



รายวิชา 040563105 การวิเคราะห์เชิงสถิติสำหรับข้อมูลขนาดใหญ่ (Statistical Analysis for Big Data)

ชื่อสถาบันอุดมศึกษา มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
วิทยาเขต/คณะ/ภาควิชา/สาขาวิชา คณะวิทยาศาสตร์ประยุกต์ ภาควิชาสถิติประยุกต์
สาขาวิชาสถิติประยุกต์สำหรับวิทยาการวิเคราะห์ธุรกิจและอุตสาหกรรม

หมวดที่ 1 ข้อมูลทั่วไป

1. รหัสและชื่อรายวิชา

040563105 การวิเคราะห์เชิงสถิติสำหรับข้อมูลขนาดใหญ่ (Statistical Analysis for Big Data)

2. จำนวนหน่วยกิต

3 หน่วยกิต (3-0-6)

3. หลักสูตรและประเภทของรายวิชา

หลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาวิทยาศาสตร์ข้อมูลและการวิเคราะห์เชิงสถิติ

เป็นรายวิชาเฉพาะ กลุ่มวิชาเลือก

4. อาจารย์ผู้รับผิดชอบรายวิชาและอาจารย์ผู้สอน

อาจารย์ผู้รับผิดชอบรายวิชา รองศาสตราจารย์ ดร.วิกานดา ผาพันธุ์

อาจารย์ผู้สอน อาจารย์ ดร.ภัสดี พุทธาพิพัฒน์

5. ภาคการศึกษา/ชั้นปีที่เรียน

ภาคการศึกษา 2 ของชั้นปีที่ 3

6. รายวิชาบังคับก่อน (Pre-requisite) (ถ้ามี)

ไม่มี

7. รายวิชาที่ต้องเรียนพร้อมกัน (Co-requisites) (ถ้ามี)

ไม่มี

8. สถานที่เรียน

คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

9. ข้อมูลประกอบการประกันคุณภาพการศึกษา

- ✓ การเรียนการสอนในรายวิชานี้มีส่วนที่ได้รับการพัฒนาขึ้นใหม่หรือปรับปรุงจากที่สอนเมื่อครั้งก่อน เช่น การปรับปรุงวิธีการสอน หรือการปรับปรุงเนื้อหา การจัดแบ่งเนื้อหา หรือวิธีการประเมินผลการเรียนรู้



รายวิชานี้มีการให้ผู้มีประสบการณ์ทางวิชาการหรือวิชาชีพจากหน่วยงานหรือชุมชนภายนอกเข้ามามีส่วนร่วมในกระบวนการเรียนการสอน



รายวิชานี้มีการบูรณาการกระบวนการวิจัยหรืองานสร้างสรรค์กับการจัดการเรียนการสอน หรือมีการจัดการเรียนรู้ที่พัฒนาจากการวิจัย หรือจากกระบวนการจัดการความรู้เพื่อพัฒนาการเรียนการสอน



รายวิชานี้มีการบูรณาการงานบริการทางวิชาการแก่สังคมกับการเรียนการสอน



รายวิชานี้มีการบูรณาการงานด้านทำนุบำรุงศิลปและวัฒนธรรมกับการจัดการเรียนการสอนและกิจกรรมนักศึกษา

10. วันที่จัดทำหรือปรับปรุงรายละเอียดของรายวิชาครั้งล่าสุด

16 มิถุนายน 2569

หมวดที่ 2 ลักษณะและการดำเนินการ

1. คำอธิบายรายวิชา

สถาปัตยกรรมและลักษณะเฉพาะของระบบข้อมูลขนาดใหญ่ ระบบนิเวศ Hadoop และ Spark การสืบค้นข้อมูลขนาดใหญ่ด้วย SQL และ DataFrame
คณิตศาสตร์และสถิติที่ใช้ในการวิเคราะห์ข้อมูลขนาดใหญ่ ได้แก่ PCA การสุ่มตัวอย่าง และการทดสอบสมมติฐาน การเรียนรู้ของเครื่องแบบมีผู้สอนและไม่มีผู้สอนบนแพลตฟอร์ม Spark กรณีสึกษาการประยุกต์ใช้ข้อมูลขนาดใหญ่ในสาขาโทรคมนาคม IoT ค่าปลีก และธนาคาร

