UDACITY

Natural Language Processing

UDACITY

Overview

This program will enhance your existing machine learning and deep learning skills with the addition of natural language processing and speech recognition techniques. These skills can be used in various applications such as part of speech tagging and machine translation, among others. You'll develop the skills you need to start applying natural language processing techniques to real-world challenges and applications.

The program is comprised of 3 courses and 3 projects. Each project you build will be an opportunity to demonstrate what you've learned in your lessons. Your completed projects become part of a career portfolio that will demonstrate your mastery of natural language processing to prospective employers.

In order to succeed in this program, we recommend having significant experience with Python, and entry-level experience with probability and statistics, and deep learning architectures. Specifically, we expect you to be able to write a class in Python, and to add comments to your code for others to read. Also, you should be familiar with "neural networks," and understand the differential math that drives backpropagation. If you feel you need to add to your Python and statistics skills, we suggest our **Intro to Machine Learning with PyTorch** or **TensorFlow** Nanodegree programs. If you'd like to learn more about neural networks and backpropagation, consider our **Deep Learning** Nanodegree program.
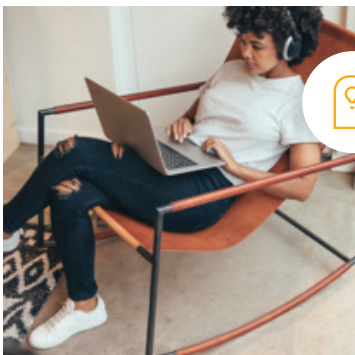
IN COLLABORATION WITH

amazon alexa   IBM **Watson**®

Estimated Time:
3 Months at
10-15 hours / week

Prerequisites:
Python, Statistics,
Machine Learning,
& Deep Learning

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

Need Help?
udacity.com/advisor
Discuss this program
with an enrollment
advisor.

Course 1: Introduction to Natural Language Processing

Learn text processing fundamentals, including stemming and lemmatization. Explore machine learning methods in sentiment analysis. Build a speech tagging model.

Course Project
Part of Speech Tagging

Use several techniques, including table lookups, n-grams, and hidden Markov models, to tag parts of speech in sentences, and compare their performance.

This project demonstrates text processing techniques that allow you to build a part of speech tagging model. You will work with a simple lookup table, and progressively add more complexity to improve the model using probabilistic graphical models. Ultimately you'll be using a Python package to build and train a tagger with a hidden Markov model, and you will be able to compare the performances of all these models in a dataset of sentences.

	LEARNING OUTCOMES	
LESSON ONE	**Intro to NLP**	• Learn the main techniques used in natural language processing. • Get familiarized with the terminology and the topics in the class. • Build your first application with IBM Watson.
LESSON TWO	**Text Processing**	• See how text gets processed in order to use it in models. • Learn techniques such as tokenization, stemming, and lemmatization. • Get started with part of speech tagging and named entity recognition.
LESSON THREE	**Part of Speech Tagging with Hidden Markov Models**	• Learn how hidden Markov models are defined. • Train HMMs with the Viterbi and the Baum-Welch algorithms. • Use HMMs to build a part of speech tagging model.

Course 2: Computing with Natural Language

Learn advanced techniques like word embeddings, deep learning attention, and more. Build a machine translation model using recurrent neural network architectures.

Course Project
Machine Translation

Build a deep neural network that functions as part of an end-to-end machine translation pipeline. Your completed pipeline will accept English text as input and return the French translation. You'll be able to explore several recurrent neural network architectures and compare their performance.

First you will preprocess the data by converting text to sequence of integers. Then you will build several deep learning models for translating the text into French. As a final step, you will run this models on English test to analyze their performance.

	LEARNING OUTCOMES	
LESSON ONE	**Feature Extraction and Embeddings**	• Learn to extract features from text. • Learn the most used embedding algorithms, such as Word2Vec and Glove. • Use the FastText model for word embeddings.
LESSON TWO	**Mopdeling**	• Learn about the main uses of deep learning models in NLP. • Learn about machine translation, topic models, and sentiment analysis.
LESSON THREE	**Deep Learningg Attention**	• Learn about attention, the advanced deep learning method empowering applications like Google Translate. • Learn about additive and multiplicative attention in applications like machine translation, text summarization, and image captioning. • Learn about cutting edge deep learning models like The Transformer that extend the use of attention to eliminate the need for RNNs.
LESSON FOUR	**Information Systems**	• Learn about information extraction and information retrieval systems. • Learn about question answering and its applications.

Course 3: Communicating with Natural Language

Learn voice user interface techniques that turn speech into text and vice versa. Build a speech recognition model using deep neural networks.

Course Project
Speech Recognizer

Build a deep neural network that functions as part of an end-to-end automatic speech recognition (ASR) pipeline. The model will convert raw audio into feature representations, which will then turn them into transcribed text.

