UDACITY

NANODEGREE PROGRAM SYLLABUS

AWS Machine Learning Engineer

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.

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)

Educational Objectives

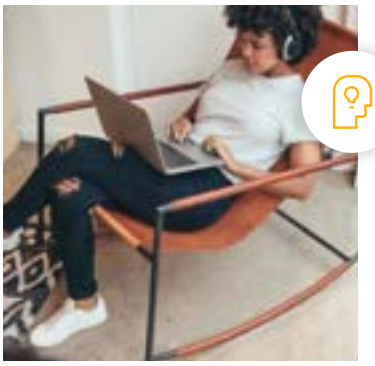
- Create machine learning models in Sagemaker on datasets cleaned using AWS tools
- Deploy machine learning models to an API endpoint and integrate it into a full workflow
- Solve computer vision and natural language problems using fine-tuned deep neural networks
- Operationalize a machine learning pipeline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

IN COLLABORATION WITH

aws

Estimated Time:
5 Months

Prerequisites:
Basic knowledge of machine learning algorithms and python programming



Flexible Learning:
Self-paced, so you can learn on the schedule that works best for you.

Technical Mentor Support:
Our knowledgeable mentors guide your learning and are focused on answering your questions, motivating you and keeping you on track

*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, you'll start learning what machine learning is by being introduced to the high level concepts through AWS SageMaker. You'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, you'll build new ML workflows with highly sophisticated models such as XGBoost and AutoGluon.

Course Project :
Predict Bike Sharing Demand with AutoGluon

In this project, students 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, students 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.

		LEARNING OUTCOMES
LESSON ONE	**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 TWO	**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 THREE	**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 FOUR	**Algorithms and 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 you will learn how to create general machine learning workflows on AWS.

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

Course Project :
Build a ML Workflow on SageMaker

In this project, students will develop an end-to-end ML Workflow on SageMaker, Lambda, and Step Functions. Students 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, students will be able to demonstrate building a scalable ML workflow on SageMaker.

		LEARNING OUTCOMES
LESSON ONE	**Introduction to MLE**	• Understand the prerequisites • Describe key business stakeholders • Understand the history of MLE • Describe when to use MLE
LESSON TWO	**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 THREE	**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 FOUR	**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 and 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 you will learn how to train, finetune and deploy deep learning models using Amazon SageMaker.

You'll begin by learning what deep learning is, where it is used and the tools used by deep learning engineers. Next we will learn about artificial neurons and neural networks and how to train them. After that we will learn about advanced neural network architectures like convolutional neural networks and BERT as well as how to finetune them for specific tasks. Finally, you will learn about Amazon SageMaker and you will take everything you learned and do them in SageMaker Studio.

Course Project :
Image Classification using AWS SageMaker

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

	LEARNING OUTCOMES	
LESSON ONE	**Introduction to Deep Learning Topics within Computer Vision and 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 TWO	**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 THREE	**Common Model Architecture Types and Fine-Tuning**	• Understand how advanced neural network architectures like convolutional neural networks and transformer-based models work • Finetune a pretrained model on a different task • Understand the important of hyperparameter tuning for training (and fine-tuning) deep neural networks
LESSON FOUR	**Deploy Deep Learning Models on SageMaker**	• Finetune 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. You will learn how to maximize output while decreasing costs. You will also learn how to deploy projects that can handle high traffic, and how to work with especially large datasets.

Course Project : Operationalizing an AWS ML Project

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

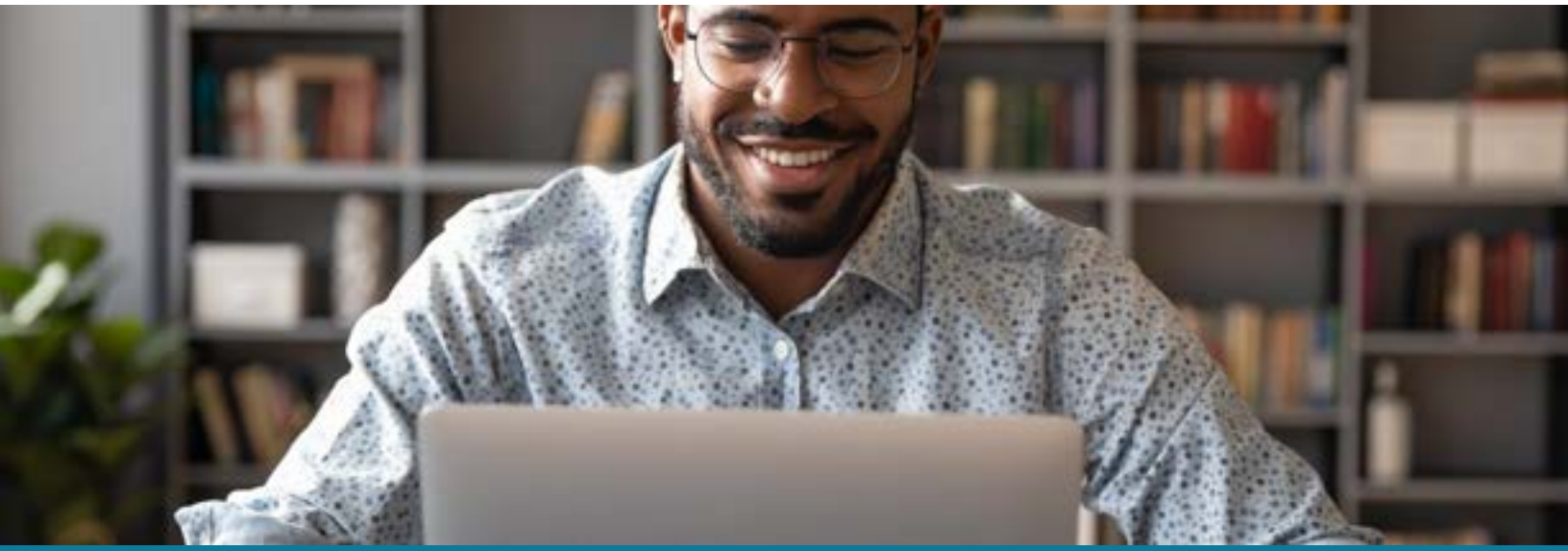
	LEARNING OUTCOMES	
LESSON ONE	**Manage compute resources in AWS accounts to ensure efficient utilization**	• Keep costs low in AWS machine learning projects • Use spot instances for efficiency • Turn off resources when they're not being used • Check costs to ensure they remain low
LESSON TWO	**Train 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 THREE	**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 FOUR	**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

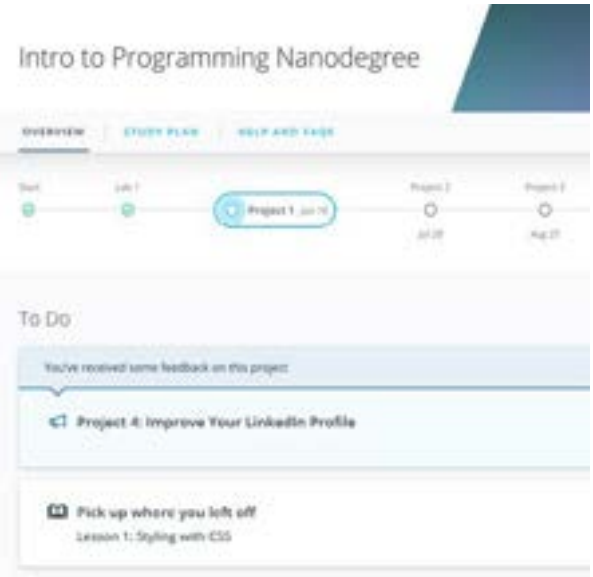
CAPSTONE PROJECT: Inventory Monitoring at Distribution Centers

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, students 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, students 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.



Our Classroom Experience



REAL-WORLD PROJECTS

Build your skills through industry-relevant projects. Get personalized feedback from our network of 900+ project reviewers. Our simple interface makes it easy to submit your projects as often as you need and receive unlimited feedback on your work.

