



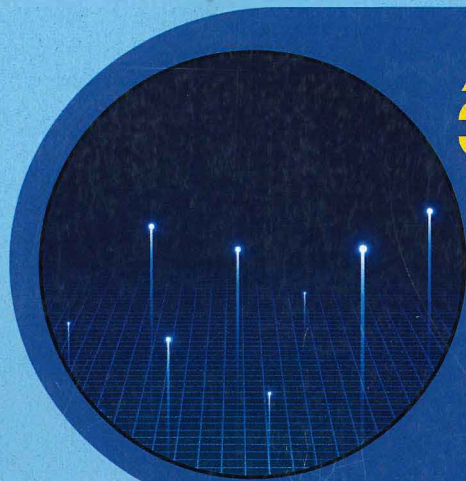
มหาวิทยาลัยสุโขทัยธรรมาธิราช
บัณฑิตศึกษา สาขาวิชาวิทยาศาสตร์และเทคโนโลยี



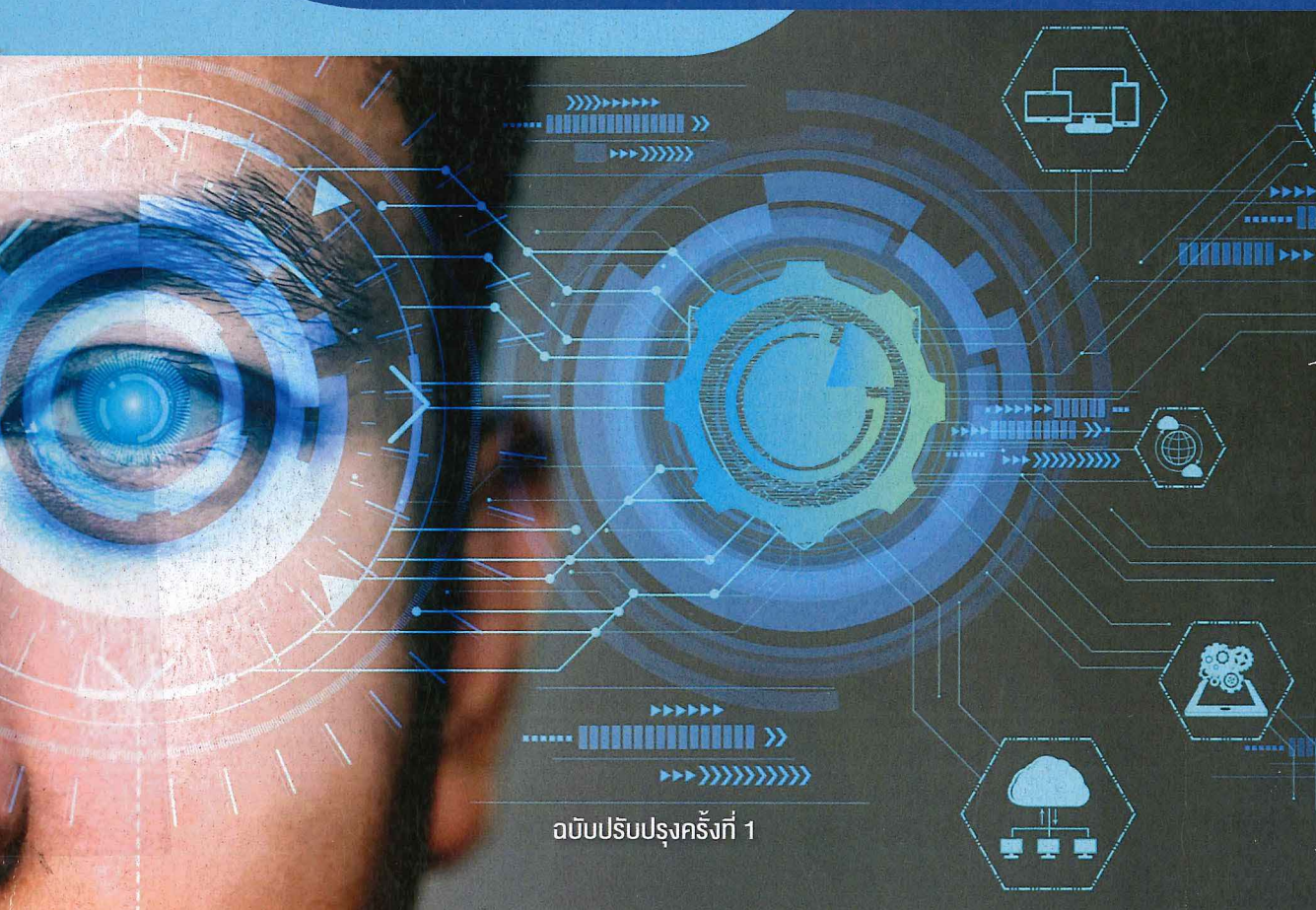
99711

ประมวลสาระชุดวิชา

หน่วยที่ 1-7



วิทยาการข้อมูล และข้อมูลขนาดใหญ่ DATA SCIENCE AND BIG DATA



ฉบับปรับปรุงครั้งที่ 1



มหาวิทยาลัยสุโขทัยธรรมาธิราช
บัณฑิตศึกษา สาขาวิชาวิทยาศาสตร์และเทคโนโลยี

ประมวลสาระชุดวิชา
วิทยาการข้อมูลและข้อมูลขนาดใหญ่

Data Science and Big Data

99711

หน่วยที่ 1-7

(ฉบับปรับปรุงครั้งที่ 1)



คณะกรรมการปรับปรุงชุดวิชาวิทยาการข้อมูลและข้อมูลขนาดใหญ่

รองศาสตราจารย์ณัฐพร เห็นเจริญเลิศ
รองศาสตราจารย์ ดร.วณูชาย์ ร่มสายหยุด
รองศาสตราจารย์ ดร.วรัญญา ปุณณวัฒน์
ผู้ช่วยศาสตราจารย์ ดร.วราภรณ์ วิทยานนท์
ผู้ช่วยศาสตราจารย์ ดร.ธงชัย แก้วกิริยา
ผู้ช่วยศาสตราจารย์ณรงค์ ทีปประชัย
อาจารย์ ดร.วชิระ พรหมวงศ์
อาจารย์ ดร.เดชรัตน์ สอนปี เพี้ยชัย
นายมนตรี ชะฎาแก้ว

ประธานกรรมการ
กรรมการและบรรณาธิการ
กรรมการด้านเนื้อหา
กรรมการด้านเนื้อหา
กรรมการด้านเนื้อหา
กรรมการด้านออกแบบระบบการสอน
กรรมการด้านออกแบบระบบการสอน
กรรมการและเลขานุการ
ผู้ช่วยเลขานุการ

