

Hide menu

Lecture: Logistic Regression

- ✔ **Video:** Welcome to the NLP Specialization
4 min
- ✔ **Video:** Welcome to Course 1
1 min
- ✔ **Reading:** Acknowledgement - Ken Church
10 min
- ✔ **Video:** Week Introduction
35 sec
- ✔ **Video:** Supervised ML & Sentiment Analysis
2 min
- ✔ **Reading:** Supervised ML & Sentiment Analysis
2 min
- ✔ **Video:** Vocabulary & Feature Extraction
2 min
- ✔ **Reading:** Vocabulary & Feature Extraction
2 min
- ✔ **Video:** Negative and Positive Frequencies
2 min
- ✔ **Video:** Feature Extraction with Frequencies
2 min
- ✔ **Reading:** Feature Extraction with Frequencies
10 min
- ⏮ **Video:** Preprocessing
3 min
- 📖 **Reading:** Preprocessing
10 min
- 📅 **Lab:** Natural Language preprocessing
1h
- ⏮ **Video:** Putting it All Together
2 min
- 📖 **Reading:** Putting it all together
10 min
- 📅 **Lab:** Visualizing word frequencies
1h
- ⏮ **Video:** Logistic Regression Overview
3 min
- 📖 **Reading:** Logistic Regression Overview
10 min
- ⏮ **Video:** Logistic Regression: Training
1 min
- 📖 **Reading:** Logistic Regression: Training
10 min
- 📅 **Lab:** Visualizing tweets and Logistic Regression models
1h
- ⏮ **Video:** Logistic Regression: Testing
4 min

Week 1 > Preprocessing

Previous Next

Preprocessing

When preprocessing, you have to perform the following:

1. Eliminate handles and URLs
2. Tokenize the string into words.
3. Remove stop words like "and, is, a, on, etc."
4. Stemming- or convert every word to its stem. Like dancer, dancing, danced, becomes 'danc'. You can use porter stemmer to take care of this.
5. Convert all your words to lower case.

For example the following tweet "@YMcourri and @AndrewYNg are tuning a GREAT AI model at https://deeplearning.ai!!!" after preprocessing becomes



Preprocessed tweet:
[tun, great, ai, model]

[*tun, great, ai, model*]. Hence you can see how we eliminated handles, tokenized it into words, removed stop words, performed stemming, and converted everything to lower case.

Mark as completed

Like Dislike Report an issue

