



เริ่มต้นกับ Machine Learning

Getting Started with Machine Learning

จันทรเจ้า มงคลนาวัน



เริ่มต้นกับ Machine Learning

เริ่มต้นกับ Machine Learning

จันทร์เจ้า มงคลนาวิน

199.-

ชื่อหนังสือ : เริ่มต้นกับ Machine Learning

ชื่อผู้แต่ง : จันท์เจ้า มงคลนาวิน

ปกและภาพประกอบปก : ธนัชชา วงศ์หาญ

ข้อมูลทางบรรณานุกรมของหอสมุดแห่งชาติ

National Library of Thailand Cataloging in Publication Data

จันท์เจ้า มงคลนาวิน.

เริ่มต้นกับ Machine Learning.-- กรุงเทพฯ : [ม.ป.พ.], 2564.

210 หน้า.

1. โปรแกรมคอมพิวเตอร์. 2. ไพธอน (ภาษาคอมพิวเตอร์). I. ชื่อเรื่อง.

005.3

ISBN (e-book) 978-616-586-683-5

สงวนลิขสิทธิ์ตามพระราชบัญญัติลิขสิทธิ์ (ฉบับเพิ่มเติม) พ.ศ. 2558

ราคา 199 บาท

จัดพิมพ์โดย ผู้ช่วยศาสตราจารย์ ดร.จันท์เจ้า มงคลนาวิน

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ถนนพญาไท แขวงวังใหม่ เขตปทุมวัน กรุงเทพฯ 10330

โทร. 0-2218-5674 www.cbs.chula.ac.th

e-mail: janjao@cbs.chula.ac.th

CUP6501-014(D)

สั่งซื้อได้ที่ ศูนย์หนังสือจุฬาลงกรณ์มหาวิทยาลัย

ถนนพญาไท เขตปทุมวัน กรุงเทพฯ 10330

<http://www.chulabook.com>

โทร. 086 323 3703 4, 02 255 4433

customer@cubook.chula.ac.th, info@cubook.chula.ac.th

Apps: CU-eBook Store

ในช่วงเวลาที่ผ่านมา ผู้เขียนรับผิดชอบการสอนรายวิชาการค้นพบองค์ความรู้และการทำเหมืองข้อมูล ซึ่งเป็นวิชาที่เกี่ยวข้องกับการเรียนรู้ของเครื่อง หรือ Machine learning และได้เห็นวิวัฒนาการที่รวดเร็วของศาสตร์ โดยเฉพาะอย่างยิ่งตั้งแต่ ค.ศ. 2015 ที่มีการตื่นตัวในด้านการเรียนรู้เชิงลึก (Deep learning) ซึ่งส่งผลให้การเรียนรู้ของเครื่องได้ถูกนำไปประยุกต์ด้านธุรกิจอย่างเป็นรูปธรรมมากยิ่งขึ้น

ด้วยเหตุผลดังกล่าว การเรียนรู้ของเครื่องจึงได้รับความสนใจจากบุคคลทั่วไปในวงกว้าง ไม่ว่าจะเป็น ผู้บริหาร บุคลากรภาคธุรกิจ เช่น นักบัญชี หรือนักการตลาด และบุคลากรทางเทคนิค เช่น นักวิเคราะห์ระบบ หรือโปรแกรมเมอร์ อย่างไรก็ตาม ผู้เขียนเชื่อว่า เพื่อให้ตอบวัตถุประสงค์การใช้งาน ผู้สนใจแต่ละกลุ่มอาจไม่จำเป็นต้องมีความรู้และความเข้าใจเกี่ยวกับการเรียนรู้ของเครื่องในระดับเดียวกัน

