

สรุปบทเรียนจากการพัฒนาความรู้

หลักสูตร ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูล สำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ

โดย นายฉัตรชัย เจริญสรรพสุข



Big Data คือ ข้อมูลขนาดใหญ่มีทั้งแบบโครงสร้างปกติและโครงสร้างข้อมูลที่ไม่มีรูปแบบ ซึ่งทั้งหมดเป็นข้อมูลที่ใช้ในเชิงธุรกิจ มักจะถูกใช้กับงานที่ต้องวิเคราะห์ข้อมูลที่มีความซับซ้อนและไม่สามารถประเมินขนาดข้อมูลได้

รูปแบบของข้อมูล Big Data สามารถเป็นไปได้หลากหลาย ได้แก่

- **Behavioral Data** ได้แก่ ข้อมูลเชิงพฤติกรรมการใช้งานต่างๆ เช่น Server Log พฤติกรรมการคลิกดูข้อมูล หรือข้อมูลการใช้ ATM

- **Image and Sounds** ตัวอย่างเช่น ภาพถ่ายวิดีโอ รูปจาก Google Street View ภาพถ่ายภาพทางการแพทย์ ลายมือ ข้อมูลเสียงที่ถูกบันทึกไว้

- **Languages** ตัวอย่างเช่น Text Message ข้อความที่ถูก Tweet เนื้อหาต่างๆ ในเว็บไซต์

- **Records** ตัวอย่างเช่น ข้อมูลทางการแพทย์ข้อมูลผลสำรวจที่มีขนาดใหญ่ ข้อมูลทางภาษี

- **Sensors** ตัวอย่างเช่น ข้อมูลอุณหภูมิ Accelerometer ข้อมูลทางภูมิศาสตร์

ลักษณะสำคัญของ Big Data จะต้องมียุคสำคัญ 4V ดังต่อไปนี้

- ๑) **Volume** คือ ข้อมูลมีขนาดใหญ่มีปริมาณข้อมูลมากซึ่งสามารถเป็นได้ทั้งข้อมูลแบบออฟไลน์หรือออนไลน์

- ๒) **Variety** คือ ข้อมูลมีความหลากหลาย สามารถเป็นได้ทั้งที่มีโครงสร้าง และข้อมูลที่ไม่สามารถจับ Pattern ได้

- ๓) **Velocity** คือ ข้อมูลมีการเปลี่ยนแปลงตลอดเวลา อย่างรวดเร็ว มีการส่งผ่านข้อมูลอย่างต่อเนื่อง ในลักษณะ Streaming ทำให้เกิดการวิเคราะห์ข้อมูลแบบ Manual มีข้อจำกัด

- ๔) **Veracity** คือ ข้อมูลมีความไม่ชัดเจน

กว่าจะมาเป็น Big Data

ในอดีตการจัดเก็บข้อมูล วิเคราะห์ข้อมูลจะอาศัย Data Lake และ Data Warehouse ในการทำงาน ซึ่งมีทั้งข้อดี-ข้อเสีย แตกต่างกันไป

Data Lake เป็นการนำเอาข้อมูลจากแหล่งข้อมูลภายนอกองค์กรข้อมูลจากเครือข่ายข้อมูลที่กระจายไปทั่วโลกมาใช้งานมากขึ้น ปริมาณข้อมูลจากแหล่งภายนอกเพิ่มอย่างต่อเนื่องและมีแนวโน้มที่จะเติบโตแบบก้าวกระโดดมากขึ้น สามารถแก้ไขข้อจำกัดหลายอย่างของ Data Warehouse ที่ใช้กันมานาน ข้อมูลที่จัดเก็บ คือ ข้อมูล

ดิบจำนวนมากและมีขนาดใหญ่ ข้อมูลที่ไม่มีรูปแบบที่แน่นอน การเข้าถึงข้อมูลไม่สามารถเข้าถึงได้ง่าย ต้องใช้ความสามารถของเจ้าหน้าที่วิเคราะห์ข้อมูล

ความแตกต่างระหว่าง Data Lake เมื่อเทียบกับ Data Warehouse

- เก็บข้อมูลทั้งหมด
- สนับสนุนข้อมูลทุกชนิดไม่ใช่เพียงข้อมูลแบบ Structure
- ผู้ใช้ทุกประเภทสามารถใช้งานได้
- ประมวลและวิเคราะห์ข้อมูลได้เร็วกว่า

