daydev depa

หลักสูตร
นักวิทยาศาสตร์ข้อมูล
Data Scientist

ภายใต้โครงการความเป็นเลิศด้านวิทยาศาสตร์ข้อมูล: ปลดล็อกศักยภาพ สร้างอาชีพ

Data Science Excellence Initiative: Unlocking Potential, Creating Careers

สารบัญ

โครงการความเป็นเลิศด้านวิทยาศาสตร์ข้อมูล: ปลดล็อกศักยภาพ สร้างอาชีพ
Data Science Excellence Initiative: Unlocking Potential, Creating Careers

รายละเอียดโครงการโดยย่อ

ในยุคดิจิทัลที่ข้อมูลมีบทบาทสำคัญในการขับเคลื่อนธุรกิจและอุตสาหกรรม การมีทักษะด้านในยุคดิจิทัลที่ข้อมูลมีบทบาทสำคัญในการขับเคลื่อนธุรกิจและอุตสาหกรรม การมีทักษะด้านการจัดการและวิเคราะห์ข้อมูลจึงเป็นสิ่งที่จำเป็นอย่างยิ่ง โครงการนักวิทยาศาสตร์ข้อมูล (Data Scientist) นี้ถูกออกแบบมาเพื่อเตรียมความพร้อมให้กับนักศึกษาที่กำลังสำเร็จการศึกษาและผู้ที่สนใจในสายงานด้านข้อมูล ให้มีความรู้และทักษะที่จำเป็นในการเป็นนักวิทยาศาสตร์ข้อมูลที่มีประสิทธิภาพ

หลักสูตรนี้ครอบคลุมเนื้อหาที่หลากหลาย ตั้งแต่การจัดการข้อมูล การวิเคราะห์ข้อมูล การพัฒนาเทคโนโลยีปัญญาประดิษฐ์ (AI) และการนำ AI ไปประยุกต์ใช้ในธุรกิจและอุตสาหกรรม นอกจากนี้ยังเน้นจริยธรรมและความปลอดภัยของข้อมูล เพื่อให้ผู้เรียนสามารถออกแบบและพัฒนาโซลูชัน AI ที่มีประสิทธิภาพ ยั่งยืน และรับผิดชอบต่อสังคม

กลุ่มเป้าหมายของโครงการนี้คือ นักศึกษาที่กำลังสำเร็จการศึกษาในสาขาที่เกี่ยวข้องกับเทคโนโลยีสารสนเทศ วิทยาการคอมพิวเตอร์ วิศวกรรมศาสตร์ หรือสาขาอื่น ๆ ที่ต้องการพัฒนาทักษะด้านการจัดการและวิเคราะห์ข้อมูล รวมถึงบุคลากรที่ทำงานในสายงานด้านเทคโนโลยี นักวิเคราะห์ข้อมูล (Data Analyst) นักวิทยาศาสตร์ข้อมูล (Data Scientist) วิศวกร AI (AI Engineer) นักพัฒนาโปรแกรม (Software Developer) และผู้ที่ต้องการนำ AI ไปประยุกต์ใช้ในองค์กร

ปลายทางของโครงการนี้คือการเตรียมความพร้อมให้ผู้เรียนมีความรู้และทักษะที่จำเป็นในการเป็นนักวิทยาศาสตร์ข้อมูลที่มีประสิทธิภาพ สามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ในงานจริง และมีความพร้อมในการหางานในสายงานด้านข้อมูลและเทคโนโลยี นอกจากนี้ยังมีโอกาสในการสร้างเครือข่ายกับผู้เชี่ยวชาญในวงการ และได้รับคำแนะนำในการพัฒนาทักษะและการหางานจากผู้เชี่ยวชาญในสายงาน

โครงการนี้ไม่เพียงแต่จะช่วยเพิ่มพูนความรู้และทักษะของผู้เรียน แต่ยังเป็นก้าวสำคัญในการเตรียมความพร้อมสำหรับการเข้าสู่ตลาดแรงงานในยุคดิจิทัลอย่างมั่นใจและมีประสิทธิภาพ

วัตถุประสงค์โครงการ

1. เพื่อเตรียมความพร้อมให้กับนักศึกษาที่กำลังสำเร็จการศึกษาและผู้ที่สนใจในสายงานด้านข้อมูล ให้มีความรู้และทักษะที่จำเป็นในการเป็นนักวิทยาศาสตร์ข้อมูลที่มีประสิทธิภาพ
2. เพื่อส่งเสริมการพัฒนาเทคโนโลยีปัญญาประดิษฐ์ (AI) และการนำ AI ไปประยุกต์ใช้ในธุรกิจและอุตสาหกรรม
3. เพื่อเน้นจริยธรรมและความปลอดภัยของข้อมูลในการออกแบบและพัฒนาโซลูชัน AI ที่มีประสิทธิภาพ ยั่งยืน และรับผิดชอบต่อสังคม
4. สร้างโอกาสในการจ้างงาน เปิดโอกาสให้นักศึกษาได้พบกับบริษัทที่กำลังมองหาบุคลากรที่มีความสามารถ เพื่อเพิ่มโอกาสในการจ้างงานและสร้างเครือข่ายทางอาชีพ