You'll begin by investigating a dataset, that will be used to train and evaluate your models. Your algorithm will first convert any raw audio to feature representations that are commonly used for ASR. You will then build neural networks that map these features to transcribed text.

|		LEARNING OUTCOMES
LESSON ONE	**Intro to Voice User Interfaces**	• Learn the basics of how computers understand spoken words. • Get familiar with the most common VUI applications. • Set up your AWS account and build Alexa skill with an existing template.
LESSON TWO	**Alexa History SKill**	• Learn the basics of Amazon AWS. • Create your own fully functional Alexa skill using Amazon's API. • Deploy your skill for everyone to use it.
LESSON THREE	**Introduction to Speech Recognition**	• Learn the pipeline used for speech recognition. • Learn to process and extract features from sound signals. • Learn to build probabilistic and machine learning Language models in order to extract words and grammar from sound signals.

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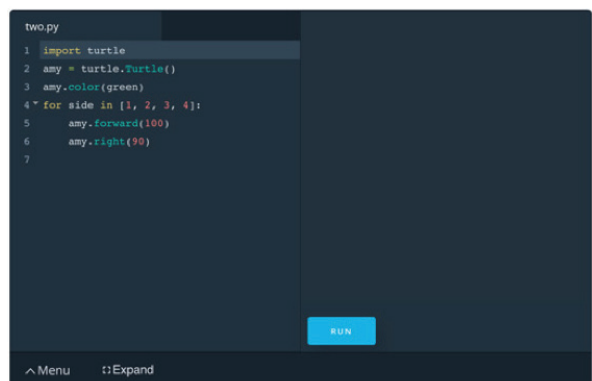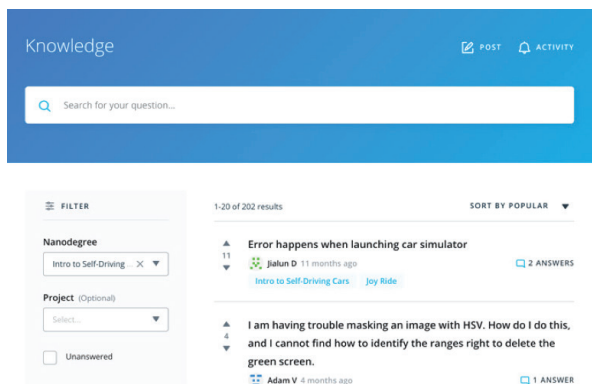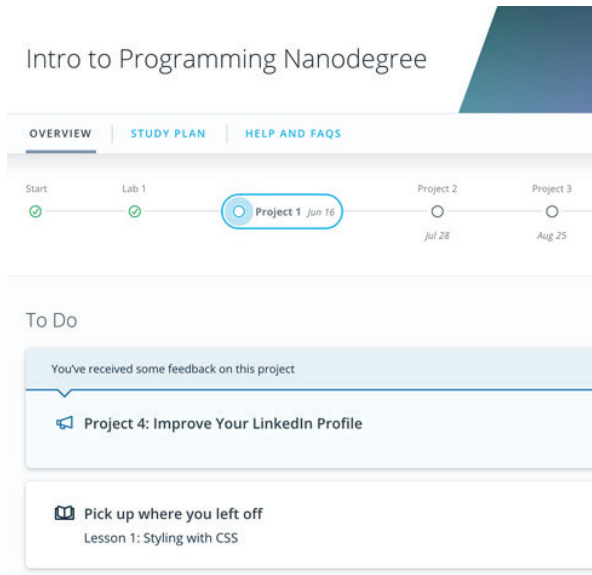
REAL-WORLD PROJECTS
Build your skills through industry-relevant projects. Get personalized feedback from our network of 900+ project reviewers. Our simple interface makes it easy to submit your projects as often as you need and receive unlimited feedback on your work.

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.

STUDENT HUB
Leverage the power of community through a simple, yet powerful chat interface built within the classroom. Use Student Hub to connect with your fellow students in your Executive Program.

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
Preschedule your study times and save them to your personal calendar to create a custom study plan. Program regular reminders to keep track of your progress toward your goals and completion of your program.

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

Learn with the Best

Luis Serrano
CURRICULUM LEAD

Luis was formerly a Machine Learning Engineer at Google. He holds a PhD in mathematics from the University of Michigan, and a Postdoctoral Fellowship at the University of Quebec at Montreal.

Jay Alammar
INSTRUCTOR

Reto Meier has been part of Android Developer Relations at Google since 2009, and leads Google's Scalable Developer Advocacy team. He is the author of the "Professional Android Application Development" book series from Wrox.

Arpan Chakraborty
INSTRUCTOR

Arpan is a computer scientist with a PhD from North Carolina State University. He teaches at Georgia Tech (within the Masters in Computer Science program), and is a coauthor of the book Practical Graph Mining with R.

Dana Sheahen
CONTENT DEVELOPER

Dana is an electrical engineer with a Masters in Computer Science from Georgia Tech. Her work experience includes software development for embedded systems in the Automotive Group at Motorola, where she was awarded a patent for an onboard operating system.

