UDACITY

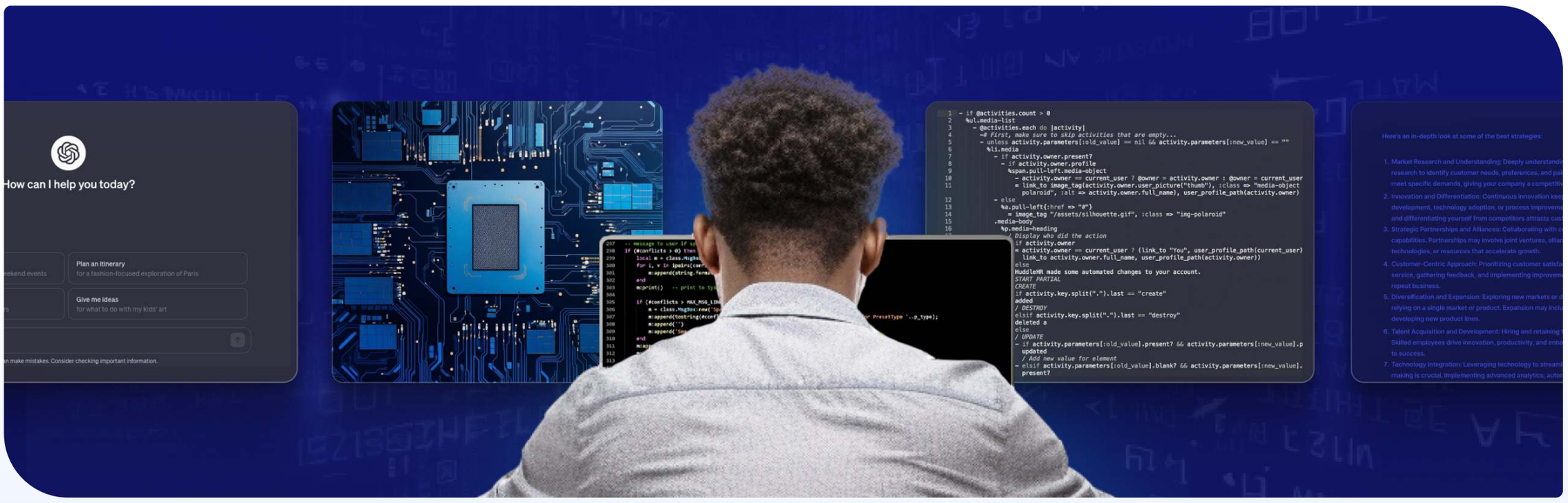
Generative AI for Software Developers

Nanodegree program syllabus

UDACITY.COM

Overview

In an era where generative AI is reshaping the landscape of software development, the Generative AI Nanodegree program equips professionals with essential skills to apply these cutting-edge technologies to enhance their software applications. The program addresses the growing demand for expertise in adapting generative foundation models, crafting custom applications using large language models (LLMs), and employing advanced techniques like Stable Diffusion. If you aspire to supercharge your software development and stay at the forefront of artificial intelligence and deep learning advancements, this program is for you!



Prerequisites

A WELL-PREPARED LEARNER:

- Uses Python at an intermediate level
- Uses SQL at an intermediate level

Educational Objectives

A GRADUATE OF THIS PROGRAM WILL BE ABLE TO:

- Situate generative AI within the broader history, context, and applications of artificial intelligence and deep learning
- Adapt generative foundation models to perform tasks in novel contexts
- Use LLMs and prompt engineering to create a custom chatbot
- Use image generation models such as Stable Diffusion to perform image inpainting
- Build applications that use LLMs, implement semantic search with vector databases, and apply retrieval augmented generation techniques

LENGTH OF PROGRAM*:

4 months

SKILL LEVEL:

Intermediate

SOFTWARE/HARDWARE AND VERSION REQUIREMENTS:

For this Nanodegree program, you will need access to the Internet.

Additional software such as Python and common deep learning libraries (e.g. PyTorch and Hugging Face) will be required, but the program includes Udacity Workspaces with all of the relevant packages installed, so students will not need to download any additional software.

