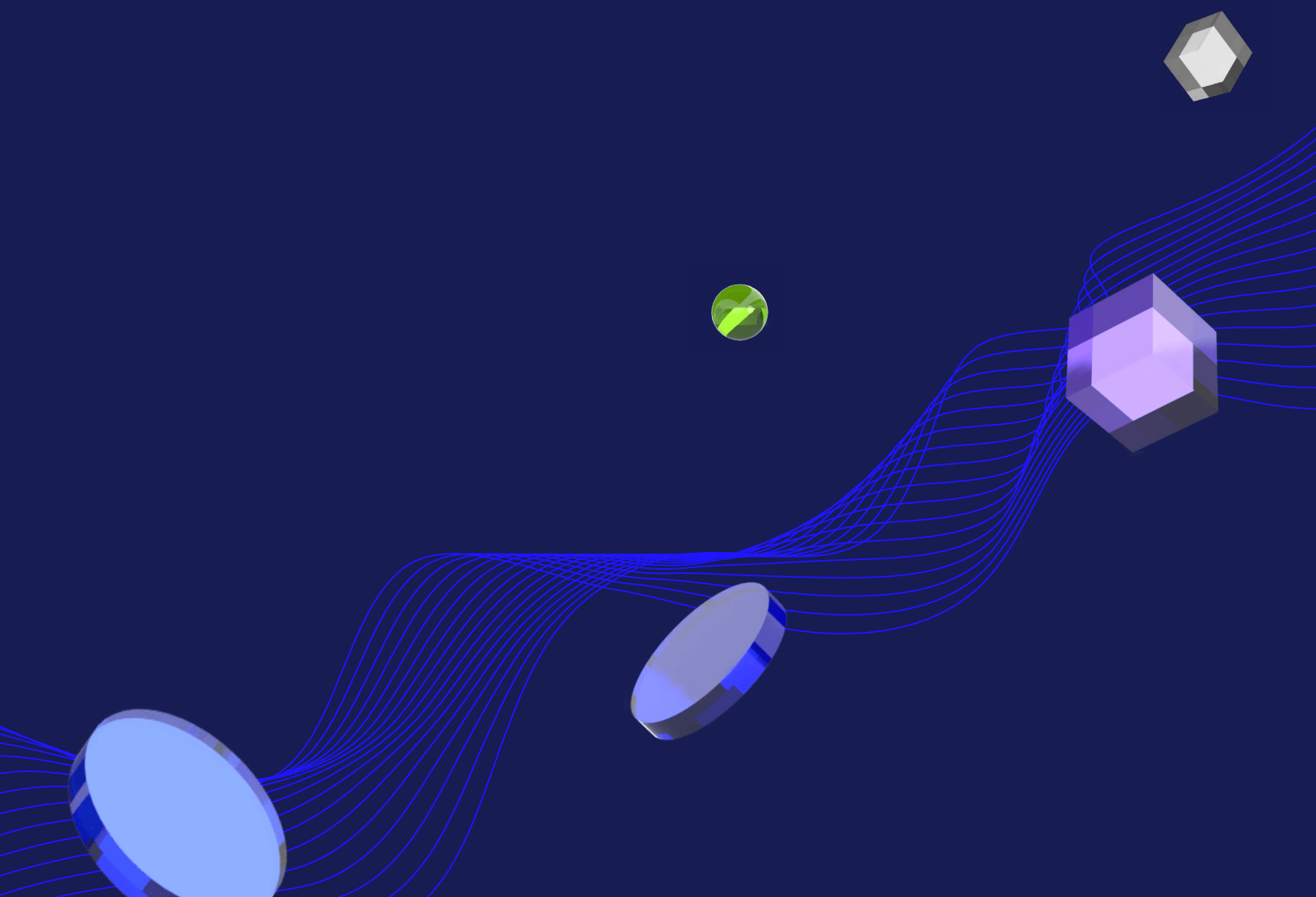
INDIVIDUAL LEARNERS

UDACITY

SCHOOL OF ARTIFICIAL INTELLIGENCE

AWS Machine Learning Engineer

Nanodegree Program Syllabus



Overview

The goal of the AWS Machine Learning Engineer (MLE) Nanodegree program is to equip software developers/data scientists with the data science and machine learning skills required to build and deploy machine learning models in production using Amazon SageMaker. This program will focus on the latest best practices and capabilities that are enabled by Amazon SageMaker, including new model design/deployment features and case studies in which they can be applied to.

