Scotch Whisky Recommender

A 'Neat' Approach Using NLP

Hello!

I'm Shannon McDonnell

I am a recent graduate of Metis's Data Science Bootcamp.
You can find me at:

✉ *mcdonnell.sc@gmail.com*
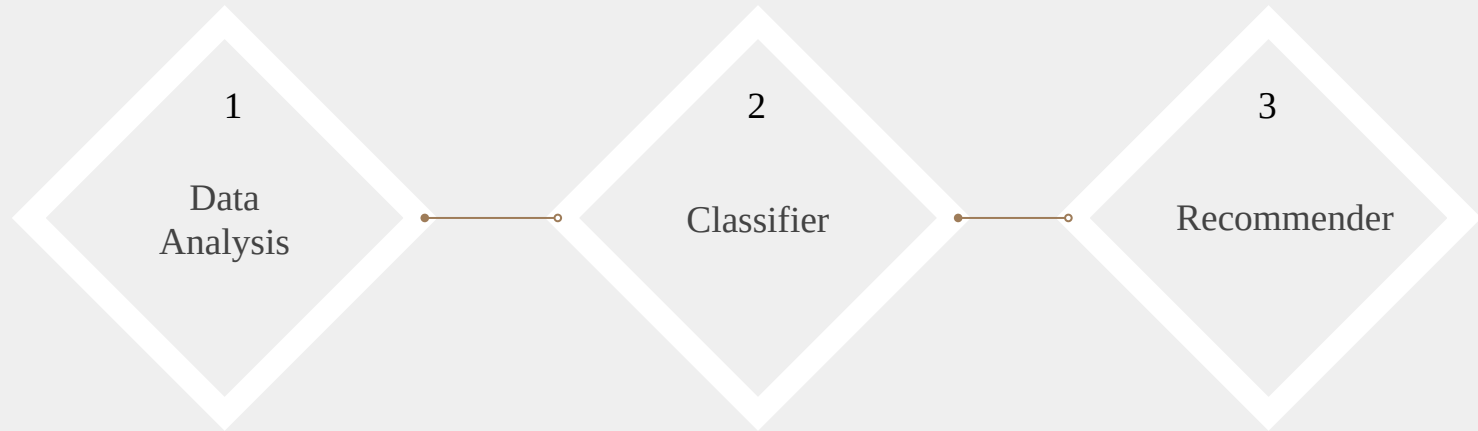in *shannon-mcdonnell*
sconsm

Goals & Objectives

◆

The purpose of this project is to explore the flavors, notes and finishes of bottles of Scotch by analyzing text reviews with NLP

Project Workflow: 3 Separate Parts

1

Data
Analysis

2

Classifier

3

Recommender

EDA

Text: *Review Descriptions*

◆ Word frequencies
◆ Text preprocessing
  ◇ Removing useless words:
    ◇ Bottle, casket, notes, palate, mouth
◆ Looking for trends or nuances

A **TEXT** sample of a review description

◦—◇—◦

"It is deep, dark, and quite powerful with a hefty malt foundation and sweeter notes of toffee and vanilla combined with resinous oak, complex fruit….tobacco, polished leather, and heavy peat smoke."
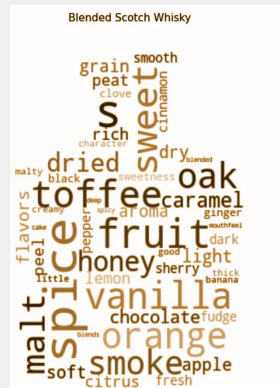
(Johnnie Walker Blue Label; Blended)

EDA

Categories

1. Single Malt
2. Blended
3. Blended Malt
4. Single Grain
5. Grain

Blended Scotch Whisky

Blended

Toffee, spice, fruit, malt, vanilla, orange
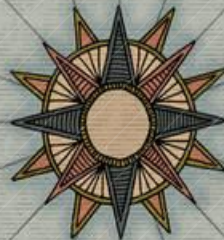
The Scotch Classifier