2. จำนวนชั่วโมงที่ใช้ต่อสัปดาห์

ทฤษฎี (ชั่วโมง)	ฝึกปฏิบัติ (ชั่วโมง)	การศึกษาด้วยตนเอง (ชั่วโมง)
45 ชั่วโมง (3 ชั่วโมง/สัปดาห์)	ไม่มี	90 ชั่วโมง (6 ชั่วโมง/สัปดาห์)

ลักษณะรายวิชา ☒ บรรยาย ☐ ปฏิบัติการ

การวัดและประเมินผล ☒ A-F ☐ S/U ☐ P

3. จำนวนชั่วโมงต่อสัปดาห์ที่จะให้คำปรึกษาและแนะนำทางวิชาการแก่นักศึกษา

- ให้คำปรึกษาและแนะนำทางวิชาการแก่นักศึกษาอย่างน้อย 1 ชั่วโมง/สัปดาห์
- ให้คำปรึกษาแนะนำผ่านเทคโนโลยีสารสนเทศ เช่น แอปพลิเคชันไลน์

4. ผลลัพธ์การเรียนรู้ของรายวิชา (Course Learning Outcomes: CLOs) : นักศึกษาสามารถ

CLO 1. อธิบายและประยุกต์ใช้แนวคิด สถาปัตยกรรม และระบบนิเวศของ Hadoop และ Spark
สำหรับการจัดการข้อมูลขนาดใหญ่ได้



CLO 2. วิเคราะห์และประมวลผลข้อมูลขนาดใหญ่โดยใช้เทคนิคทางสถิติ คณิตศาสตร์
และการเรียนรู้ของเครื่องบนแพลตฟอร์ม Spark ได้

CLO 3.

สื่อสารและนำเสนอผลการวิเคราะห์ข้อมูลขนาดใหญ่ผ่านการสร้างภาพข้อมูลและรายงานได้อย่างมีประสิทธิภาพ

CLO 4. ทำงานร่วมกับผู้อื่นในโครงการด้านข้อมูลขนาดใหญ่ มีความรับผิดชอบต่องานที่ตนรับผิดชอบ
และมีจรรยาบรรณวิชาชีพ

- ความสอดคล้องของผลลัพธ์การเรียนรู้ที่คาดหวังของหลักสูตร (Expected Learning Outcomes: ELOs)
และผลลัพธ์การเรียนรู้ของรายวิชา (Course Learning Outcomes: CLOs)

ตารางที่ 5.1 ความสอดคล้องของ ELOs และ CLOs

ELOs/CLOs	CLO 1	CLO 2	CLO 3	CLO 4
ELO 1 สามารถปฏิบัติงานตามหลักวิชาชีพและนำเสนอข้อมูลตามข้อเท็จจริง มีความซื่อสัตย์ มีวินัย ตรงต่อเวลา และมีความรับผิดชอบต่องานที่ตนรับผิดชอบ				✓
ELO 2 มีความสามารถในการประยุกต์ใช้ความรู้พื้นฐานทางศึกษาทั่วไปและคว ามรู้พื้นฐานทางวิทยาศาสตร์ในสถานการณ์ต่างๆ ได้อย่างถูกต้อง	✓			
ELO 3 สามารถใช้องค์ความรู้ทางด้านการแยกประเภทข้อมูลและวิเคราะห์อย างเป็นระบบและสื่อสารได้อย่างมีประสิทธิภาพ	✓	✓		
ELO 4 สามารถคิดวิเคราะห์อย่างเป็นระบบ สามารถวินิจฉัยตามเหตุตามผล และสามารถนำเทคโนโลยีมาประยุกต์ใช้ได้ทันต่อเหตุการณ์และพัฒนา ศักยภาพอย่างสร้างสรรค์	✓	✓	✓	
ELO 6 สามารถทำงานร่วมกับกลุ่มคนได้หลากหลาย รู้จักบทบาทหน้าที่และความรับผิดชอบของตนเอง มีจิตสำนึกสาธารณะ เสียสละเพื่อส่วนรวม และเป็นที่พึ่งทางวิชาการ				✓
ELO 8 มีความสามารถในการใช้เทคโนโลยีในการสืบค้น เก็บรวบรวมข้อมูล วิเคราะห์ แยกแยะข้อมูลเชิงสถิติ	✓	✓	✓	