เนื้อหาเกี่ยวกับการเรียนรู้ของเครื่องรวมถึงตัวอย่างโปรแกรมมีแพร่หลายทั่วไปในอินเทอร์เน็ต จากประสบการณ์สอนของผู้เขียน พบว่า ผู้ประยุกต์บางส่วนมีความเข้าใจว่าโปรแกรมที่ทำงานได้คือโปรแกรมที่สร้างผลลัพธ์ที่สามารถนำไปใช้งานได้ ความเข้าใจดังกล่าวไม่ผิดทั้งหมดเพราะโปรแกรมที่ทำงานได้เป็นเงื่อนไขเบื้องต้นของการสร้างผลลัพธ์ แต่การที่จะสร้างผลลัพธ์ที่สามารถใช้ตอบคำถามที่ตั้งไว้เกี่ยวข้องกับหลายปัจจัย ไม่ว่าจะเป็น ความรู้ความเข้าใจในข้อมูล ความเข้าใจในเทคนิคที่เลือกใช้ รวมถึงการเลือกใช้วิธีวัดประสิทธิผลของผลลัพธ์ที่สอดคล้องกับวัตถุประสงค์ของการวิเคราะห์

หนังสือเล่มนี้จึงมีวัตถุประสงค์เพื่อให้ผู้อ่านได้รับความรู้เบื้องต้นของศาสตร์ด้านการเรียนรู้ของเครื่องผ่านตัวอย่างเทคนิคที่ได้รับความนิยมและตัวอย่างการประยุกต์เทคนิค

เหล่านั้น โดยสามารถนำไปทดลองปฏิบัติตามได้ ตลอดจนทราบวิธีวัดประสิทธิผลของผลลัพธ์ของแต่ละเทคนิคและวิธีแปลผล

ผู้เขียนขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร.วิรัตน์ จาริวงศ์ไพบูลย์ และผู้ช่วยศาสตราจารย์ ดร.จรัสศรี รุ่งรัตนอุบล เป็นอย่างสูงที่ให้ความอนุเคราะห์ในการตรวจสอบคุณภาพเนื้อหา และใคร่ขอขอบคุณ อาจารย์ชัยสุทธิ์ จงถาวรวิทยา และ ดร.ชูชาติ หฤไชยะศักดิ์ สำหรับความเห็นและคำแนะนำที่ทำให้หนังสือสมบูรณ์มากยิ่งขึ้น หนังสือเล่มนี้จะไม่เสร็จสมบูรณ์หากปราศจากกำลังใจและการสนับสนุนจากครอบครัวและบุคคลใกล้ชิด โดยเฉพาะอย่างยิ่ง คุณพ่อคุณแม่ คุณมงคล – คุณจุไรรัตน์ สุติวราพันธ์ ผู้ซึ่งให้ความรัก ความเข้าใจ และสนับสนุนผู้เขียนในทุกด้านเสมอมา

หากผู้อ่านมีข้อคิดเห็นหรือข้อเสนอแนะเกี่ยวกับเนื้อหาในหนังสือเล่มนี้ สามารถส่งอีเมลมาที่ janjao.mongkolnavin@gmail.com

