



# บันทึกข้อความ

สำนักดิจิทัลเพื่อการศึกษา	1137
วันที่รับ	17 ส.ค. 2565
ผู้รับ	<i>[Signature]</i>

ส่วนราชการ สำนักดิจิทัลเพื่อการศึกษา โทร. ๕๙๒๔ โทรสาร. ๕๙๒๔

ที่ อว.๐๖๑๒.๑๖.๐๑.๐๒/๐๗๗ วันที่ ๑๖ สิงหาคม ๒๕๖๕

เรื่อง รายงานผลการเข้าร่วมอบรม Practical Data Science and Machine Learning (Workshop with RapidMiner TurboPrep and AutoModel) รุ่นที่ ๕

เรียน ผู้อำนวยการสำนักดิจิทัลเพื่อการศึกษา

ตามที่ Data Cube by Cube Analytics Consulting Co.,Ltd ได้จัดอบรมหลักสูตร Practical Data Science and Machine Learning with RapidMiner TurboPrep and AutoModel รุ่นที่ ๕ ผ่านระบบ Zoom ในวันที่ ๒๕ - ๒๗ กรกฎาคม ๒๕๖๕ เวลา ๐๙.๐๐ - ๑๖.๓๐ น. โดยมีผู้ช่วยศาสตราจารย์ ดร. กัลยา ใจรักษ์ พร้อมบุคลากรในสำนักดิจิทัลเพื่อการศึกษา ได้เข้าร่วมอบรมดังกล่าว

บัดนี้การอบรมได้เสร็จสิ้นแล้ว จึงขอส่งรายงานการเข้าร่วมอบรม ดังเอกสารแนบมาพร้อมนี้

จึงเรียนมาเพื่อโปรดพิจารณา

*[Signature]*

(นายธีระพงษ์ ใจคำมา)

นักวิชาการคอมพิวเตอร์ชำนาญการ

เรียน ผู้อำนวยการสำนักดิจิทัลเพื่อการศึกษา

เพื่อโปรดทราบ

เพื่อโปรดพิจารณา

เรียนควร.....

*[Signature]*

ลงชื่อ.....

นางสาวนันทวี คุณศิริขันธ์

เจ้าหน้าที่บริหารงานทั่วไป  
17 ส.ค. 2565

*[Signature]*

*[Signature]*

นางสาวอ้ออานาจ ไกรวรรณ

ผู้อำนวยการสำนักดิจิทัลเพื่อการศึกษา

*[Signature]* 17 ส.ค. 2565

นางชนมาตุ/เปี่ยมภค  
17 ส.ค. 2565

## รายงานผลการเข้าร่วม อบรม/สัมมนา

เรียน ผู้อำนวยการสำนักดิจิทัลเพื่อการศึกษา

ชื่อ - นามสกุล นายธีระพงษ์ ใจคำมา ตำแหน่ง นักวิชาการคอมพิวเตอร์ชำนาญการ

หัวข้อการอบรม/สัมมนา Practical Data Science and Machine Learning (Workshop with RapidMiner TurboPrep and AutoModel) รุ่นที่5