กลุ่มเป้าหมาย จำนวน 70 คน/โครงการฯ

1. นักศึกษาชั้นปีสุดท้าย
2. บัณฑิตจบใหม่ (ไม่เกิน 2 ปีหลังสำเร็จการศึกษา) ที่ยังไม่มีงานทำ
3. ผู้สมัครต้องมี เกรดเฉลี่ยสะสม (GPA) ไม่น้อยกว่า 2.90
4. ศึกษาในสาขาวิชาที่เกี่ยวข้อง กับหลักสูตรอบรมและตำแหน่งงานที่ต้องการ

คุณสมบัติผู้เข้ารับการอบรม

1. นักศึกษาชั้นปีสุดท้าย (เกรดเฉลี่ยไม่ต่ำกว่า 2.90) หรือผู้สำเร็จการศึกษาไม่เกิน 2 ปี
2. ศึกษาในสาขาวิชาที่เกี่ยวข้องกับวิทยาการข้อมูล (Data Science) หรือสาขาใกล้เคียง
3. ผ่านเกณฑ์การทดสอบ Screen Test จากโครงการ
4. สามารถเข้าร่วมโครงการได้ตั้งแต่แรก จนจบโครงการฯ

สิ่งที่ผู้เข้ารับการอบรมจะได้รับ

1. อบรมฟรี มูลค่า 50,000 บาท (ไม่มีค่าใช้จ่าย)
2. อัพสกิล Data Scientis ครบวงจร
3. อบรมออนไลน์ 10 วัน (60 ชั่วโมง) พร้อม Workshop ฝึกปฏิบัติจริง
4. รับ 2 ใบประกาศ
   - ใบรับรองสากล IT Specialist: Data Analytics Certification
   - ใบประกาศนียบัตรจาก depa
5. เตรียมพร้อมสู่การทำงานจริง → ทำ Resume, เข้าร่วม Job Matching, สัมภาษณ์กับบริษัทพันธมิตร
6. มีโอกาสได้งานทันที ในตำแหน่ง นักวิทยาศาสตร์ข้อมูล (Data Scientist), นักวิเคราะห์ข้อมูล (Data Analyst), วิศวกร AI (AI Engineer), วิศวกรการเรียนรู้ของเครื่อง (Machine Learning Engineer), นักพัฒนาโปรแกรม (Software Developer) รวมถึงตำแหน่งที่เกี่ยวข้องกับการประยุกต์ใช้ข้อมูลและ AI ในธุรกิจ

ขั้นตอนการเข้าร่วมโครงการ

1. ลงทะเบียนผ่านเว็บไซต์โครงการ
2. โครงการคัดเลือกผู้สมัคร และประกาศผล
3. อบรมออนไลน์
4. สอบใบรับรองมาตรฐานสากล IT Specialist: Data Analytics Certification
5. ผู้อบรมจัดทำและส่ง Resume
6. กิจกรรม Job Matching (ออนไลน์)
7. บริษัทพันธมิตรนัดสัมภาษณ์
8. บริษัทคัดเลือกและตกลงจ้างงาน
9. โครงการติดตามและรายงานผล
10. สิ้นสุดโครงการ

กำหนดการฝึกอบรม (ออนไลน์) เดือนพฤศจิกายน 2568

เวลา: 09.00 – 16.00 น.

อาทิตย์	จันทร์	อังคาร	พุธ	พฤหัสบดี	ศุกร์	เสาร์
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29

สรุปวันอบรม เดือนพฤศจิกายน 2568 รวม 10 วัน ดังนี้

สัปดาห์ที่ 1 → 5, 6, 8
สัปดาห์ที่ 2 → 9, 10
สัปดาห์ที่ 1 → 19, 20,
สัปดาห์ที่ 2 → 26, 27, 29,

ตารางแผนงานและช่วงเวลาดำเนินโครงการ

ขอบเขตและแผนการดำเนินงาน	พ.ศ. 2568				พ.ศ. 2569							
	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.	มิ.ย.
1. การวางแผนและการประชาสัมพันธ์โครงการ	★	★										
2. รับสมัครผู้เข้าร่วมโครงการ	★	★										
3. กิจกรรมการฝึกอบรม		★	★	★								
4. กิจกรรม Job Matching ออนไลน์					★	★	★					
5. การติดตามผู้อบรมหลังการสัมภาษณ์							★	★				
6. การจัดทำรายงานส่งมอบงาน									★	★	★	★

รูปแบบและระยะเวลาการอบรม

1. อบรมรูปแบบออนไลน์
2. ระยะเวลา 10 วัน (60 ชั่วโมง)

ทักษะที่ได้รับ

- ☑ Programming/Coding
- ☑ Numerical Analysis
- ☑ Statistical Analysis
- ☑ Database Management
- ☑ Data Visualization
- ☑ Data Analytics
- ☑ Data Governance
- ☑ Data Architecture
- ☑ Cloud Computing
- ☑ Artificial Intelligence

ช่องทางการรับสมัคร

สมัครผ่านเว็บไซต์โครงการ https://datasci-levelup.com
ภายในวันที่ 17 ตุลาคม 2568
ประกาศผลผู้ผ่านการคัดเลือกเข้ารับอบรม วันที่ 24 ตุลาคม 2568

ผู้บริหารโครงการ

บริษัท เออาร์ไอที จำกัด
1023 อาคารเอ็มเอสสยาม ชั้น 8 ถ.พระราม 3 ช่องนนทรี ยานนาวา กรุงเทพฯ 10120

ฝ่ายประสานงานโครงการ

คุณรัชชานนท์ แป้นแก้ว
098-386-6488
Daydev.dslevelup@gmail.com

หลักสูตร นักวิทยาศาสตร์ข้อมูล (Data Scientist)

