

โครงการอบรม Big Data Analytics and Machine Learning

1. หลักการและเหตุผล

ในปัจจุบันรัฐบาลได้ตระหนักถึงความจำเป็นเร่งด่วนในการใช้เทคโนโลยีดิจิทัลมาเป็นเครื่องมือสำคัญในการพัฒนาและปฏิรูปประเทศไทยไปสู่ความมั่นคง มั่งคั่ง และยั่งยืน จึงได้จัดทำแผนพัฒนาดิจิทัลเพื่อเศรษฐกิจและสังคมของประเทศ เพื่อเป็นกรอบแนวทางการดำเนินการตามนโยบายเศรษฐกิจและสังคมดิจิทัลของรัฐบาลให้เกิดการนำเทคโนโลยีที่ทันสมัยและหลากหลายมาเปลี่ยนเป็นวิธีการดำเนินธุรกิจ การดำเนินชีวิตของประชาชน และการดำเนินงานของภาครัฐ ซึ่งจะส่งผลให้เกิดความมั่งคั่งทางเศรษฐกิจที่แข่งขันได้ในเวทีโลก และความมั่นคงทางสังคมของประเทศต่อไป

ในส่วนของ การดำเนินการของภาครัฐ การปฏิรูปหน่วยงานไปสู่การเป็นรัฐบาลดิจิทัล โดยการนำเอาเทคโนโลยีดิจิทัลต่างๆ มาปรับใช้ในการปฏิบัติงานในทุกรูปแบบ ทำให้ข้อมูลในรูปแบบดิจิทัล มีการขยายตัวเพิ่มขึ้นอย่างมหาศาล อย่างมีนัยสำคัญ ทั้งในเชิงปริมาณ รูปแบบ ความซับซ้อน และความรวดเร็วในการเปลี่ยนแปลง เช่น ข้อมูลรายการ (Transaction) ที่เกิดขึ้นจากกระบวนการที่ถูกปรับเปลี่ยนมาอยู่ในรูปแบบดิจิทัล ข้อมูลจากไฟล์เอกสารต่าง ๆ ที่ใช้แทนที่กระดาษ ข้อมูลที่เกิดจากการเชื่อมโยงบูรณาการข้อมูลระหว่างหน่วยงาน (Data Integration) ข้อมูลแบบเปิด (Open Data) จากภาครัฐและเอกชนทั้งในและต่างประเทศ ข้อมูลจาก Social Media เป็นต้น ดังนั้น การเตรียมบุคลากรของกระทรวงการคลัง ให้มีความรู้ความเชี่ยวชาญในการบริหารจัดการข้อมูลขนาดใหญ่และการวิเคราะห์ข้อมูลเชิงลึก (Big Data & Analytics) ของข้อมูลมหาศาลดังกล่าว จึงมีความสำคัญและจำเป็นอย่างยิ่งยวด เพื่อให้สามารถบริหารจัดการข้อมูลที่มีขนาดใหญ่และมีความหลากหลาย (Big Data) ได้อย่างเป็นระบบ รวมทั้งสามารถใช้เครื่องมือวิเคราะห์ข้อมูลเชิงลึก (Analytics) เพื่อนำข้อมูลไปต่อยอด ทำเข้าใจข้อมูลอย่างลึกซึ้ง และนำข้อมูลไปใช้วิเคราะห์คาดการณ์ให้สร้างคุณค่าให้เกิดผลในที่สุด (Insight, Action, Outcome) สอดคล้องกับยุทธศาสตร์ที่ 5 ของแผนพัฒนาดิจิทัลเพื่อเศรษฐกิจและสังคม ในเรื่องการพัฒนากำลังคนให้พร้อมเข้าสู่เศรษฐกิจและสังคมดิจิทัล เพื่อให้สามารถใช้ประโยชน์จากเทคโนโลยีดิจิทัลได้อย่างรอบรู้ เท่าทันการเปลี่ยนแปลงของเทคโนโลยีในยุคปัจจุบัน และอนาคต