ELOs/CLOs	CLO 1	CLO 2	CLO 3	CLO 4
สามารถสื่อสารในยุคดิจิทัลได้และสามารถเลือกใช้ได้อย่างเหมาะสมกับสถานการณ์				

ตารางที่ 5.2 ความสอดคล้องของคุณลักษณะพื้นฐานร่วมกันของบัณฑิตที่พึงประสงค์ มจพ. และ CLOs

คุณลักษณะพื้นฐานร่วมกันของบัณฑิตที่พึงประสงค์ มจพ./CLOs	CLO 1	CLO 2	CLO 3	CLO 4
มีความรู้ความสามารถในวิชาชีพ และมีทักษะด้านความคิดสร้างสรรค์ (Professional and Thinking Skills)	✓	✓	✓	
ซื่อสัตย์ รับผิดชอบ มีคุณธรรม จริยธรรม ทำประโยชน์เพื่อสังคมและเป็นที่พึ่งทางวิชาการ (Social Responsibility)				✓
มีฐานคิดและความเป็นผู้ประกอบการด้านนวัตกรรมและเทคโนโลยี (Innovative and Technopreneur Mindset)	✓		✓	
สามารถแข่งขันได้ในระดับชาติและนานาชาติ (Global Competence)	✓	✓	✓	✓

หมวดที่ 3 การพัฒนานักศึกษาตามผลลัพธ์การเรียนรู้ที่คาดหวัง

วิธีการจัดประสบการณ์การเรียนรู้เพื่อพัฒนาความรู้หรือทักษะ และการวัดผลลัพธ์การเรียนรู้ของรายวิชา
ที่สอดคล้องกับผลลัพธ์การเรียนรู้ที่คาดหวังของรายวิชา (CLOs) ในหมวดที่ 2 ข้อ 4

ผลลัพธ์การเรียนรู้ที่คาดหวังของ รายวิชา (CLOs)	วิธีการจัดการสอน/ประสบการณ์ การเรียนรู้ตาม CLOs	วิธีการวัดผลลัพธ์การเรียนรู้ตาม CLOs
CLO 1	บรรยาย มอบหมายงานและแบบฝึกหัดปฏิบัติด้วย PySpark บน Databricks	- การสอบกลางภาค - การประเมินแบบฝึกหัดที่ได้รับมอบหมาย
CLO 2	บรรยาย กรณีศึกษาจากข้อมูลจริง มอบหมายงานวิเคราะห์ข้อมูลขนาดใหญ่	- การสอบปลายภาค - การประเมินแบบฝึกหัดที่ได้รับมอบหมาย - การประเมินผลงานโครงงาน
CLO 3	บรรยายและยกตัวอย่างการสร้างภาพข้อมูลด้วย PySpark	- การสอบปลายภาค - การประเมินการนำเสนอผลงาน



