



กิจกรรมประจำชุดวิชา
99423 ข้อมูลขนาดใหญ่และการประยุกต์
ภาคการศึกษา 2/2568

สาขาวิชาวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยสุโขทัยธรรมมาธิราช

คำนำ

เนื่องด้วยมหาวิทยาลัยสุโขทัยธรรมาธิราช มุ่งให้ผู้เรียนได้มีส่วนร่วมในกระบวนการศึกษาเล่าเรียนครบวงจร ตั้งแต่ก่อนเรียน ระหว่างเรียน และหลังจากเรียนเสร็จสิ้นไปแล้ว โดยจัดระบบการประเมินครบทั้ง 3 ส่วน ทั้งการประเมินก่อนเรียน ระหว่างเรียน และประเมินผลสุดท้าย

การประเมินกิจกรรมเป็นส่วนหนึ่งของการวัดผลสัมฤทธิ์ทางการเรียนสุดท้าย จึงให้นักศึกษาทำกิจกรรมภาคปฏิบัติตามที่กำหนดให้ โดยมีวัตถุประสงค์เพื่อให้นักศึกษามีความสามารถ ดังนี้

1. สรุปหรือประมวลเนื้อหาสาระของเอกสารการสอนทั้งชุดวิชาหรือกลุ่มเนื้อหาในกลุ่มใดกลุ่มหนึ่ง
2. ประยุกต์ความรู้จากเอกสารการสอนเพื่อจัดทำโครงการพัฒนางานอย่างใดอย่างหนึ่งที่นักศึกษาทำ
3. พัฒนาระบบ โครงการ ชิ้นงาน ฯลฯ ตามกระบวนการหรือขั้นตอนที่แสดงไว้ในหน่วยใดหน่วยหนึ่งของเอกสารการสอน
4. คิด วิเคราะห์ นำเสนอข้อมูลและความคิดในเชิงสร้างสรรค์

นอกจากนี้การทำกิจกรรมประจำชุดวิชายังทำให้นักศึกษาได้ศึกษาเอกสารการสอนตั้งแต่ต้นภาคการศึกษา และจากการวิจัยพบว่านักศึกษาที่ทำกิจกรรมจะมีโอกาสสอบผ่านในปลายภาคมากกว่านักศึกษาที่ไม่ทำกิจกรรม

คณะกรรมการบริหารชุดวิชาข้อมูลขนาดใหญ่และการประยุกต์ ขอให้นักศึกษาทุกท่านประสบความสำเร็จในการศึกษาชุดวิชานี้ และสามารถนำความรู้ไปเป็นประโยชน์ต่อการดำเนินชีวิตและการประกอบอาชีพสืบไป

คณะกรรมการบริหารชุดวิชา

99423 ข้อมูลขนาดใหญ่และการประยุกต์

1. การประเมินผล

เกณฑ์การให้คะแนนกิจกรรมพิจารณาจากการตอบที่ตรงประเด็นคำถาม การครอบคลุมความถูกต้องของคำตอบ ความชัดเจนของการนำเสนอ และความละเอียดประณีตของชิ้นงาน

มหาวิทยาลัยไม่บังคับให้นักศึกษาทุกคนต้องทำกิจกรรม นักศึกษาอาจเลือกทำหรือไม่ทำก็ได้ โดยการประเมินปลายภาคสำหรับชุดวิชานี้ แบ่งออกเป็น 2 กรณี

กรณีที่ 1 นักศึกษาทำกิจกรรม ในกรณีนี้มหาวิทยาลัยแบ่งคะแนนออกเป็น 2 ส่วน ส่วนแรกจากคะแนนสอบปลายภาคคิดร้อยละ 80 และส่วนที่สองจากคะแนนกิจกรรมคิดร้อยละ 20 โดยคะแนนกิจกรรมจะนำไปใช้ในการประเมินทั้งการสอบไล่และสอบซ่อม นักศึกษาที่ไม่ได้ส่งกิจกรรมในการสอบไล่ จะส่งกิจกรรมเพื่อเป็นคะแนนในการสอบซ่อมไม่ได้

กรณีที่ 2 นักศึกษาไม่ทำกิจกรรม ในกรณีนี้มหาวิทยาลัยประเมินผลจากการสอบปลายภาคเพียงอย่างเดียว

