



Next

< Previous

#### **≔** Hide menu

### **Lecture: Logistic** Regression

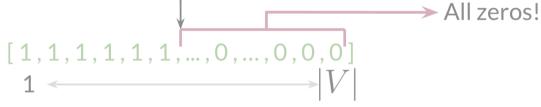
- Video: Welcome to the **NLP Specialization** 4 min
- Video: Welcome to Course 1 1 min
- Reading: Acknowledgement - Ken Church 10 min
- Video: Week Introduction 35 sec
- Video: Supervised ML & Sentiment Analysis 2 min
- **Reading:** Supervised ML & Sentiment Analysis 2 min
- Video: Vocabulary & **Feature Extraction** 2 min
- Reading: Vocabulary & Feature Extraction
- 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

### 

## Vocabulary & Feature Extraction

Given a tweet, or some text, you can represent it as a vector of dimension V, where V corresponds to your vocabulary size. If you had the tweet "I am happy because I am learning NLP", then you would put a 1 in the corresponding index for any word in the tweet, and a 0 otherwise.

# I am happy because I am learning NLP



As you can see, as V gets larger, the vector becomes more sparse. Furthermore, we end up having many more features and end up training  $\theta$  V parameters. This could result in larger training time, and large prediction time.

Go to next item







**□** Dislike

☐ Report an issue