หลักสูตรนักวิทยาศาสตร์ข้อมูล (Data Scientist) มุ่งพัฒนาความรู้และทักษะด้านการจัดการข้อมูล การวิเคราะห์ข้อมูล และการพัฒนาเทคโนโลยีปัญญาประดิษฐ์ เพื่อตอบโจทย์ความต้องการของอุตสาหกรรมดิจิทัล ครอบคลุมตั้งแต่พื้นฐานข้อมูล เครื่องมือและเทคโนโลยี AI การเรียนรู้ของเครื่อง (Machine Learning) และการเรียนรู้เชิงลึก (Deep Learning) จนถึงการประยุกต์ใช้ในธุรกิจและอุตสาหกรรม เน้นทั้งภาคทฤษฎีและปฏิบัติ เพื่อให้ผู้เรียนสามารถออกแบบและพัฒนาโซลูชันที่มีประสิทธิภาพ พร้อมตระหนักถึงจริยธรรม ความปลอดภัย และความรับผิดชอบต่อสังคมในทุกขั้นตอนของการใช้ AI

วัตถุประสงค์

1. เพื่อให้ผู้เรียนมีมีความรู้พื้นฐานและแนวคิดสำคัญเกี่ยวกับ การจัดการข้อมูล Data Management), การวิเคราะห์ข้อมูล (Data Analytics), และการพัฒนา AI (Artificial Intelligence)
2. เพื่อให้ผู้เรียนมีทักษะการใช้เครื่องมือและเทคโนโลยี AI ในการออกแบบและพัฒนาโมเดลปัญญาประดิษฐ์ที่มีประสิทธิภาพ
3. เพื่อให้ผู้เรียนตระหนักด้านจริยธรรมและความปลอดภัยของ AI
4. พัฒนาทักษะเพื่อให้สามารถนำ AI และ Data Science ไปใช้ได้จริงในองค์กรธุรกิจ และอุตสาหกรรม

รูปแบบการฝึกอบรม

บรรยาย และสาธิตพร้อมให้ผู้เรียนฝึกปฏิบัติกับไฟล์ตัวอย่าง (Workshop)

เนื้อหาการอบรม วันที่ 1

- การเขียนโปรแกรมเชิงประสิทธิภาพ
  - Big-O Notation และการวิเคราะห์ประสิทธิภาพโค้ด
    - หลักการวิเคราะห์ประสิทธิภาพของอัลกอริธึม
    - ตัวอย่างอัลกอริธึมที่มีประสิทธิภาพและไม่มีประสิทธิภาพ
    - การเลือกใช้โครงสร้างข้อมูลให้เหมาะสม
  - เทคนิคการเพิ่มความเร็วโปรแกรม
    - การใช้ Multi-threading และ Multi-processing
    - การทำ Caching ด้วย functools.lru_cache และ Memoization
    - การใช้ NumPy และ Pandas ให้มีประสิทธิภาพ
- การเขียนโปรแกรมเชิงวัตถุ (OOP) ขั้นสูง
  - Design Patterns ที่สำคัญสำหรับ Data Science
    - Factory Pattern, Singleton, Observer, Strategy
    - ตัวอย่างการนำไปใช้จริงในงาน Data Science

        - Metaclasses และ Dynamic Code Generation
            - การสร้างคลาสแบบไดนามิก
            - การใช้ Metaclasses เพื่อควบคุมพฤติกรรมของคลาส
        - การจัดการหน่วยความจำใน Python
            - การทำ Garbage Collection
            - การใช้ Weak References และ Slots
            - การจัดการหน่วยความจำด้วย __slots__
    - การพัฒนาโค้ดที่ทนทานและปรับขยายได้
        - หลักการเขียนโค้ดแบบ Clean Code
            - การตั้งชื่อตัวแปรและฟังก์ชันให้สื่อความหมาย
            - การลดความซับซ้อนของโค้ด (Code Complexity)
            - การเขียนโค้ดที่อ่านง่ายและดูแลรักษาง่าย
        - การจัดการข้อผิดพลาดและ Exception Handling ที่มีประสิทธิภาพ
            - Best Practices ในการใช้ Try-Except
            - การสร้าง Custom Exception
        - การพัฒนาโค้ดแบบ Modular และ Reusable
            - การแยกโค้ดเป็นโมดูล
            - การใช้ Dependency Injection
            - การทำ Unit Testing และ Test-Driven Development (TDD)
    - การพัฒนาโครงการจริง
        - การออกแบบระบบวิเคราะห์ข้อมูลที่ซับซ้อน
            - ออกแบบโครงสร้างโค้ดที่ปรับขยายได้
            - ใช้ OOP และ Design Patterns ในโครงการ
        - การปรับปรุงโค้ดจาก Performance Bottleneck
            - วิเคราะห์และแก้ไขปัญหาประสิทธิภาพในโค้ด
            - ใช้เครื่องมือ Profiler เช่น cProfile และ line_profiler
        - การพัฒนา API ขนาดเล็กสำหรับ Data Science
            - สร้าง REST API ด้วย FastAPI
            - การทำให้ API รองรับงานหนักโดยใช้ Asynchronous Programming