การวิเคราะห์ข้อมูล Big Data ในปัจจุบัน ทำให้มีข้อมูลที่เป็นข้อเท็จจริง ซึ่งผ่านการวิเคราะห์อย่างเป็นระบบ เพื่อใช้ประกอบการตัดสินใจ โดยระดับของการวิเคราะห์ก็เป็นไปได้หลากหลาย แล้วแต่รูปแบบการนำไปใช้งาน

- **Descriptive Analytics** เป็นการวิเคราะห์ในระดับที่บอกว่าเกิดอะไรขึ้น จำนวนเท่าไร ความถี่แค่ไหน เกิดเหตุการณ์สำคัญตอนไหน ตรงไหนบ้าง

- **Predictive Analytics** เป็นการวิเคราะห์ในลักษณะที่ซับซ้อนขึ้นไปอีกขั้นหนึ่ง เป็นการประเมินว่าจะเกิดอะไรขึ้นต่อไป มีการให้ข้อมูลตัวชี้วัดของผลลัพธ์ที่อาจจะเกิดขึ้น ถ้าแนวโน้มยังเป็นไปอย่างนี้ต่อไป

- **Prescriptive Analytics** เป็นรูปแบบการวิเคราะห์ข้อมูลที่มีความซับซ้อนและยากที่สุดเพราะไม่เพียงพยากรณ์หรือทำนายว่าเกิดอะไรขึ้นยังให้คำแนะนำในทางเลือกต่างๆ และผลของทางเลือกต่างๆ ว่าจะมีผลดีและผลเสียอย่างไร โมเดลของ prescriptive analytic จะสามารถปรับเปลี่ยนได้ตามข้อมูลที่เพิ่มเติมเข้ามามากขึ้น และยังเป็นการใช้ข้อมูลมากที่สุดที่เกี่ยวกับเรื่องของ Big Data เป็นอย่างมาก

Data Mining คือการวิเคราะห์ข้อมูลจาก ข้อมูลจำนวนมาก Big Data เพื่อหาความสัมพันธ์ของข้อมูลที่ซ่อนอยู่โดยทำการจำแนกประเภทรูปแบบเชื่อมโยงข้อมูลที่มีความสัมพันธ์กันและหาความน่าจะเป็นที่จะเกิดขึ้นเพื่อให้ได้องค์ความรู้ใหม่ที่สามารถนำไปใช้ประกอบการตัดสินใจในด้านต่างๆ เช่น ตลาดหลักทรัพย์ ทางธุรกิจ ทางด้านการแพทย์ยุทธศาสตร์ทหาร เป็นต้น

Text Mining เป็นเทคนิคเพื่อค้นหารูปแบบจากข้อมูลจำนวนมหาศาลโดยอัตโนมัติ โดยใช้ขั้นตอนวิธีจากวิชาสถิติ การเรียนรู้ของเครื่อง และการรู้จำรูปแบบ หรือในอีกนิยามหนึ่ง การทำเหมืองข้อความ คือกระบวนการที่กระทำกับข้อความเพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อความนั้นโดยอาศัยหลักสถิติการรู้จำการเรียนรู้ของเครื่องหลักคณิตศาสตร์หลักการประมวลเอกสาร Document Processing หลักการประมวลผลข้อความ Text processing และหลักการประมวลผลภาษาธรรมชาติ Natural Language Processing

Machine Learning เป็นศาสตร์ของการสร้างโมเดลคณิตศาสตร์ มุ่งเน้นที่จะสร้างองค์ความรู้จากข้อมูล โดยเริ่มจากการสร้างโมเดลทางคณิตศาสตร์ที่มีความยืดหยุ่น และสามารถปรับตัวเองเข้ากับข้อมูลที่ได้รับเพื่อที่จะสามารถทำนายอนาคตได้เทคโนโลยีที่อยู่เบื้องหลังหุ่นยนต์อัจฉริยะต่างๆหรือเทคโนโลยีการบินแบบไร้คนขับหรือระบบวิเคราะห์ภาพเสียงภาพเคลื่อนไหวและภาษามนุษย์

