

รายงานสรุปเนื้อหาและการนำไปใช้ประโยชน์จากการเข้าอบรม สัมมนา หรือประชุมวิชาการ

ข้าพเจ้า นางหนึ่งหทัย ชัยอากร ตำแหน่ง ผู้ช่วยศาสตราจารย์ สังกัด สาขาวิชาสถิติคณะวิทยาศาสตร์ ขอเสนอรายงานสรุปเนื้อหาและการนำไปใช้ประโยชน์จากการเข้าอบรมเชิงปฏิบัติการเรื่อง หลักสูตรการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining โดยซอฟต์แวร์ RapidMiner Studio ๗ (ขั้นพื้นฐานและปานกลาง) รุ่นที่ ๑๒ เมื่อวันที่ ๑๗ มีนาคม ๒๕๕๙ ถึงวันที่ ๑๙ มีนาคม ๒๕๕๙ ณ โรงแรม เค ยู โฮม มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตบางเขน ตามหนังสือขออนุญาตเดินทางไปราชการ เลขที่ ศธ. ๐๕๒๓.๔.๓/๒๓ ลงวันที่ ๒๑มกราคม ๒๕๕๙ ซึ่งการเข้าร่วมการอบรมปฏิบัติการดังกล่าวข้าพเจ้าได้เลือกใช้งบประมาณการพัฒนาบุคลากรตามกรณีที่ ๔ ดังนั้นจึงขอเสนอสรุปเนื้อหาและการนำไปใช้ประโยชน์ของสัมมนาดังต่อไปนี้

จากการที่เข้าร่วมการอบรมเชิงปฏิบัติการ เรื่อง หลักสูตรการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining โดยซอฟต์แวร์ RapidMiner Studio ๗ (ขั้นพื้นฐานและปานกลาง) รุ่นที่ ๑๒ ณ โรงแรม เค ยู โฮม มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตบางเขน ระหว่างวันที่ ๑๗ มีนาคม ๒๕๕๙ ถึงวันที่ ๑๙ มีนาคม ๒๕๕๙ สามารถสรุปเป็นด้าน ๆ ดังนี้

ด้านเนื้อหา

การอบรมเชิงปฏิบัติการดังกล่าว แบ่งเนื้อหาออกเป็นหลัก ๆ ได้ดังนี้

ดร.เอกสิทธิ์ พัทธวงศ์ศักดิ์ ซึ่งเป็นวิทยากรประจำการอบรมครั้งนี้ ได้บรรยายและนำฝึกปฏิบัติ ดังนี้ ในการอบรมเชิงปฏิบัติการในครั้งนี้มีเนื้อหาที่เกี่ยวข้องของจำนวนทั้งหมด ๙ เนื้อหา ได้แก่

บทที่ ๑ Introduction to Data Mining เป็นการบอกถึงวัตถุประสงค์ของ Data Mining ว่าเกี่ยวกับการวิเคราะห์ข้อมูลเพื่อหารูปแบบ (patterns) หรือความสัมพันธ์ (relation) ระหว่างข้อมูลในฐานข้อมูลขนาดใหญ่ และเป็นกระบวนการดึงข่าวสารที่น่าสนใจ และมีประโยชน์มาใช้

บทที่ ๒ Introduction to CRISP-DM Cross-Industry Standard Process for Data Mining เป็นโปรแกรมที่ถูกพัฒนาขึ้นโดย ๓ บริษัทได้แก่ SPSS DaimlerChrysler และ NCR ประกอบด้วย ๖ ขั้นตอน คือ ๑) Business Understanding ๒) Data Understanding ๓) Data Preparation ๔) Modeling ๕) Evaluation ๖) Deployment

บทที่ ๓ Introduction to RapidMiner Studio ๗ เป็นการกล่าวถึงโปรแกรมนี้ว่า เป็นซอฟต์แวร์ไม่ต้องเสียค่าใช้จ่าย (Open source software) และเป็นซอฟต์แวร์ที่ได้รับความนิยมในการทำดาต้าไมนิง

บทที่ ๔ Pre -Processing เป็นการเตรียมข้อมูล หรือขั้นตอนการเตรียมข้อมูล ก่อนนำมาวิเคราะห์ เป็นการจัดการข้อมูลต่างๆ เช่น นำข้อมูลจากไฟล์ต่างๆ เข้ามาใช้งาน Import Data และ Data Management เป็นต้น

บทที่ ๕ Association Rules กฎการหาความสัมพันธ์ ในการหากฎความสัมพันธ์มี ๒ ขั้นตอนใหญ่ เทคนิคที่นิยมใช้คือเทคนิค Apriori และ เทคนิค FP-Growth

บทที่ ๖ Clustering การแบ่งกลุ่มข้อมูล ข้อมูลที่มีลักษณะคล้ายๆกัน อยู่กลุ่มเดียวกัน และข้อมูลที่อยู่คนละกลุ่มจะมีลักษณะที่แตกต่างกันมาก แต่ละกลุ่มจะเรียกว่า คลัสเตอร์ มีหลายเทคนิค และ Clustering validity เป็นการวัดประสิทธิภาพของ Clustering เพื่อดูว่าเทคนิคใดสามารถทำให้การแบ่งกลุ่มมีประสิทธิภาพสูงที่สุด และควรจัดข้อมูลออกมาเป็นกี่กลุ่ม

