

รายงานผลการอบรม/สัมมนา/พัฒนาความรู้/ประชุมเชิงปฏิบัติการ /และเป็นวิทยากร
กองสำรวจดินและวิจัยทรัพยากรดิน กรมพัฒนาที่ดิน

ส่วนที่ 1 ข้อมูลทั่วไป

คำนำหน้า นาย นาง นางสาว อื่น..... ชื่อ-นามสกุล พิชามณต์ อินตะโม
ตำแหน่ง นักสำรวจดินชำนาญการ..... กลุ่ม/ฝ่าย กลุ่มมาตรฐานการสำรวจจำแนกดิน
หลักสูตร/หัวข้อข้อมูลเรื่องอบรม/สัมมนา/พัฒนาความรู้ การใช้โปรแกรมดิจิทัลเพื่อการวิเคราะห์ข้อมูล
สถานที่อบรม / สัมมนา /พัฒนาความรู้ E-learning สำนักงาน.ก.พ.
ตั้งแต่วันที่ 1 เมษายน 2567 ถึงวันที่ 16 สิงหาคม 2567

ส่วนที่ 2 สิ่งที่ได้รับจากการอบรม/สัมมนา/พัฒนาความรู้

2.1 รายงานสรุปเนื้อหาในการการอบรม

2.1.1 การวิเคราะห์ข้อมูลเบื้องต้น หลักการวิเคราะห์ข้อมูล กระบวนการที่นำมาค้นหาข้อมูลเชิงลึก ที่นำมาช่วยในการตัดสินใจ ทั้งในระดับองค์กร และการใช้ในชีวิตประจำวันได้

(1) ขั้นตอนของการวิเคราะห์ ระบุปัญหา (define the problem) รวบรวมข้อมูล (collect the data) การเตรียมข้อมูลเบื้องต้น (pre-process) วิเคราะห์ข้อมูล (analyze) ตีความหมายของผลลัพธ์ (interpret the data)

(2) เครื่องมือในการวิเคราะห์ข้อมูล เช่น excel และ google sheet ใช้ในการวิเคราะห์ข้อมูลอย่างง่าย SPSS Strata และ SAS ใช้ในการวิเคราะห์ข้อมูลทางสถิติเชิงลึก กลุ่มเครื่องมือในการทำเหมืองข้อมูล (data mining tools) ประกอบด้วยโปรแกรม RapidMiner กลุ่ม Business intelligence tools เช่น Power BI และ Tableau เป็นเครื่องมือที่นำมาสร้างกราฟการแสดงผลในรูปแบบของ dashboard ที่สามารถมีกราฟหลายๆ อัน และสามารถเชื่อมโยงกันได้ และกลุ่ม Programming languages เช่น โปรแกรม R และ Python ใช้วิเคราะห์ข้อมูลเชิงลึกที่ต้องมีการเขียนคำสั่งในการวิเคราะห์ข้อมูล เป็นเครื่องมือที่ไม่มีค่าใช้จ่าย

2.1.2 โปรแกรม RapidMiner คือซอฟต์แวร์ data science ใช้สำหรับการเตรียมข้อมูล การเรียนรู้เครื่องมือ การเรียนรู้เชิงลึก การทำเหมืองข้อมูลและการวิเคราะห์การทำนาย (predictive analysis) เป็นซอฟต์แวร์ที่ช่วยในการจัดส่งข้อมูล และลดข้อผิดพลาดโดยไม่ต้องเขียนโค้ดเพิ่ม RapidMiner มีขั้นตอนพร้อมสำหรับการทำ data mining และ machine learning ซึ่งรวมไปถึงการไหลและการแปลงข้อมูล การประมวลผลล่วงหน้าและการวาดภาพจากข้อมูล การวิเคราะห์เชิงพยากรณ์และการสร้างแบบจำลองทางสถิติ ซึ่งมีการจัดการข้อมูลเบื้องต้น ประกอบด้วยขั้นตอนการทำงานย่อยที่เปลี่ยนข้อมูลดิบให้กลายเป็นความรู้ ประกอบด้วย ขั้นตอนดังนี้