ผลลัพธ์การเรียนรู้ที่คาดหวังของ รายวิชา (CLOs)	วิธีการจัดการสอน/ประสบการณ์ การเรียนรู้ตาม CLOs	วิธีการวัดผลลัพธ์การเรียนรู้ตาม CLOs
	มอบหมายงานนำเสนอผลการวิเคราะห์	
CLO 4	มอบหมายโครงงานกลุ่ม อภิปรายกลุ่ม เรียนรู้จากการใช้โครงงาน (Project-based learning)	- การประเมินการนำเสนอและผลงานของโครงงานกลุ่ม - นักศึกษาประเมินเพื่อนร่วมกลุ่มในการท ำงานเป็นทีม - อาจารย์สังเกตการทำงานในแต่ละกลุ่ม

หมวดที่ 4 แผนการสอนและการประเมินผล

1. แผนการสอน

ลำดับที่	หัวข้อ/รายละเอียด	CLOs	จำนวน ชั่วโมง	กิจกรรมการเรียนการสอน/ สื่อที่ใช้ (ถ้ามี)	ผู้สอน
1	บทที่ 1: ความรู้เบื้องต้นเกี่ยวกับข้อมูลขนาดใหญ่ (Introduction to Big Data) แนวคิด ลักษณะเฉพาะ และสถาปัตยกรรมของระบบข้อมูลขนาดใหญ่	1, 2	3	บรรยาย ยกตัวอย่างกรณีศึกษา	อ.ดร.ภัสต์ พุทราพิพัฒน์
2	บทที่ 2: ระบบนิเวศ Hadoop และ Spark (Hadoop & Spark Ecosystem) HDFS, MapReduce, Spark Core, RDD, DataFrame	1, 2	3	บรรยาย กรณีศึกษา ปฏิบัติ PySpark บน Databricks	อ.ดร.ภัสต์ พุทราพิพัฒน์
3	บทที่ 3: ทบทวน Python, SQL และ Bash สำหรับข้อมูลขนาดใหญ่	1, 2, 3	3	ปฏิบัติ Python มอบหมายแบบฝึกหัด	อ.ดร.ภัสต์ พุทราพิพัฒน์
4	บทที่ 4: การสืบค้นข้อมูลขนาดใหญ่ (Querying Big Data) SQL, DataFrame API, Spark SQL	1, 2, 3	3	ปฏิบัติ Spark SQL บน Databricks มอบหมายแบบฝึกหัด	อ.ดร.ภัสต์ พุทราพิพัฒน์



ลำดับที่	หัวข้อ/รายละเอียด	CLOs	จำนวน ชั่วโมง	กิจกรรมการเรียนการสอน น/สื่อที่ใช้ (ถ้ามี)	ผู้สอน
5	บทที่ 5: คณิตศาสตร์สำหรับข้อมูลขนาดใหญ่ (Mathematics for Big Data) เมทริกซ์ ค่าเฉพาะ PCA การสุ่มตัวอย่าง	1, 2	3	บรรยาย Lecture + ปฏิบัติ PySpark มอบหมายแบบฝึกหัด	อ.ดร.ภัสดี พุทธาพิพัฒน์
6	บทที่ 6: ขั้นตอนวิธีสำหรับข้อมูลขนาดใหญ่ (Algorithms for Big Data) Streaming, Graph, การทดสอบสมมติฐาน	1, 2	3	บรรยาย กรณีศึกษา ปฏิบัติ PySpark มอบหมายแบบฝึกหัด	อ.ดร.ภัสดี พุทธาพิพัฒน์
7	บทที่ 7: การแสดงผลข้อมูลขนาดใหญ่ (Big Data Visualisation) Matplotlib, Seaborn, Plotly บน Spark	1, 2, 3	3	ปฏิบัติสร้างภาพข้อมูล ด้วย PySpark บน Databricks	อ.ดร.ภัสดี พุทธาพิพัฒน์
8	นำเสนอข้อเสนอโครงการ (Project Proposal) และทบทวนเนื้อหาก่อนสอบกลางภาค	1, 2, 3, 4	3	Workshop นำเสนอ อภิปรายกลุ่ม	อ.ดร.ภัสดี พุทธาพิพัฒน์
9	สอบกลางภาค	-	-	-	-
10	บทที่ 8: การเรียนรู้ของเครื่องแบบมีผู้สอน (Supervised ML for Big Data) Linear/Logistic Regression, Decision Tree, Random Forest บน Spark MLlib	1, 2, 3	3	บรรยาย กรณีศึกษา ปฏิบัติ Spark MLlib บน Databricks	อ.ดร.ภัสดี พุทธาพิพัฒน์
11	บทที่ 9: การเรียนรู้ของเครื่องแบบไม่มีผู้สอน (Unsupervised ML for Big Data) K- Means, PCA, Association Rules บน Spark MLlib	1, 2, 3	3	บรรยาย กรณีศึกษา ปฏิบัติ Spark MLlib บน Databricks	อ.ดร.ภัสดี พุทธาพิพัฒน์