เนื้อหาการอบรม วันที่ 2

- บทนำและแนวคิดสำคัญของ Numerical Analysis
  - Overview of Numerical Analysis
    - ความสำคัญของ Numerical Analysis ใน Data Science และวิศวกรรม
    - ปัญหาที่สามารถแก้ไขได้ด้วย Numerical Methods
    - ความสัมพันธ์ระหว่าง Numerical Analysis และ Machine Learning
  - Error Analysis & Stability
    - ชนิดของข้อผิดพลาด (Truncation Error, Round-off Error)
    - Forward, Backward และ Mixed Error
    - วิธีลดข้อผิดพลาดในการคำนวณเชิงตัวเลข
  - Conditioning & Stability in Computation
    - Well-conditioned vs. Ill-conditioned problems
    - ความเสถียรของอัลกอริธึมทางตัวเลข
    - ตัวอย่างการใช้ Python/NumPy ในการศึกษาปัญหาความเสถียร
- เทคนิคการแก้ปัญหาทางตัวเลขขั้นสูง
  - Root Finding Methods (การหาค่ารากของสมการเชิงตัวเลข)
    - Bisection Method, Newton-Raphson Method, Secant Method
    - ความเร็วของการลู่เข้า (Convergence Rate)
    - การนำไปใช้ในปัญหาทาง Data Science เช่น Optimization
  - Numerical Linear Algebra (พีชคณิตเชิงเส้นเชิงตัวเลข)
    - Gaussian Elimination & LU Decomposition
    - Eigenvalues & Eigenvectors (Power Method)
    - Singular Value Decomposition (SVD) และการประยุกต์ใช้ใน Machine Learning
  - Optimization Techniques in Numerical Analysis (การหาค่าต่ำสุด/สูงสุดของฟังก์ชันเชิงตัวเลข)
    - Gradient Descent และ Variants (Momentum, Adam, RMSProp)
    - Constrained vs. Unconstrained Optimization
    - Convex vs. Non-convex Optimization
    - ตัวอย่างจาก Deep Learning และ Reinforcement Learning
- การประมาณค่าและการแก้สมการเชิงอนุพันธ์
  - Numerical Interpolation & Approximation (การประมาณค่าเชิงตัวเลข)
    - Polynomial Interpolation (Lagrange, Newton)
    - Splines & B-Splines
    - Least Squares Approximation และการใช้ PCA ใน Data Science

- Numerical Differentiation & Integration (การหาอนุพันธ์และปริพันธ์เชิงตัวเลข)
  - Finite Difference Methods (Forward, Backward, Central)
  - Newton-Cotes Integration (Trapezoidal & Simpson's Rule)
  - Monte Carlo Integration
- Solving Ordinary & Partial Differential Equations (ODEs & PDEs) (การแก้สมการเชิงอนุพันธ์เชิงตัวเลข)
  - Euler's Method, Runge-Kutta Methods (RK4)
  - Finite Difference & Finite Element Methods สำหรับ PDEs
  - ตัวอย่างการใช้ ODE/PDE ในการจำลองทางฟิสิกส์และวิทยาการข้อมูล

เนื้อหาการอบรม วันที่ 3

- การทดสอบสมมติฐานขั้นสูง (Advanced Hypothesis Testing)
  - ทบทวนแนวคิดการทดสอบสมมติฐาน
    - หลักการของ p-value, Confidence Interval
    - ประเภทของข้อผิดพลาด (Type I & Type II Errors)
    - ปัญหาของ Multiple Comparisons และวิธีแก้ไข
  - เทคนิคการทดสอบสมมติฐานขั้นสูง
    - Permutation Test & Bootstrap: การสร้าง Distribution โดยไม่อิง Normality
    - Bayesian Hypothesis Testing: การใช้แนวทางแบบ Bayesian
    - Monte Carlo Simulation: การสร้างตัวอย่างแบบสุ่มเพื่อทดสอบสมมติฐาน
  - การเปรียบเทียบมากกว่าสองกลุ่ม
    - ANOVA ขั้นสูง (Repeated Measures ANOVA, MANOVA)
    - Kruskal-Wallis Test & Friedman Test (ทางเลือกสำหรับ Non-parametric)
    - Post-hoc Analysis (Tukey, Bonferroni, Scheffé, Holm-Bonferroni)
- การวิเคราะห์การถดถอยขั้นสูง (Advanced Regression Analysis)
  - การถดถอยเชิงเส้นขั้นสูง (Multiple & Generalized Linear Regression)
    - การวิเคราะห์ Multicollinearity (VIF, Condition Index)
    - การตรวจสอบ Heteroscedasticity และการแก้ไข
    - Robust Regression และการใช้ Weighted Least Squares (WLS)
  - การถดถอยเชิงไม่เชิงเส้น (Non-Linear Regression)
    - Polynomial Regression
    - Spline Regression
    - Generalized Additive Models (GAMs)
  - การถดถอยโลจิสติกและการขยายผล (Logistic & Advanced Classification Models)

- Multinomial & Ordinal Logistic Regression
- Regularized Regression (Ridge, Lasso, Elastic Net)
- Bayesian Regression