จันทร์เจ้า มงคลนาวิน

กันยายน 2564

คำแนะนำการใช้หนังสือ

ใครควรอ่านหนังสือเล่มนี้

หนังสือเล่มนี้มีวัตถุประสงค์ให้ผู้อ่านได้ความรู้เบื้องต้นของศาสตร์ด้านการเรียนรู้ของเครื่อง ผ่านตัวอย่างเทคนิคที่ได้รับความนิยม และได้เห็นตัวอย่างการประยุกต์เทคนิคเหล่านั้นผ่านเครื่องมือที่สามารถนำไปทดลองปฏิบัติตามได้ ตลอดจนทราบวิธีการวัดประสิทธิผลของผลลัพธ์ของแต่ละเทคนิคและวิธีแปลผล ดังนั้น หนังสือเล่มนี้จึงเหมาะสมกับบุคคลทั่วไปที่เริ่มต้นศึกษาเกี่ยวกับการเรียนรู้ของเครื่อง ไม่ว่าจะเป็น นักเรียน นักศึกษา หรือผู้สนใจการเรียนรู้ของเครื่องไปประยุกต์กับงานในระดับต่าง ๆ โดยที่ไม่ต้องมีความรู้เกี่ยวกับการเรียนรู้ของเครื่องมาก่อน และไม่จำเป็นต้องมีความรู้ทางคณิตศาสตร์ขั้นสูง โดยหากผู้อ่านเข้าใจการทำงานของภาษาไพธอน (Python) ในเบื้องต้นก็จะเป็นประโยชน์ต่อการทำความเข้าใจการทำงานของโปรแกรมตัวอย่าง สำหรับผู้อ่านที่ประสงค์จะทำความเข้าใจเกี่ยวกับการเรียนรู้ของเครื่อง แต่ไม่ได้ประสงค์ที่จะพัฒนาโปรแกรมด้วยตนเอง อาจข้าม บทที่ 6 บทที่ 8 และบทที่ 10 ซึ่งเป็นตัวอย่างโปรแกรมที่ประยุกต์เทคนิคที่กล่าวถึงในหนังสือได้

โครงสร้างของหนังสือ

หนังสือเล่มนี้ ประกอบด้วย 5 ส่วน คือ

- ส่วนที่ 1 ภาพรวมของการเรียนรู้ของเครื่อง
- ส่วนที่ 2 การสร้างตัวแบบทำนาย (Predictive Modeling)
- ส่วนที่ 3 การแบ่งกลุ่มข้อมูล (Data Segmentation)
- ส่วนที่ 4 การวิเคราะห์ความสัมพันธ์ (Association Analysis)
- ส่วนที่ 5 ส่วนอื่นของการเรียนรู้ของเครื่อง

โดยแต่ละส่วนประกอบด้วยบทที่มีเนื้อหาเกี่ยวข้องกันและบทที่เป็นตัวอย่างโปรแกรมที่ประยุกต์เทคนิคที่เกี่ยวข้องกับเนื้อหาในแต่ละส่วน ผู้อ่านที่สนใจสามารถดาวน์โหลดตัวอย่างโปรแกรมในรูปแบบ Jupyter notebook ไปทดลองทำงานได้จากยูอาร์แอลภายใต้ [gsw-ml.site/](https://www.gsw-ml.site/) ที่ระบุไว้ในส่วนต้นของบทที่เป็นตัวอย่างโปรแกรม

เครื่องมือที่ใช้

เพื่อให้ผู้อ่านสามารถทดลองโปรแกรมตัวอย่างได้สะดวก เครื่องมือที่ใช้ในโปรแกรมตัวอย่างเป็นเครื่องมือที่สามารถใช้งานเบื้องต้นโดยไม่มีค่าใช้จ่าย ได้แก่

ภาษาไพธอน

ภาษาไพธอน (Python programming language) หรือที่มักเรียกกันว่า ไพธอน เป็นภาษาระดับสูงซึ่งสร้างโดย Guido van Rossum เปิดให้ใช้งานใน ค.ศ. 1991 โดยมีมุ่งเน้นให้เป็นภาษาคอมพิวเตอร์ที่ผู้ใช้โปรแกรมสามารถอ่านชุดคำสั่งได้ง่าย ไพธอนรองรับการเขียนโปรแกรมหลายรูปแบบ เช่น การเขียนโปรแกรมแบบมีโครงสร้าง (Structured programming) การเขียนโปรแกรมเชิงวัตถุ (Object-oriented programming) และการเขียนโปรแกรมเชิงฟังก์ชัน (Functional programming) ไพธอนเป็นภาษาคอมพิวเตอร์ที่มีไลบรารีมาตรฐานจำนวนมาก เช่น ไลบรารีสำหรับการจัดการกับโครงสร้างข้อมูล เช่น pandas ไลบรารีสำหรับการคำนวณทางคณิตศาสตร์ เช่น NumPy ไลบรารีสำหรับนำเสนอข้อมูล เช่น Matplotlib และไลบรารีสำหรับการเรียนรู้ของเครื่อง เช่น scikit-learn จึงมีความครอบคลุมสำหรับงานด้านการวิเคราะห์ข้อมูล และได้รับความนิยมนำมาใช้งานด้านการเรียนรู้ของเครื่องในปัจจุบัน