ลำดับที่	หัวข้อ/รายละเอียด	CLOs	จำนวน ชั่วโมง	กิจกรรมการเรียนการสอน /สื่อที่ใช้ (ถ้ามี)	ผู้สอน
12	กรณีศึกษา: ข้อมูลขนาดใหญ่ในอุตสาหกรรมโทรคมนาคม (Telco Use Case) Customer Churn, Network Analysis	1, 2, 3, 4	3	กรณีศึกษา อภิปรายกลุ่ม มอบหมายโครงการ	อ.ดร.ภัสต์ พุทราพิพัฒน์
13	กรณีศึกษา: ข้อมูลขนาดใหญ่ใน IoT (IoT Use Case) Sensor Data, Real-time Analytics	1, 2, 3, 4	3	กรณีศึกษา อภิปรายกลุ่ม	อ.ดร.ภัสต์ พุทราพิพัฒน์
14	กรณีศึกษา: ข้อมูลขนาดใหญ่ในอุตสาหกรรมค้าปลีกและธนาคาร (Retail & Banking Use Case) Fraud Detection, Recommendation System	1, 2, 3, 4	3	กรณีศึกษา อภิปรายกลุ่ม	อ.ดร.ภัสต์ พุทราพิพัฒน์
15	นำเสนอโครงงานปลายภาค (Final Project Presentation)	1, 2, 3, 4	3	นำเสนอโครงงาน อภิปรายกลุ่ม	อ.ดร.ภัสต์ พุทราพิพัฒน์
16	นำเสนอโครงงานปลายภาค (ต่อ) (Final Project Presentation cont'd)	1, 2, 3, 4	3	นำเสนอโครงงาน อภิปรายกลุ่ม	อ.ดร.ภัสต์ พุทราพิพัฒน์
17	สอบปลายภาค	-	-	-	-
รวม			45		

2. แผนการประเมินตามผลลัพธ์การเรียนรู้ที่คาดหวังของรายวิชา

ผลลัพธ์การเรียนรู้ที่คาดหวัง ของรายวิชา (CLOs)	กิจกรรมการประเมินผลการเรียนรู้ของ ผู้เรียน	กำหนดการประเมิน (ลำดับที่)	สัดส่วนของการ ประเมินผล
CLO 1, 2, 3, 4	แบบฝึกหัดที่มอบหมาย (Practical Assignments)	3, 4, 5, 6, 10	20%
CLO 1, 2, 3, 4	การสอบกลางภาค (Midterm Exam)	9	20%
CLO 1, 2, 3, 4	โครงงานกลุ่มและการนำเสนอ (Class Project)	8, 12-16	25%



ผลลัพธ์การเรียนรู้ที่คาดหวัง ของรายวิชา (CLOs)	กิจกรรมการประเมินผลการเรียนรู้ของ ผู้เรียน	กำหนดการประเมิน (สัปดาห์ที่)	สัดส่วนของการ ประเมินผล
CLO 1, 2, 3, 4	การสอบปลายภาค (Final Exam)	17	35%

หมวดที่ 5 ทรัพยากรประกอบการเรียนการสอน

ตำราและเอกสารที่ใช้ประกอบการเรียนการสอน

1. เอกสารและตำราหลัก

- Damji, J. S., Wenig, B., Das, T., & Lee, D. Learning Spark (2nd ed.). O'Reilly Media.
- Chambers, B. & Zaharia, M. Spark: The Definitive Guide. O'Reilly Media.
- Leskovec, J., Rajaraman, A. & Ullman, J. D. Mining of Massive Datasets. Cambridge University Press.

