



Kelly Slatery  
US-DSI-10  
Capstone  
03.13.2020

# ShowMeTunes: A Musical Recommender

## TABLE OF CONTENTS

# 01

### PROBLEM

Who's interested?  
Why?

# 02

### PROCESS

Collection  
Cleaning  
Preprocessing  
Deployment



# 03

### PRODUCT

Recommender System

# 04

### FUTURE

Next Steps

# PROBLEM

01

Who's interested?

Why?

QUESTION

What's your go-to  
song?



## QUESTION

# What's your go-to song?

...when you're:  
happy?  
sad?  
frustrated?  
stressed?



## UNDERSTANDING THE PROBLEM



### MOOD

People like to listen to different music depending on their moods.



### NEW MUSIC

Some older musicals need some love. Some newer, more popular musicals get old.



### CHOICES

People don't always want endless choices. Sometimes, we just want direction.

## OUR SOLUTIONS

Collect a list of  
popular musical  
titles and their  
synopses

**POPULAR MUSICALS**

Use NLP to find  
musicals with  
similar synopses &  
sentiments to the  
user's input

**MATCH USER INPUT TO  
SYNOPSIS**

Output the names  
of three new  
musicals for the  
user to listen to!

**RECOMMEND 3  
MUSICALS**

# PROCESS

# 02

Collection  
Cleaning  
Preprocessing  
Deployment



# DATA COLLECTION



WIKIPEDIA  
*The Free Encyclopedia*

Wikipedia

186

synopses

- Wikipedia API
- BeautifulSoup

Ranker.com

Musicals to recommend

196 musical names

>100,000 votes

36 Rerankers

- Selenium
- BeautifulSoup



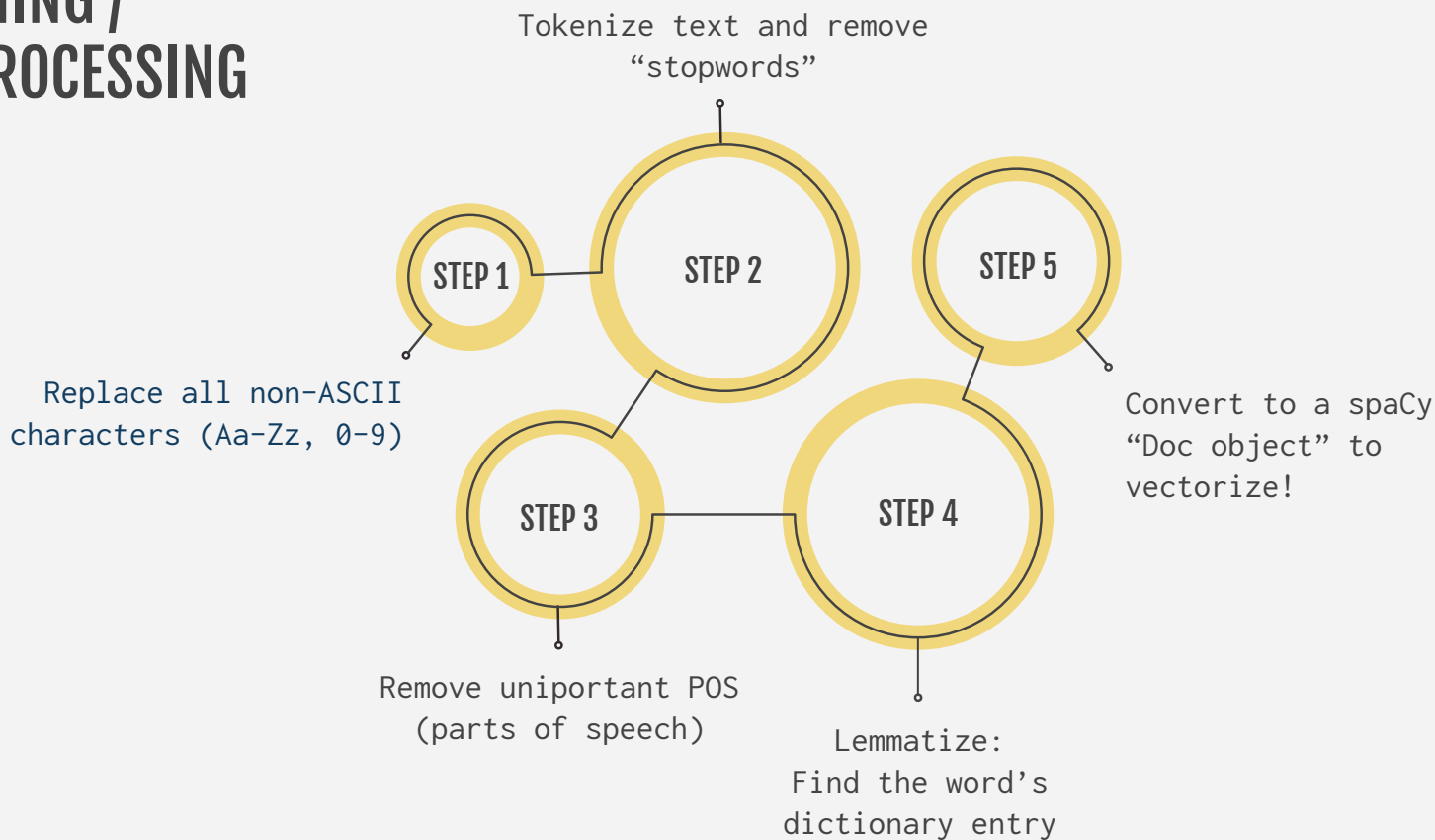
All Musicals

