



กิจกรรมสอบประเมินสมรรถนะของบุคคลตามมาตรฐานอาชีพ สาขาวิทยาศาสตร์ข้อมูล อาชีพนักวิทยาศาสตร์ข้อมูล (Data Scientist) ระดับ 5

คุณสมบัติผู้เข้ารับการประเมิน

1. เป็นผู้สำเร็จการศึกษาระดับปริญญาตรี และมีประสบการณ์บริหารจัดการข้อมูลอย่างน้อย 1 โครงการ หรือ
2. มีประสบการณ์ทำงานเกี่ยวกับงานด้านวิทยาศาสตร์ข้อมูลไม่น้อยกว่า 2 ปี และมีประสบการณ์บริหารจัดการข้อมูลอย่างน้อย 1 โครงการ รวมถึงโครงการระหว่างศึกษา หรือ
3. ได้รับการรับรองคุณวุฒิวิชาชีพ อาชีพนักวิเคราะห์ข้อมูล (ระดับ 4) หรือ อาชีพผู้ปฏิบัติงานด้านวิศวกรรมข้อมูล (ระดับ 4) และมีประสบการณ์บริหารจัดการข้อมูลอย่างน้อย 1 โครงการ รวมถึงโครงการระหว่างศึกษา

หน่วยสมรรถนะที่ใช้ในการประเมิน

1. 70401 เลือกเทคนิควิธีสำหรับการเรียนรู้แบบมีผู้สอน (Select Modelling Techniques for Supervised learning)
2. 70402 เลือกเทคนิควิธีสำหรับการเรียนรู้แบบไม่มีผู้สอน (Select Modelling Techniques for Unsupervised learning)
3. 70403 สร้างแบบจำลองสำหรับเทคนิคการวิเคราะห์แบบมีผู้สอน (Build Model for Supervised learning)
4. 70404 สร้างแบบจำลองสำหรับเทคนิคการวิเคราะห์แบบไม่มีผู้สอน (Build Model for Unsupervised learning)

ระยะเวลาการสอบประเมิน

กิจกรรมสอบประเมิน ประกอบด้วย ข้อเขียน สัมภาษณ์ และสาธิตการปฏิบัติ
รวมระยะเวลา 3 ชั่วโมง เวลา 09:00 – 12:00 น.

เกณฑ์การผ่านการประเมิน

- "ผ่านการประเมิน"
หมายถึง ผู้เข้ารับการประเมิน มีผลคะแนน ภาคทฤษฎีไม่น้อยกว่าร้อยละ 70 และภาคปฏิบัติไม่น้อยกว่า ร้อยละ 80
- "ไม่ผ่านการประเมิน"
หมายถึง ผู้เข้ารับการประเมิน มีผลคะแนน ภาคทฤษฎีน้อยกว่าร้อยละ 70 และภาคปฏิบัติน้อยกว่าร้อยละ 80



กำหนดการจัดกิจกรรมอบรม

09:00 – 12:00 น.

กิจกรรมสอบประเมิน

แนะนำ และอธิบายกระบวนการ พร้อมสอบประเมิน ข้อเขียน สัมภาษณ์ และสาธิตการปฏิบัติ ภายใต้หน่วย
สมรรถนะและเกณฑ์การปฏิบัติงานต่อไปนี้

70401 เลือกเทคนิควิธีสำหรับการเรียนรู้แบบมีผู้สอน (Select Modelling Techniques for Supervised learning)

- **70401.01 จำแนกเทคนิคการเรียนรู้แบบมีผู้สอนแบบต่าง ๆ**

- 1. ระบุความแตกต่างของเทคนิคการวิเคราะห์แบบมีผู้สอนแบบต่าง ๆ ได้
- 2. ระบุลักษณะข้อมูลที่เหมาะสมกับการเรียนรู้แบบมีผู้สอนได้
- 3. ระบุความแตกต่างของคำตอบหรือผลลัพธ์ที่จะได้ระหว่างเทคนิคต่าง ๆ สำหรับการเรียนรู้แบบมีผู้สอนได้

