

Open in app ↗

Sign up

Sign in

Medium

🔍 Search

✍ Write



NLP90 : Self-learn NLP in 90 hours

Kushal Shah  · Follow

3 min read · Aug 19, 2022



215



1



Pre-requisites : [Basics of Machine Learning](#)

The content is designed so that you spend 6hrs per week for around 15 weeks making it 90 hrs (assuming good familiarity with general ML algorithms and Python). Of course, you are free to speed up or take it easy!

Week 1 : General Reading

<https://towardsdatascience.com/your-guide-to-natural-language-processing-nlp-48ea2511f6e1>

<https://medium.com/@ageitgey/natural-language-processing-is-fun-9a0bff37854e>

<https://www.nltk.org/book/ch01.html>

Week 2 : Word Tokenization and Sentence Segmentation

<https://stanfordnlp.github.io/stanza/tokenize.html#start-with-pretokenized-text>

<https://www.nltk.org/api/nltk.tokenize.html>

<https://www.guru99.com/tokenize-words-sentences-nltk.html>

Take 10 paragraphs from any source like Wikipedia and check if you are able to use word tokenization and sentence segmentation on this data. Also compare these results of Stanza with the NLTK. Can you spot any important differences or patterns?

Week 3 : Stemming and Lemmatization

<https://www.guru99.com/stemming-lemmatization-python-nltk.html>

<https://www.nltk.org/book/ch03.html>

Week 4 : N-gram models

<https://medium.com/swlh/language-modelling-with-nltk-20eac7e70853>

What Are n-grams and How to Implement Them in Python?

Nithyashree V - Published On September 13, 2021 and Last Modified On July 25th, 2022 This article was published as a...

www.analyticsvidhya.com

Week 5 : Naive Bayes & Sentiment Classification

Sentiment Analysis (Introduction to Naive Bayes Algorithm)

Data Set: <https://www.kaggle.com/c/sentiment-analysis-on-movie-reviews/data>

towardsdatascience.com

<https://www.analyticsvidhya.com/blog/2021/07/performing-sentiment-analysis-with-naive-bayes-classifier/>

Week 6 : Sentiment Classification using POS Tagging and Logistic Regression

Fictometer : A simple and explainable algorithm for sentiment analysis

Our ability to solve NLP problems using AI algorithms has advanced a lot, but our understanding of human language is...

medium.com

Week 7 : Text Classification with Logistic Regression

Build Your First Text Classifier in Python with Logistic Regression - Kavita Ganesan, PhD

Text classification is the automatic process of predicting one or more categories given a piece of text. For example...

kavita-ganesan.com

Week 8 : Word Embeddings

<https://medium.com/@phylypo/a-survey-of-the-state-of-the-art-language-models-up-to-early-2020-aba824302c6>

How to Predict Sentiment from Movie Reviews Using Deep Learning (Text Classification) - Machine...

Sentiment analysis is a natural language processing problem where text is understood, and the underlying intent is...

machinelearningmastery.com

Week 9 : Recurrent Neural Networks (RNNs)

<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

<https://www.youtube.com/watch?v=WCUNPb-5EYI>

Dropout in RNNs:

<https://adriangcoder.medium.com/a-review-of-dropout-as-applied-to-rnns-72e79ecd5b7b>

RNN Regularization:

<https://arxiv.org/abs/1409.2329>

RNN Regularization: Which Component to Regularize?

I am building an RNN for classification (there is a softmax layer after the RNN). There are so many options for what to...

stackoverflow.com

Week 10 : Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU)

<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

<http://blog.echen.me/2017/05/30/exploring-lstms/>

Stacked LSTMs:

<https://machinelearningmastery.com/stacked-long-short-term-memory-networks/>

<https://machinelearningmastery.com/return-sequences-and-return-states-for-lstms-in-keras/>

LSTM Regularization:

<https://machinelearningmastery.com/use-weight-regularization-lstm-networks-time-series-forecasting/>

Week 11 & 12 : The Attention Mechanism & Transformers

<https://www.youtube.com/watch?v=TQQlZhbC5ps>

https://www.youtube.com/watch?v=OyFJWRnt_AY

<http://nlp.seas.harvard.edu/2018/04/03/attention.html>

<https://www.youtube.com/watch?v=Osj0Z6rwJB4&list=PLEJK-H61XlwxfpVzt3oDLQ8vr1XiEhev&index=2>

https://kazemnejad.com/blog/transformer_architecture_positional_encoding/

<https://towardsdatascience.com/master-positional-encoding-part-i-63c05d90a0c3>

<https://towardsdatascience.com/https-medium-com-chaturangarajapakshe-text-classification-with-transformer-models-d370944b50ca>

Week 13 : BERT

<https://medium.com/@mromerocalvo/dissecting-bert-part1-6dcf5360b07f>

<https://www.youtube.com/c/ChrisMcCormickAI/videos>

<http://jalammar.github.io/illustrated-bert/>

<https://www.thepythoncode.com/article/finetuning-bert-using-huggingface-transformers-python>

Sentiment Analysis with BERT and Transformers by Hugging Face using PyTorch and Python

TL;DR In this tutorial, you'll learn how to fine-tune BERT for sentiment analysis. You'll do the required text...

curiously.com

Week 14 : NER using BERT

<https://www.depends-on-the-definition.com/named-entity-recognition-with-bert/>

<https://towardsdatascience.com/named-entity-recognition-with-bert-in-pytorch-a454405e0b6a>

Week 15 : Text Classification with BERT

https://www.tensorflow.org/text/tutorials/classify_text_with_bert

<https://towardsdatascience.com/text-classification-with-bert-in-pytorch-887965e5820f>

Machine Learning

Naturallanguageprocessing

Artificial Intelligence

Transformers