- การวิเคราะห์ข้อมูลหลายตัวแปร (Multivariate Data Analysis)
  - การวิเคราะห์องค์ประกอบหลัก (PCA - Principal Component Analysis)
    - หลักการลดมิติข้อมูล
    - การเลือกจำนวนองค์ประกอบที่เหมาะสม
    - การตีความผลลัพธ์ของ PCA
  - การวิเคราะห์กลุ่มข้อมูล (Cluster Analysis)
    - K-Means, Hierarchical Clustering
    - Gaussian Mixture Model (GMM)
    - วิธีวัดความเหมาะสมของกลุ่ม (Silhouette Score, Davies-Bouldin Index)
  - เทคนิคการลดมิติข้อมูลอื่น ๆ
    - t-SNE & UMAP
    - Factor Analysis & Latent Variable Models
- การวิเคราะห์อนุกรมเวลา (Time Series Analysis)
  - หลักการพื้นฐานของอนุกรมเวลา
    - Stationarity & Differencing
    - ACF/PACF & การเลือก Lag ที่เหมาะสม
  - โมเดลเชิงสถิติสำหรับอนุกรมเวลา
    - ARIMA, SARIMA, Prophet
    - การวิเคราะห์ Seasonal & Trend Components
  - การพยากรณ์เชิงลึก (Deep Time Series Forecasting)
    - LSTM & GRU
    - Hybrid Models (SARIMA + ML Models)
- Machine Learning เชิงสถิติ (Statistical Machine Learning)
  - หลักการของโมเดลเชิงสถิติใน Machine Learning
    - เปรียบเทียบ Parametric vs. Non-parametric Approaches
    - Bias-Variance Tradeoff
  - เทคนิคสำคัญในการเลือกโมเดล
    - Cross-Validation & Bootstrapping
    - Feature Selection (RFE, Boruta, Mutual Information)
  - โมเดลสถิติใน ML ยอดนิยม
    - Bayesian Networks
    - Hidden Markov Models
    - Gaussian Processes

เนื้อหาการอบรมวันที่ 4

- แนวคิดเชิงสถาปัตยกรรมของฐานข้อมูล (Database Architecture)
  - โครงสร้างภายในของ RDBMS (Relational Database Management System)
  - สถาปัตยกรรม Client-Server และ Distributed Database
  - ฐานข้อมูลในระบบ Cloud (Cloud Databases) เช่น AWS RDS, Google Cloud SQL
- การออกแบบฐานข้อมูลขั้นสูง (Advanced Database Design)
  - การออกแบบ Normalization และ Denormalization
  - การเลือกใช้ Index อย่างมีประสิทธิภาพ
  - Partitioning และ Sharding – วิธีแบ่งข้อมูลในฐานข้อมูลขนาดใหญ่
  - การจัดการความสัมพันธ์ระหว่างตาราง (Foreign Key, Cascade, Referential Integrity)
- การเพิ่มประสิทธิภาพ (Performance Tuning)
  - เทคนิคการปรับปรุง Query ให้ทำงานได้เร็วขึ้น
  - การใช้ Explain Plan และ Query Execution Plan
  - การใช้ Materialized Views, Caching และ Read Replicas
  - การลด Locking และ Deadlocks
- การจัดการฐานข้อมูลสำหรับ Big Data
  - หลักการของฐานข้อมูล NoSQL (Key-Value Store, Document Store, Columnar, Graph)
  - เปรียบเทียบฐานข้อมูล SQL vs. NoSQL
  - การเลือกใช้ฐานข้อมูลให้เหมาะกับงาน เช่น MySQL, PostgreSQL, MongoDB, Cassandra
- ความปลอดภัยและการสำรองข้อมูล (Database Security & Backup)
  - การเข้ารหัสข้อมูล (Encryption) และการจัดการสิทธิ์ผู้ใช้ (Role-based Access Control)
  - การสำรองและกู้คืนฐานข้อมูล (Backup & Recovery)
  - การป้องกัน SQL Injection และภัยคุกคามทางไซเบอร์
  - การบันทึก Log และ Audit Trail
- การบริหารจัดการฐานข้อมูลในระบบที่ต้องรองรับการขยายตัว (Scaling & High Availability)
  - การออกแบบ High Availability ด้วย Replication และ Failover
  - การใช้ Load Balancer ในการจัดการทราฟฟิกฐานข้อมูล
  - การออกแบบ Multi-region Database Architecture
  - Case Study และแนวปฏิบัติที่ดีที่สุด

เนื้อหาการอบรม วันที่ 5

- หลักการออกแบบ Data Visualization ที่มีประสิทธิภาพ
  - หลักจิตวิทยาของการมองเห็น (Visual Perception)
    - Gestalt Principles และการออกแบบให้ดึงดูดสายตา
    - Pre-attentive attributes: Color, Shape, Position
  - การเลือกประเภทของกราฟที่เหมาะสม (Chart Selection)
    - กราฟเปรียบเทียบ (Comparison): Bar Chart, Line Chart
    - กราฟแสดงสัดส่วน (Proportion): Pie Chart, Donut Chart
    - กราฟแนวโน้ม (Trend): Line Chart, Area Chart
    - กราฟความสัมพันธ์ (Relationship): Scatter Plot, Bubble Chart
    - กราฟเชิงลึก (Complex): Heatmap, Treemap, Network Diagram
  - แนวทางป้องกันการใช้กราฟที่ทำให้เข้าใจผิด (Misleading Visualization)
- การสร้าง Data Visualization ขั้นสูงด้วย Python
  - Matplotlib และ Seaborn สำหรับการสร้างกราฟขั้นสูง
    - ปรับแต่งกราฟขั้นสูงด้วย Matplotlib (Customizing Plots)
    - การใช้ Seaborn สำหรับ Data Visualization เชิงสถิติ
    - การสร้าง Multi-Panel Plots และ Subplots
    - การใช้ Color Palette และการเลือกสีให้เหมาะสม
    - การเพิ่ม Annotations และ Labels เพื่อเน้นข้อมูลสำคัญ
  - การสร้าง Interactive Visualization ด้วย Plotly และ Bokeh
    - การสร้าง Interactive Line Chart, Scatter Plot และ Heatmap
    - การใช้ Hover Effects และ Dynamic Updates
    - การทำ Animation เพื่อแสดงการเปลี่ยนแปลงของข้อมูลตามเวลา
- การออกแบบ Dashboard ขั้นสูง
  - การสร้าง Dashboard ด้วย Tableau หรือ Power BI
    - การเชื่อมต่อกับแหล่งข้อมูล
    - การใช้ Calculated Fields และ Parameter เพื่อสร้างกราฟแบบ Dynamic
    - การสร้าง Filters, Slicers และ Interactive Elements
    - การใช้ Map Visualization สำหรับข้อมูลเชิงพื้นที่
    - การออกแบบ Storytelling Dashboard เพื่อเล่าเรื่องด้วยข้อมูล
  - การสร้าง Web-based Dashboard ด้วย Python (Dash/Streamlit)
    - พื้นฐานของ Dash และ Streamlit
    - การสร้าง Interactive Dashboard ด้วย Python
    - การนำข้อมูลแบบ Real-time มาแสดงผล
    - การ Deploy Dashboard ให้พร้อมใช้งาน