The Python Software Foundation (PSF) ซึ่งเป็นองค์กรไม่แสวงผลกำไร ทำหน้าที่ส่งเสริมการพัฒนาเทคโนโลยีโอเพนซอร์สที่เกี่ยวข้องกับไพธอน และสนับสนุนการขยายชุมชนนักพัฒนาโปรแกรมไพธอน ผู้สนใจสามารถศึกษาข้อมูลเกี่ยวกับไพธอนและชุมชนนักพัฒนาโปรแกรมไพธอนได้ที่ <https://www.python.org/>

ผู้สนใจศึกษาการเขียนโปรแกรมภาษาไพธอน สามารถศึกษาเพิ่มเติมได้จากแหล่งข้อมูลต่าง ๆ ในอินเทอร์เน็ต สำหรับแหล่งข้อมูลที่เป็นภาษาไทย สามารถศึกษาจาก “Python ๑๐๑”

ซึ่งจัดทำโดย กิตติภณ พละการ, กิตติภพ พละการ, สมชาย ประสิทธิ์จูตระกูล และ สุกรี สินธุภิญโญ คณาจารย์ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย (<https://www.cp.eng.chula.ac.th/books/python101/>)

scikit-learn

scikit-learn (Pedregosa F. , et al., 2011) เป็นไลบรารีโอเพนซอร์ส (open source library) สำหรับการเรียนรู้ของเครื่องทั้งในส่วนของการเรียนรู้แบบมีผู้สอน (Supervised learning) และการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) รวมถึงเครื่องมือต่าง ๆ สำหรับการปรับตัวแบบให้เหมาะสม (Model fitting) การทำนาย (Model predicting) การเลือกและการประเมินตัวแบบ (Model selection and evaluation) และโปรแกรมอรรถประโยชน์ (Utility programs) ที่เกี่ยวข้องอื่น ๆ ผู้สนใจสามารถศึกษาข้อมูลเกี่ยวกับ scikit-learn ได้ที่ <https://scikit-learn.org/>

Jupyter Notebook

Jupyter Notebook เป็นโปรแกรมประยุกต์ทางเว็บที่มีลักษณะโอเพนซอร์สที่ใช้สร้างเอกสารที่เป็นซอร์สโค้ด (source code) สมการ กราฟรูปแบบต่าง ๆ และคำอธิบายประกอบ (comments) จัดเป็น IDE (Integrated Development Environment) หรือเครื่องมือพัฒนาโปรแกรมตัวหนึ่ง ซึ่งสามารถสั่งให้ซอร์สโค้ดทำงานเฉพาะส่วนที่ต้องการ โดยไม่จำเป็นต้องทำงานตามลำดับที่ปรากฏในเอกสาร ผลลัพธ์ที่ได้จะถูกแสดงไว้ใน Jupyter Notebook เดียวกับซอร์สโค้ด ซึ่งเหมาะสำหรับใช้พัฒนาโปรแกรมด้านการเรียนรู้ของเครื่อง ซึ่งประกอบด้วยกระบวนการย่อย โดยที่ซอร์สโค้ดของแต่ละกระบวนการย่อยสามารถทำงานแยกส่วนและบ่อยครั้งต้องสอบทานผลลัพธ์เพื่อใช้ประกอบการตัดสินใจที่เกี่ยวข้อง เช่น เปรียบเทียบผลลัพธ์จากเทคนิคเดียวกันที่ได้จากการกำหนดไฮเพอร์พารามิเตอร์ (Hyperparameter) ที่แตกต่างกัน เพื่อเลือกไฮเพอร์พารามิเตอร์ที่เหมาะสมที่สุด