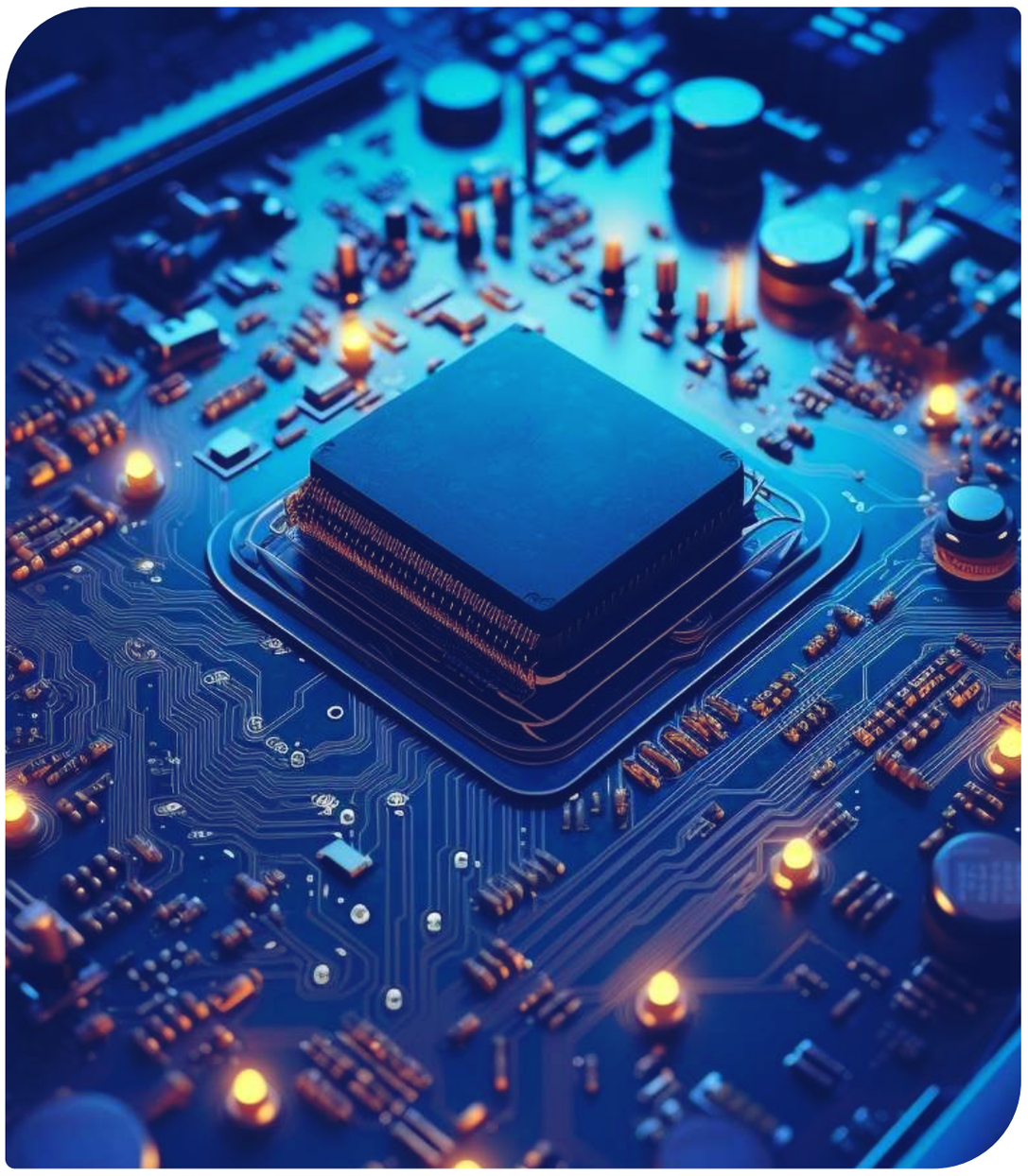
---

*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

GenAI Fundamentals

A "crash course" in deep learning and how generative AI compares and contrasts with previous AI techniques. Then some hands-on practice with model adaptation.