177

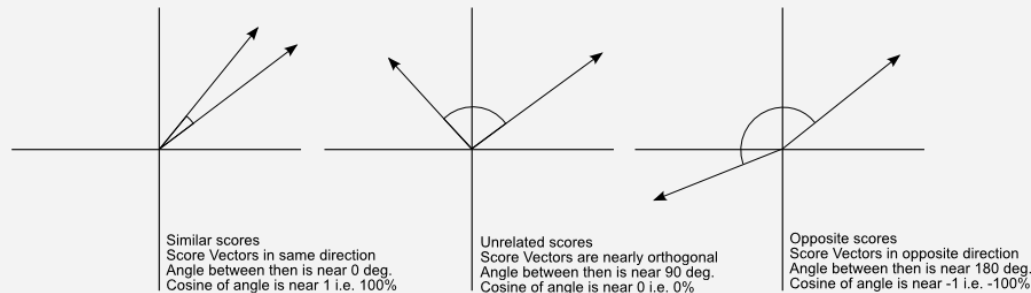
synopses

- Requests
- BeautifulSoup

# CLEANING / PREPROCESSING



# How does a “vector” represent a word?



- **300 dimensions** (list of 300 numbers)
- Each dimension **represents** some aspect of that word
- Communicates its **meaning/relationship** to other words
- Vectors are **unique** to each word, as it's used in **different contexts**
- Produced using Word2Vec algorithms (shallow, two-layer neural networks)
- AKA “**word embeddings**”

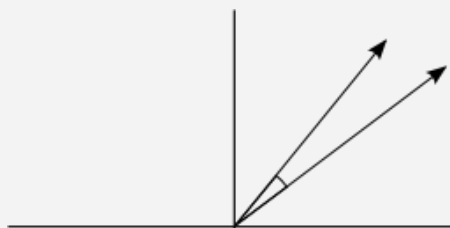
## Example:

```
Les Misérables = [-2.55308696e-03  1.25066981e-01  
                  -3.78664001e-03  ...  3.62479091e-02  
                  1.23177804e-02 -4.90440764e-02]
```

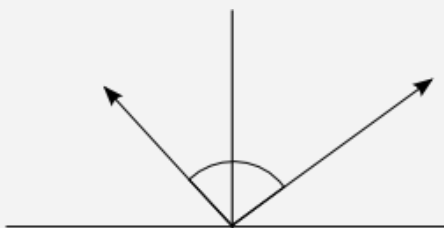
# DOCUMENT SIMILARITY

- “cosine similarity”  $[0,1]$
- Doc vector = Average of all Token vectors

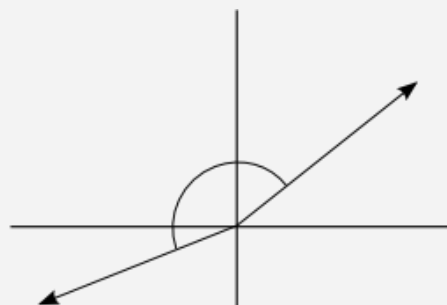
## 2-D Vectors



Similar scores  
Score Vectors in same direction  
Angle between them is near 0 deg.  
Cosine of angle is near 1 i.e. 100%



Unrelated scores  
Score Vectors are nearly orthogonal  
Angle between them is near 90 deg.  
Cosine of angle is near 0 i.e. 0%



Opposite scores  
Score Vectors in opposite direction  
Angle between them is near 180 deg.  
Cosine of angle is near -1 i.e. -100%

# SENTIMENT ANALYSIS

- Token sentiment: value representing its sentiment
  - Ex. “awesome”=1, “awful”=-1
- Doc sentiment: Average of all Token sentiments