- **70401.02 เลือกใช้เทคนิคการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิค Decision Tree**

- 1. ระบุหลักการของเทคนิคการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิค Decision Tree ได้
- 2. บอกถึงผลลัพธ์ที่จะได้รับจากการสร้างสมการหรือต้นแบบจำลองด้วยเทคนิค Decision Tree ได้
- 3. กำหนดเครื่องมือช่วยในการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิค Decision Tree ได้
- 4. ตั้งสมมติฐานที่คาดว่าจะได้รับได้

- **70401.03 เลือกใช้เทคนิคการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิค SVM**

- 1. ระบุหลักการของเทคนิคการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิค SVM ได้
- 2. บอกถึงผลลัพธ์ที่จะได้รับจากการสร้างสมการหรือแบบจำลองด้วยเทคนิค SVM ได้
- 3. กำหนดเครื่องมือช่วยในการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิค SVM ได้
- 4. ตั้งสมมติฐานที่คาดว่าจะได้รับได้

- **70401.04 เลือกใช้เทคนิคการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิค ANN**

- 1. ระบุหลักการของเทคนิคการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิค ANN ได้
- 2. บอกถึงผลลัพธ์ที่จะได้รับจากการสร้างสมการหรือแบบจำลองด้วยเทคนิค ANN ได้
- 3. กำหนดเครื่องมือช่วยในการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification) ด้วยเทคนิค ANN ได้
- 4. ตั้งสมมติฐานที่คาดว่าจะได้รับได้



- **70401.05 เลือกใช้เทคนิคการวิเคราะห์แบบมีผู้สอนแบบการวิเคราะห์การถดถอย (Regression) ด้วยเทคนิค Linear Regression**

- 1. ระบุหลักการของเทคนิคการวิเคราะห์แบบมีผู้สอนแบบการวิเคราะห์การถดถอย (Regression) ด้วยเทคนิค Linear Regression ได้
- 2. บอกถึงผลลัพธ์ที่จะได้รับจากการสร้างสมการหรือแบบจำลองด้วยเทคนิค Linear Regression ได้
- 3. กำหนดเครื่องมือช่วยในการวิเคราะห์แบบมีผู้สอนแบบการวิเคราะห์การถดถอย (Regression) ด้วยเทคนิค Linear Regression ได้
- 4. ตั้งสมมติฐานที่คาดว่าจะได้รับได้

- **70401.06 สื่อสารในระหว่างการดำเนินงาน**

- 1. รับและส่งสารในระหว่างการดำเนินงานได้
- 2. สอบถามและตอบข้อมูลในระหว่างการดำเนินงาน ได้
- 3. สามารถสื่อสารผ่านสื่อสารสนเทศช่องทางต่าง ๆ ได้

70402 เลือกเทคนิควิธีสำหรับการเรียนรู้แบบไม่มีผู้สอน (Select Modelling Techniques for Unsupervised learning)

- **70402.01 การจำแนกเทคนิคการวิเคราะห์แบบไม่มีผู้สอนแบบต่าง ๆ**

- 1. ระบุเทคนิคการวิเคราะห์แบบไม่มีผู้สอนแบบต่าง ๆ ได้
- 2. ระบุถึงแหล่งที่มาของข้อมูลที่เหมาะสมกับเทคนิคการวิเคราะห์แบบไม่มีผู้สอนได้
- 3. บอกความแตกต่างของคำตอบหรือผลลัพธ์ที่จะได้ระหว่างเทคนิคต่าง ๆ ที่จัดอยู่ในกลุ่มนี้ได้

- **70402.02 การเลือกใช้เทคนิคการวิเคราะห์แบบไม่มีผู้สอนแบบการแบ่งกลุ่มข้อมูล (Clustering) ด้วยเทคนิค k-Mean**

