่างข้อมูลเปรียบเทียบค่าคำตอบที่แท้จริงกับผลการทำนายของตัวแบบการจำแนกประเภท.....	183
9.3 ตัวอย่างผลลัพธ์จากการจัดกลุ่มข้อมูลและค่าเซนทรอยด์ที่ได้.....	187
9.4 ค่าเฉลี่ยรวม (ค่า C) ของข้อมูลทั้งหมด.....	188

บทที่ 1

บทนำ

การเรียนรู้ของเครื่อง (machine learning: ML) เป็นสาขาย่อยของปัญญาประดิษฐ์ (artificial intelligence: AI) ที่มีเป้าหมายในการพัฒนาเทคนิคและวิธีการให้คอมพิวเตอร์สามารถเรียนรู้จากข้อมูลหรือสภาพแวดล้อม เพื่อปรับปรุงประสิทธิภาพและความถูกต้องแม่นยำในการทำงานต่าง ๆ ที่ต้องการ ในบทนี้จะกล่าวถึงความรู้ทั่วไปเกี่ยวกับการเรียนรู้ของเครื่องเพื่อเป็นข้อมูลพื้นฐานก่อนจะได้ศึกษาในรายละเอียดเกี่ยวกับพื้นฐานการคำนวณและการนำไปประยุกต์ใช้ในบทต่อ ๆ ไป

1.1 ความหมายและความสำคัญของการเรียนรู้ของเครื่อง

ในหัวข้อนี้จะกล่าวถึงความหมายของการเรียนรู้ของเครื่อง รวมถึงแสดงให้เห็นถึงความสำคัญและผลกระทบที่มีต่อการใช้ชีวิตของทุกคนในโลกยุคปัจจุบัน ดังนี้

1.1.1 ความหมายของการเรียนรู้ของเครื่อง

การเรียนรู้ของเครื่อง เป็นศาสตร์ที่ถูกกล่าวถึงและถูกนำไปใช้อย่างกว้างขวางในปัจจุบัน จึงทำให้เกิดนิยามความหมายที่หลากหลาย ยกตัวอย่างได้ดังนี้

- ทอม มิตเชลล์ (Mitchell, T., 1997) กล่าวว่า “การเรียนรู้ของเครื่อง คือ การศึกษาขั้นตอนวิธีการทางคอมพิวเตอร์ที่สามารถพัฒนาและปรับปรุงได้โดยอัตโนมัติโดยอาศัยข้อมูลและประสบการณ์ที่ได้รับ”
- อิสซัม เอล นากา และ มาร์ติน เจ. เมอร์ฟี (El Naqa, I. and Murphy, M.J., 2015) กล่าวว่า “การเรียนรู้ของเครื่อง เป็นแขนงหนึ่งของขั้นตอนวิธีการคำนวณที่ถูกออกแบบมาเพื่อเลียนแบบสติปัญญาของมนุษย์การเรียนรู้จากสภาพแวดล้อมโดยรอบ”
- โมห์รี เมห์รียาร์ และคณะ (Mehryar, M. และคณะ, 2018) กล่าวว่า “การเรียนรู้ของเครื่องสามารถนิยามอย่างกว้าง ๆ ได้ว่า เป็นวิธีการคำนวณที่ใช้ประสบการณ์เพื่อปรับปรุงประสิทธิภาพหรือเพื่อคาดการณ์ที่มีความถูกต้องแม่นยำ”
- แบตตา มาเฮช (Mahesh, B., 2020) กล่าวว่า “การเรียนรู้ของเครื่อง คือ การศึกษาทางวิทยาศาสตร์เกี่ยวกับขั้นตอนวิธีและตัวแบบทางสถิติที่ระบบคอมพิวเตอร์ใช้ดำเนินการโดยไม่จำเป็นต้องระบุความต้องการที่ชัดเจน”

จากนิยามข้างต้น สามารถสรุปได้ว่า “การเรียนรู้ของเครื่อง” เป็นการศึกษาขั้นตอนวิธีการทางคอมพิวเตอร์เพื่อให้คอมพิวเตอร์สามารถเรียนรู้ได้โดยอัตโนมัติจากประสบการณ์และข้อมูลที่ได้รับจาก