2. เอกสารและข้อมูลแนะนำ

- เว็บไซต์ PySpark Documentation: <https://spark.apache.org/docs/latest/api/python/>
- Databricks Community Edition: <https://community.cloud.databricks.com/>
- Kaggle Datasets: <https://www.kaggle.com/>
- MLlib Guide: <https://spark.apache.org/docs/latest/ml-guide.html>

หมวดที่ 6 การประเมินและปรับปรุงการดำเนินการของรายวิชา

กลยุทธ์การประเมินประสิทธิผลของรายวิชาโดยนักศึกษา

- ☒ แบบประเมินรายวิชา
- ☒ การสนทนากลุ่มระหว่างผู้สอนและผู้เรียน
- ☒ การสะท้อนคิดจากพฤติกรรมของผู้เรียน
- ☒ ข้อเสนอแนะผ่านช่องทางออนไลน์ที่อาจารย์ผู้สอนได้จัดทำเป็นช่องทางการสื่อสารกับนักศึกษา
- ☐ อื่นๆ (ระบุ)

กลยุทธ์การประเมินการจัดการเรียนรู้

- ☒ แบบประเมินผู้สอน
- ☒ ผลการสอบ
- ☒ การทวนสอบผลประเมินผลลัพธ์การเรียนรู้
- ☒ การประเมินโดยคณะกรรมการประเมินข้อสอบ



หลักสูตรวิทยาศาสตรบัณฑิต
ระดับปริญญาตรี

ภาควิชาสถิติประยุกต์/สาขาวิชาวิทยาศาสตร์ข้อมูลและการวิเคราะห์เชิงสถิติ
คณะวิทยาศาสตร์ประยุกต์

- ☐ การสังเกตการณ์สอนของผู้ร่วมทีมการสอน
- ☐ อื่นๆ (ระบุ)

กลไกการปรับปรุงการจัดการเรียนรู้

- ✓ สัมมนาการจัดการเรียนการสอน
- ✓ การวิจัยในและนอกชั้นเรียน
- ☐ อื่นๆ (ระบุ)

กระบวนการทวนสอบผลลัพธ์การเรียนรู้ของรายวิชาของนักศึกษา

- ✓ มีการตั้งคณะกรรมการในสาขาวิชา ตรวจสอบผลการประเมินผลลัพธ์การเรียนรู้ของนักศึกษา โดยตรวจสอบข้อสอบรายงาน วิธีการให้คะแนนสอบ และการให้คะแนนพฤติกรรม
- ✓ การทวนสอบการให้คะแนนการตรวจผลงานของนักศึกษาโดยกรรมการวิชาการประจำภาควิชาและคณะ
- ✓ การทวนสอบการให้คะแนนจากการสุ่มตรวจผลงานของนักศึกษาโดยอาจารย์ หรือผู้ทรงคุณวุฒิอื่นๆ ที่ไม่ใช่อาจารย์ประจำหลักสูตร
- ☐ อื่นๆ (ระบุ)

การดำเนินการทบทวนและการวางแผนปรับปรุงประสิทธิผลของรายวิชา

- ✓ ปรับปรุงรายวิชาในแต่ละปี ตามข้อเสนอแนะและผลการทวนสอบตามข้อ 4
- ✓ ปรับปรุงรายวิชาในแต่ละปี ตามผลการประเมินผู้สอนโดยนักศึกษา
- ☐ อื่นๆ (ระบุ)