ในการประเมินผลปลายภาค นักศึกษากลุ่มที่ทำกิจกรรมและไม่ทำกิจกรรม จะได้รับการประเมินผลโดยใช้ข้อสอบฉบับเดียวกัน นักศึกษากลุ่มที่ทำกิจกรรมมีคะแนนเต็ม 80 คะแนน ส่วนนักศึกษากลุ่มที่ไม่ทำกิจกรรมจะมีคะแนนเต็ม 100 คะแนน

สำหรับนักศึกษาที่ทำกิจกรรม มหาวิทยาลัยจะพิจารณาให้นักศึกษาได้ประโยชน์สูงสุด โดยการนำคะแนนสอบปลายภาคของนักศึกษาเพียงอย่างเดียว มาเปรียบเทียบกับความคิดคะแนนสอบปลายภาครวมกับคะแนนกิจกรรม แล้วนำคะแนนส่วนที่มากกว่าไปใช้ในการตัดสินผลการสอบให้นักศึกษา ดังตัวอย่างต่อไปนี้

ตัวอย่างที่ 1 นักศึกษาได้คะแนนกิจกรรม 18 คะแนน และทำข้อสอบได้ 70 ข้อ (คิดเป็น $\frac{70}{120} \times 80$ เท่ากับ 46.67 คะแนน) นักศึกษาจะได้คะแนนกิจกรรมรวมกับคะแนนสอบปลายภาค $18 + 46.67$ เท่ากับ 64.67 คะแนน นั่นคือ “สอบผ่าน” กรณีที่นักศึกษาไม่ได้ทำกิจกรรมประจำชุดวิชา จะคิดคะแนนจากการสอบปลายภาคเพียงอย่างเดียว นักศึกษาทำข้อสอบได้ 70 ข้อ จะได้ $70 \times \frac{100}{120}$ เท่ากับ 58.33 คะแนน นั่นคือ “สอบไม่ผ่าน”

ดังนั้นมหาวิทยาลัยจะเลือกให้นักศึกษาได้คะแนน 64.67 คะแนน นั่นคือ “สอบผ่าน”

ตัวอย่างที่ 2 นักศึกษาได้คะแนนกิจกรรม 13 คะแนน และทำข้อสอบได้ 92 ข้อ (คิดเป็น $\frac{92}{120} \times 80$ เท่ากับ 61.33 คะแนน) นักศึกษาจะได้คะแนนกิจกรรมรวมกับคะแนนสอบปลายภาค $13 + 61.33$ เท่ากับ 74.33 คะแนน กรณีคิดคะแนนจากการสอบปลายภาคเพียงอย่างเดียว นักศึกษาจะได้ $92 \times \frac{100}{120}$ เท่ากับ 76.67 คะแนน

ดังนั้นมหาวิทยาลัยจะเลือกให้นักศึกษาได้คะแนนมากที่สุด ได้แก่ 76.67 คะแนน นั่นคือ “สอบผ่าน”

2. การส่งกิจกรรมประจำชุดวิชา

ให้นักศึกษาดำเนินการดังนี้

1. ให้นักศึกษาส่งกิจกรรมประจำชุดวิชาฉบับจริงไปยังมหาวิทยาลัย และสำเนากิจกรรมที่ทำเสร็จแล้วไว้ 1 ชุด เพื่อเก็บไว้เป็นหลักฐาน
2. กำหนดการส่งกิจกรรมประจำชุดวิชาภายใน วันที่ 30 เมษายน 2569
3. ให้จัดทำหน้าปกรายงานให้มีข้อความตามตัวอย่างที่แนบมา
4. ส่งกิจกรรมที่ทำเสร็จเรียบร้อยได้ 2 วิธี ทั้งส่งทางไปรษณีย์ หรือส่งผ่านช่องทางออนไลน์

- 4.1 ถ้าเลือกส่งทางไปรษณีย์ ให้ส่งมาที่ สำนักบริการการศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช หรือส่งทางไปรษณีย์ลงทะเบียน โดยจำหน้าซองดังนี้

ศูนย์บริการการสอนทางไปรษณีย์

สำนักบริการการศึกษา

มหาวิทยาลัยสุโขทัยธรรมาธิราช

ตำบลบางพูด อำเภอปากเกร็ด

จังหวัดนนทบุรี 11120

(กิจกรรมประจำชุดวิชา 99423 ข้อมูลขนาดใหญ่และการประยุกต์)