Google Colaboratory

Google Colaboratory หรือเรียกสั้น ๆ ว่า "Google Colab" เป็นเครื่องมือสำหรับการศึกษาและวิจัยทางด้านการเรียนรู้ของเครื่อง ที่ไม่ต้องเสียค่าใช้จ่ายในการใช้งานเบื้องต้น ให้บริการ Jupyter notebook environment ผ่านเว็บเบราว์เซอร์ (web browser) โดยที่ไม่ต้องติดตั้ง สามารถจัดเก็บ Jupyter notebook ไว้ใน Google Drive ของตัวเองได้ และให้บริการการใช้งาน Jupyter notebook ร่วมกัน รวมถึงสามารถใช้ทรัพยากรการคำนวณในคลาวด์ หรือทรัพยากรการคำนวณที่เชื่อมต่อผ่านระบบอินเทอร์เน็ต ปัจจุบันเหล่านี้ส่งผลให้ Google Colab เหมาะกับการทำงานเป็นทีม เนื่องจากสมาชิกในทีมไม่จำเป็นต้องใช้เวลากับการติดตั้งสภาพแวดล้อม (environment) ของโปรแกรมให้เหมือนกัน อีกทั้ง การที่สามารถใช้ทรัพยากรการคำนวณในคลาวด์ทำให้ Google Colab เหมาะสมกับงานด้านการเรียนรู้ของเครื่อง ซึ่งบ่อยครั้งเทคนิคที่ใช้มีการคำนวณที่ซับซ้อนและต้องทำงานกับข้อมูลขนาดใหญ่ ซึ่งต้องใช้ทรัพยากรในการคำนวณสูง

ลักษณะตัวอักษรที่ใช้ในหนังสือ

ตัวเอียง	ใช้ระบุข้อความที่เป็น ยูอาร์แอล
Calibri	ใช้แสดงส่วนของโปรแกรม

สารบัญ

คำนำ.....	v
คำแนะนำการใช้หนังสือ	vii
สารบัญ	xi
บทที่ 1 ภาพรวมของการเรียนรู้ของเครื่อง.....	1
1.1 การเรียนรู้ของเครื่องและความสัมพันธ์กับศาสตร์ที่เกี่ยวข้อง.....	1
1.1.1 การเรียนรู้ของเครื่องและศาสตร์ที่เกี่ยวข้อง.....	1
1.1.2 การเรียนรู้ของมนุษย์และการเรียนรู้ของเครื่อง.....	5
1.2 ขั้นตอนทั่วไปในการประยุกต์การเรียนรู้ของเครื่องทางธุรกิจ	7
1.3 ตำแหน่งงานที่เกี่ยวข้องกับการเรียนรู้ของเครื่องและทักษะที่เกี่ยวข้อง	11
1.4 กลุ่มของเทคนิคการเรียนรู้ของเครื่อง	13
1.4.1 การสร้างตัวแบบทำนาย (Predictive Modeling)	15
1.4.2 การแบ่งกลุ่มข้อมูล (Data Segmentation).....	15
1.4.3 การวิเคราะห์ความสัมพันธ์ (Association Analysis)	16
1.5 สรุป	16
บทที่ 2 ลักษณะของข้อมูลและกระบวนการเตรียมข้อมูล	17
2.1 ลักษณะและประเภทของข้อมูล	17
2.2 ทำไมต้องเตรียมข้อมูล.....	19
2.3 การแปลงข้อมูล (Data transformation).....	21
2.4 การเลือกฟีเจอร์ (Feature selection).....	22