สภาพแวดล้อม สามารถนำมาใช้ในการปรับปรุงประสิทธิภาพหรือคาดการณ์ผลลัพธ์ให้มีความถูกต้องแม่นยำมากขึ้นโดยไม่ต้องโปรแกรมล่วงหน้า

1.1.2 ความสำคัญของการเรียนรู้ของเครื่อง

ปัจจุบัน การเรียนรู้ของเครื่องถูกนำไปประยุกต์ใช้ในการพัฒนาและปรับปรุงงานต่าง ๆ อย่างกว้างขวาง ตัวอย่างเช่น

- *ด้านการเงินการธนาคาร*: ธนาคารและสถาบันการเงินสามารถวิเคราะห์ความเสี่ยงด้านต่าง ๆ ที่อาจเกิดขึ้น เช่น ความเสี่ยงต่อการอนุมัติสินเชื่อให้กับลูกค้า การทำนายความผันผวนของตลาดการเงิน วิเคราะห์ปัจจัยสำคัญและศึกษาความเชื่อมโยงของปัจจัยต่าง ๆ ที่มีผลต่อสภาพคล่อง เป็นต้น เพื่อนำมาใช้สำหรับการวางแผนบริหารจัดการความเสี่ยงต่าง ๆ เหล่านี้ที่อาจเกิดขึ้นได้อย่างมีประสิทธิภาพต่อไป (Leo, M. และคณะ, 2019)
- *ด้านการค้าอิเล็กทรอนิกส์*: อุตสาหกรรมการค้าอิเล็กทรอนิกส์สามารถนำมาใช้ในการวิเคราะห์ความนิยมของสินค้าและพฤติกรรมการซื้อของลูกค้า เพื่อการแนะนำสินค้าและโปรโมชั่นที่เหมาะสมให้กับลูกค้า และมีประสิทธิภาพในการเพิ่มยอดขายให้กับเจ้าของกิจการ (Rabhi, F.A. และคณะ, 2020)
- *ด้านสุขภาพ*: คลินิกและโรงพยาบาลสามารถนำมาช่วยในการวินิจฉัยโรค วางแผนการรักษา การจัดการทรัพยากรทางสาธารณสุข รวมถึงการวิเคราะห์ข้อมูลผู้ป่วยเพื่อคาดการณ์ความเสี่ยงของการเกิดภาวะแทรกซ้อนและการหาแนวทางการป้องกัน (Ghassemi, M. และคณะ, 2020)
- *ด้านกระบวนการผลิต*: อุตสาหกรรมการผลิตสามารถนำมาใช้ปรับปรุงกระบวนการผลิต การควบคุมคุณภาพสินค้า รวมถึงการบำรุงรักษาอุปกรณ์เครื่องจักร ซึ่งนำไปสู่การเพิ่มประสิทธิภาพการผลิตและลดค่าใช้จ่าย (Kang, Z. และคณะ, 2020)
- *ด้านการขนส่งและโลจิสติกส์*: อุตสาหกรรมการขนส่งและโลจิสติกส์สามารถนำมาช่วยในการคาดการณ์ความต้องการการขนส่ง วางแผนเส้นทางการขนส่ง และปรับปรุงการจัดการคลังสินค้า เพื่อให้การขนส่งมีประสิทธิภาพมากขึ้น (Chaudhary, K. และคณะ, 2018)
- *ด้านการประกันภัย*: บริษัทประกันภัยสามารถนำมาใช้เพื่อการประเมินความเสี่ยง การกำหนดราคากรมธรรม์ และตรวจสอบการอ้างสิทธิ์ประกันภัยได้อย่างรวดเร็วและถูกต้องแม่นยำ (Hanafy, M. and Ming, R., 2021)

ความสำคัญของการเรียนรู้ของเครื่องในอุตสาหกรรมต่าง ๆ ที่กล่าวมาข้างต้นไม่ว่าจะเป็นทางการเงิน การค้าอิเล็กทรอนิกส์ สุขภาพ การผลิต การขนส่งและโลจิสติกส์ การประกันภัย หรือเรื่องอื่น ๆ ที่นอกเหนือจากนี้ ทำให้เห็นได้ว่าการเรียนรู้ของเครื่องกลายเป็นเทคโนโลยีที่มีความสำคัญต่อการพัฒนาทาง