ผู้ร่วมปรับปรุง

รองศาสตราจารย์นันทพร ชื่นปัญญาชนะ

วิธีการศึกษา

หน่วยที่ 1	หลักการของวิทยาการข้อมูล	1-1
	ตอนที่ 1.1 ภาพรวมเกี่ยวกับวิทยาการข้อมูล.....	1-3
	เรื่องที่ 1.1.1 ความหมายและความสำคัญของวิทยาการข้อมูล.....	1-4
	เรื่องที่ 1.1.2 องค์ความรู้และทักษะของวิทยาการข้อมูล	1-11
	เรื่องที่ 1.1.3 กระบวนการทางวิทยาการข้อมูลและการทำเหมืองข้อมูล	1-16
	ตอนที่ 1.2 เทคโนโลยีและเทคนิคเกี่ยวกับวิทยาการข้อมูล	1-27
	เรื่องที่ 1.2.1 ความรู้ที่เกี่ยวข้องกับวิทยาการข้อมูล.....	1-29
	เรื่องที่ 1.2.2 เทคโนโลยีที่เกี่ยวข้องกับวิทยาการข้อมูล.....	1-43
	เรื่องที่ 1.2.3 ตำแหน่งงานด้านวิทยาการข้อมูลและตัวอย่างการใช้ประโยชน์ จากข้อมูล.....	1-54
	บรรณานุกรม.....	1-61
 หน่วยที่ 2	 สถิติและการเขียนโปรแกรมคอมพิวเตอร์สำหรับนักวิทยาการข้อมูล.....	 2-1
	ตอนที่ 2.1 สถิติสำหรับการวิเคราะห์ข้อมูล.....	2-4
	เรื่องที่ 2.1.1 แนวคิดเกี่ยวกับสถิติสำหรับการวิเคราะห์ข้อมูล	2-6
	เรื่องที่ 2.1.2 การวิเคราะห์ข้อมูลโดยใช้สถิติเชิงพรรณนา	2-15
	เรื่องที่ 2.1.3 การวิเคราะห์ข้อมูลโดยใช้สถิติเชิงอ้างอิงหรืออนุมาน.....	2-37
	ตอนที่ 2.2 การโปรแกรมคอมพิวเตอร์สำหรับนักวิทยาการข้อมูล.....	2-73
	เรื่องที่ 2.2.1 การโปรแกรมคอมพิวเตอร์ด้วยภาษาไพธอน	2-74
	เรื่องที่ 2.2.2 ไลบรารีสำหรับการวิเคราะห์ข้อมูล	2-101
	เรื่องที่ 2.2.3 การโปรแกรมคอมพิวเตอร์เพื่อการวิเคราะห์ข้อมูล	2-116
	บรรณานุกรม.....	2-146
	ภาคผนวก	2-149

หน่วยที่ 3 การวิเคราะห์เชิงทำนายสำหรับวิทยาการข้อมูล	3-1
ตอนที่ 3.1 แนวคิดเชิงทำนายข้อมูลสำหรับวิทยาการข้อมูล	3-4
เรื่องที่ 3.1.1 หลักการวิเคราะห์ข้อมูลเชิงทำนาย	3-5
เรื่องที่ 3.1.2 การเตรียมข้อมูลนำเข้าสำหรับการวิเคราะห์เชิงทำนาย	3-14
ตอนที่ 3.2 การวิเคราะห์เชิงทำนายแบบการถดถอยเชิงสถิติ	3-25
เรื่องที่ 3.2.1 การวิเคราะห์การถดถอยเชิงเส้น	3-26
เรื่องที่ 3.2.2 การวิเคราะห์การถดถอยพหุคูณ	3-35
เรื่องที่ 3.2.3 การวิเคราะห์การถดถอยโลจิสติก	3-41
ตอนที่ 3.3 การวิเคราะห์เชิงทำนายด้วยการแยกประเภท	3-51
เรื่องที่ 3.3.1 การวิเคราะห์ด้วยนาอ็อฟเบย์	3-52
เรื่องที่ 3.3.2 การวิเคราะห์ด้วยต้นไม้ตัดสินใจ	3-59
เรื่องที่ 3.3.3 การวิเคราะห์ด้วยซัพพอร์ตเวกเตอร์แมชชีน	3-67
บรรณานุกรม	3-76
 หน่วยที่ 4 หลักการ เทคโนโลยีและเครื่องมือของข้อมูลขนาดใหญ่	 4-1
ตอนที่ 4.1 แนวคิดและหลักการพื้นฐานของข้อมูลขนาดใหญ่	4-4
เรื่องที่ 4.1.1 ความหมาย คุณลักษณะ และประโยชน์ของข้อมูลขนาดใหญ่	4-6
เรื่องที่ 4.1.2 ประเภทของข้อมูลขนาดใหญ่	4-14
เรื่องที่ 4.1.3 สถาปัตยกรรมของข้อมูลขนาดใหญ่	4-21
เรื่องที่ 4.1.4 หลักการทำงานของข้อมูลขนาดใหญ่	4-34
ตอนที่ 4.2 เทคโนโลยีและเครื่องมือของข้อมูลขนาดใหญ่	4-42
เรื่องที่ 4.2.1 เทคโนโลยีการเชื่อมต่อข้อมูลขนาดใหญ่	4-44
เรื่องที่ 4.2.2 เทคโนโลยีและเครื่องมือการรวมข้อมูลขนาดใหญ่	4-54
เรื่องที่ 4.2.3 การบริหารจัดการข้อมูลขนาดใหญ่	4-59
บรรณานุกรม	4-63