- Data cleaning เป็นขั้นตอนสำหรับการคัดข้อมูลที่ไมเกี่ยวข้องออกไป
- Data integration เป็นขั้นตอนการรวมข้อมูลที่มีหลายแหล่งให้เป็นข้อมูลชุดเดียวกัน
- Data selection เป็นขั้นตอนการดึงข้อมูลสำหรับการวิเคราะห์จากแหล่งที่บันทึกไว้
- Data transformation เป็นขั้นตอนการแปลงข้อมูลให้เหมาะสมสำหรับการใช้งาน
- Data mining เป็นขั้นตอนการค้นหารูปแบบที่เป็นประโยชน์จากข้อมูลที่มีอยู่
- Pattern evaluation เป็นขั้นตอนการประเมินรูปแบบที่ได้จากการทำเหมืองข้อมูล
- Knowledge representation เป็นขั้นตอนการนำเสนอความรู้ที่ค้นพบ โดยใช้เทคนิคในการนำเสนอเพื่อให้เข้าใจง่าย

2.1.3 การใช้โปรแกรม RapidMiner การทำงานของ RapidMiner เป็นแบบ drag and drop สำหรับการวิเคราะห์ในขั้นพื้นฐาน RapidMiner ให้ผู้ใช้งานได้ออกแบบ workflow ในการวิเคราะห์ข้อมูลในรูปแบบของรูปภาพ ซึ่งทุกอย่างจะเกิดขึ้น

ในหน้า design view การติดตั้ง RapidMiner Studio สามารถดาวน์โหลด จากการค้นหาใน www.google.com แล้วใช้คำว่า “download RapidMiner Studio” RapidMiner มีองค์ประกอบที่สำคัญ คือ

(1) Repository มีลักษณะคล้าย folder ในคอมพิวเตอร์ ซึ่งเป็นพื้นที่จัดเก็บข้อมูล process และผลลัพธ์

(2) Operators เป็นองค์ประกอบสำคัญ โดยจะเป็นตัวการในการทำงานต่อจาก input ที่ได้รับ ซึ่งเหล่า operators มีการทำงานที่มากกว่าหลายร้อยเลยทีเดียว ตั้งแต่การทำความสะอาดข้อมูล cleansing การทำโมเดล modeling รวมไปถึงการ blend ข้อมูลเข้าด้วยกัน

(3) Ports เป็นส่วนของ output ของ operators การเชื่อมต่อ port ต้องตรวจสอบก่อนว่า output กับ input ของ Operator อีกตัวหนึ่งเข้ากันได้

(4) Process คือ workflow ที่เชื่อมต่อ operators ต่างๆ และสามารถนำไปวิเคราะห์ข้อมูลได้ด้วยอย่าง process ที่เข้าใจง่ายคือ

(5) Parameters พารามิเตอร์นี้เปรียบเสมือนการตั้งค่าการทำงานของแต่ละ operators ซึ่งผู้ใช้งานจะสามารถกำหนดค่าของแต่ละ operators ที่แตกต่างกัน

2.2 ประโยชน์ที่ได้รับ/ประยุกต์ใช้กับหน่วยงาน

1. โปรแกรม RapidMiner เป็นโปรแกรมที่ใช้งานง่าย สามารถนำมาประยุกต์ใช้ในการจัดการข้อมูลงานวิจัยเพื่อให้ได้ผลวิเคราะห์ที่ถูกต้อง แม่นยำ และมีการแสดงผลที่เข้าใจง่าย เหมาะกับโครงการที่ได้รับมอบหมายในปัจจุบัน