SKILLS COVERED INCLUDE:

- What is GenAI?
- Examples of GenAI
- Deep Learning Fundamentals
- Foundation Models
- Domain Adaptation
- ChatGPT & BERT

PROJECT #1

Apply Lightweight Fine-Tuning to a Foundation Model

In this project, learners will explore advanced techniques in AI model fine-tuning with Hugging Face and PyTorch tools.

To pass this project, learners will need to execute the following core tasks:

- Load a foundation model
- Identify and load a Hugging Face dataset for your particular task
- Utilize a state-of-the-art technique to adjust the foundation model's weights to meet the needs of your task, using a lightweight (AKA parameter-efficient) fine-tuning technique that improves performance faster and more efficiently

Course #2

LLMs & Text Generation

Dive deeper into how computers understand and create language, and build a custom chatbot using unsupervised machine learning and prompt engineering.

SKILLS COVERED INCLUDE:

- Describe business applications of LLMs
- Design prompts for LLMs
- Custom Datasets for LLMs
- Perform prompt engineering with Python
- NLP Fundamentals
- Transformers & Attention Mechanisms
- Retrieval Augmented Generation

PROJECT #2

Build a Custom Chat Bot

Apply retrieval-augmented generation (RAG) to create a custom OpenAI chatbot without fine-tuning.

To pass this project, learners will need to execute the following core tasks:

- Find and prepare a dataset that augments a foundation model's knowledge, from a source such as APIs, web scraping, or documents on hand
- Create a semantic search pipeline by implementing a custom Python vector similarity search algorithm to match user questions to relevant parts of the custom dataset
- Compose a custom query by combining the semantic search results with the user's question and send it to the foundation model

Course #3

Computer Vision & Generative AI

Learn how computers process and understand image data, then harness the power of the latest Generative AI models to create new images.

SKILLS COVERED INCLUDE:

- Describe key image processing techniques
- Image Generation & Generative Adversarial Networks (GANs)
- Transformer-Based Computer Vision Models
- Diffusion Models
- Build a simple diffusion algorithm