2. วัตถุประสงค์

- 2.1 เพื่อเสริมสร้างความเข้าใจในหลักการ และองค์ประกอบต่าง ๆ ของการบริหารจัดการข้อมูล ที่มีขนาดใหญ่และวิเคราะห์ข้อมูลเชิงลึก จากการเรียนรู้ทั้งในภาคทฤษฎีและปฏิบัติ
- 2.2 สามารถบริหารจัดการข้อมูลที่มีขนาดใหญ่และมีความหลากหลายได้อย่างเป็นระบบ
- 2.3 เรียนรู้กระบวนการวิเคราะห์ข้อมูลขนาดใหญ่ เพื่อนำไปใช้ประโยชน์ได้อย่างมีประสิทธิภาพ
- 2.4 สามารถสร้างตัวแบบ(Model) เพื่อใช้ในการทำนายหรือพยากรณ์ ตามหลักการของ Data Science ได้อย่างถูกต้อง
- 2.5 สร้างผู้เชี่ยวชาญดิจิทัลด้าน Big Data & Analytics ให้กับกระทรวงการคลัง

3. รายละเอียดโครงการ

การอบรมหลักสูตรนี้จะสร้างความเข้าใจเกี่ยวกับเทคโนโลยี Big Data และแนวคิดด้านการทำ Modernizing Data Warehouse ด้วย Data Lake ในส่วนการอบรมภาคปฏิบัติ มุ่งเน้นให้ผู้เรียนสามารถประมวลผลข้อมูลในแบบ Distributed and Parallel Processing โดยใช้ MapReduce ในแบบ in-memory processing ที่ใช้ภาษา Python และ Spark ทำการประมวลผลข้อมูลที่อยู่ใน Hadoop ได้ ผู้สอนจะถ่ายทอดประสบการณ์ด้าน Big Data Analytics จากการปฏิบัติงานในองค์กร ด้วยชุดข้อมูล (Dataset) ที่หลากหลาย โดยเฉพาะอย่างยิ่ง Unstructure Data กับกระบวนการปฏิบัติตามหลักการของ Data Science เพื่อให้ผู้เรียนสามารถนำความรู้ไปปฏิบัติได้จริง โดยมีรายละเอียดเนื้อหาการอบรม ดังนี้

วิทยากร:

1. รศ.ดร.ธนชาติ นุ่มนนท์
 - ผู้อำนวยการสถาบัน ไอเอ็มซี
 - กรรมการที่ปรึกษาด้านเทคโนโลยีสารสนเทศ ธนาคาร ธนชาติ
 - อดีตผู้อำนวยการเขตอุตสาหกรรมซอฟต์แวร์ประเทศไทย
 - Certified Sun Java Instructor
2. คุณเอกอนันต์ ทองแท้
 - วิทยากรผู้เชี่ยวชาญด้าน Big Data ของสถาบัน ไอเอ็มซี
 - อดีตผู้จัดการส่วนสถาปัตยกรรมและต้นแบบ สำนักงานรัฐบาลอิเล็กทรอนิกส์ องค์การมหาชน (EGA)
 - ที่ปรึกษาด้านระบบการทำธุรกิจอัจฉริยะ (Business Intelligence System) บริษัทป่าตอง เบย์เวเคชั่นคลับ จำกัด

รายเอียดหลักสูตร

การอบรมเชิงปฏิบัติการ วันที่ 1

Big Data Architecture

- Modernizing Data Warehouse & Data Lake
- Introduction to Big Data Technology
- NoSQL
- Big Data Architecture
- What is Hadoop?
- Hadoop Eco-system
- Big Data Architecture Use Cases
- Data Science in Era of Big Data

การอบรมเชิงปฏิบัติการ วันที่ 2

Introduction to Hadoop

- Lecture: Introduction to Hadoop
- Lecture: Hadoop Architecture
- Hands-On: Installing Hadoop and Software Components
- Lecture: HDFS - Hadoop's Storage
- Hands-On: HDFS operations
- Lecture: Introduction to Sqoop/Flume
- Hands-On: Sqoop/Flume Examples

