

Overview

A. Industry applications of natural language processing

Natural language processing (NLP) is one of the most important tasks in industry, and it makes heavy use of machine learning. NLP deals with anything related to using machines to process and understand human text/speech (i.e. “natural language”). Many companies, including giants like Google and Amazon, have teams dedicated to NLP research and application in products.

B. How NLP works

In this course, we focus more on the text aspect of NLP, since speech-related tasks will utilize many disciplines outside of machine learning (e.g. signal processing, linguistics, etc.). With written text, it is much simpler to process the data into a feasible input for a machine learning model.

Text data is normally processed with relation to a *vocabulary*. The vocabulary just is the set of unique words that appear across all the text in the *corpus* (the set of documents used to train the model). There are many different ways to process text data, but in this course you’ll learn how to convert sentences/documents into *embedding vectors*.

After processing the text data, we feed it into a particular type of neural network called a *recurrent neural network* (RNN). RNNs are great for dealing with sequential data like text, and in this course you’ll be using the long short-term memory (LSTM) variation of RNNs.

The LSTM model can be adjusted to perform various NLP tasks, ranging from text classification to text generation. For different NLP tasks, the model will have different outputs.

C. What will this course provide?

After taking this course, you’ll be able to process text data, train different

After taking this course, you will be able to process text data, train different LSTM models on the data, and use models to perform a variety of NLP tasks.

Specifically, you will be able to:

- Process documents of text into embedding vectors
- Build a variety of different LSTM models for tasks ranging from text classification to text generation and machine translation