Data Driven Business

- สินทรัพย์ในทางธุรกิจ ข้อมูลถือเป็นสินทรัพย์ที่มีมูลค่า จะแสดงให้เห็นทั้งจุดอ่อนและจุดแข็งจากการดำเนินงานที่ผ่านมาและช่วยให้เกิดการพัฒนาในเชิงธุรกิจมากยิ่งขึ้น
- เข้าถึงลูกค้าได้ดียิ่งขึ้น การเก็บข้อมูลของลูกค้าเพื่อให้ได้มาซึ่งบริการและสินค้าที่ตรงต่อความต้องการของลูกค้า ได้แก่ความต้องการช่องทางซื้อขายและช่องทางการชำระเงินที่สะดวก
- นอกจากนี้ ยังทำให้แต่ละบริษัทเกิดความเปลี่ยนแปลงในด้านของการจัดการโดยเฉพาะข้อมูลที่จะต้องมีการพัฒนาระบบจัดเก็บและระบบป้องกันรักษาข้อมูล นอกจากนี้ ทางบริษัทยังต้องตื่นตัวในการสร้างและปรับปรุงนโยบายที่เกี่ยวกับการใช้ระบบข้อมูลให้มีความรัดกุมและทันสมัยอยู่เสมอ
- พัฒนาประสิทธิภาพและการทำงาน Big Data มีความสามารถที่จะช่วยด้านการพัฒนาประสิทธิภาพและการทำงานภายในของธุรกิจเกือบทุกประเภท สามารถช่วยพัฒนา ในด้านทรัพยากรบุคคลและการจ้างงานได้ด้วยนอกจากนี้ ทางบริษัทยังต้องตื่นตัวในการสร้างและปรับปรุงนโยบาย ที่เกี่ยวกับการใช้ระบบข้อมูลให้มีความรัดกุมและทันสมัยอยู่เสมอ
- พัฒนาความพึงพอใจของลูกค้าและผลิตภัณฑ์ สามารถใช้ข้อมูลที่เก็บมาได้จากลูกค้าในการพัฒนาผลิตภัณฑ์และประสบการณ์การใช้สินค้าได้ เพื่อคาดการณ์ถึงปัญหาที่อาจจะเกิดขึ้นในอนาคต

Big Data analytic กับการบริหารภาครัฐ

- **รับฟังความเห็นรวบรวมข้อมูลและปรึกษากับผู้มีส่วนได้ส่วนเสีย** ก่อนที่จะเริ่มการสร้างระบบ Big Data analytic นั้น องค์กรที่รับผิดชอบจะต้องทำการรับฟังความคิดเห็นรวบรวมข้อมูลและปรึกษากับผู้มีส่วนได้ส่วนเสีย ทุกภาคส่วนเกี่ยวกับศักยภาพการเพิ่มประสิทธิภาพและการกำหนดนิยามทางเทคนิคที่เกี่ยวข้องรวมถึงเป้าหมายที่จะทำให้เกิดประโยชน์ในการนำเอาข้อมูลมารวมใช้อย่างมีประสิทธิภาพในรูปแบบ open source ให้มากขึ้น
- **วางแผนการลงทุนในการจัดโครงสร้าง** รวบรวมข้อมูล และการบริหารจัดการข้อมูลโดยอยู่บนพื้นฐานของความเชื่อมั่นต่อผู้มีส่วนได้ส่วนเสีย ซึ่งการจัดการข้อมูลต่างๆ จะต้องเป็นไปตามกฎหมายและหลักจริยธรรมที่เกี่ยวข้องรวมทั้งต้องสื่อสารให้ทุกภาคส่วนรวมทั้งประชาชนให้มีความเข้าใจและมั่นใจในเรื่องความเป็นส่วนตัวของข้อมูล privacy
- **มีความเข้าใจและมีทักษะทางธุรกิจและทักษะทางเทคนิค** ความท้าทายที่เกิดขึ้นจากพลังของการวิเคราะห์ข้อมูล Big Data ในระดับชาตินั้นผู้บริหารและเจ้าหน้าที่ในหน่วยงานภาครัฐจะต้องมีความเข้าใจและมีทักษะทางธุรกิจและทักษะทางเทคนิคอย่างมืออาชีพโดยจะต้องทำให้ Big Data สามารถเชื่อมโยงประโยชน์ให้เกิดขึ้นกับผู้มีส่วนได้ส่วนเสียทุกกลุ่มโดยเฉพาะอย่างยิ่งประชาชนด้วยการสร้างพันธมิตรในรูปแบบ Win-Win เพื่อที่จะเกิดความสำเร็จได้อย่างรวดเร็ว
- **เตรียมพร้อมภายใต้การเปลี่ยนแปลงของเทคโนโลยี** เพื่อตอบสนองความต้องการของผู้มีส่วนได้ส่วนเสีย ที่มีภาวะเปราะบางที่หลากหลายจึงทำให้การออกแบบระบบ Big Data analytics จะต้องใช้สถาปัตยกรรมด้านไอทีแบบผสมผสานที่สามารถปรับขนาดได้ Scalability และมีความยืดหยุ่น flexibility โดยองค์กรภาครัฐที่ทำหน้าที่ดูแลระบบจะต้องเตรียมพร้อมภายใต้การเปลี่ยนแปลงของเทคโนโลยีอย่างรวดเร็ว