◈

Using Machine Learning Models to Predict the Category of a Scotch

2

Classifier
Methodology

NLP Text Data

Machine Learning Model

Perform Logistic Regression

Predict Accuracy

Word2Vec

Neural Network Model

Perform LSTM

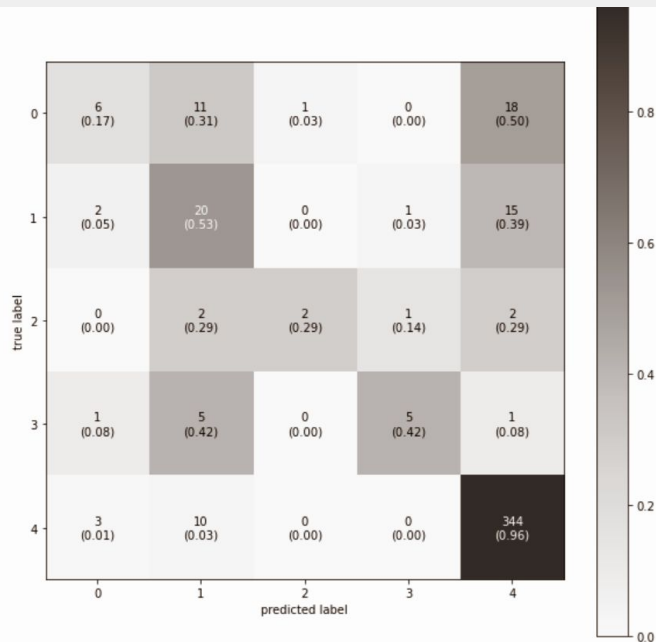Predict Accuracy

GloVe
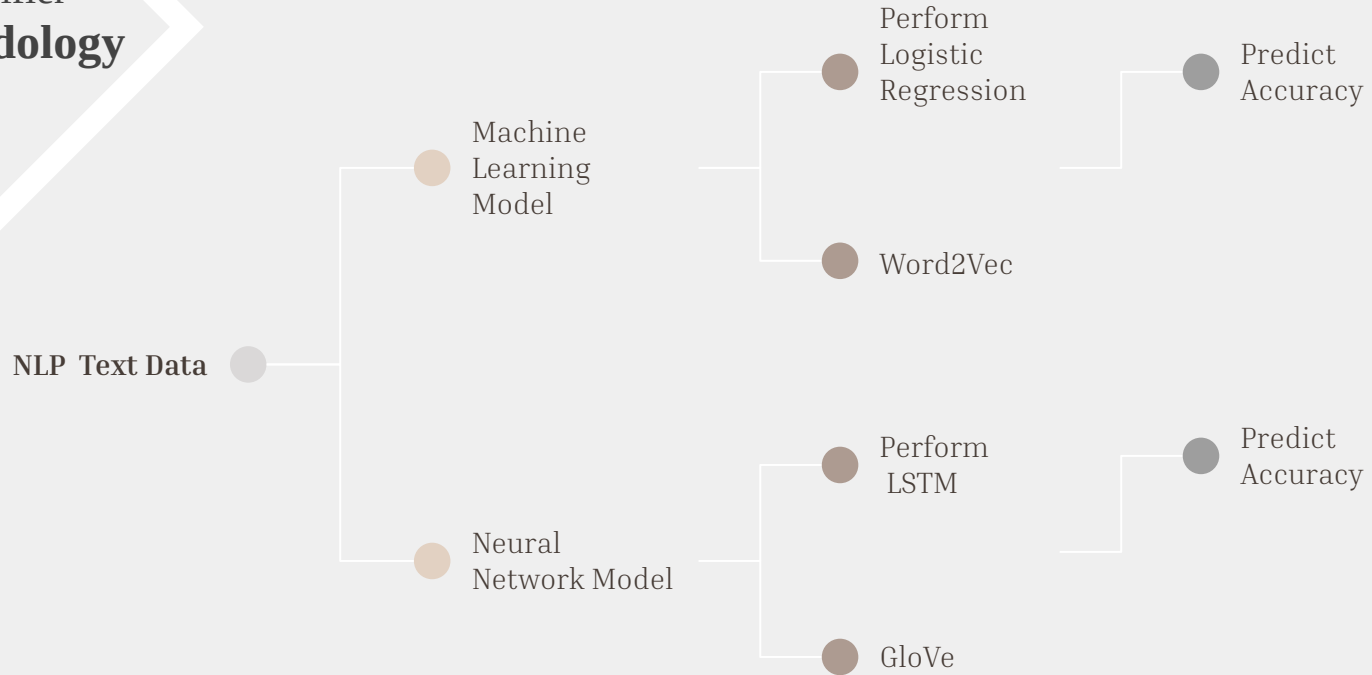
Results

Logistic Regression Accuracy:

84%

LSTM Accuracy:

Test: 81%

Train: 82%

3.
The Recommender

❖

The system in action! Streamlit App Demo

Recommendation System:
NLP & NMF Topic Modeling

3

Recommender

Objective

Use NLP to create a Scotch recommendation from reviews

Tools & Algorithms

- ◆ Nltk
- ◆ Text preprocessing
- ◆ TFIDF Vectorizer
- ◆ NMF Topic Modeling
- ◆ Pairwise distances (cosine similarity)

Results

Yields fairly reasonable results upon examination of feature correlations

NMF Topics

Winey

Bordeaux, grape, berry, wine, red, currant, sherry, port, raspberry

Fruity

Sweet, fresh, apple, strawberry, water, light

Peaty

Earthy, smokey, seaweed, brine, linger, pepper, salt

Woody/ Nutty

Honey, vanilla, creamy, chocolate, spice, caramel

Feinty

Leather, dark, plain, cherry, fig, raisin, oloroso

Spices

Ginger, spice, chocolate, oak, vanilla, bitter

Recommender In Action:

1. Choose between different topic words (aka flavors, notes)

2. Pick the # of bottle recommendations you want

3. Wait for the results!

Let's give it a try!

Streamlit App

Scotch Whisky Recommender

A 'Neat' Approach to Find Your Ideal Scotch Using NLP

This Application Displays EDA and NMF Text Recommendations for Scotch Whiskies From A *Whisky Advocate* Kaggle Dataset

Recommendation System in Action!

Interact with Exploratory Data Analyses!

Compare Results

◇

Input: "chocolate", "orange," "I like soft smokey whisky"

Output: Johnnie Walker Blue

"... sweeter notes of toffee and vanilla...complex fruit (both bright citrus notes and darker pit fruits)… leather, and heavy peat smoke."

(Johnnie Walker Blue Label; Blended)

HIGHLANDS

FLAVOURS: dry, spicy, honey

MEDIUM & FRUITY
ROBUST & COMPLEX
MEDIUM & SPICY
ROBUST & SMOKY

SPEYSIDE

FLAVOURS: honey, fruity, hint of spice, oak dryness and light smoke

MEDIUM & FRUITY
ROBUST & COMPLEX

Applications & Future Work

Practical Usage

- Predictive ML modeling can help businesses understand where they need to focus their attentions.
  - Knowing which category of whisky to produce
- Rec systems benefit both user and provider, providing an ability to predict user preferences
  - Gather informative whisky drinker data, while providing useful suggestions to users

Continued Work

- Scraping for more whisky types
  - American Bourbon, Irish Whiskey
  - Flavors, smells, tastes
- Collaborative and content filtering
  - Getting user info data
- Finish predicting review scores
  - Given a text, can you predict the rating score
- Finish radar plots for topic variable visualization

Thanks!

Any Questions?

◆

Feel free to reach out!
Contact me at:
mcdonnell.sc@gmail.com
linkedin.com/in/shannon-mcdonnell

End.

Appendix Following

For Fun:

Try and match category to region!

A map of Scotland's single malt whiskies

NSLC

ISLANDS
FLAVOURS: big, bold, sweet
MEDIUM & FRUITY   MEDIUM & SPICY   ROBUST & COMPLEX

SPEYSIDE
FLAVOURS: honey, fruity, hint of spice, oak dryness and light smoke
MEDIUM & FRUITY   ROBUST & COMPLEX

HIGHLANDS
FLAVOURS: dry, spicy, honey
MEDIUM & FRUITY   ROBUST & COMPLEX
MEDIUM & SPICY   ROBUST & SMOKY

ISLAY
FLAVOURS: bold, buttery, very smoky
ROBUST & SMOKY

LOWLAND
FLAVOURS: light, grassy, mildly floral
MEDIUM & FRUITY   ROBUST & COMPLEX

CAMPBELTOWN
FLAVOURS: briny, smoky
LIGHT & FRESH   MEDIUM & FRUITY

Project Workflow

EDA

Classifier

Recommender

Scotch reviews: raw text descriptions from users.