-1

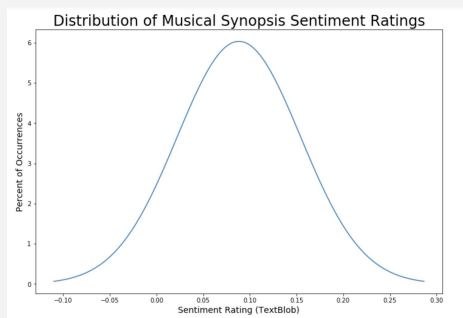
negative  
sentiment

0

neutral  
sentiment

+1

positive  
sentiment



Mean = 0.088

Standard Deviation  
= 0.066

# RECOMMENDER SYSTEM



Rank similarity  
scores between user  
input and each  
musical synopsis



Of the top ten,  
sort by magnitude  
of distance from  
sentiment rating  
of user input



Reveal the top  
three musicals!

**PRODUCT**

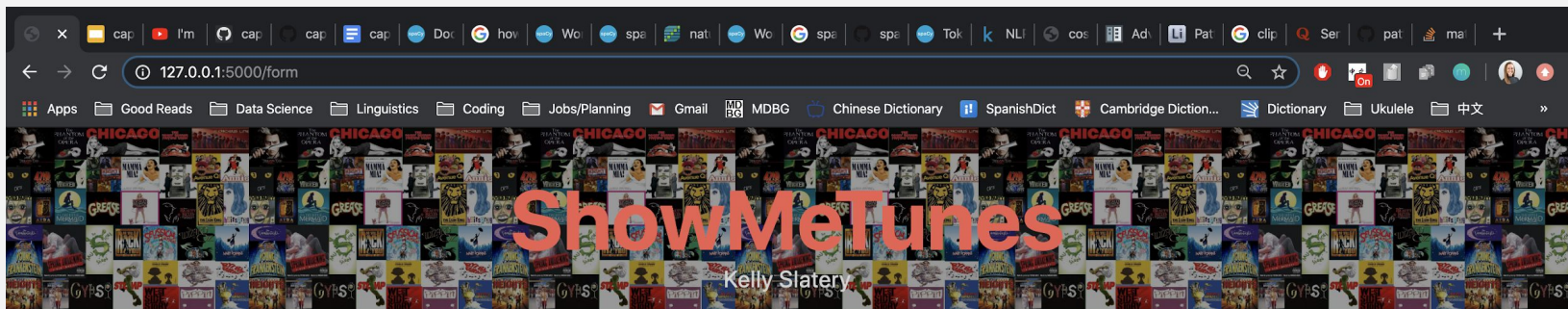
03

Recommender System

Let's give it a try!

Switch to terminal /  
Local Host





*Musical fans come in all shapes and sizes and levels of intensity, but the one thing we all have in common is: we love musicals. But no matter how much of a Broadway buff you are, none of us knows ALL the popular musicals' soundtracks word-for-word. And sometimes you just can't match Rent's "One Song Glory" energy. Sometimes, all we want is for someone (or something) to tell us what musical to listen to in the moment. That's where ShowMeTunes comes in: you input a short description of your current state (mood, what's going on in your head and in your life), and ShowMeTunes will show you the show tunes you should give a listen to right when you need it. Now, you'll never have to go another minute wondering, "What musicals am I missing out on?" or simply, "What should I play next?" So what are you waiting for? Give it a try below!*

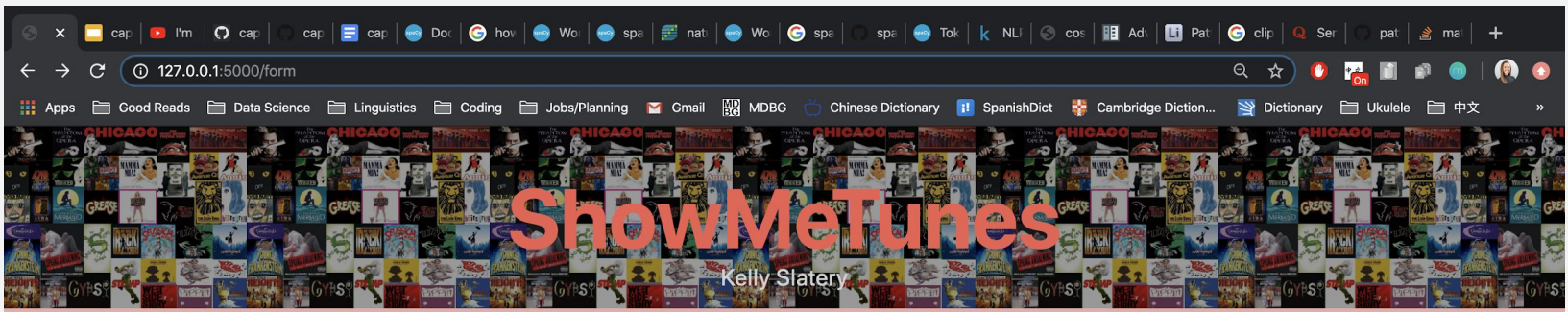
**How are you feeling today?  
What's going on with you?**

