**หลักสูตร Big Data with Hadoop**

***รายละเอียดหลักสูตร***

หลักสูตรนี้มุ่งหมายให้ผู้เรียนเข้าใจกระบวนการทำงานของระบบ Hadoop และสามารถติดตั้งระบบ Hadoop Cluster เพื่อใช้งานรวมถึงเข้าใจเครื่องมือแต่ละตัวและประยุกต์ใช้งานซอฟต์แวร์เหล่านั้นได้ ในเนื้อหาเป็นการลงมือปฏิบัติคอนฟิกเครื่องเซิร์ฟเวอร์คลัสเตอร์ให้ทำงานร่วมกัน และศึกษาส่วนประกอบหลัก ๆ ของ Hadoop ไล่ไปทีละส่วน ตั้งแต่ส่วนของระบบไฟล์แบบกระจายที่เรียกว่า Hadoop Distributed File System (HDFS) การประมวลผลข้อมูลด้วย Map Reduce รวมถึงซอฟต์แวร์แวดล้อมที่มาทำงานบนระบบ Map Reduce อย่าง Pig และ Hive เพื่อใช้จัดการกับข้อมูลในรูปภาษาสคริปต์ และภาษาในลักษณะ SQLตามลำดับ นอกจากนั้นยังได้หัดใช้ Sqoop เพื่อเชื่อมต่อกับซอฟต์แวร์ฐานข้อมูล (DBMS) รวมถึงการติดตั้งและใช้งาน Hue, impala และ spark ผู้เรียนจะได้ศึกษาไปทีละขึ้น รวมถึงจะได้เรียนรู้คำสั่งจำเป็นต่อการดูแลระบบ การอ่านและวิเคราะห์ Log File

***วัตถุประสงค์***

1. เพื่อเสริมสร้างความเข้าใจด้านเทคโนโลยี Big Data และมองภาพการนำไปใช้ได้อย่างมีประสิทธิภาพ
2. เพื่อรู้และเข้าใจโครงสร้างสถาปัตยกรรมของระบบ Hadoop
3. เพื่อให้สามารถติดตั้งระบบ Hadoop และนำไปประยุกต์ใช้งานเพื่อบริหารจัดการข้อมูลขนาดใหญ่ (Big Data) และตอบโจทย์เชิงธุรกิจได้

***ความรู้พื้นฐาน***

ผู้เข้าอบรมควรมีความสามารถในการใช้งานคำสั่งลีนุกซ์ (Linux) พื้นฐาน และควรมีความเข้าใจเรื่องระบบเครือข่ายและไฟล์วอล์

***สายงาน/ตำแหน่งงาน ที่เหมาะสม***

* Database Administrator
* System Administrator / System Engineer
* Network Administrator / Network Engineer
* IT Support / HelpDesk / Programmer
* บุคคลทั่วไปที่สนใจและต้องการศึกษา Big Data ด้วย Hadoop ตั้งแต่ระดับเบื้องต้น

***ประโยชน์ที่ผู้อบรมได้รับ***

1. ได้รับความรู้ความเข้าใจเกี่ยวกับการบริหารจัดการข้อมูลขนาดใหญ่ (Big Data) ด้วยระบบ Hadoop
2. สามารถติดตั้งระบบ Hadoop เพื่อรองรับการจัดการข้อมูลขนาดใหญ่ (Big Data) และตอบโจทย์ธุรกิจและเพิ่มประสิทธิภาพการทำงานขององค์กรได้

***ซอฟต์แวร์ที่ใช้***

* Cloudera Hadoop (CDH5)
* JDK-1.8
* CentOS-7 x86\_64
* Virtual Box

***ฮาร์ดแวร์ที่ใช้***

เครื่องคอมพิวเตอร์ที่มีหน่วยความจำไม่น้อยกว่า 8 GB (ในการอบรมจะใช้ Virtual Machine ด้วยซอฟต์แวร์ Virtual Box)

***ระยะเวลาการอบรม***

5 วัน (40 ชั่วโมง)

***เนื้อหาหลักสูตร***

**วันที่ 1**

|  |  |  |
| --- | --- | --- |
| **เวลา** | **เนื้อหา** | **หมายเหตุ** |
| 08:30 – 09:00 | Course Introduction |  |
| 09:00 – 10:30 | Big Data Technology & Use cases |  |
| **10:30 – 10:45** | **พักรับประทานอาหารว่าง** |  |
| 10:45 – 12:00 | Big Data Hadoop Architecture & Ecosystem |  |
| **12:00 – 13:00** | **พักรับประทานอาหารกลางวัน ตามอัธยาศัย** |  |
| 13:30 – 14:45 | Hadoop Installation |  |
| **14:45 – 15:00** | **พักรับประทานอาหารว่าง** |  |
| 15:00 – 17:30 | Lab :Pre-Installation |  |