Learning Objectives

A graduate of this program will be able to:

- Create machine learning models in Sagemaker on datasets cleaned using AWS tools.
- Deploy machine learning models to an API endpoint and integrate it into a full workflow.
- Solve computer vision and natural language problems using fine-tuned deep neural networks.
- Operationalize a machine learning pipeline using SageMaker to allow for training and deployment on industry-scale problems.
- Select a machine learning challenge and propose a possible solution.

Built in collaboration with:

Program information

Estimated Time

5 months at 5-10hrs/week*

Skill Level

Intermediate

Prerequisites

Python programming knowledge, including:

- At least 40 hours of programming experience.
- Familiarity with data structures like dictionaries and lists.
- Experience with libraries like NumPy and pandas.
- Knowledge of functions, variables, loops, and classes.
- Exposure to Python through Jupyter Notebooks is recommended.
- Experience with constructing and calling HTTP API endpoints is recommended.

Basic knowledge of machine learning algorithms, including:

- Basic understanding of the machine learning workflow.
- Basic theoretical understanding of ML algorithms such as linear regression, logistic regression, and neural network.
- Basic understanding of model training and testing processes.
- Basic knowledge of commonly used metrics for ML models evaluation such as accuracy, precision, recall, and mean square error (MSE).

Required Hardware/Software

There are no software and version requirements to complete this Nanodegree program.

*The length of this program is an estimation of total hours the average student may take to complete all required coursework, including lecture and project time. If you spend about 5-10 hours per week working through the program, you should finish within the time provided. Actual hours may vary.

Course 1

Introduction to Machine Learning

In this course, learners will be introduced to high level concepts through AWS SageMaker. They'll begin by using SageMaker Studio to perform exploratory data analysis. Know how and when to apply the basic concepts of machine learning to real world scenarios. Create machine learning workflows, starting with data cleaning and feature engineering, to evaluation and hyperparameter tuning. Finally, build new ML workflows with highly sophisticated models such as XGBoost and AutoGluon.

Course Project

Predict Bike Sharing Demand with AutoGluon

In this project, learners will apply the knowledge and methods they learned in the Introduction to Machine Learning course to compete in a Kaggle competition. Using the AutoGluon framework, learners will first train a baseline model, then improve their model through feature engineering and hyperparameter tuning. Finally, they'll submit their optimized model for a public Kaggle rank and write a report on their findings to showcase their work.

Lesson 1

Exploratory Data Analysis

- Use AWS SageMaker Studio to access datasets from S3 and perform data analysis functions using AWS tools.
- Perform data analysis and feature engineering with Data Wrangler.
- Perform data analysis and feature engineering with Pandas in SageMaker Studio.
- Label new data for a dataset with Sagemaker ground truth.

Lesson 2

Machine Learning Concepts

- Design a domain, model, and data outline for a case study.
- Build a ML lifecycle and apply it to a dataset.
- Differentiate between supervised and unsupervised models and apply them to an appropriate dataset.
- Differentiate between regression and classification methods and apply them to an appropriate dataset.

Lesson 3

Model Deployment Workflow

- Load new dataset, create 3 data set types, and identify features/values in SageMaker.
- Clean or create new features from a dataset.
- Train (fit) a regression/classification model using scikit learn.
- Evaluate a trained model using methods like mse, rmse, r2, accuracy, f1, and precision.
- Tune a model's hyper parameters to achieve a better result.

Lesson 4

Algorithms & Tools

- Train, test, and optimize a linear model, tree-based model, XGBoost model, and AutoGluon Tabular prediction model.
- Create a model using Sagemaker Jumpstart.

Course 2

Developing Your First ML Workflow

In order to execute on machine learning's versatile capabilities, we need to have the infrastructure to execute our ML operations. With the easy availability of managed infrastructure from AWS, we can dynamically create the necessary resources to train, deploy, and evaluate our models. In this course, learners will create general machine learning workflows on AWS.

Begin with an introduction to the general principles of machine learning engineering. From there, learn the fundamentals of SageMaker to train, deploy, and evaluate a model. Following that, learn how to create a machine learning workflow on AWS utilizing tools like Lambda and Step Functions. Finally, learn how to monitor machine learning workflows with services like Model Monitor and Feature Store. With all this, learners will have all the information they need to create an end-to-end machine learning pipeline.

Course Project

Build a ML Workflow on SageMaker

In this project, learners will develop an end-to-end ML Workflow on SageMaker, Lambda, and Step Functions. Learners will showcase their model deployment capabilities with SageMaker Model Endpoints and Lambda, and their workflow monitoring capabilities with SageMaker Model Monitor and Step Functions. At the end of the project, learners will be able to demonstrate building a scalable ML workflow on SageMaker.

Lesson 1

Introduction to MLE

- Understand the prerequisites.
- Describe key business stakeholders.
- Understand the history of MLE.
- Describe when to use MLE.

Lesson 2

SageMaker Essentials

- Launch training jobs within SageMaker.
- Deploy an endpoint that can perform inference on live data.
- Evaluate datasets with batch transform jobs.
- Perform custom processing jobs on raw data.

Lesson 3

Designing Your Own Workflow

- Create Lambda functions.
- Trigger Lambda functions utilizing both the SDK and other AWS Services.
- Design and execute a workflow utilizing State Machines.
- Learn about the use cases for SageMaker Pipelines.

Lesson 4

Monitoring a ML Workflow

- Use SageMaker Feature Store to serve and monitor model data.
- Configure SageMaker Model Monitor to generate and track metrics about our models.
- Use Clarify to explain model predictions and surface biases in models.

Course 3

Deep Learning Topics within Computer Vision & NLP

As more machine learning products are being deployed, machine learning engineering is becoming a very important and sought after skill in the industry. Building infrastructures for training, deployment, and monitoring of deep learning models is different from building other software systems. In this course, learners will train, fine-tune, and deploy deep learning models using Amazon SageMaker.

Begin by finding out what deep learning is, where it is used, along with the tools used by deep learning engineers. Next we will learn about artificial neurons and neural networks and how to train them. After that, learners will explore advanced neural network architectures like convolutional neural networks and BERT and fine-tune them for special tasks. Finally learners will explore Amazon SageMaker and apply new skills in SageMaker Studio.

Course Project

Image Classification using AWS SageMaker

In this project, learners will be using AWS Sagemaker to finetune a pretrained model that can perform image classification. Learners will have to use Sagemaker profiling, debugger, hyperparameter tuning and other good ML engineering practices to finish this project. To finish this project, learners will have to perform tasks and use tools that a typical ML Engineer does as a part of their job.

Lesson 1

Introduction to Deep Learning Topics within Computer Vision & NLP

- Understand the need and importance of deep learning.
- Learn the history of deep learning and the business stakeholders in a deep learning project.
- Learn the tools used by deep learning engineers.

Lesson 2

Introduction to Deep Learning

- Understand the workings of artificial neurons and neural networks.
- Understand how to set cost functions and optimizers to train neural networks.
- Build and train a neural network on an image classification task.

Lesson 3

Common Model Architecture Types & Fine-Tuning

- Understand how advanced neural network architectures like convolutional neural networks and transformer-based models work.
- Fine-tune a pretrained model on a different task.
- Understand the important of hyperparameter tuning for training (and fine-tuning) deep neural networks.

Lesson 4

Deploy Deep Learning Models on SageMaker

- Fine-tune models for image and text classification using SageMaker JumpStart
- Debug and profile training jobs using SageMaker Debugger.
- Tune hyperparameters when training a model.
- Package a model in a Dockerfile for deployment.

Course 4

Operationalizing Machine Learning Projects on SageMaker

This course covers advanced topics related to deploying professional machine learning projects on SageMaker. It also covers security applications. Learners will find out how to maximize output while decreasing costs. They will also learn how to deploy projects that can handle high traffic and work with especially large datasets.

Course Project

Operationalizing an AWS ML Project

In this project, learners will start with a machine learning project that accomplishes computer vision tasks. Learners will deploy the project on AWS and add several important features: cost minimization, security, and redeployment on a separate server. This project will prepare learners to successfully deploy professional projects in industrial applications.

Lesson 1

Manage Compute Resources in AWS Accounts to Ensure Efficient Utilization

- Keep costs low in AWS machine learning projects.
- Use spot instances for efficiency.
- Turn off resources when they're not being used.
- Check costs to ensure they remain low.

Lesson 2

Train Models on Large-Scale Datasets Using Distributed Training

- Perform multi-instance training.
- Use distributed data to improve performance.
- Create and interpret manifest files.
- Choose the best data stores for projects.

Lesson 3

Construct Pipelines for High-Throughput, Low-Latency Models

- Set up Lambda functions for AWS projects.
- Configure endpoints for auto-scaling.
- Set up concurrency for Lambda functions.
- Create feature stores for data imports.

Lesson 4

Design Secure Machine Learning Projects in AWS

- Resolve security issues using IAM settings.
- Set up a virtual private cloud for security.
- Manage security in SageMaker.

Capstone Project

Inventory Monitoring at Distribution Centers

Distribution centers often use robots to move objects as a part of their operations. Objects are carried in bins where each bin can contain multiple objects. In this project, learners will have to build a model that can count the number of objects in each bin. A system like this can be used to track inventory and make sure that delivery consignments have the correct number of items.

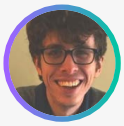
To build this project, learners will have to use AWS Sagemaker and good machine learning engineering practices to fetch data from a database, preprocess it and then train a machine learning model. This project will serve as a demonstration of end-to-end machine learning engineering skills that will be an important piece of their job-ready portfolio.

Meet your instructors.

Matt Maybeno

Principal Software Engineer at SOCi

Matt Maybeno is a principal software engineer at SOCi. With a master's degree in bioinformatics from SDSU, he utilizes his cross domain expertise to build solutions in NLP and predictive analytics.



Joseph Nicolls

Senior Machine Learning Engineer at Blue Hexagon

Joseph Nicolls is a senior machine learning scientist at Blue Hexagon. With a major in biomedical computation from Stanford University, he currently utilizes machine learning to build malware detecting solutions at Blue Hexagon.

Charles Landau

Technical Lead at Guidehouse

Charles Landau is a developer at Guidehouse, a management consulting company. Charles holds a MPA from George Washington University, where he focused on econometrics and regulatory policy, and holds a BA from Boston University. At Guidehouse, he supports data scientists and developers working on internal and client-facing ML platforms.

Soham Chatterjee

Multi-Cloud Engineer

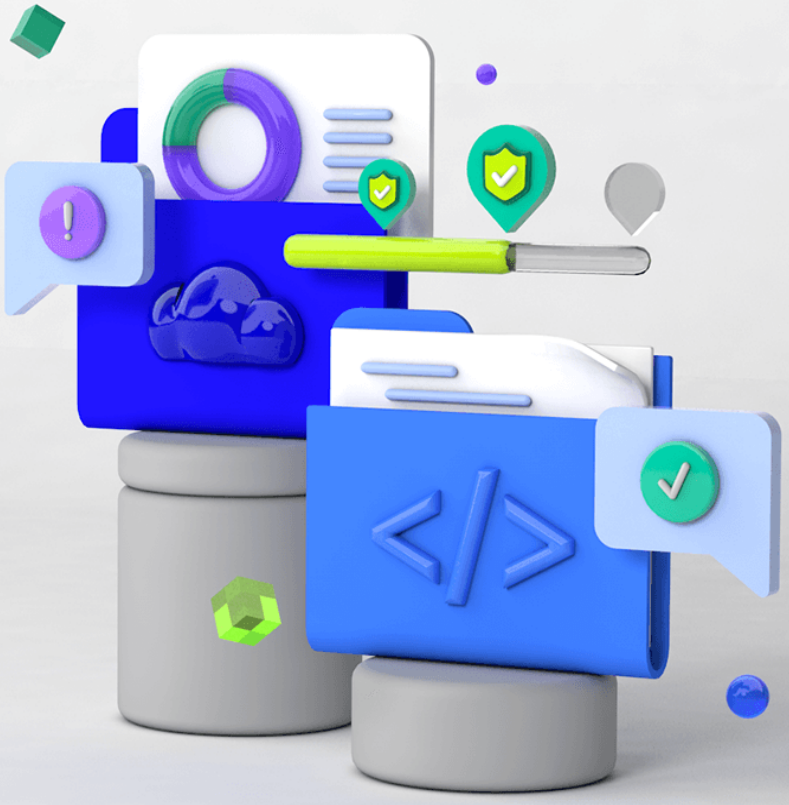
Soham is an Intel® software innovator and a former deep learning researcher at Saama Technologies. He is currently a masters by research student at NTU, Singapore. His research is on Edge Computing, IoT, and neuromorphic hardware.