บทที่ ๗ Classification การจำแนกประเภทข้อมูล เป็นการนำข้อมูลเดิมที่มีคำตอบที่เราสนใจ มาสร้างเป็นโมเดล เพื่อหาคำตอบให้กับข้อมูลใหม่ การประมาณค่าข้อมูล (regression) การสร้างโมเดลและการวัดประสิทธิภาพของโมเดล โดยการดูค่าความแม่นยำว่า โมเดลใดให้ความแม่นยำในการทายข้อมูลได้ถูกมากที่สุด

บทที่ ๘ Introduction to Text Mining ข้อมูลทั่วไปจะแบ่งเป็นมีโครงสร้าง (Structure) และไม่มีโครงสร้าง (Unstructure) ดังนั้นข้อมูลส่วนใหญ่จะอยู่ในรูปของข้อความ รูปภาพ และเสียง ดังนั้นหัวข้อนี้จึงเป็นประโยชน์อย่างมากในการจัดการข้อมูลประเภทนี้

บทที่ ๙ Introduction to Image Mining เป็นการจัดการข้อมูลที่มีลักษณะที่เป็นรูปภาพ โดยการใช้การแยกสี ทำให้จัดเป็นหมวดหมู่ของรูปภาพชัดเจน

ด้านการนำไปใช้ประโยชน์

จากการอบรมเชิงปฏิบัติการเกี่ยวกับหลักสูตรการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining โดยซอฟต์แวร์ RapidMiner Studio ๗ (ขั้นพื้นฐานและปานกลาง) รุ่นที่ ๑๒ สามารถนำไปใช้ประโยชน์ในด้านต่างๆ ได้ดังนี้

ด้านโปรแกรม เป็นซอฟต์แวร์ไม่ต้องเสียค่าใช้จ่าย (Open source software) ดังนั้นจึงสะดวกในการใช้งาน และในปัจจุบันข้อมูลต่างๆ มีค่อนข้างมาก ดังนั้นเทคนิค Data Mining จึงถูกนำมาใช้อย่างแพร่หลายในการจัดการข้อมูล ซึ่งโปรแกรมสามารถทำการจัดกลุ่มข้อมูลได้ โดยมีเทคนิคหลายวิธีและแต่ละวิธีก็จะมีการวัดประสิทธิภาพให้เหมาะสมกับข้อมูลนั้น ทั้งยังสามารถทำการทำนายข้อมูลในอนาคตได้ ด้วยวิธี Classification ซึ่งมีเทคนิคหลายวิธีและแต่ละวิธีก็จะมี การวัดประสิทธิภาพออกมาเป็น เปอร์เซนต์ความถูกต้องของข้อมูล

ด้านการเรียนการสอนและการวิจัยสามารถนำโปรแกรม และการอบรมเชิงปฏิบัติการดังกล่าวมาใช้ประกอบการเรียนการสอนในรายวิชา Multivariate และใช้ในการวิเคราะห์งานวิจัยได้ค่อนข้างมาก

ณัทพร นพรัตน์

(นางหนึ่งทัย ชัยอาร)

๒๓ มีนาคม ๒๕๕๙

ความคิดเห็นของผู้บังคับบัญชาชั้นต้น(ประธานหลักสูตร/เลขานุการคณะ/หัวหน้างาน)

မှတ်ချက်များ - ပုံစံအတိုင်း ဖြည့်စွက်ရန်
အောက်ပါအတိုင်း ဖြည့်စွက်ရန်

Pr. 21
(o. n. m. n. e. a. n. d. i. n. t. s.)
23 / 24 / 59

ความคิดเห็นของคุณบดีคณะวิทยาศาสตร์หรือผู้แทน

(.....)

...../...../.....

หมายเหตุ : แบบฟอร์มเป็นรูปแบบเพื่อเสนอการรายงาน เนื้อหาอาจไม่เพียงพอสำหรับการกรอกข้อมูล
สามารถขยายหรือเพิ่มเติมตามความเหมาะสม

ภาพการเข้าร่วมการอบรมเชิงปฏิบัติการ

เรื่อง “หลักสูตรการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining โดยซอฟต์แวร์ RapidMiner Studio ๓/ (ชั้น
พื้นฐานและปานกลาง) รุ่นที่ ๑๒”

วันที่ ๑๓ มีนาคม ๒๕๕๙ ถึงวันที่ ๑๙ มีนาคม ๒๕๕๙

ณ โรงแรม เค ยู โฮม มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตบางเขน



(data)³

base | warehouse | mining
Store. Analyze. Predict.

Data Cube

ordinary partnership

presents this certificate of completion to

NEUNGHATAI CHAIARPORN

in recognition of successfully completing the

Practical Data Mining with RapidMiner Studio 7
(batch 12)

17-19 March 2016

E. Pacharawongsakda

EAKASIT PACHARAWONGSAKDA, Ph.D.

Co-founder of Data Cube

DATAcube-11-1603-074

certificate number