- Best Practices และ Case Study
  - Best Practices ในการออกแบบ Data Visualization
    - การใช้สี, ฟอนต์ และ Layout ที่เหมาะสม
    - การทำ Data Storytelling อย่างมีประสิทธิภาพ
    - การหลีกเลี่ยง Pitfalls ในการสร้าง Visualization
  - Workshop Case Study
    - ออกแบบ Visualization สำหรับ Business Dashboard
    - การนำเสนอข้อมูลทางสถิติสำหรับงานวิจัย
    - สร้าง Data Storytelling สำหรับสื่อและการตลาด

เนื้อหาการอบรม วันที่ 6

- การเตรียมข้อมูลสำหรับการวิเคราะห์ (Data Preparation)
  - แนวคิดเกี่ยวกับคุณภาพของข้อมูล (Data Quality)
    - การตรวจสอบข้อมูลขาดหาย (Missing Data) และวิธีการจัดการ
    - การตรวจจับและแก้ไขข้อมูลผิดปกติ (Outlier Detection & Handling)
    - การแปลงข้อมูล (Data Transformation) และการปรับขนาด (Feature Scaling)
  - การลดมิติข้อมูล (Dimensionality Reduction)
    - หลักการของการลดมิติ (Feature Selection vs. Feature Extraction)
    - เทคนิค PCA (Principal Component Analysis)
    - เทคนิค t-SNE และ UMAP สำหรับข้อมูลที่มีมิติมาก
- การวิเคราะห์เชิงสถิติขั้นสูง (Advanced Statistical Analysis)
  - การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร (Correlation & Causation)
    - Pearson, Spearman, Kendall Correlation
    - Granger Causality Test
  - การวิเคราะห์ความแปรปรวน (ANOVA & MANOVA)
    - One-way ANOVA, Two-way ANOVA
    - Multivariate ANOVA (MANOVA)
  - การสร้างแบบจำลองทางสถิติขั้นสูง (Advanced Regression)
    - Multiple Regression & Interaction Effects
    - Generalized Linear Models (GLM)
    - Time Series Analysis (ARIMA, SARIMA, Exponential Smoothing)
- การใช้ Machine Learning เพื่อการวิเคราะห์ข้อมูล (Machine Learning for Data Analytics)
  - การเลือกโมเดลที่เหมาะสมกับข้อมูล (Model Selection)
    - Supervised Learning vs. Unsupervised Learning

- การเลือกค่า Hyperparameters และการปรับแต่งโมเดล (Hyperparameter Tuning)
- การจัดกลุ่มข้อมูล (Clustering Techniques)
  - K-Means Clustering, DBSCAN
  - Hierarchical Clustering
- การพยากรณ์ข้อมูล (Predictive Analytics)
  - Decision Trees, Random Forest
  - Gradient Boosting (XGBoost, LightGBM, CatBoost)

- การวิเคราะห์ข้อความและข้อมูลที่ไม่มีโครงสร้าง (Text & Unstructured Data Analytics)
  - พื้นฐานการวิเคราะห์ข้อความ (Text Analytics)
    - Tokenization, Stopword Removal, Lemmatization
    - การสร้างเวกเตอร์ข้อความ (TF-IDF, Word2Vec, FastText)
  - การวิเคราะห์ความรู้สึก (Sentiment Analysis)
    - การใช้ Naïve Bayes และ LSTM สำหรับวิเคราะห์อารมณ์ของข้อความ
    - การนำผลไปใช้งาน เช่น วิเคราะห์รีวิวสินค้า หรือ Social Media Monitoring
  - การสกัดคุณสมบัติจากข้อมูลที่ไม่มีโครงสร้าง (Feature Extraction from Unstructured Data)
    - การวิเคราะห์ข้อมูลเสียงและภาพเบื้องต้น

เนื้อหาการอบรม วันที่ 7

- พื้นฐานและแนวคิดหลักของธรรมาภิบาลข้อมูล
  - ความหมายของธรรมาภิบาลข้อมูล
    - ความแตกต่างระหว่าง Data Governance, Data Management และ Data Stewardship
    - ทำไม Data Governance จึงเป็นเรื่องสำคัญในองค์กร
  - องค์ประกอบสำคัญของ Data Governance
    - Data Policies & Standards
    - Data Ownership & Stewardship
    - Data Quality & Integrity
    - Data Security & Privacy
  - กรณีศึกษา: ความล้มเหลวของ Data Governance ในองค์กรชั้นนำ
- กรอบการดำเนินงานของ Data Governance
  - Data Governance Frameworks ที่ได้รับความนิยม
    - DAMA DMBOK (Data Management Body of Knowledge)
    - EDM Council DCAM (Data Capability Assessment Model)
    - CMMI Data Management Maturity (DMM)