**วันที่ 2**

|  |  |  |
| --- | --- | --- |
| **เวลา** | **เนื้อหา** | **หมายเหตุ** |
| 08:30 – 09:00 | Course Introduction |  |
| 09:00 – 10:30 | Principles of HDFS & MapReduce2 |  |
| **10:30 – 10:45** | **พักรับประทานอาหารว่าง** |  |
| 10:45 – 12:00 | Principles of Hive, Impala, Pig & Hue |  |
| **12:00 – 13:00** | **พักรับประทานอาหารกลางวัน ตามอัธยาศัย** |  |
| 13:30 – 14:45 | Lab :HDFS & Map Reduce Programing (Python) |  |
| **14:45 – 15:00** | **พักรับประทานอาหารว่าง** |  |
| 15:00 – 17:00 | Lab :Hive,Impala & Pig |  |
| 17:30 – 17:30 | Lab Wrap Up |  |

**วันที่ 3**

|  |  |  |
| --- | --- | --- |
| **เวลา** | **เนื้อหา** | **หมายเหตุ** |
| 08:30 – 09:00 | Course Introduction |  |
| 09:00 – 10:30 | Principles of Sqoop & Flume |  |
| **10:30 – 10:45** | **พักรับประทานอาหารว่าง** |  |
| 10:45 – 12:00 | Principles of Spark |  |
| **12:00 – 13:00** | **รับประทานอาหารกลางวัน ตามอัธยาศัย** |  |
| 13:00 – 14:45 | Lab :Sqoop & Flume |  |
| **14:45 – 15:00** | **พักรับประทานอาหารว่าง** |  |
| 15:00 – 17:.00 | Lab :Spark |  |
| 17:00 – 17:30 | Lab Wrap Up |  |

**วันที่ 4**

|  |  |  |
| --- | --- | --- |
| **เวลา** | **เนื้อหา** | **หมายเหตุ** |
| 08:30 – 09:00 | Course Introduction |  |
| 09:00 – 10:30 | Principles of Machine Learning |  |
| **10:30 – 10:45** | **พักรับประทานอาหารว่าง** |  |
| 10:45 – 12:00 | Visualization Tools |  |
| **12:00 – 13:00** | **รับประทานอาหารกลางวัน ตามอัธยาศัย** |  |
| 13:00 – 14:45 | Lab : Machine Learning |  |
| **14:45 – 15:00** | **พักรับประทานอาหารว่าง** |  |
| 15:00 – 17:00 | Lab : Visualization |  |
| 17:00 – 17:30 | Lab Wrap Up |  |

**วันที่ 5** **Big Data Hack Day**

|  |  |  |
| --- | --- | --- |
| **เวลา** | **เนื้อหา** | **หมายเหตุ** |
| 08:30 – 09:00 | Course Introduction |  |
| 09:00 – 09:30 | Mini Project :Select Topic |  |
| 09:30 – 10:30 | Mini Project : Collaboration Time |  |
| **10:30 – 10:45** | **พักรับประทานอาหารว่าง** |  |
| 10:45 – 12:00 | Mini Project : Collaboration Time |  |
| **12:00 – 13:00** | **รับประทานอาหารกลางวัน ตามอัธยาศัย** |  |
| 13:00 – 14:45 | Mini Project : Collaboration Time |  |
| **14:45 – 15:00** | **พักรับประทานอาหารว่าง** |  |
| 15:00 – 16:15 | Mini Project :Presentation |  |
| 16:15 – 17:30 | Wrap Up |  |

**หมายเหตุ** หลักสูตรและกำหนดการอบรมอาจมีการเปลี่ยนแปลงได้ตามความเหมาะสม

***เกณฑ์การคัดเลือก***

* ข้อสอบวัดพื้นฐานด้านความรู้ทาง Big Data และความรู้ทางด้านการใช้งานข้อมูล

***เกณฑ์การวัดผล***

* ใบงาน และเวลาของการเข้าอบรม

***รายชื่อวิทยากร***

1. ดร.มณีรัตน์ วงษ์ซิ้ม อาจารย์ประจำคณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม