Title: AI/ML Lab: AWS Foundations for AI Development
Lecturers: Nicholas Beaudoin
https://ctme.caltech.edu/educator/beaudoin-nicholas/

Description:
This fast-paced, three-day learning lab complements Caltech CTME’s AI/ML offerings by focusing on the Amazon Web Services (AWS) ecosystem to enhance AI and Machine Learning skills. Participants will explore cloud-based architectures for both low-code and coding development environments, deepening their coding abilities and understanding of model deployment and machine learning operations. This lab prepares attendees for the AWS Machine Learning Specialty certification and introduces advanced approaches to machine learning solutions.

Target Audience:
This lab is designed for proficient engineers and citizen developers who possess a solid foundation in machine learning and Python coding. It offers an introduction to cloud-based development while preparing participants for the AWS Machine Learning Specialty certification.

Learning Outcomes:
You will:
- Gain awareness of the fundamentals of AWS tools, enhancing technical proficiency in cloud-based AI and ML deployments
- Develop the ability to efficiently set up and manage cloud environments for machine learning applications ready for industry
- Gain insights into low-code development environments to streamline project workflows
- Enhance confidence in handling complex datasets and machine learning models within AWS
- Prepare for the AWS Machine Learning Specialty certification
- Foster an innovative mindset through hands-on experience with advanced AWS functionalities
- Cultivate strategic insights into cost control, scalability, and security measures in cloud-based projects

Topics:
This lab focuses on example products from AWS AI/ML portfolio, which are subject to change.
- Amazon SageMaker: Comprehensive training in building, training, and deploying scalable machine learning models
- AWS Bedrock: Insights into simplifying Large Language Model (LLM) development, enhancing model management efficiency
- AWS JumpStart: Utilization of pre-built machine learning models and solutions, accelerating project timelines
- Foundational AWS Components:
  - Security: Key strategies including IAM, encryption, and security
  - Cost Control: Techniques for cost optimization in large-scale ML projects
  - Scalability: Approaches to scaling AI and ML solutions efficiently
  - Serverless Computing: Leveraging AWS Lambda and other serverless technologies

Training Delivery:
Participants will be graded pass/fail based on participation in in-class discussions and exercises. There are no exams or out-of-class assignments. Learners will have access to AWS products for use in this short course.