การอบรมเชิงปฏิบัติการ วันที่ 3**Big Data Processing**

- Lecture: Understanding MapReduce Processing
- Lecture: Hive Programming
- Hands-On: Hive Programming
- Lecture: Introduction to Impala
- Hands-On: Impala Programming
- Lecture: Spark Core Engine using Python API.
- Hands-On: Extraction, Loading and Transforming Data Spark
- Hands-On: Descriptive Analytics using Spark
- Lecture: Introduction to SparkSQL

การอบรมเชิงปฏิบัติการ วันที่ 4**Introduction to Data Science**

- Lecture: Data Science & Machine Learning
- Lecture: Machine Learning Algorithm
- Introduction of MLlib/ML Pipeline (Spark Machine Learning)
- Hands-On: Basic Programming for Classification and Cluster

การอบรมเชิงปฏิบัติการ วันที่ 5**Large-Scale Machine Learning**

- Hands-On: Credit Scoring (Classification/Regression)
- Hands-On: Prediction of Loan Payment (Classification)
- Hands-On: Analysis of Network Intrusion (Clustering)

4. ประโยชน์ที่ได้รับ

- 4.1 บุคลากรกระทรวงการคลังที่รับการอบรมมีความรู้ความเข้าใจในการบริหารจัดการข้อมูลขนาดใหญ่ Big Data
- 4.2 สามารถวางกลยุทธ์ในการนำเอาเทคโนโลยี Big Data ไปประยุกต์ใช้งานกับข้อมูลปริมาณมากและหลากหลาย ในหน่วยงานได้อย่างเหมาะสม และมีประสิทธิภาพ
- 4.3 สามารถใช้ประโยชน์ใช้ประโยชน์จากการวิเคราะห์ข้อมูลขนาดใหญ่ เพื่อประกอบการวางแผนและการตัดสินใจอย่างถูกต้อง ทันสถานการณ์
- 4.4 บุคลากรกระทรวงการคลังมีความเชี่ยวชาญดิจิทัลด้าน Big Data & Analytics

5. ผู้เข้ารับการฝึกอบรมสัมมนา

- 5.1 ข้าราชการในสายงานนักวิชาการคอมพิวเตอร์ระดับปฏิบัติการ หรือสูงกว่า
- 5.2 เป็นผู้ที่มีหน้าที่รับผิดชอบในการดูแลและพัฒนาระบบคลังข้อมูลของหน่วยงาน
- 5.3 เป็นผู้มีความรู้ในการเขียนโปรแกรมคอมพิวเตอร์ เช่น C หรือ Java เป็นอย่างน้อย

6. จำนวนผู้เข้ารับการฝึกอบรมสัมมนา

จำนวน 35 คน

7. เกณฑ์การผ่านการฝึกอบรมสัมมนา

ผู้เข้ารับการอบรมต้องมีเวลาเข้ารับการฝึกอบรมไม่น้อยกว่า 80% ของระยะเวลาการฝึกอบรมทั้งหมดตามที่กำหนดไว้ในหลักสูตร

8. รูปแบบการดำเนินการฝึกอบรมสัมมนา

บรรยาย สัมมนากลุ่มย่อย แลกเปลี่ยนความรู้ กรณีศึกษา และเรียนรู้เสริมสร้างความเข้าใจ การทำความเข้าใจและแลกเปลี่ยนความเห็นภายในห้อง การนำเอาประเด็นปัญหาจริงในงานมาปรับปรุงพัฒนาการปฏิบัติงานในอนาคต

9. ระยะเวลาดำเนินการ

- วันจันทร์ที่ 26 – วันศุกร์ที่ 30 มีนาคม 2561 (เวลา 09.00-16.00 น.)
- รวมระยะเวลาจำนวน 5 วัน

10. สถานที่จัดฝึกอบรมสัมมนา

สถาบัน ไอเอ็มซี อาคารสกลไทย สุรวงศ์ ทาวเวอร์ ถนนสุรวงศ์ กรุงเทพมหานคร

.....