หน่วยที่ 5	ฐานข้อมูลโนเอสคิวแอล	5-1
ตอนที่ 5.1	แนวคิดเกี่ยวกับฐานข้อมูลโนเอสคิวแอล	5-4
เรื่องที่ 5.1.1	ความหมายและคุณสมบัติของฐานข้อมูลโนเอสคิวแอล	5-6
เรื่องที่ 5.1.2	ประเภทของฐานข้อมูลโนเอสคิวแอล	5-14
เรื่องที่ 5.1.3	ปัจจัยในการเลือกฐานข้อมูล	5-23
ตอนที่ 5.2	การจัดการฐานข้อมูลโนเอสคิวแอล	5-27
เรื่องที่ 5.2.1	ฐานข้อมูลแบบคีย์-แวลู	5-29
เรื่องที่ 5.2.2	ฐานข้อมูลแบบเอกสาร	5-38
เรื่องที่ 5.2.3	ฐานข้อมูลแบบคอลัมน์	5-50
เรื่องที่ 5.2.4	ฐานข้อมูลแบบกราฟ	5-59
บรรณานุกรม		5-69
 หน่วยที่ 6	 หลักการและวิธีการวิเคราะห์ข้อมูลขนาดใหญ่	 6-1
ตอนที่ 6.1	หลักการวิเคราะห์ข้อมูลขนาดใหญ่	6-4
เรื่องที่ 6.1.1	แนวคิดเกี่ยวกับหลักการวิเคราะห์ข้อมูลขนาดใหญ่	6-5
เรื่องที่ 6.1.2	แนวคิดเกี่ยวกับการเรียนรู้ของเครื่อง	6-9
ตอนที่ 6.2	วิธีการวิเคราะห์ข้อมูลขนาดใหญ่ด้วยการเรียนรู้ของเครื่อง	6-15
เรื่องที่ 6.2.1	การวิเคราะห์ข้อมูลขนาดใหญ่ด้วยวิธีเรียนรู้แบบมีผู้สอน	6-16
เรื่องที่ 6.2.2	การวิเคราะห์ข้อมูลขนาดใหญ่ด้วยวิธีเรียนรู้แบบไม่มีผู้สอน	6-33
เรื่องที่ 6.2.3	การวิเคราะห์ข้อมูลขนาดใหญ่ด้วยวิธีเรียนรู้แบบเสริมแรง	6-49
บรรณานุกรม		6-57

หน่วยที่ 7 หลักการเหมืองข้อมูล การเรียนรู้ของเครื่อง การเรียนรู้เชิงลึก

และปัญญาประดิษฐ์	7-1
ตอนที่ 7.1 ความรู้เกี่ยวกับการทำเหมืองข้อมูล	7-5
เรื่องที่ 7.1.1 แนวคิดเกี่ยวกับการทำเหมืองข้อมูล	7-7
เรื่องที่ 7.1.2 กระบวนการ และอัลกอริทึมการทำเหมืองข้อมูล	7-13
ตอนที่ 7.2 ความรู้เกี่ยวกับการเรียนรู้ของเครื่อง	7-20
เรื่องที่ 7.2.1 แนวคิดเกี่ยวกับการเรียนรู้ของเครื่อง	7-22
เรื่องที่ 7.2.2 กระบวนการ และอัลกอริทึมการเรียนรู้ของเครื่อง	7-28
ตอนที่ 7.3 ความรู้เกี่ยวกับการเรียนรู้เชิงลึก	7-38
เรื่องที่ 7.3.1 แนวคิดเกี่ยวกับการเรียนรู้เชิงลึก	7-40
เรื่องที่ 7.3.2 กระบวนการและอัลกอริทึมการเรียนรู้เชิงลึก	7-47
ตอนที่ 7.4 ความรู้เกี่ยวกับปัญญาประดิษฐ์	7-55
เรื่องที่ 7.4.1 แนวคิดเกี่ยวกับปัญญาประดิษฐ์	7-57
เรื่องที่ 7.4.2 กระบวนการและอัลกอริทึมปัญญาประดิษฐ์	7-63
บรรณานุกรม	7-70