- บทบาทและหน้าที่ใน Data Governance
  - Chief Data Officer (CDO)
  - Data Owners, Data Stewards และ Data Custodians
  - Data Governance Council
- แนวทางการกำหนดนโยบาย Data Governance ให้สอดคล้องกับเป้าหมายองค์กร

○ คุณภาพและการจัดการข้อมูล
- Data Quality Dimensions
  - Accuracy, Completeness, Consistency
  - Timeliness, Validity, Uniqueness
- แนวทางการตรวจสอบและปรับปรุงคุณภาพข้อมูล
  - Data Profiling & Data Cleansing
  - Master Data Management (MDM)
- เครื่องมือที่ใช้ในการจัดการคุณภาพข้อมูล (เช่น Talend, Informatica, IBM InfoSphere)

○ ความปลอดภัยและการปฏิบัติตามกฎหมายด้านข้อมูล
- มาตรฐานและกฎหมายที่เกี่ยวข้องกับข้อมูล
  - GDPR (General Data Protection Regulation)
  - PDPA (Personal Data Protection Act)
  - HIPAA (Health Insurance Portability and Accountability Act)
- Data Classification & Access Control
  - การจำแนกประเภทข้อมูล (Sensitive Data, Public Data, Restricted Data)
  - Role-Based Access Control (RBAC)
- Data Encryption & Masking
  - แนวทางการเข้ารหัสข้อมูล
  - เทคนิคการทำ Data Anonymization

○ การนำ Data Governance ไปใช้ในองค์กร
- กระบวนการนำ Data Governance ไปใช้ (Implementation Roadmap)
  - ขั้นตอนการวางกลยุทธ์ Data Governance
  - การบริหารความเปลี่ยนแปลง (Change Management)
- เครื่องมือและแพลตฟอร์มที่ใช้ในการบริหาร Data Governance
  - Collibra, Alation, Informatica Axon
  - Case Study: องค์กรที่ประสบความสำเร็จในการนำ Data Governance ไปใช้

เนื้อหาการอบรม วันที่ 8

- พื้นฐานและแนวคิดหลักของ Data Architecture
  - บทบาทของ Data Architecture ในองค์กรและความสำคัญ
  - องค์ประกอบสำคัญของ Data Architecture (Data Ingestion, Storage, Processing, Access)
  - รูปแบบของ Data Architecture แบบดั้งเดิมและแบบใหม่
    - Monolithic vs. Distributed Data Architecture
    - On-Premises vs. Cloud-based Architecture
  - การเลือกใช้สถาปัตยกรรมข้อมูลให้เหมาะกับองค์กร
- Data Warehouse vs. Data Lake vs. Lakehouse
  - แนวคิดของ Data Warehouse
    - โครงสร้างแบบ Schema-on-Write
    - การจัดการ ETL และ ELT
    - ตัวอย่างเครื่องมือ: Amazon Redshift, Google BigQuery, Snowflake
  - แนวคิดของ Data Lake
    - โครงสร้างแบบ Schema-on-Read
    - การรองรับข้อมูลกึ่งโครงสร้าง (Semi-Structured) และไม่มีโครงสร้าง (Unstructured)
    - ตัวอย่างเครื่องมือ: AWS S3, Azure Data Lake Storage, Apache Hadoop
  - แนวคิดของ Lakehouse Architecture
    - ข้อดีและข้อเสียของ Data Lake และ Data Warehouse
    - การรวมข้อดีของทั้งสองสถาปัตยกรรม
    - ตัวอย่างเครื่องมือ: Databricks Delta Lake, Apache Iceberg
- แนวทางใหม่ของ Data Architecture – Data Mesh & Data Fabric
  - Data Mesh คืออะไร?
    - แนวคิดการกระจายอำนาจของข้อมูล (Decentralized Data Ownership)
    - การเปลี่ยนจาก Data Lake ไปสู่แนวคิด Data-as-a-Product
    - บทบาทของ Domain-Oriented Data Teams
    - ตัวอย่างองค์กรที่ใช้ Data Mesh
  - Data Fabric คืออะไร?
    - การใช้ AI/ML ช่วยบริหารจัดการข้อมูลอัตโนมัติ
    - การผสานรวมข้อมูลจากหลายแหล่งอย่างไร้รอยต่อ
    - เปรียบเทียบ Data Fabric กับ Data Mesh
  - การออกแบบสถาปัตยกรรมข้อมูลในองค์กรที่ซับซ้อน
    - การเลือกใช้ Data Mesh หรือ Data Fabric
    - การจัดการ Metadata และ Data Governance ในระบบใหม่