- 1. ระบุหลักการของเทคนิคการวิเคราะห์แบบไม่มีผู้สอนแบบการแบ่งกลุ่มข้อมูล (Clustering) ด้วยเทคนิค k-Mean ได้
- 2. บอกถึงผลลัพธ์ที่จะได้รับจากการสร้างสมการหรือแบบจำลองด้วยเทคนิค k-Mean ได้
- 3. กำหนดเครื่องมือช่วยในการวิเคราะห์แบบไม่มีผู้สอนแบบการแบ่งกลุ่มข้อมูล (Clustering) ด้วยเทคนิค k-Mean ได้
- 4. ตั้งสมมติฐานที่คาดว่าจะได้รับได้

- **70402.03 การเลือกใช้เทคนิคการวิเคราะห์แบบไม่มีผู้สอนแบบการแบ่งกลุ่มข้อมูล (Clustering) ด้วยเทคนิค Hierarchical**

- 1. ระบุหลักการของเทคนิคการวิเคราะห์แบบไม่มีผู้สอนแบบการแบ่งกลุ่มข้อมูล (Clustering) ด้วยเทคนิค Hierarchical ได้
- 2. บอกผลลัพธ์ที่จะได้รับจากการสร้างสมการหรือแบบจำลองด้วยเทคนิค Hierarchical ได้
- 3. กำหนดเครื่องมือช่วยในการวิเคราะห์แบบไม่มีผู้สอนแบบการแบ่งกลุ่มข้อมูล (Clustering) ด้วยเทคนิค Hierarchical ได้
- 4. ตั้งสมมติฐานที่คาดว่าจะได้รับได้

- **70402.04 การเลือกใช้เทคนิคการวิเคราะห์แบบไม่มีผู้สอนแบบการค้นหากฎความสัมพันธ์ (Association) ด้วยเทคนิค Apriori**

- 1. ระบุหลักการของเทคนิคการวิเคราะห์แบบไม่มีผู้สอนแบบการค้นหากฎความสัมพันธ์ (Association) ด้วยเทคนิค Apriori ได้
- 2. บอกถึงผลลัพธ์ที่จะได้รับจากการสร้างสมการหรือแบบจำลองด้วยเทคนิค Apriori ได้



- 3. กำหนดเครื่องมือช่วยในการวิเคราะห์แบบไม่มีผู้สอนแบบการค้นหากฎความสัมพันธ์ (Association) ด้วยเทคนิค Apriori ได้
- 4. ตั้งสมมติฐานที่คาดว่าจะได้รับได้

70403 สร้างแบบจำลองสำหรับเทคนิคการวิเคราะห์แบบมีผู้สอน (Build Model for Supervised learning)