วัน / เดือน / ปี 25 - 27 กรกฎาคม 2565

ผู้จัดการอบรม/สัมมนา

สถานที่อบรม/สัมมนา อบรมออนไลน์ผ่านโปรแกรม Zoom Meeting

### รายละเอียดและกำหนดการ การเข้าร่วมอบรม/สัมมนา

วันที่ 25 กรกฎาคม 2565	
09.00-10.30 น.	<ul style="list-style-type: none"> <li>● แนะนำการวิเคราะห์ข้อมูลด้วยเทคนิค Data Science และ Machine Learning และการใช้ประโยชน์ในด้านต่างๆ</li> <li>● อธิบายแนวคิดพื้นฐานในการเก็บข้อมูลในลักษณะต่างๆ ได้แก่                             <ul style="list-style-type: none"> <li>◆ พื้นฐานการเก็บข้อมูลในรูปแบบฐานข้อมูล (Database)</li> <li>◆ พื้นฐานการเก็บข้อมูลในรูปแบบคลังข้อมูล (Data Warehouse)</li> </ul> </li> <li>● แนะนำกระบวนการ CRISP-DM เบื้องต้นสำหรับการวิเคราะห์ข้อมูล</li> </ul>
10.30-10.45 น.	พักเบรก
10.45-12.00 น.	<ul style="list-style-type: none"> <li>● แนะนำส่วนต่างๆ ของซอฟต์แวร์ RapidMiner Studio 9</li> <li>● การนำข้อมูลไฟล์ Excel, CSV เข้ามาใช้ใน RapidMiner Studio 9</li> <li>● ลักษณะของแอตทริบิวต์ (attribute) ต่างๆ ในชุดข้อมูล</li> </ul>
12.00-13.00 น.	พักรับประทานอาหารกลางวัน
13.00-14.30 น.	<ul style="list-style-type: none"> <li>● การจัดการข้อมูลด้วย RapidMiner Turbo Prep                             <ul style="list-style-type: none"> <li>◆ การเลือกแอตทริบิวต์ (attribute) ที่ต้องการ</li> <li>◆ การเลือกข้อมูลตัวอย่างตามเงื่อนไขที่กำหนด</li> <li>◆ การ replace ค่าที่ผิดพลาด</li> <li>◆ การแก้ไขค่าว่าง (Replace Missing Value) แบบต่างๆ</li> <li>◆ การสร้างแอตทริบิวต์หรือตัวแปรใหม่ (Generate Attributes)</li> <li>◆ การแบ่งช่วงข้อมูล (Discretization) แบบต่างๆ</li> <li>◆ การ join ข้อมูล 2 ชุดเข้าด้วยกัน</li> <li>◆ การสร้างตาราง Pivot Table</li> </ul> </li> </ul>
14.30-14.45 น.	พักเบรก

14.45-16.30 น.	<ul style="list-style-type: none"> <li>● Workshop การจัดการข้อมูล transaction การซื้อสินค้า online ด้วย RapidMiner TurboPrep</li> <li>● อธิบายแนวคิดของการวิเคราะห์ข้อมูลเพื่อสนับสนุนการตัดสินใจแบบต่างๆ ได้แก่ <ul style="list-style-type: none"> <li>◆ การวิเคราะห์เชิงพรรณนา (Descriptive Analytics)</li> <li>◆ การวิเคราะห์เชิงวินิจฉัย (Diagnostic Analytics)</li> <li>◆ การวิเคราะห์เชิงพยากรณ์ (Predictive Analytics)</li> <li>◆ การวิเคราะห์เชิงแนะนำ (Prescriptive Analytics)</li> </ul> </li> </ul>
<b>วันที่ 26 กรกฎาคม 2565</b>	
09.00-10.30 น.	<ul style="list-style-type: none"> <li>● พื้นฐานทางสถิติ (Statistics) และความน่าจะเป็น (Probability) <ul style="list-style-type: none"> <li>◆ การวัดค่าสถิติต่างๆ เช่น การหาค่าเฉลี่ย (Mean) การหาค่ากลาง (Median) การหาค่าฐานนิยม (Mode)</li> <li>◆ การคำนวณค่าเบี่ยงเบนมาตรฐาน (Standard Deviation)</li> <li>◆ Rank Order Statistics</li> <li>◆ การหา Outlier ด้วยวิธีการ IQR (Inter Quartile Range)</li> </ul> </li> <li>● พื้นฐานความน่าจะเป็น (Probability) <ul style="list-style-type: none"> <li>◆ การแสดงกราฟ (Chart) ในรูปแบบต่างๆ ได้แก่</li> <li>◆ กราฟแท่ง (Bar Chart)</li> <li>◆ กราฟวงกลม (Pie Chart)</li> <li>◆ กราฟเส้น (Line Chart)</li> <li>◆ การกระจายตัว (Scatter Plot)</li> </ul> </li> </ul>
10.30-10.45 น.	พักเบรก
10.45-12.00 น.	<ul style="list-style-type: none"> <li>● การแบ่งกลุ่มข้อมูลด้วยวิธีการ RFM Segmentation</li> <li>● ตัวอย่างการแบ่งกลุ่ม RFM Segmentation ด้วย RapidMiner TurboPrep</li> <li>● Workshop การแบ่งกลุ่ม RFM Segmentation กับข้อมูล online transactions ด้วย RapidMiner TurboPrep</li> </ul>
12.00-13.00 น.	พักรับประทานอาหารกลางวัน
13.00-14.30 น.	<ul style="list-style-type: none"> <li>● แนวคิดพื้นฐานของเทคนิค Machine Learning</li> <li>● เทคนิค Machine Learning ประเภทต่างๆ ได้แก่ <ul style="list-style-type: none"> <li>◆ Unsupervised Learning หรือการเรียนรู้แบบไม่มีผู้สอน</li> <li>◆ Supervised Learning หรือการเรียนรู้แบบมีผู้สอน</li> <li>◆ Semi-supervised Learning หรือการเรียนรู้แบบกึ่งมีผู้สอน</li> </ul> </li> </ul>
14.30-14.45 น.	พักเบรก
14.45-16.30 น.	<ul style="list-style-type: none"> <li>● อธิบายแนวคิดของการแบ่งกลุ่มข้อมูลด้วยเทคนิค K-Means Clustering</li> <li>● Workshop การแบ่งกลุ่มข้อมูลด้วย RapidMiner AutoModel</li> </ul>
<b>วันที่ 27 กรกฎาคม 2565</b>	
09.00-10.30 น.	<ul style="list-style-type: none"> <li>● อธิบายแนวคิดพื้นฐานของการทำ Classification และ Regression</li> </ul>