KNOWLEDGE

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

WORKSPACES

See your code in action. Check the output and quality of your code by running them on workspaces that are a part of our classroom.

QUIZZES

Check your understanding of concepts learned in the program by answering simple and auto-graded quizzes. Easily go back to the lessons to brush up on concepts anytime you get an answer wrong.

CUSTOM STUDY PLANS

Create a custom study plan to suit your personal needs and use this plan to keep track of your progress toward your goal.

PROGRESS TRACKER

Stay on track to complete your Nanodegree program with useful milestone reminders.

Learn with the Best

Matt Maybeno

PRINCIPAL SOFTWARE ENGINEER
DATA SCIENCE AND MACHINE LEARNING

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

Joseph Nicolls

SENIOR MACHINE LEARNING ENGINEER, 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,
AI/ML 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.

All Our Nanodegree Programs Include:

EXPERIENCED PROJECT REVIEWERS

REVIEWER SERVICES

- Personalized feedback & line by line code reviews
- 1600+ Reviewers with a 4.85/5 average rating
- 3 hour average project review turnaround time
- Unlimited submissions and feedback loops
- Practical tips and industry best practices
- Additional suggested resources to improve

TECHNICAL MENTOR SUPPORT

MENTORSHIP SERVICES

- Questions answered quickly by our team of technical mentors
- 1000+ Mentors with a 4.7/5 average rating
- Support for all your technical questions

PERSONAL CAREER SERVICES

CAREER SUPPORT

- Github portfolio review
- LinkedIn profile optimization

Frequently Asked Questions

PROGRAM OVERVIEW

WHY SHOULD I ENROLL?

This program is designed to help you take advantage of the growing need for skilled machine learning professionals. Prepare to meet the demand for qualified engineers that can build and deploy machine learning models in production.

WHAT JOBS WILL THIS PROGRAM PREPARE ME FOR?

The skills you will gain from this Nanodegree program will qualify you for jobs in several industries as countless companies are trying to incorporate machine learning into their practices.

HOW DO I KNOW IF THIS PROGRAM IS RIGHT FOR ME?

The course is for individuals who are looking to advance their engineering careers with cutting-edge machine learning skills.

ENROLLMENT AND ADMISSION

DO I NEED TO APPLY? WHAT ARE THE ADMISSION CRITERIA?

No. This Nanodegree program accepts all applicants regardless of experience and specific background.

WHAT ARE THE PREREQUISITES FOR ENROLLMENT?

A well-prepared learner has:

- 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

Students should also have:

- 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)

FAQs Continued

IF I DO NOT MEET THE REQUIREMENTS TO ENROLL, WHAT SHOULD I DO?

Students who do not feel comfortable in the above may consider taking Udacity's Introduction to Programming or Intermediate Python to obtain prerequisite skills.

TUITION AND TERM OF PROGRAM

HOW IS THIS NANODEGREE PROGRAM STRUCTURED?

The AWS Machine Learning Engineer Nanodegree program consists of content and curriculum to support 5 projects. We estimate that students can complete the program in 5 months working 5-10 hours per week.

Each project will be reviewed by the Udacity reviewer network. Feedback will be provided and if you do not pass the project, you will be asked to resubmit the project until it passes.

HOW LONG IS THIS NANODEGREE PROGRAM?

Access to this Nanodegree program runs for the length of time specified above. If you do not graduate within that time period, you will continue learning with month to month payments. See the **Terms of Use** and **FAQs** for other policies regarding the terms of access to our Nanodegree programs.

CAN I SWITCH MY START DATE? CAN I GET A REFUND?

Please see the Udacity Nanodegree program **FAQs** for policies on enrollment in our programs.

SOFTWARE AND HARDWARE

WHAT SOFTWARE AND VERSIONS WILL I NEED IN THIS PROGRAM?

There are no software and version requirements to complete this Nanodegree program. All coursework and projects can be completed via Student Workspaces in the Udacity online classroom. Udacity's basic tech requirements can be found at **https://www.udacity.com/tech/requirements**.