2. ได้อัพเดทความรู้ทางด้านโปรแกรมที่ใช้วิเคราะห์ข้อมูลเพิ่มขึ้นอีก 1 โปรแกรม ซึ่งทำให้สามารถเลือกใช้โปรแกรมการวิเคราะห์ข้อมูลได้หลากหลายมากยิ่งขึ้น

3. จากการเรียนรู้เรื่องการใช้เครื่องมือดิจิทัลในการจัดการข้อมูลนั้น ยิ่งทำให้รู้ว่าพัฒนาการด้านข้อมูลมีความก้าวหน้าไปไวมาก ซึ่งช่วยกระตุ้นให้ผู้เรียนได้ตระหนักและเรียนรู้สิ่งใหม่ๆ อยู่เสมอ เพื่อให้สามารถก้าวไปพร้อมกับเครื่องมือต่างๆ ที่พัฒนาขึ้นในปัจจุบัน

2.3 ปัญหาและอุปสรรคในการอบรม

ไม่มีอุปสรรคในการเข้ารับการอบรม

2.4 ข้อคิดเห็นและข้อเสนอแนะ

ไม่มีข้อเสนอแนะ

ลงชื่อ..... *ปิณฑุ์ อินตะโม*

(นางสาวพิชามณูญ์ อินตะโม)

ตำแหน่ง นักสำรวจดินชานาญการ

ผู้รายงาน

วันที่ 16 สิงหาคม 2567

ส่วนที่ 3 ความเห็นผู้บังคับบัญชา

ทราบ

ลงชื่อ..... *กฤษณ์ ดวงกมล*

(นายกฤษณ์ โสภณ ดวงกมล)

ตำแหน่ง ผู้อำนวยการกลุ่มมาตรฐานการสำรวจจำแนกดิน

รายงานผลการอบรม/สัมมนา/พัฒนาความรู้/ประชุมเชิงปฏิบัติการ /และเป็นวิทยากร
กองสำรวจดินและวิจัยทรัพยากรดิน กรมพัฒนาที่ดิน

ส่วนที่ ๑ ข้อมูลทั่วไป

คำนำหน้า นาย นาง นางสาว อื่น..... ชื่อ-นามสกุล.....นางสาวสุวิทย์.สกุลเจริญมิตร.....
ตำแหน่ง.....นักสำรวจดินปฏิบัติการ.....กลุ่ม/ฝ่าย.....กลุ่มมาตรฐานการสำรวจจำแนกดิน.....
หลักสูตร/หัวข้อข้อมูลเรื่องอบรม/สัมมนา/พัฒนาความรู้ฯ.....การใช้โปรแกรมดิจิทัลเพื่อการวิเคราะห์ข้อมูล.....
สถานที่อบรม / สัมมนา / พัฒนาความรู้ฯ e-learning.สำนักงาน.ภ.พ.....
ตั้งแต่วันที่๑๙.กรกฎาคม.๒๕๖๗.....ถึงวันที่๑๓.สิงหาคม.๒๕๖๗.....

ส่วนที่ ๒ สิ่งที่ได้รับจากการอบรม/สัมมนา/พัฒนาความรู้

๒.๑ รายงานสรุปเนื้อหาในการการอบรม

Data Mining คือ การทำเหมืองข้อมูล เก็บรวบรวมข้อมูลเพื่อวิเคราะห์ออกมาเป็นชุดข้อมูลที่เอามาใช้ประโยชน์กับองค์กร โดยผ่านกระบวนการคัดแยก แบ่งกลุ่ม และกระบวนการอื่น ๆ เพื่อจะได้ชุดข้อมูลที่เหมาะสมกับการนำไปใช้งานสำหรับมาตรฐานการทำ Data Mining มีขั้นตอน ดังนี้

๑. วิเคราะห์วัตถุประสงค์ว่าจะนำผลลัพธ์ของข้อมูลนี้ไปใช้เพื่ออะไร เพื่อให้การดำเนินการเป็นไปอย่างราบรื่น ไม่ผิดวิธีและเสียเวลาไปโดยไม่ตอบโจทย์