- **70403.01 การกำหนดขั้นตอนต่าง ๆ ของแบบจำลอง**
 - 1. ระบุวิธีการสร้างและทดสอบแบบจำลองได้
 - 2. กำหนดขั้นตอน กระบวนการ ในการทดสอบแบบจำลองได้
- **70403.02 การแบ่งข้อมูลสำหรับชุดฝึกการเรียนรู้ (Training set, Validation and Test set)**
 - 1. ระบุวิธีการแบ่งข้อมูลเพื่อทำการทดสอบแบบจำลองได้
 - 2. อธิบายความแตกต่างของวิธีการแบ่งข้อมูลเพื่อทำการทดสอบแบบ Self-Consistency Test, Split Test และ Cross validation Test ได้
 - 3. กำหนดขั้นตอนการแบ่งข้อมูลเพื่อทำการทดสอบแบบจำลองโดยวิธี Self-Consistency Test, Split Test และ Cross validation Test ได้
 - 4. สามารถเลือกใช้วิธีการแบ่งข้อมูลเพื่อทำการทดสอบแบบจำลองกับข้อมูลได้
- **70403.03 การสร้างแบบจำลองโดยใช้เครื่องมือ (Tools) สำหรับการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification)**
 - 1. ใช้เครื่องมือช่วย (Tools) เพื่อสร้างแบบจำลองได้
 - 2. สามารถกำหนดค่าตัวแปรสำหรับในเครื่องมือช่วย (Tools) สำหรับเทคนิค การจำแนกข้อมูล (Classification) ที่เลือกใช้ได้
 - 3. จัดทำเอกสารรายงานประกอบการกำหนดค่าตัวแปรได้
 - 4. สามารถดำเนินการฝึกการเรียนรู้ (Training) กับข้อมูลที่เลือกได้
- **70403.04 การทดสอบแบบจำลองโดยใช้เครื่องมือ (Tools) สำหรับการวิเคราะห์แบบมีผู้สอนแบบการจำแนกข้อมูล (Classification)**
 - 1. สามารถดำเนินการตรวจสอบประสิทธิภาพ (Validation) เทคนิคที่เลือกได้
 - 2. สามารถดำเนินการทดสอบแบบจำลอง (Testing) เทคนิคที่เลือกได้
 - 3. สามารถตรวจสอบสถานะแบบจำลองที่เกิดจากการปรับแก้ปัญหา Overfitting หรือ Underfitting ได้
 - 4. สามารถปรับค่าตัวแปรของแบบจำลอง (regularization) ได้
- **70403.05 สร้างแบบจำลองโดยใช้เครื่องมือ (Tools) สำหรับการวิเคราะห์แบบมีผู้สอนแบบการวิเคราะห์ถดถอย (Regression)**
 - 1. ใช้เครื่องมือช่วย (Tools) เพื่อสร้างแบบจำลองได้
 - 2. สามารถกำหนดค่าตัวแปรสำหรับในเครื่องมือช่วย (Tools) สำหรับเทคนิคการวิเคราะห์แบบมีผู้สอนแบบการวิเคราะห์ถดถอย (Regression) ที่เลือกใช้ได้
 - 3. จัดทำเอกสารรายงานประกอบการกำหนดค่าตัวแปรได้
 - 4. สามารถดำเนินการฝึกการเรียนรู้ (Training) กับข้อมูลที่เลือกได้
- **70403.06 ทดสอบแบบจำลองโดยใช้เครื่องมือ (Tools) สำหรับการวิเคราะห์แบบมีผู้สอนแบบการวิเคราะห์ถดถอย (Regression)**
 - 1. สามารถดำเนินการตรวจสอบประสิทธิภาพ (Validation) เทคนิคที่เลือกได้
 - 2. สามารถดำเนินการทดสอบแบบจำลอง (Testing) เทคนิคที่เลือกได้



- 3. สามารถตรวจสอบสถานะแบบจำลองที่เกิดจากการปรับแก้ปัญหา Overfitting หรือ Underfitting ได้
- 4. สามารถปรับค่าตัวแปรของแบบจำลอง (regularization) ได้
- **70403.07 สรุปผลแบบจำลอง**
 - 1. บันทึกการตั้งค่าตัวแปรสำหรับเทคนิคที่เลือกใช้ได้
 - 2. อธิบายคุณลักษณะเฉพาะของแบบจำลองเพื่อใช้ต่อไปได้
 - 3. สามารถสรุปผลการใช้แบบจำลองกับข้อมูล และผลลัพธ์ที่ได้

70404 สร้างแบบจำลองสำหรับเทคนิคการวิเคราะห์แบบไม่มีผู้สอน (Build Model for Unsupervised learning)