2.5 การจัดการกับค่าขาดหาย (Data imputation)	23
2.6 การลดมิติของข้อมูล (Dimensionality reduction)	24
2.7 การแปลงประเภทข้อมูล	24
2.7.1 การแปลงพีเชอร์ประเภท discrete ให้เป็นประเภท continuous	25
2.7.2 การแปลงพีเชอร์ประเภท continuous ให้เป็นประเภท discrete	25
2.8 ผลลัพธ์ของการเตรียมข้อมูล	27
2.9 การแบ่งชุดข้อมูลสำหรับใช้สร้างตัวแบบทำนาย	27
2.10 สรุป	30
บทที่ 3 การประเมินตัวแบบทำนาย	35
3.1 ทำไมต้องประเมินตัวแบบ	35
3.2 การแลกเปลี่ยนของ Bias และ Variance ในการเรียนรู้ของเครื่อง	36
3.3 ประเภทของตัวแบบทำนาย	37
3.4 คำวัดที่ใช้ประเมิน Classification model	38
3.4.1 Confusion matrix	38
3.4.2 Log loss/Binary cross entropy/Categorical cross entropy	41
3.4.3 Receiver operating characteristic curve (ROC) และ Area under the ROC curve (AUC-ROC)	44
3.5 คำวัดที่ใช้ประเมิน Regression model	45
3.5.1 Mean absolute error (MAE)	45
3.5.2 Mean absolute percentage error (MAPE)	46
3.5.3 Mean squared error (MSE)	46

3.5.4 Root mean squared error (RMSE)	47
3.5.5 R Squared (R^2).....	47
3.6 วิธีประเมินตัวแบบทำนาย	48
3.6.1 วิธีประเมินความเหมาะสมของตัวแบบทำนาย	48
3.6.2 วิธีประเมินประสิทธิภาพของตัวแบบทำนาย	49
3.7 สรุป	51
บทที่ 4 เทคนิคต้นไม้การตัดสินใจ (Decision Tree).....	53
4.1 ลักษณะของต้นไม้การตัดสินใจ	53
4.2 ขั้นตอนวิธีเรียนรู้ต้นไม้การตัดสินใจ	56
4.3 Purity	58
4.4 ค่าวัด Purity	60
4.5 Pruning	65
4.5.1 Pre-pruning	65
4.5.2 Post-pruning	67
4.6 การแปลผลต้นไม้การตัดสินใจ	68
4.7 สรุป	70
บทที่ 5 การเรียนรู้แบบหลายตัวแบบ (Ensemble Methods).....	71
5.1 ทำไมต้องใช้การเรียนรู้แบบหลายตัวแบบ.....	71
5.2 Bagging (Bootstrap aggregation)	73
5.3 Boosting	75
5.4 Random forest	77

5.5 สรุป	79
บทที่ 6 ตัวอย่างโปรแกรมสำหรับสร้างตัวแบบทำนาย	81
6.1 ชุดข้อมูลสำหรับใช้ในตัวอย่าง	81
6.2 ตัวอย่างโปรแกรม	82
6.2.1 การนำข้อมูลเข้าและการเตรียมข้อมูล	82
6.2.2 การสร้างตัวแบบด้วยเทคนิคต้นไม้การตัดสินใจ	87
6.2.3 การสร้างตัวแบบด้วยเทคนิค Bagging	88
6.2.4 การสร้างตัวแบบด้วยเทคนิค Boosting	89
6.2.5 การสร้างตัวแบบด้วยเทคนิค Random Forest	92
6.2.6 การประเมินตัวแบบ	94
6.3 สรุป	98
บทที่ 7 การแบ่งกลุ่มข้อมูล (Data Segmentation)	101
7.1 การวัดความคล้ายของข้อมูล	101
7.1.1 การวัดระยะห่างของข้อมูลที่เป็นค่าต่อเนื่อง	102
7.1.2 การวัดระยะห่างของข้อมูลที่เป็นค่าไม่ต่อเนื่อง	104
7.1.3 การวัดระยะห่างด้วยฟังก์ชันที่ผู้ใช้กำหนด	104
7.2 การเตรียมข้อมูลสำหรับการแบ่งกลุ่มข้อมูล	106
7.3 K-Means Clustering	107
7.3.1 K-Means clustering algorithm	107
7.3.2 การกำหนดค่า k	108
7.4 Agglomerative Clustering	109