๒. เก็บข้อมูล เมื่อทราบวัตถุประสงค์แล้วว่ามีเป้าหมายอะไรในการทำ Data Mining ต้องทำการเก็บข้อมูล โดยข้อมูลที่มีคุณภาพนั้นต้องเป็นข้อมูลที่แม่นยำและน่าเชื่อถือ โดยที่ไม่ขัดต่อหลักความปลอดภัย ความเป็นส่วนตัว และข้อจำกัดอื่น ๆ ซึ่งบางครั้งข้อมูลมีจำนวนเยอะ และมีขนาดใหญ่มาก กลุ่มก้อนข้อมูลเหล่านี้จะถูกเรียกว่า Big Data เปรียบได้กับเหมืองขนาดใหญ่ที่มีแร่หลายชนิด

๓. เตรียมข้อมูล กระบวนการนี้คือการนำข้อมูลที่มีอยู่มากัดเฉพาะส่วนที่จำเป็นต้องใช้ในประเด็นย่อยต่าง ๆ เมื่อสกัดข้อมูลมาได้แล้วนำมาแบ่งหมวดหมู่และจัดเรียงภายใต้มาตรฐานเดียวกัน ตรวจสอบข้อผิดพลาด จนถึงประเมินขนาดของข้อมูลก่อนนำไปสู่ขั้นตอนถัดไป

๔. เจียระไนข้อมูล หลังจากเตรียมข้อมูลเสร็จ นำข้อมูลเหล่านั้นมาทำให้พร้อมนำไปใช้งานได้มากขึ้นด้วยการเข้าโปรแกรมประมวลผลผ่าน Algorithm ที่ซับซ้อนเพื่อทำให้เห็นว่าข้อมูลเหล่านี้มีความสัมพันธ์หรือรูปแบบอย่างไร

๕. นำเสนอข้อมูล ผลลัพธ์ที่ได้จะถูกนำมาสร้างเป็นโมเดลที่ทำได้คนทั่วไปเห็นภาพและเข้าใจข้อมูลมากขึ้น

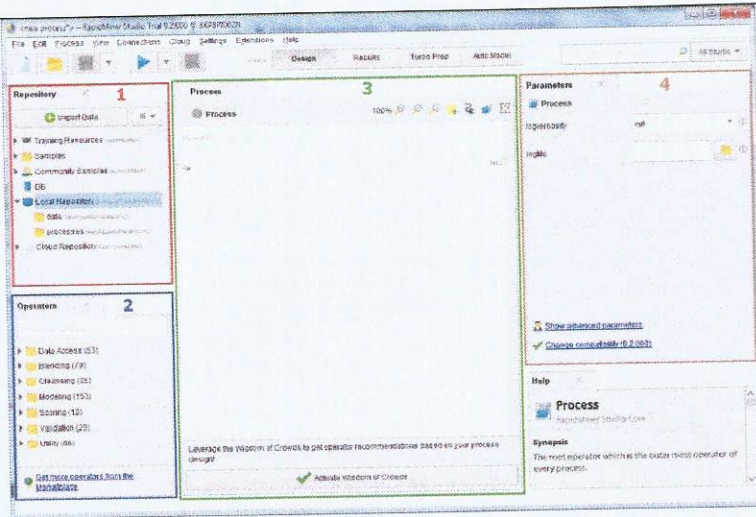
Rapidminer คือซอฟต์แวร์ Data Science ใช้สำหรับการเตรียมข้อมูล และการวิเคราะห์การทำนาย สามารถใช้ในการทำ Data mining (ชุดข้อมูล) และ Machine learning ซึ่งรวมไปถึงการโหลดและการแปลงข้อมูล (ETL) การประมวลผลล่วงหน้าและการวาดภาพจากข้อมูล การวิเคราะห์เชิงพยากรณ์และการสร้างแบบจำลองทางสถิติ การประเมินผลและการปรับใช้ต่าง ๆ ประกอบไปด้วยส่วนสำคัญ ดังนี้