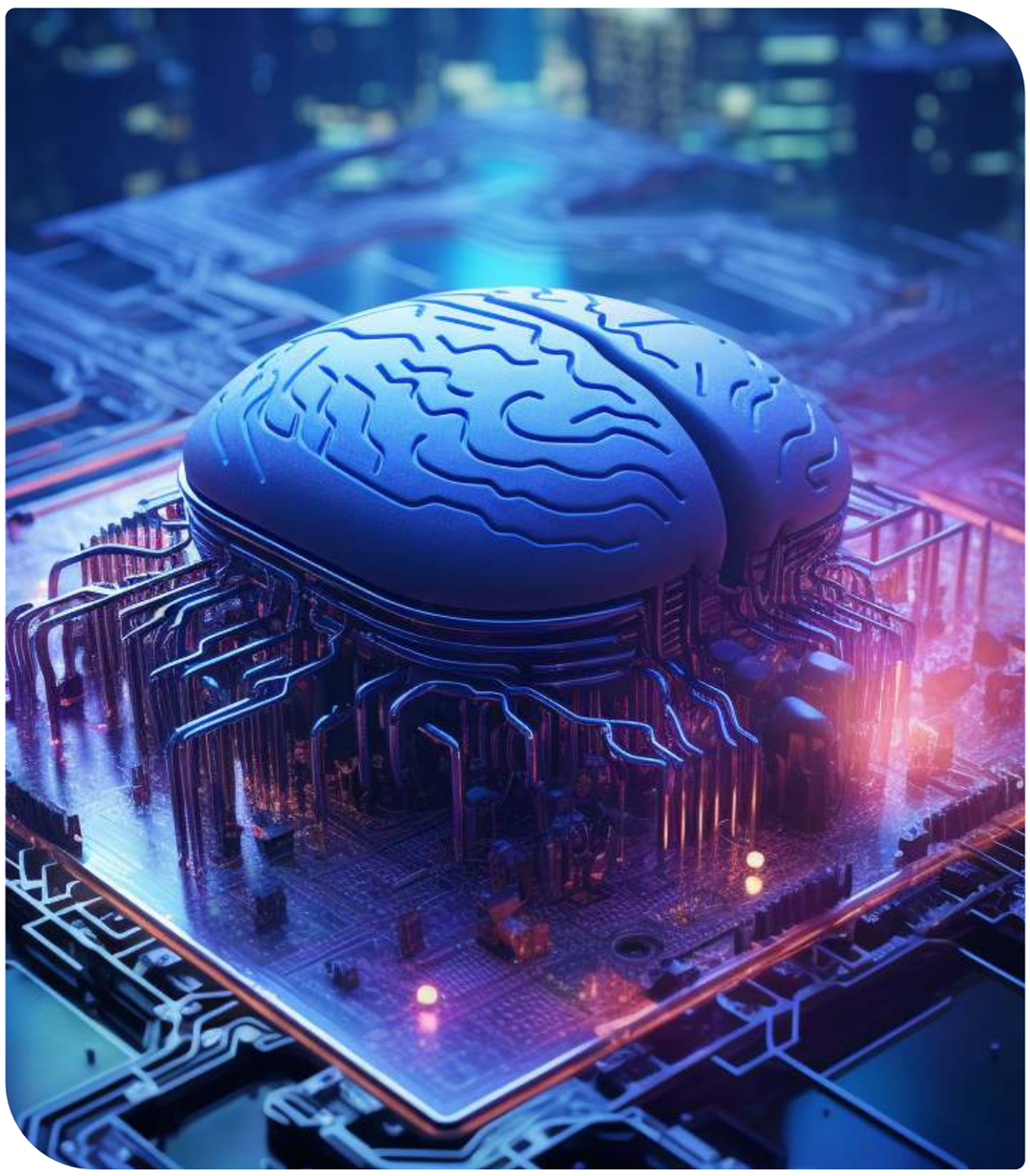
PROJECT #3

AI Photo Editing with Inpainting

Use the Segment Anything Model (SAM) and Stable Diffusion to replace parts of images with AI-generated content.

To pass this project, learners will need to execute the following core tasks:

- Create a segmentation mask by differentiating between the subject and background of an image and create a matrix of pixels indicating the locations of these two components
- Given a text prompt and the pixel locations of the subject or background, replace part of the image with an AI-generated image
- Connect your inpainting pipeline to a web interface that allows users to upload their own images and specify their own text prompts



Course #4

GenAI Solutions

Learn how to design and build applications that use generative AI models. Build applications and AI agents using the OpenAI API, vector databases and semantic search, the LangChain framework, and techniques for recommendation, personalization, and retrieval augmented generation (RAG).

SKILLS COVERED INCLUDE:

- OpenAI API
- Prompt design
- Vector databases
- Semantic Search
- LangChain
- Retrieval Augmented Generation

PROJECT #4

Personalized Real Estate Agent

Build a “real estate agent” application that uses LLMs for content generation, vector databases, semantic search and RAG techniques to transform standard real estate listings into personalized narratives.

To pass this project, learners will need to execute the following core tasks:

- Generate synthetic data using LLMs
- Embed property listing data in a vector database
- Perform semantic search over property listings against user preferences
- Design prompts and use RAG techniques to deliver personalized recommendations

Meet your instructors

Sergei Kozyrenko
SENIOR STAFF ENGINEER

SoFi

Chuyi Shang
MACHINE LEARNING RESEARCHER

Berkeley UNIVERSITY OF CALIFORNIA

Giacomo Vianello
DISTINGUISHED SCIENTIST

CAPE ANALYTICS

Annabel Ng
MACHINE LEARNING RESEARCHER

Berkeley UNIVERSITY OF CALIFORNIA

Emily McMilin
RESEARCH SCIENTIST

Meta

Derek Xu
MACHINE LEARNING RESEARCHER

Berkeley UNIVERSITY OF CALIFORNIA

Brian Cruz
HEAD OF AI ENGINEERING, ADVOCATE

Nathaniel Haynam
MACHINE LEARNING RESEARCHER

BAIR BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

Erick Galinkin
AI SECURITY RESEARCHER

NVIDIA

Valerie Scarlata
TECHNICAL CONTENT DEVELOPER

UDACITY

Victor Geislinger
MACHINE LEARNING ENGINEER

rocketship.vc

Chang She
CEO & CO-FOUNDER

LanceDB

Jason Lin
RESEARCH ENGINEER

Google DeepMind

Learn more at

WWW.UDACITY.COM