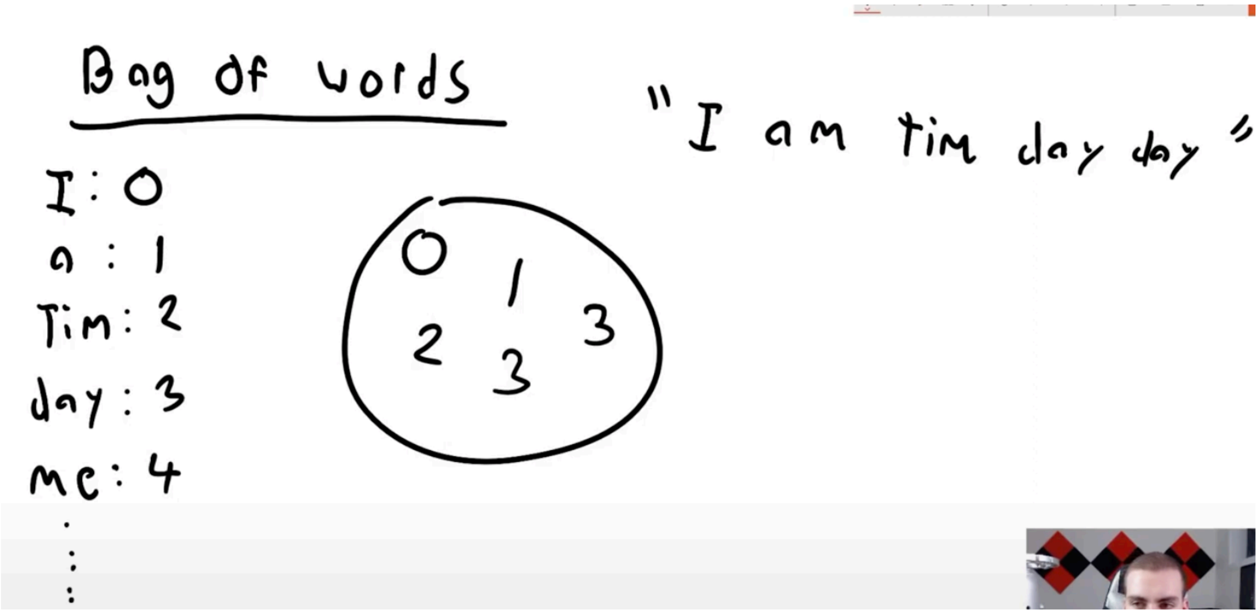
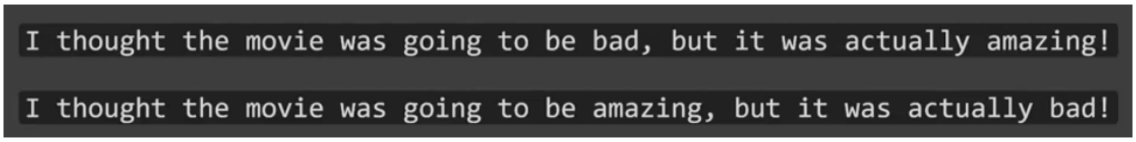
- ***Natural language processing with recurrent neural networks***
    - -> NLP <- using words to do machine learning
    - -> "natural language" because it's not a computer programming language
    - -> recurrent neural networks
        - examples are -> spell checks, chat bots, text data, words
        - recurrent neural networks are used for natural language processing
    - -> this is an example of sentiment analysis
- ***Examples in this module***
    - ***Sentiment analysis***
        - -> using NLP to determine how popular a movie was
    - ***Making a play using NLP***
        - -> an NLP model to generate the next character in a sequence of text
            - -> doing this to generate an entire play
- ***Textual vs numeric data***
    - -> turning a block of text into information which the neural network can understand and process
    - -> turning words into numbers
    - ***bag of words***
        - ***the bag of words technique***
            - -> this is a method to convert text into numbers
            - -> we look at the entire training dataset and create a dictionary lookup of the vocabulary
            - -> making a block of text into a dictionary and giving a number to every word in that dictionary
            - -> there can be millions of words -> but we are only keeping track of the words which are present and the frequency they appear
            - -> taking a sentence and creating a histogram of each of the words and the amount of times they appear
                - -> in other words, it's a "bag of words"
            - -> loosing the ordering the words appear in
            - -> this method is ineffective for more complex datasets