๑) Repositor: ส่วนที่ใช้ในการจัดการไฟล์ต่าง ๆ ของ RapidMiner โดยจะเก็บไฟล์ข้อมูล ไว้ใน Folder เพื่อความสะดวกในการเรียกใช้งานครั้งถัดไป

๒) Operators: ส่วนที่เก็บ Operators ในการใช้งานต่าง ๆ ไว้เป็นกลุ่ม ซึ่งสามารถค้นหา Operators ที่ต้องการได้ในช่อง Search เพื่อความสะดวกในการเรียกใช้งาน

๓) Process: ส่วนที่สำคัญของ RapidMiner เพราะเป็นการนำเอา Operators ต่าง ๆ มาเชื่อมกันให้เป็น Process ขึ้นมาใช้งาน

๔) Parameters: ส่วนที่เป็นส่วนที่แสดงพารามิเตอร์ (parameter) ที่เกี่ยวข้องกับแต่ละ Operator เพื่อให้ผู้ใช้งานปรับแต่งตามที่ต้องการ



ภาพที่ ๑ โปรแกรม RapidMiner

ต้นไม้ตัดสินใจ (Decision Tree) เป็นหนึ่งในวิธีการเรียนรู้ซึ่งใช้ในสถิติ การเรียนรู้ของเครื่อง (machine learning) และการทำเหมืองข้อมูล เป็นโมเดลทางคณิตศาสตร์ที่ใช้ทำนายประเภทของวัตถุโดยพิจารณาจากลักษณะของวัตถุ ก้านภายใน (inner node) ของต้นไม้จะแสดงตัวแปร ส่วนกิ่งจะแสดงค่าที่เป็นไปได้ของตัวแปร ส่วนใบ (leaf node) จะแสดงประเภทของวัตถุ ต้นไม้ตัดสินใจสร้างขึ้นเพื่อช่วยการตัดสินใจเพื่อใช้ในการสร้างแผนงาน และนิยมใช้มากในการบริหารความเสี่ยง (risk management) ซึ่งต้นไม้ตัดสินใจเป็นส่วนหนึ่งของทฤษฎีการตัดสินใจ ซึ่ง เราสามารถใช้ Rapidminer สร้างต้นไม้ตัดสินใจวิเคราะห์การทำนายในวัตถุประสงค์ต่าง ๆ ได้

๒.๒ ประโยชน์ที่ได้รับ/ประยุกต์ใช้กับหน่วยงาน

.....สามารถนำความรู้มาเป็นแนวทางในพัฒนาระบบฐานการเก็บข้อมูลในองค์กร ช่วยให้องค์กรใช้ทรัพยากรได้อย่างมีประสิทธิภาพมากขึ้น

๒.๓ ปัญหาและอุปสรรคในการอบรม

.....เป็นเครื่องมือใหม่ที่นำมาใช้ในการจัดการข้อมูล ต้องใช้ทักษะและความชำนาญในการใช้เครื่องมือ

๒.๔ ข้อคิดเห็นและข้อเสนอแนะ

.....เนื้อหาที่มีความน่าสนใจ เหมาะกับการนำมาใช้กับองค์กรในยุคดิจิทัล

ลงชื่อ.....^{ศรีวิทย์}
 (.....นางสาวสุภัทรา สกลอารีมิตร) ₁
 ตำแหน่ง นักสำรวจดินปฏิบัติการ
 ผู้รายงาน
 วันที่ 20 ก.ค. 67

ส่วนที่ ๓ ความเห็นผู้บังคับบัญชา

ทราบ

ลงชื่อ.....^{มนตรี}
 (.....นายฤทธิโสภณ ดวงกลม)

ตำแหน่ง ผู้อำนวยการกลุ่มมาตรฐานการสำรวจจำแนกดิน