Bradford Tuckfield

Independent Consultant

Bradford does independent consulting for machine learning projects related to manufacturing, law, pharmaceutical operations, and other fields. He also writes technical books about programming, algorithms, and data science.

Udacity's learning experience



Hands-on Projects

Open-ended, experiential projects are designed to reflect actual workplace challenges. They aren't just multiple choice questions or step-by-step guides, but instead require critical thinking.

Quizzes

Auto-graded quizzes strengthen comprehension. Learners can return to lessons at any time during the course to refresh concepts.

Knowledge

Find answers to your questions with Knowledge, our proprietary wiki. Search questions asked by other students, connect with technical mentors, and discover how to solve the challenges that you encounter.

Custom Study Plans

Create a personalized study plan that fits your individual needs. Utilize this plan to keep track of movement toward your overall goal.

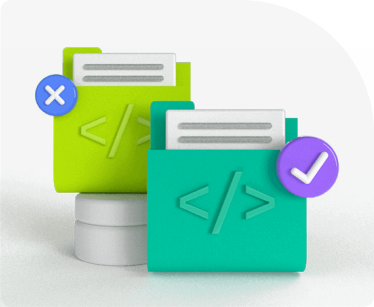
Workspaces

See your code in action. Check the output and quality of your code by running it on interactive workspaces that are integrated into the platform.

Progress Tracker

Take advantage of milestone reminders to stay on schedule and complete your program.

Our proven approach for building job-ready digital skills.



Experienced Project Reviewers

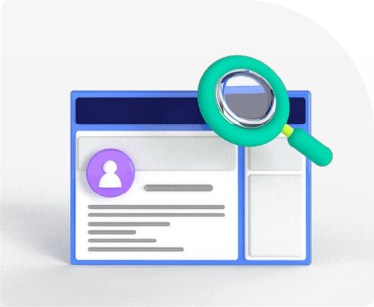
Verify skills mastery.

- Personalized project feedback and critique includes line-by-line code review from skilled practitioners with an average turnaround time of 1.1 hours.
- Project review cycle creates a feedback loop with multiple opportunities for improvement—until the concept is mastered.
- Project reviewers leverage industry best practices and provide pro tips.

Technical Mentor Support

24/7 support unblocks learning.

- Learning accelerates as skilled mentors identify areas of achievement and potential for growth.
- Unlimited access to mentors means help arrives when it's needed most.
- 2 hr or less average question response time assures that skills development stays on track.



Personal Career Services

Empower job-readiness.

- Access to a Github portfolio review that can give you an edge by highlighting your strengths, and demonstrating your value to employers.*
- Get help optimizing your LinkedIn and establishing your personal brand so your profile ranks higher in searches by recruiters and hiring managers.

Mentor Network

Highly vetted for effectiveness.

- Mentors must complete a 5-step hiring process to join Udacity's selective network.
- After passing an objective and situational assessment, mentors must demonstrate communication and behavioral fit for a mentorship role.
- Mentors work across more than 30 different industries and often complete a Nanodegree program themselves.

*Applies to select Nanodegree programs only.

UDACITY

Learn more at

www.udacity.com/online-learning-for-individuals →

12.09.22 | V1.0