7.4.1 การวัดระยะห่างระหว่างกลุ่มข้อมูล.....	110
7.4.2 Agglomerative clustering algorithm.....	110
7.5 การประเมินผลการแบ่งกลุ่ม	112
7.5.1 การประเมินจำนวนกลุ่มที่เหมาะสม.....	113
7.5.2 การประเมินคุณลักษณะของกลุ่มที่เป็นผลลัพธ์.....	116
7.6 สรุป	117
บทที่ 8 ตัวอย่างโปรแกรมสำหรับแบ่งกลุ่มข้อมูล	119
8.1 ชุดข้อมูลสำหรับใช้ในตัวอย่าง.....	119
8.2 ตัวอย่างโปรแกรม.....	120
8.2.1 การนำข้อมูลเข้า	120
8.2.2 การเตรียมข้อมูล	121
8.2.3 การแบ่งกลุ่มข้อมูลด้วยเทคนิค K-Means clustering	124
8.2.4 การประเมินคุณลักษณะของกลุ่มที่เป็นผลลัพธ์ของเทคนิค K-Means clustering	128
8.2.5 การแบ่งกลุ่มข้อมูลด้วยเทคนิค Agglomerative clustering.....	132
8.2.6 การประเมินคุณลักษณะของกลุ่มที่เป็นผลลัพธ์ของเทคนิค Agglomerative clustering.....	136
8.3 สรุป	139
บทที่ 9 การวิเคราะห์ความสัมพันธ์ (Association Analysis).....	143
9.1 กฎความสัมพันธ์ (Association rules).....	143
9.2 คำวัดที่เกี่ยวข้องกับกฎความสัมพันธ์	144

9.2.1 Support.....	144
9.2.2 Confidence	145
9.2.3 Lift.....	145
9.3 ข้อมูลสำหรับการวิเคราะห์กฎความสัมพันธ์	146
9.4 Apriori algorithm	148
9.5 Sequential analysis	151
9.6 สรุป	153
บทที่ 10 ตัวอย่างโปรแกรมสำหรับวิเคราะห์ความสัมพันธ์.....	155
10.1 ชุดข้อมูลสำหรับใช้ในตัวอย่าง.....	155
10.2 ตัวอย่างโปรแกรม	156
10.2.1 การนำเข้าข้อมูล.....	156
10.2.2 การตรวจสอบและแก้ไขข้อมูล	156
10.2.3 การเตรียมข้อมูลให้อยู่ในรูปแบบสำหรับใช้งานกับเครื่องมือ	159
10.2.4 การสร้าง Frequent itemsets	161
10.2.5 การสร้าง Association rules.....	162
10.2.6 ตัวอย่างการเลือก Association rules สำหรับนำไปประยุกต์	163
10.3 สรุป	164
บทที่ 11 ประเด็นอื่น ๆ ของการเรียนรู้ของเครื่อง.....	167
11.1 การเรียนรู้แบบอื่น.....	167
11.2 ประเด็นด้านเทคนิค	168
11.2.1 เทคนิคกลุ่มการสร้างตัวแบบทำนาย	168

11.2.2 เทคนิคกลุ่มการแบ่งกลุ่มข้อมูล	171
11.2.3 เทคนิคกลุ่มการวิเคราะห์ความสัมพันธ์	173
11.3 ประเด็นด้านข้อมูล	173
11.4 ประเด็นด้านการประยุกต์	176
11.5 แหล่งความรู้เพิ่มเติม	176
11.6 บทสรุป	177
บรรณานุกรม	181
ดัชนี	186

“We cannot teach people anything; we can only help them
discover it within themselves.”

Galileo Galilei