	<ul style="list-style-type: none"> <li>◆ การแบ่งข้อมูลเพื่อทำการทดสอบประสิทธิภาพของโมเดล Classification และ Regression</li> <li>◆ ตัววัดประสิทธิภาพของโมเดล Classification และ Regression ต่างๆ ได้แก่ <ul style="list-style-type: none"> <li>▫ Confusion Matrix</li> <li>▫ Precision</li> <li>▫ Recall</li> <li>▫ F-Measure</li> <li>▫ Accuracy</li> <li>▫ ROC Curve และ AUC (Area Under Curve)</li> <li>▫ Mean Absolute Error (MAE)</li> <li>▫ Root Mean Square Error (RMSE)</li> </ul> </li> </ul>
10.30-10.45 น.	พักเบรก
10.45-12.00 น.	<ul style="list-style-type: none"> <li>● อธิบายแนวคิดของเทคนิค Decision Tree</li> <li>● ตัวอย่างการประยุกต์ใช้งานด้วย RapidMiner Auto Model</li> <li>● อธิบายแนวคิดของเทคนิค Naive Bayes</li> <li>● ตัวอย่างการประยุกต์ใช้งานด้วย RapidMiner Auto Model</li> <li>● อธิบายแนวคิดของเทคนิค Linear Regression</li> <li>● ตัวอย่างการประยุกต์ใช้งานด้วย RapidMiner Auto Model</li> </ul>
12.00-13.00 น.	พักรับประทานอาหารกลางวัน
13.00-14.30 น.	<ul style="list-style-type: none"> <li>● อธิบายแนวคิดของเทคนิค Logistic Regression</li> <li>● ตัวอย่างการประยุกต์ใช้งานด้วย RapidMiner Auto Model</li> <li>● อธิบายแนวคิดของเทคนิค Support Vector Machines (SVM)</li> <li>● ตัวอย่างการประยุกต์ใช้งานด้วย RapidMiner Auto Model</li> </ul>
14.30-14.45 น.	พักเบรก
14.45-16.30 น.	<ul style="list-style-type: none"> <li>● อธิบายแนวคิดของเทคนิค Gradient Boosted Tree</li> <li>● อธิบายแนวคิดของเทคนิค Deep Learning</li> <li>● ตัวอย่างการประยุกต์ใช้งานด้วย RapidMiner AutoModel</li> <li>● • แนะนำการสอบ certificate ของ RapidMiner</li> </ul>