            - -> not caring where the words are in relation to each other -> just the amount of times they appear in the block of text
        - ***where this approach breaks down***
            - -> the same word can have multiple meanings -> e.g for movie reviews

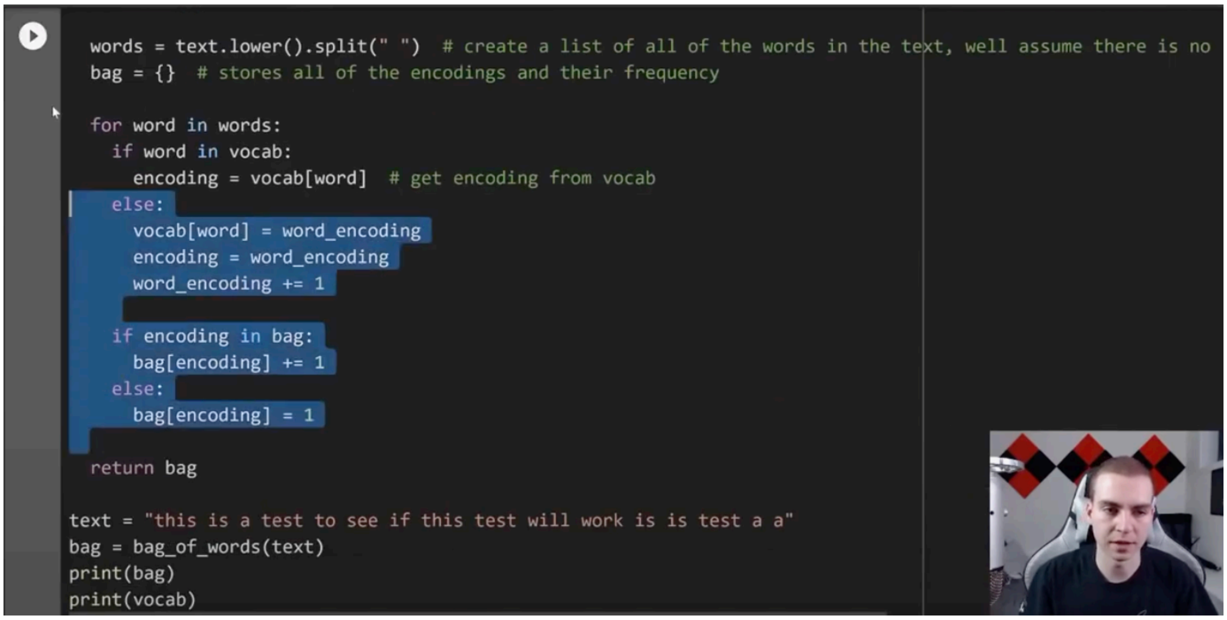


```
I thought the movie was going to be bad, but it was actually amazing!

I thought the movie was going to be amazing, but it was actually bad!
```

  - <u>-> in this example the words are the same but are in a different order</u>

- ***example code for bag of words***



```
  words = text.lower().split(" ")  # create a list of all of the words in the text, well assume there is no
  bag = {}  # stores all of the encodings and their frequency

  for word in words:
    if word in vocab:
      encoding = vocab[word]  # get encoding from vocab
    else:
      vocab[word] = word_encoding
      encoding = word_encoding
      word_encoding += 1

    if encoding in bag:
      bag[encoding] += 1
    else:
      bag[encoding] = 1

  return bag

text = "this is a test to see if this test will work is is test a a"
bag = bag_of_words(text)
print(bag)
print(vocab)
```

- -> doing natural language processing based only off of of the frequency which the different words appear in
- -> <u>nuatural Language Processing is a branch of artificial intelligence that...: deals with how computers understand and process natural/human languages</u>