- **เจ้าหน้าที่ภาครัฐจะต้องปรับ Mindset** ในการเข้าร่วมกับทุกภาคส่วน ภาครัฐมีความท้าทายอย่างมากที่จะต้องปรับตัวให้องค์กรภาครัฐเป็นส่วนหนึ่งของระบบนิเวศอย่างกลมกลืนและจะต้องไม่เป็นผู้สร้างอุปสรรคให้เกิดขึ้นในการสร้างระบบ Big Data analytic ดังนั้นเจ้าหน้าที่ภาครัฐจะต้องปรับ mindset ในการเข้าร่วมกับทุกภาคส่วนในลักษณะพันธมิตร (Partnership) ให้ได้อย่างรวดเร็ว

- **ปรับปรุงวิธีคิดและกระบวนการเพื่อทำให้เกิดการแลกเปลี่ยนข้อมูลและการใช้ข้อมูลร่วมกันระหว่างหน่วยงานภาครัฐ** ภาครัฐต้องปรับปรุงวิธีคิดและกระบวนการเพื่อทำให้เกิดการ แลกเปลี่ยนข้อมูลและการใช้ข้อมูลร่วมกันระหว่างหน่วยงานภาครัฐกันเองอย่างมีประสิทธิภาพ โดยไม่หวงข้อมูลว่าเป็นข้อมูลของหน่วยงานตนเองซึ่งการแลกเปลี่ยนข้อมูลจะเป็นกุญแจสำคัญที่นำไปสู่การมองเห็นความรู้ใหม่ๆ ที่เกิดจากข้อมูลในมิติใหม่และมีผลต่อการสร้างแหล่งข้อมูล Big Data ได้อย่างรวดเร็วโดยรัฐบาลจะต้องสร้างสภาพแวดล้อมได้เทคโนโลยีที่ปลอดภัยแต่สร้างกรอบการ กำกับดูแลที่เหมาะสมในด้านการแลกเปลี่ยนข้อมูลและใช้ข้อมูลร่วมกันซึ่งการวางนโยบาย Open Government เพื่อเป็นการเปิดเผยข้อมูลภาครัฐให้มากขึ้นเพื่อนำไปสู่การทำ Big Data ให้ประสบความสำเร็จอย่างแท้จริง

- **กำหนดแนวทางและการบริการให้คำปรึกษาในด้าน Big Data Analytics แก่ทุกภาคส่วน** เป็นอีกหนึ่งปัจจัยแห่งความสำเร็จซึ่งองค์กรภาครัฐจะต้องมีส่วนร่วมให้มากขึ้นกว่าที่เป็นมาในอดีตในรูปแบบการให้การปรึกษา โดยจะต้องสร้างแนวทางและมาตรการเพื่อให้เกิดความร่วมมือในการให้ข้อมูลเชิงลึกก็จะทำให้ประโยชน์จาก Big Data Analytics เห็นผลและเกิดประสิทธิภาพมากยิ่งขึ้น

แหล่งที่มา

สรุปบทเรียนจากการเรียนรู้ผ่านระบบเรียนออนไลน์ e-Learning

หลักสูตร : ความรู้พื้นฐานเพื่อการวิเคราะห์ข้อมูลสำหรับข้าราชการและบุคลากรภาครัฐทุกระดับ

บรรยายโดย : รองศาสตราจารย์ ดร.วฤชาย รมสหายหยุด

สาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช

สถาบัน/หน่วยงาน/ระบบ : สถาบันพัฒนาบุคลากรภาครัฐด้านดิจิทัล Thailand Digital Government Academy

รูปแบบหลักสูตร : e-Learning

ช่วงเวลาการฝึกอบรม : กรกฎาคม 2566