

Natural Language Processing Nanodegree Syllabus



Become a Natural Language Processing Expert

Welcome to the Natural Language Processing Nanodegree program!

Before You Start

Educational Objectives: In this program, you'll learn and put in practice the concepts behind natural language processing and speech recognition, including machine translation, part of speech tagging, and sentiment analysis. This course will cover the latest in deep learning architectures, which, combined with the more classical probabilistic methods used in a variety of applications, will give you a repertory of techniques widely used in the industry. With the practical skills you gain in this course, you'll be able to build your own models which will extract information from text or voice data, and enable you to solve real-world challenges.

Prerequisite Knowledge: 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 [Machine Learning](#) Nanodegree program. If you'd like to learn more about neural networks and backpropagation, consider our [Deep Learning](#) Nanodegree program.

Length of Program: This program is comprised of 1 term, lasting 3 months. We estimate students will need to work 10 hours/week on average. Make sure to set aside adequate time on your calendar for focused work.

Instructional Tools Available: Video lectures, Jupyter notebooks, personalized project reviews.

Contact Info

While going through the program, if you have questions about anything, you can reach us at nlpnd-support@udacity.com.

Nanodegree Program Info

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.

Length of Program: 120 Hours*

Number of Reviewed Projects: 3

* 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. Actual hours may vary.

Projects

One of our main goals at Udacity is to help you create a job-ready portfolio. Building a project is one of the best ways both to test the skills you've acquired, and to demonstrate your newfound abilities to prospective employers. Throughout this Nanodegree program, you'll have the opportunity to flex your skills by building the following projects:

- Part of Speech Tagging
- Machine Translation
- Speech to Text

In the sections below, you'll find a detailed description of each project, along with the course material that presents the skills required to complete the project.

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.

Introduction to Natural Language Processing

Lesson	Learning Outcomes
INTRO TO NLP	<ul style="list-style-type: none">→ 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.
TEXT PROCESSING	<ul style="list-style-type: none">→ 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.
PART OF SPEECH TAGGING WITH HIDDEN MARKOV MODELS	<ul style="list-style-type: none">→ 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.

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.

Computing with Natural Language

Lesson	Learning Outcomes
FEATURE EXTRACTION AND EMBEDDINGS	<ul style="list-style-type: none">→ Learn to extract features from text.→ Learn the most used embedding algorithms, such as Word2Vec and Glove.→ Use the FastText model for word embeddings.
MODELING	<ul style="list-style-type: none">→ Learn about the main uses of deep learning models in NLP.→ Learn about machine translation, topic models, and sentiment analysis.
DEEP LEARNING ATTENTION	<ul style="list-style-type: none">→ 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.
INFORMATION SYSTEMS	<ul style="list-style-type: none">→ Learn about information extraction and information retrieval systems.→ Learn about question answering and its applications.

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.

Communicating with Natural Language

Lesson	Learning Outcomes
INTRO TO VOICE USER INTERFACES	<ul style="list-style-type: none">→ 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.
ALEXA HISTORY SKILL	<ul style="list-style-type: none">→ 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.
INTRODUCTION TO SPEECH RECOGNITION	<ul style="list-style-type: none">→ 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.