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 COACHING

- Personal assistance in your job search
- Monthly 1-on-1 calls
- Personalized feedback and career guidance
- Interview preparation
- Resume services
- Github portfolio review
- LinkedIn profile optimization

Frequently Asked Questions

PROGRAM OVERVIEW

WHY SHOULD I ENROLL?
This program offers a deep dive into modern Natural Language Process techniques. Mastering these skills will prepare you to build applications involving written and spoken language. We've collaborated with leading innovators such as IBM and Amazon to create our world-class curriculum, and you'll learn from an instructor team comprised of experts from both Udacity and industry professionals. Massive growth is being predicted for the Natural Language Processing software market, making now the perfect time to enter this field.

WHAT JOBS WILL THIS PROGRAM PREPARE ME FOR?
In this program, you'll develop and refine specialized skills in natural language processing and voice user interfaces. The curriculum is not designed to prepare you for a specific job; instead, the goal is that you'll expand your skills in the natural language processing domain. Growth predictions are extremely high for this market, and having these in-demand skills will significantly enhance your ability to advance your AI career.

HOW DO I KNOW IF THIS PROGRAM IS RIGHT FOR ME?
If your goal is to become an expert in Natural Language Processing (NLP), this program is ideal for you. Over the course of this program, you'll become an expert in the main components of NLP, including speech recognition, sentiment analysis, and machine translation. You'll learn cutting edge probabilistic and deep learning models, code them and train them on real data, and build a career-ready portfolio as an NLP expert!

ENROLLMENT AND ADMISSION

DO I NEED TO APPLY? WHAT ARE THE ADMISSION CRITERIA?
There is no application. This Nanodegree program accepts everyone, regardless of experience and specific background.

WHAT ARE THE PREREQUISITES FOR ENROLLMENT?
To succeed in this Nanodegree program, we recommend you first take any course in Deep Learning equivalent to our Deep Learning Nanodegree program. You also need to be able to communicate fluently and professionally in written and spoken English.

Additionally, you should have the following knowledge:

Intermediate Python programming knowledge, including:
- Strings, numbers, and variables
- Statements, operators, and expressions

FAQs Continued

- Lists, tuples, and dictionaries
- Conditions & loops
- Generators & comprehensions
- Procedures, objects, modules, and libraries
- Troubleshooting and debugging
- Research & documentation
- Problem solving
- Algorithms and data structures

Basic shell scripting:
- Run programs from a command line
- Debug error messages and feedback
- Set environment variables
- Establish remote connections

Basic statistical knowledge, including:
- Populations, samples
- Mean, median, mode
- Standard error
- Variation, standard deviations
- Normal distribution

Intermediate differential calculus and linear algebra, including:
- Derivatives & Integrals
- Series expansions
- Matrix operations through eigenvectors and eigenvalues

You will need to be able to communicate fluently and professionally in written and spoken English.

IF I DO NOT MEET THE REQUIREMENTS TO ENROLL, WHAT SHOULD I DO?
We have a number of courses and programs we can recommend that will help prepare you for the program, depending on the areas you need to address. For example:
- **Intro to Machine Learning**
- **Intro to Data Science**
- **Intro to Programming Nanodegree program**
- **Data Analyst Nanodegree program**
- **Artificial Intelligence Programming with Python Nanodegree program**
- **Deep Learning Nanodegree program**
- **Machine Learning Engineer Nanodegree program**

TUITION AND TERM OF PROGRAM

HOW IS THIS NANODEGREE PROGRAM STRUCTURED?
The Natural Language Processing Nanodegree program is comprised of content and curriculum to support three (3) projects. We estimate that students can complete the program in three (3) months, working 10 hours per week.

FAQs Continued

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 in the payment card 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.

I HAVE GRADUATED FROM THE NATURAL LANGUAGE PROCESSING NANODEGREE PROGRAM BUT I WANT TO KEEP LEARNING. WHERE SHOULD I GO FROM HERE?
If you would like to explore other applications for convolutional and recurrent neural networks, and have an interest in computer vision, then consider enrolling in the Computer Vision Nanodegree program. If you are looking for additional advanced topics in AI, the Robotics Engineer and Self-Driving Car Engineer Nanodegree programs could be ideal for you. And regardless of your future career destination, you'll find that the Artificial Intelligence Nanodegree program is full of valuable content that will serve you well in almost any AI role.

SOFTWARE AND HARDWARE

WHAT SOFTWARE AND VERSIONS WILL I NEED IN THIS PROGRAM?
You will need a computer running a 64-bit operating system (most modern Windows, OS X, and Linux versions will work) with at least 8GB of RAM, along with administrator account permissions sufficient to install programs including Anaconda with Python 3.5 and supporting packages. Your network should allow secure connections to remote hosts (like SSH). We will provide you with instructions to install the required software packages. Udacity does not provide any hardware or software.