*3-5 sentences*

I have no confidence in technology. I want to quit my job because of politics. My family thinks I am lazy at computers. I am tired of being disappointed in the country. I have no idea what I will do in retirement. I am uncertain of the future and it makes me anxious.

Give me my musicals!





*Musical fans come in all shapes and sizes and levels of intensity, but the one thing we all have in common is: we love musicals. But no matter how much of a Broadway buff you are, none of us knows ALL the popular musicals' soundtracks word-for-word. And sometimes you just can't match Rent's "One Song Glory" energy. Sometimes, all we want is for someone (or something) to tell us what musical to listen to in the moment. That's where ShowMeTunes comes in: you input a short description of your current state (mood, what's going on in your head and in your life), and ShowMeTunes will show you the show tunes you should give a listen to right when you need it. Now, you'll never have to go another minute wondering, "What musicals am I missing out on?" or simply, "What should I play next?" So what are you waiting for? Give it a try below!*

**How are you feeling today?  
What's going on with you?**

3-5 sentences

I'm okay

Give me my musicals!





# MOST COMMON 25 WORDS



act, asks,  
away, day,  
**family,**  
**father,** *find,*  
**girl,**



*home,* *leave,*  
life, love,  
**man,** **mother,**  
new, old,  
people,



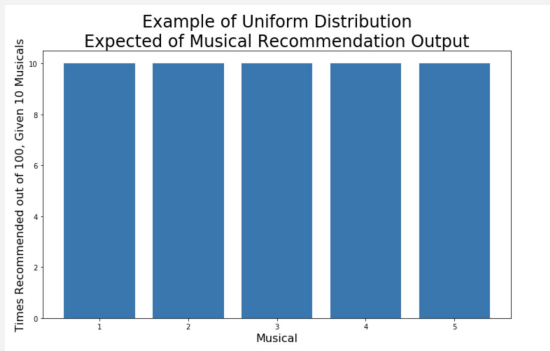
reprise, song,  
soon, tells,  
*time,* **wife,**  
**woman,** **young**

**FUTURE**

01

Next Steps

# EVALUATION



**Musicals Recommended Equally**

*Future Development:*  
Write a Python script to  
collect site data - count  
how many times each musical  
is recommended!



0,1,2

3,4,5

6,7,8,9

## 10 Clusters

*Expectation:*  
2/3 or 3/3 musicals  
recommended are from the  
same cluster

# NEXT STEPS

- Rank musicals by popularity myself:
- Wikipedia API: get lists of musicals by category
  - Spotify API: get track popularities

**Ranking Algorithm**

Define algorithm to detect different features, such as:

- intensity/chill level
- nostalgic/modern
- happy/sad

**Data Collection**

Collect as many synopses/plot summaries as possible to add specific training to spaCy model

**Sentiment Analysis**

**Deployment**

- Create evaluation tools
- Create export playlist feature
- Modernize Flask style
- Debug & Optimize
- Ensure copyrights
- Create App



# THANK YOU

Questions?

[kelly.slatery@gmail.com](mailto:kelly.slatery@gmail.com)

[linkedin.com/in/kellyslatery/](https://www.linkedin.com/in/kellyslatery/)



# Instructions for use

In order to use this template, you must credit [Slidesgo](#) by keeping the Credits slide.

## You are allowed to:

- Modify this template.
- Use it for both personal and commercial projects.

## You are not allowed to:

- Sublicense, sell or rent any of Slidesgo Content (or a modified version of Slidesgo Content).
- Distribute Slidesgo Content unless it has been expressly authorized by Slidesgo.
- Include Slidesgo Content in an online or offline database or file.
- Offer Slidesgo templates (or modified versions of Slidesgo templates) for download.
- Acquire the copyright of Slidesgo Content.

For more information about editing slides, please read our FAQs:

<https://slidesgo.com/faqs>

# Fonts & colors used

This presentation has been made using the following fonts:

## **Inconsolata**

(<https://fonts.google.com/specimen/Inconsolata>)

## **Fjalla One**

(<https://fonts.google.com/specimen/Fjalla+One>)

#434343

#f1d77b

#f3f3f3