หมายเหตุ: กำหนดการอาจจะมีการเปลี่ยนแปลงตามความเหมาะสม

สาระ/สิ่ง ที่ได้จากการเข้าร่วมอบรม/สัมมนา

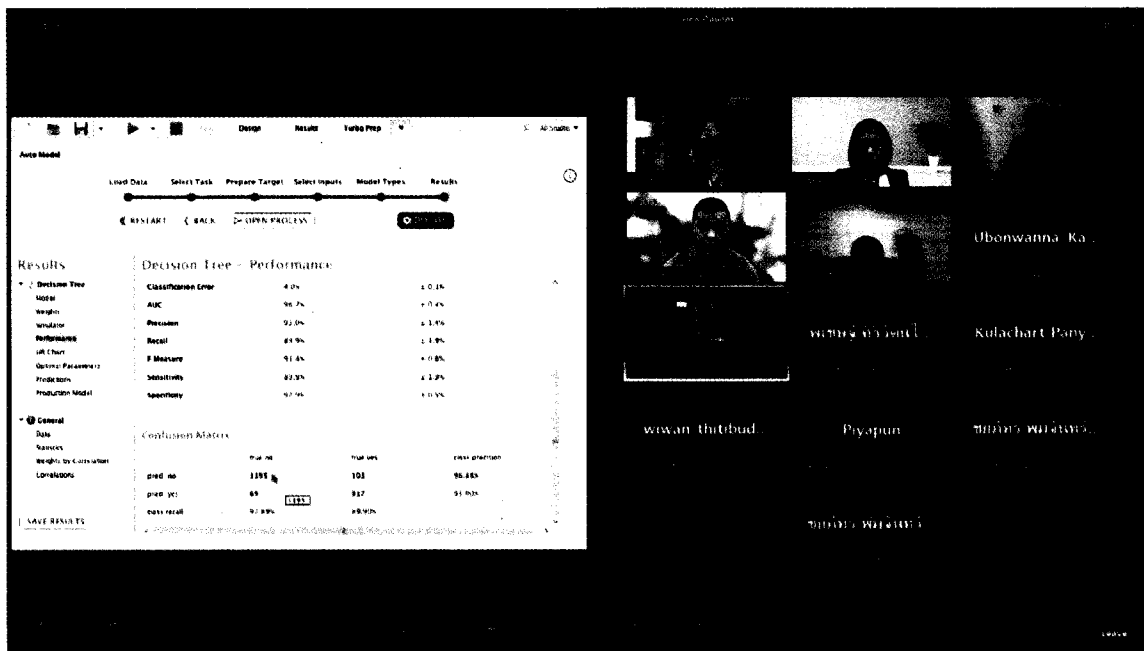
ได้แนวคิดและวิธีการวิเคราะห์ข้อมูลด้วยเทคนิค Data Science และ Machine Learning และการใช้ประโยชน์ในด้านต่างๆ. ซึ่งจะมีกระบวนการในการจัดการกับข้อมูลที่จะนำมาทำการวิเคราะห์ข้อมูล ได้แก่ การทำ Extract Transform Load (ETL)... การทำ Cleaning Data และการจัดการข้อมูลผ่านโปรแกรม

RapidMiner ซึ่งมีการใช้โมเดลในการวิเคราะห์ข้อมูลในการประมวลผล ทำนายผลข้อมูล และการวัดประสิทธิภาพของโมเดล

หลังจากได้เข้าร่วมอบรมจะนำความรู้มาพัฒนางานในปี 2565

มีการบริหารจัดการข้อมูลสารสนเทศในมหาวิทยาลัยที่มีการจัดเก็บในหลายรูปแบบ มาทำการรวบรวม ประมวลผล วิเคราะห์ เพื่อจะสร้างสารสนเทศสำหรับการบริหารจัดการในการวางแผนและคาดการณ์ในอนาคต ให้มหาวิทยาลัยมีการบริหารจัดการที่มีประสิทธิภาพและมีขีดธรรมภิบาลในการบริหารจัดการ

ภาพบรรยากาศของการเข้าร่วมอบรม/สัมมนา



chaith Patha      Kullaya Jarak      Ubonwanita Kas      Kullachart Panva

Cube Space      datacube      CUBE Analytics CONSULTING

“Theory without practice is empty,  
practice without theory is blind”  
— Kant

Free! Bundles available to Everyone

Zoom Meeting      Kullaya Jarak      Ubonwanita Kas      Kullachart Panva

# ร้านค้าที่ไม่ต้องรอคิว

- Amazon Go

Participants (13)

- CP Chaith Patha (Host)
- EP Kullachart Panva
- EM Bundie Nuanrues
- KJ Kullaya Jarak
- KP Kullachart Panva
- N Natthawit

Invite      Invite Me

Chat

Leave this meeting and rejoin later

**แชร์ให้แล้วครับ**

password ในการเปิด  
สไลด์คือ cube2022 และ  
ครับ

To: Everyone

41

Zoom Meeting

Yong ubonwanna.kas

# Data Storage: Data Warehouse

CUBE CONSULTING

The diagram illustrates a data warehouse architecture. On the left, data sources labeled 'Sales' and 'Marketing' feed into an 'Extract Transform Load (ETL)' process. This process is represented by a central stack of three cylinders. From this central stack, data is distributed to 'Business Intelligence' and 'Machine Learning' applications. A small bar chart on the right shows data for '01/2018' across categories 'A', 'B', and 'C'.

Shop	Month	Year	Total
A	01	2018	10,000
B	01	2018	2,000
C	01	2018	24,000

2016 2017 2018

Zoom Meeting

Teerapong Jak ubonwanna.kas

Turbo Prep

## Generate sale\_data

Generate a new column below. You can type a formula or drag in columns from the left and functions from the right.