- **70404.01 กำหนดขั้นตอนต่าง ๆ ของแบบจำลอง**
 - 1. ระบุนิธีการสร้างและทดสอบแบบจำลองได้
 - 2. กำหนดขั้นตอน กระบวนการ ในการทดสอบแบบจำลองได้
- **70404.02 สร้างแบบจำลองด้วยเครื่องมือ (Tools) สำหรับการแบ่งกลุ่มข้อมูล (Clustering)**
 - 1. ใช้เครื่องมือ (Tools) เพื่อสร้างและทดสอบแบบจำลองได้
 - 2. สามารถกำหนดค่าตัวแปรสำหรับในเครื่องมือ (Tools) สำหรับเทคนิคการแบ่งกลุ่มข้อมูล (Clustering) ที่เลือกใช้ได้
 - 3. จัดทำเอกสารรายงานประกอบการกำหนดค่าตัวแปรได้
 - 4. สามารถดำเนินการสร้างแบบจำลองการแบ่งกลุ่มข้อมูลได้
- **70404.03 ทดสอบแบบจำลองด้วยเครื่องมือ (Tools) สำหรับการแบ่งกลุ่ม**
 - 1. สามารถตรวจสอบประสิทธิภาพแบบจำลองได้
 - 2. สามารถปรับค่าตัวแปรของแบบจำลองได้
- **70404.04 สร้างแบบจำลองด้วยเครื่องมือ (Tools) สำหรับการค้นหากฎความสัมพันธ์ (Association)**
 - 1. ใช้เครื่องมือ (Tools) เพื่อสร้างและทดสอบแบบจำลองได้
 - 2. สามารถกำหนดค่าตัวแปรสำหรับในเครื่องมือ (Tools) สำหรับเทคนิคการค้นหากฎความสัมพันธ์ (Association) ที่เลือกใช้ได้
 - 3. จัดทำเอกสารรายงานประกอบการกำหนดค่าตัวแปรได้
 - 4. สามารถดำเนินการสร้างการค้นหากฎความสัมพันธ์ข้อมูลได้
- **70404.05 ทดสอบแบบจำลองด้วยเครื่องมือ (Tools) สำหรับการค้นหากฎความสัมพันธ์ (Association)**
 - 1. สามารถตรวจสอบประสิทธิภาพแบบจำลองได้
 - 2. สามารถปรับค่าตัวแปรของแบบจำลองได้
- **70404.06 สรุปผลแบบจำลอง**
 - 1. บันทึกการตั้งค่าตัวแปรสำหรับเทคนิคที่เลือกใช้ได้
 - 2. อธิบายและรายงานคุณลักษณะเฉพาะของแบบจำลองเพื่อใช้ต่อไปได้
 - 3. อธิบายพฤติกรรมของแบบจำลองที่เลือกใช้ได้
 - 4. สามารถวิเคราะห์ผลการใช้แบบจำลองกับข้อมูล และผลลัพธ์ที่ได้



อุปกรณ์ที่ต้องใช้สำหรับอบรมและประเมิน
อาชีพนักวิทยาศาสตร์ข้อมูล ระดับ 5 (Data Scientist)

| ลำดับ | อุปกรณ์ | จำนวน | หมายเหตุ |
|-------|--|------------------------|---|
| 1 | คอมพิวเตอร์ | 1 เครื่อง | เครื่องคอมพิวเตอร์ 1 เครื่องต่อผู้เข้ารับการประเมิน 1 คน <ul style="list-style-type: none">• มีหน่วยความจำไม่น้อยกว่า 4 GB• ติดตั้งระบบปฏิบัติการไม่ต่ำกว่า Windows 10 64-bit หรือ Window เวอร์ชันใหม่กว่า |
| 2 | ซอฟต์แวร์ที่ใช้ในการประเมินสำหรับการสร้างแบบจำลองประเภทต่าง ๆ สำหรับการวิเคราะห์ และภาษาที่ใช้เขียนโค้ด (Coding) | เลือกอย่างใดอย่างหนึ่ง | <ol style="list-style-type: none">1. ภาษาที่ใช้เขียนโค้ด (Coding) เช่น R, Python, Java เป็นต้น หรือ2. โปรแกรมการจัดเตรียมข้อมูล สร้างโมเดล และวิเคราะห์ข้อมูล เช่น RapidMiner หรือ3. โปรแกรมการจัดการข้อมูล ประมวลผล และทำเหมืองข้อมูล (Data Mining) เช่น Weka <p>หมายเหตุ : ผู้เข้าอบรมสามารถติดตั้งโปรแกรมอื่นเพิ่มเติมได้ หากต้องการพัฒนาโปรแกรมด้วยเครื่องมือ นอกเหนือจากที่กำหนด</p> |

* ศูนย์ทดสอบเลือก ซอฟต์แวร์ อย่างใดอย่างหนึ่ง ในการสอบประเมินภาคปฏิบัติ