        - เครื่องมือที่ใช้: AWS Glue, Apache Kafka, dbt, Trino
    - Cloud Data Architecture & Scalable Data Infrastructure
        - Cloud Data Architecture คืออะไร?
            - Public, Private, Hybrid Cloud สำหรับ Data Architecture
            - Multi-Cloud Strategy
            - Serverless Data Architecture และตัวอย่างการใช้งาน
        - การออกแบบระบบข้อมูลที่สามารถขยายขนาดได้ (Scalable Data Infrastructure)
            - การใช้ Containerization และ Kubernetes ใน Data Architecture
            - การประมวลผลแบบ Batch Processing vs. Stream Processing
            - เครื่องมือ: Apache Spark, Apache Flink, AWS Lambda
        - Best Practices ในการจัดการ Data Pipeline
            - การทำ Data Orchestration ด้วย Apache Airflow
            - การสร้าง Data Lineage และ Data Quality Control

เนื้อหาการอบรม วันที่ 9

- ภาพรวมของ Cloud Computing ขั้นสูง
    - แนวคิดการประมวลผลบนคลาวด์ขั้นสูง
        - การเปรียบเทียบระหว่าง Public, Private และ Hybrid Cloud
        - การออกแบบสถาปัตยกรรมแบบ Multi-Cloud และ Hybrid Cloud
        - แนวทางปฏิบัติที่ดีที่สุดในการเลือกใช้คลาวด์
    - เทคโนโลยีคลาวด์ที่สำคัญในปัจจุบัน
        - Serverless Computing และ FaaS (Function as a Service)
        - Edge Computing และการนำไปใช้ในอุตสาหกรรม
        - AI และ Machine Learning บนคลาวด์
- การจัดการโครงสร้างพื้นฐานคลาวด์
    - Infrastructure as Code (IaC)
        - Terraform และ AWS CloudFormation
        - การทำ CI/CD Pipeline บนคลาวด์
    - Containerization & Orchestration
        - Docker และ Containerized Applications
        - Kubernetes (K8s) และการบริหารจัดการคลัสเตอร์
        - การทำ Auto-scaling และ Load Balancing
    - Serverless Computing และ Microservices
        - การออกแบบแอปพลิเคชันแบบ Serverless
        - AWS Lambda, Google Cloud Functions, Azure Functions
        - Event-driven Architecture และ API Gateway

- ความปลอดภัยและการบริหารจัดการคลาวด์
  - แนวทางปฏิบัติด้านความปลอดภัยบนคลาวด์
    - การป้องกันข้อมูลรั่วไหลและการเข้ารหัสข้อมูล
    - IAM (Identity and Access Management) และ Zero Trust Model
    - การป้องกัน DDoS และการรักษาความปลอดภัยของ API
  - การบริหารจัดการต้นทุนคลาวด์
    - วิธีเพิ่มประสิทธิภาพต้นทุนบน AWS, GCP และ Azure
    - การทำ Cost Forecasting และการใช้ Reserved Instances
    - การตรวจสอบและวิเคราะห์ค่าใช้จ่ายคลาวด์

เนื้อหาการอบรม วันที่ 10

- บทนำสู่ปัญญาประดิษฐ์ขั้นสูง (Introduction to Advanced AI)
  - ความแตกต่างระหว่าง AI ทั่วไป (Traditional AI) กับ AI ขั้นสูง
  - สถาปัตยกรรมของ AI สมัยใหม่ (Modern AI Architectures)
  - ภาพรวมของเทคนิค AI ขั้นสูงที่ใช้ในปัจจุบัน
- การเรียนรู้เชิงลึกขั้นสูง (Advanced Deep Learning)]
  - โมเดลโครงข่ายประสาทเทียมขั้นสูง (Advanced Neural Networks)
    - Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), และ Gated Recurrent Units (GRU)
    - Transformer Model และการใช้งานใน NLP เช่น BERT และ GPT
  - Generative AI และโมเดลสร้างข้อมูล (Generative Models & AI)
    - Generative Adversarial Networks (GANs) และ Variational Autoencoders (VAEs)
    - Diffusion Models และ Stable Diffusion
  - เทคนิคเพิ่มประสิทธิภาพโมเดล (Optimization & Fine-Tuning)
    - Hyperparameter Tuning & Optimization Techniques
    - Transfer Learning และ Few-Shot Learning
    - Model Pruning และ Quantization
  - การพัฒนา AI ที่สามารถอธิบายได้ (Explainable AI - XAI)
    - ความสำคัญของ Explainability และ Trustworthiness ใน AI
    - เทคนิคการทำ Explainable AI (SHAP, LIME, Grad-CAM)
    - ตัวอย่างการใช้งาน XAI ในสาขาต่างๆ เช่น การแพทย์ การเงิน และกฎหมาย
- การประยุกต์ใช้ AI ในระดับอุตสาหกรรม (AI in Industry)
  - AI สำหรับงานอุตสาหกรรม (Industrial AI Use Cases)
    - AI ใน Healthcare: การวิเคราะห์ภาพทางการแพทย์ และการพยากรณ์โรค
    - AI ในการเงิน: Fraud Detection และ Algorithmic Trading
    - AI ในการผลิต: Predictive Maintenance และ Robotics

- การนำ AI ไปใช้ในระดับองค์กร (AI Deployment & MLOps)
  - การจัดการเวิร์กโฟลว์ของโมเดล AI ด้วย MLOps
  - การใช้ Kubernetes, Docker และ CI/CD pipeline สำหรับ AI
- ความปลอดภัยและจริยธรรมของ AI (AI Safety & Ethics)
  - ความเสี่ยงของ AI และวิธีป้องกัน Bias ในโมเดล
  - กฎหมายและแนวทางการกำกับดูแล AI (เช่น GDPR, AI Act)
  - AI Alignment และการออกแบบ AI ที่มีความปลอดภัยต่อมนุษย์