COMMIT GENERATE CLEAR ALL CANCEL

Name: RFM\_score X

Formula:

Preview:

RFM score	RFM score
142	3
112	1
485	2
471	2

Functions:

- Comparison
- Text Information
- Text Transformation
- Mathematical Functions
- Statistical Functions
- Trigonometric Functions
- Rounding Functions
- Constants
- Basic Constants
- A TRUE
- B FALSE
- C #

Zoom Meeting

Chait Potha | Teerapong Jaik... | Ubonwanha Kas... | Weera... | (1) | 1

Design Results Turbo Prep

Repository Process Parameters

Import Data

Expression

date\_diff (date | date\_parse | '2021-11-01' | '1000/24/6072)

Info: Expression is syntactically correct

Functions

- Logical
- Comparison
- Text Information
- Text Transformation
- Mathematical Functions
- Statistical Functions
- Trigonometric Functions

Inputs

- Regular Attributes
- Basic Constants
- Date Function Constants
- Fuzzy Matching Constants
- Macros

Apply Cancel

Chat

Chat

To: Everyone

Zoom Meeting

Chait Potha | Weera... | Weera... | (1) | 1

# Data Analytics

CUBE CONSULTING

“Data and analytics is the **management of data for all uses** (operational and analytical) and the **analysis of data to drive business processes and improve business outcomes through more effective decision making and enhanced customer experiences.**”  
— *Gartner*

“เป็นการ**บริหารจัดการข้อมูล**สำหรับการวิเคราะห์เพื่อช่วยขับเคลื่อนธุรกิจและ**สนับสนุนการตัดสินใจ**ได้อย่างมีประสิทธิภาพ”

Copyright © 2022 CUBE Analytics Consulting | 285

Chat

Chat

To: Everyone

~ ๑ ~

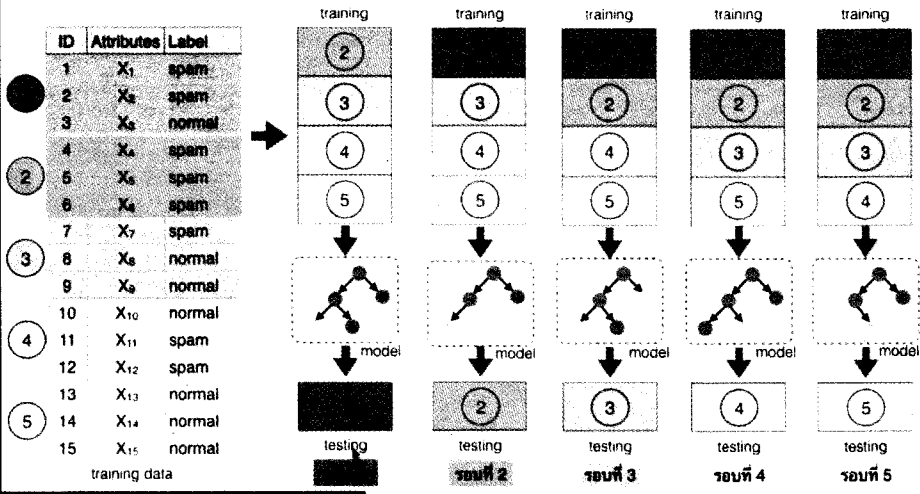




# Cross-Validation

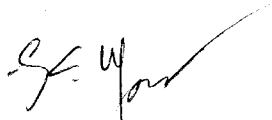


ตัวอย่างของ 5-fold cross-validation



	True no	True yes	Miss precision
pred no	3185	102	96.86%
pred yes	87	937	91.04%
Miss recall	97.85%	89.90%	

ผู้เข้าร่วมอบรม/สัมมนา



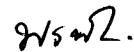
ผศ.ดร.กัลยา ใจรัมย์  
รองผู้อำนวยการสำนัก



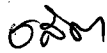
นายธีระพงษ์ ใจคำมา  
นักวิชาการคอมพิวเตอร์ชำนาญการ



นางสาวรุ่งทิวา กิตติยงกุล  
นักวิชาการคอมพิวเตอร์ปฏิบัติการ



นางพรพิมล แก้วฟุ้งรัมย์  
นักวิชาการคอมพิวเตอร์ปฏิบัติการ



นายชลิต โปธา  
นักวิชาการคอมพิวเตอร์ปฏิบัติการ



นายบัณฑิต นันทะเทศ  
นักวิชาการคอมพิวเตอร์ปฏิบัติการ



นายกุลชาติ ปัญญาดี  
นักวิชาการคอมพิวเตอร์ปฏิบัติการ