สาขาวิชาวิทยาศาสตร์และเทคโนโลยี

- 4.2 ถ้าเลือกส่งทางออนไลน์ ให้ส่งมาที่เว็บไซต์ <https://eduapp.stou.ac.th/practice/>

ในกรณีที่ส่งทางไปรษณีย์ให้เก็บสลิปหรือต้นข้าวการส่ง และถ่ายเอกสารกิจกรรมที่ส่งไปมหาวิทยาลัยไว้เป็นหลักฐาน ในการส่งกิจกรรมทุกชิ้น นักศึกษาจะต้องจัดทำหน้าปกรายงาน (ปรากฏในภาคผนวกที่ส่งมาด้วย)

5. นักศึกษาสามารถตรวจสอบว่าสำนักบริการการศึกษาได้รับกิจกรรมที่นักศึกษาส่งไปแล้วหรือยัง โดยโทรศัพท์สอบถามได้ที่หมายเลข 02-982-9633 หรือโทรศัพท์ติดต่อสำนักบริการการศึกษา หมายเลข 02-504-7621 หรือโทรศัพท์ติดต่อศูนย์สารสนเทศ หมายเลข 02-504-7888 โทรศัพท์มือถือ 084-360-4465, 084-439-9478, 084-360-5612 และ 084-360-4957 หรือที่ e-mail : ic.proffice@stou.ac.th หรือตรวจสอบผ่านระบบสารสนเทศได้ที่เว็บไซต์ <https://regis.stou.ac.th/STOU/login.jsp>

3. เนื้อหากิจกรรม

คำชี้แจง ให้นักศึกษาตอบคำถามใน 15 กิจกรรมต่อไปนี้ในกระดาษ A4 โดยให้นักศึกษาทำกิจกรรมด้วยตนเอง และเขียนด้วยลายมือตนเองเท่านั้น ถ้าตรวจสอบได้ว่าการลอกกันหรือไม่ได้ใช้ความรู้ของตนเองในการตอบคำถาม จะไม่ตรวจให้คะแนนไม่ว่ากรณีใดๆ กิจกรรมประจำชุดวิชานี้ มีจำนวน 15 ข้อ คะแนนเต็ม 60 คะแนน

กิจกรรมที่ 1 หน่วยที่ 1 ความรู้พื้นฐานเกี่ยวกับข้อมูลขนาดใหญ่

ข้อที่ 1.1 อธิบายความหมายของข้อมูลขนาดใหญ่หรือบิ๊กดาต้า (big data) พร้อมทั้งยกตัวอย่างข้อมูลขนาดใหญ่ที่มีในชีวิตประจำวัน (2 คะแนน)

ข้อที่ 1.2 สถาปัตยกรรมบิ๊กดาต้า (big data architecture) คืออะไร ประกอบด้วยคอมโพเนนต์อะไรบ้าง อธิบายเทคโนโลยี เครื่องมือ และกระบวนการของแต่ละคอมโพเนนต์มาพอสังเขป (2 คะแนน)

กิจกรรมที่ 2 หน่วยที่ 2 คลังข้อมูล

ข้อที่ 2.1 คลังข้อมูล (data warehouse) แตกต่างจากฐานข้อมูล (database) อย่างไร พร้อมทั้งบอกคุณลักษณะสำคัญของคลังข้อมูลมาพอสังเขป (2 คะแนน)

ข้อที่ 2.2 ส่วนประกอบหลักที่สำคัญของระบบคลังข้อมูลมีอะไรบ้าง พร้อมทั้งอธิบายรายละเอียดของแต่ละส่วนประกอบมาพอเข้าใจ (2 คะแนน)

กิจกรรมที่ 3 หน่วยที่ 3 การสกัด การแปลง และการนำข้อมูลเข้าระบบคลังข้อมูล

ข้อที่ 3.1 กระบวนการ ETL คืออะไร อธิบายการทำงานของแต่ละฟังก์ชันมาพอสังเขป (2 คะแนน)

[illegible]

ข้อที่ 3.2 อธิบายความแตกต่างระหว่าง ETL กับ ELT มาพอเข้าใจ (2 คะแนน)

[illegible]

กิจกรรมที่ 4 หน่วยที่ 4 เหมืองข้อมูล

ข้อที่ 4.1 เทคนิคการทำเหมืองข้อมูล แบ่งเป็นกี่กลุ่ม อะไรบ้าง อธิบายพร้อมยกตัวอย่างการประยุกต์ใช้แต่ละเทคนิคมาพอเข้าใจ (2 คะแนน)

ข้อที่ 4.2 การใช้เทคนิคการแบ่งกลุ่มแบบเคมีน (K-Mean) มาคำนวณหาระยะห่างระหว่างชุดข้อมูลกับจุดศูนย์กลาง ได้ค่าดังตารางต่อไปนี้

ชุดข้อมูลที่	ระยะห่างระหว่างจุดศูนย์กลางที่ 1	ระยะห่างระหว่างจุดศูนย์กลางที่ 2
1	2.8	4.4
2	3.5	2.4
3	2.1	1.6
4	5.7	2.6
5	3.8	4.1

จงอธิบายว่าข้อมูลจะถูกแบ่งเป็นกี่กลุ่ม และแต่ละกลุ่มประกอบด้วยชุดข้อมูลใดบ้าง (2 คะแนน)

กิจกรรมที่ 6 หน่วยที่ 6 ฐานข้อมูลโนเอสคิวแอล

ข้อที่ 6.1 ฐานข้อมูลโนเอสคิวแอล (NoSQL) ที่นิยมใช้มีกี่ประเภท อะไรบ้าง อธิบายพร้อมยกตัวอย่างมาพอเข้าใจ (2 คะแนน)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

ข้อที่ 6.2 ปัจจัยที่ใช้ในการพิจารณาเลือกฐานข้อมูล NoSQL มีอะไรบ้าง บอกมาอย่างน้อย 3 ข้อ (2 คะแนน)

[illegible]

กิจกรรมที่ 7 หน่วยที่ 7 การวิเคราะห์ข้อมูลขนาดใหญ่

ข้อที่ 7.1 กระบวนการวิเคราะห์ข้อมูล มีกี่ขั้นตอน อะไรบ้าง อธิบายแต่ละขั้นตอนมาพอสังเขป (2 คะแนน)

[illegible]

ข้อที่ 7.2 ยกตัวอย่างการวิเคราะห์ข้อมูลขนาดใหญ่ (big data analytics) ในธุรกิจ มาอย่างน้อย 2 ตัวอย่าง โดยให้เป็นตัวอย่างอื่นๆ ที่ไม่ได้กล่าวถึงในเอกสารการสอน เรื่องที่ 7.2.3 (2 คะแนน)

[illegible]

กิจกรรมที่ 8 หน่วยที่ 8 การประมวลผลข้อมูลขนาดใหญ่แบบฮาดูปและแมปรีดิวซ์

ข้อที่ 8.1 องค์ประกอบหลักของฮาดูป (Hadoop) มีอะไรบ้าง และมีหน้าที่อย่างไร (2 คะแนน)

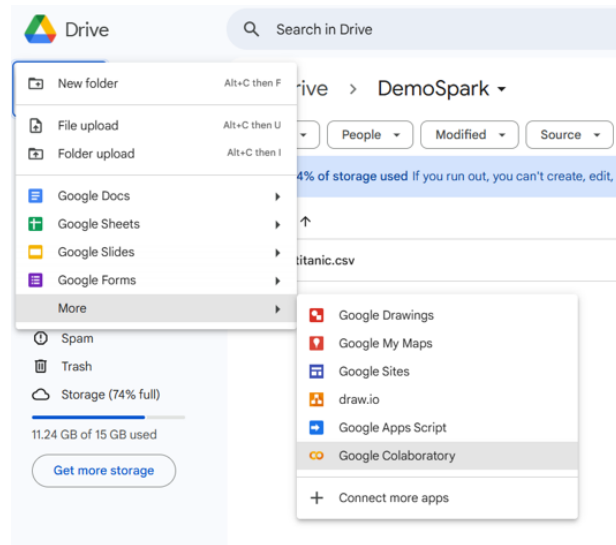
[illegible]

ข้อที่ 8.2 อธิบายขั้นตอนการโปรแกรมแมปรีดิวซ์ (MapReduce) มาพอเข้าใจ (2 คะแนน)

[illegible]

กิจกรรมที่ 9 หน่วยที่ 9 การประมวลผลข้อมูลขนาดใหญ่ด้วยสแปร์ค

ข้อที่ 9. ศึกษาการประมวลผลข้อมูลขนาดใหญ่ (big data) ผ่าน Google Colab ดังภาพ



แล้วทำการติดตั้งแพ็คเกจ PySpark พร้อมทั้งนำเข้าข้อมูลบน Spark ผ่าน Spark Context APIs หรือ Spark Session APIs โดยใช้ข้อมูลนำเข้าเป็นไฟล์ .txt, .csv หรือ .json (อย่างใดอย่างหนึ่ง) แล้ว capture หน้าจอโปรแกรมที่แสดงส่วนของ code และผลลัพธ์การนำเข้าข้อมูลมาแสดงด้านล่าง (4 คะแนน)

หมายเหตุ สามารถศึกษารายละเอียดเพิ่มเติมได้จากเอกสารการสอน เรื่องที่ 9.1.2 และ 9.2.1

กิจกรรมที่ 10 หน่วยที่ 10 การประมวลผลข้อมูลแบบคลาวด์

ข้อที่ 10. การประมวลผลข้อมูลแบบคลาวด์ (cloud computing) สามารถแบ่งประเภทโดยการนำไปใช้งานได้กี่ประเภท อะไรบ้าง อธิบายมาพอเข้าใจ (4 คะแนน)

กิจกรรมที่ 11 หน่วยที่ 11 การประยุกต์ใช้ข้อมูลขนาดใหญ่กับงานด้านสาธารณสุข

ข้อที่ 11. ยกตัวอย่างการใช้เทคโนโลยีปัญญาประดิษฐ์ (AI) ในการดูแลสุขภาพ พร้อมทั้งอธิบายมาพอสังเขป (4 คะแนน)

กิจกรรมที่ 12 หน่วยที่ 12 การประยุกต์ใช้ข้อมูลขนาดใหญ่ในการเกษตร

ข้อที่ 12. ยกตัวอย่างการประยุกต์ใช้ข้อมูลขนาดใหญ่ในการปลูกพืช พร้อมทั้งอธิบายมาพอสังเขป (4 คะแนน)

[illegible]

กิจกรรมที่ 13 หน่วยที่ 13 การประยุกต์ใช้ข้อมูลขนาดใหญ่กับคมนาคมและเมืองอัจฉริยะ

ข้อที่ 13. ยกตัวอย่างการวิเคราะห์ข้อมูลขนาดใหญ่กับคมนาคมอัจฉริยะ พร้อมทั้งอธิบายมาพอสังเขป (4 คะแนน)

[illegible]

กิจกรรมที่ 14 หน่วยที่ 14 การประยุกต์ใช้ข้อมูลขนาดใหญ่กับสื่อสังคมออนไลน์

ข้อที่ 14. อธิบายขั้นตอนการเตรียมข้อมูลสำหรับการวิเคราะห์ข้อมูลขนาดใหญ่จากสื่อสังคมออนไลน์ พร้อมทั้งยกตัวอย่างการวิเคราะห์ข้อมูลทางด้านการตลาดมาพอเข้าใจ (4 คะแนน)

กิจกรรมที่ 15 หน่วยที่ 15 การประยุกต์ใช้ข้อมูลขนาดใหญ่กับการรักษาความมั่นคงปลอดภัยไซเบอร์

ข้อที่ 15. ยกตัวอย่างการประยุกต์ใช้เทคโนโลยีข้อมูลขนาดใหญ่กับความมั่นคงปลอดภัยไซเบอร์ พร้อมทั้งอธิบายมาพอสังเขป (4 คะแนน)

ปกรายงาน

กิจกรรมประจำชุดวิชา
99423 ข้อมูลขนาดใหญ่และการประยุกต์
ภาคการศึกษา 2/2568

สาขาวิชาวิทยาศาสตร์และเทคโนโลยี

ชื่อนักศึกษา.....

รหัสประจำตัวนักศึกษา

--	--	--	--	--	--	--	--	--	--

ที่อยู่

.....

โทรศัพท์ (ถ้ามี)

ข้าพเจ้าขอยอมรับการตัดสินผลคะแนนภาคปฏิบัติจากผู้ประเมินเป็นที่สุด

ลงชื่อ.....

(.....)