Created a classifier that predicts categories of whisky.

NMF topic modeling and text recommendation system using NLP.

❖ *Whisky Advocate 2,245 data points*

❖ *Can we predict Scotch category using NLP on text reviews?*

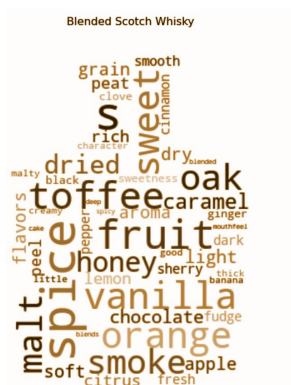❖ *Make inferences re: regions & category*

❖ *Recommends x bottles of similar whiskies, based on a user's inputted flavor preferences*

Text In Each Scotch Category

Single Malt

Vanilla, sweet, fruit, long, maturd

Blended

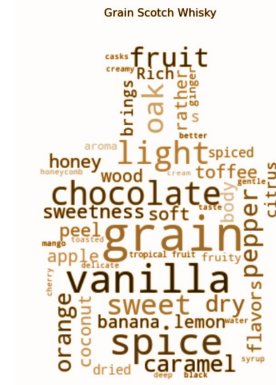Toffee, spice, fruit, malt, vanilla, orange

Blended Malt

Smoky, fruit, spice, pepper, chocolate

Single Grain

Spice, vanilla, apple, sweet, fruit

Grain

Vanilla, spice, chocolate, caramel

2

Classifier

Scotch Classifier

◦◇◦

Predicting Categories

Objective

Create a predictive ML model that accurately predicts Scotch categories

Tools & Algorithms

◈ Logistic Regression
◈ LSTM
◇ GloVe

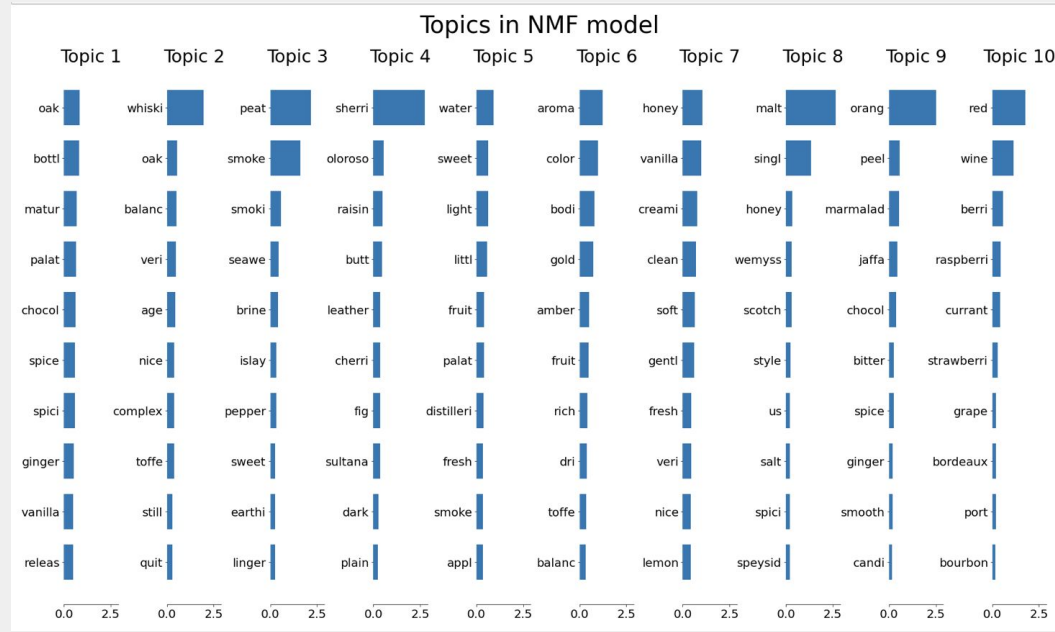## Results

◈ Classify Scotch bottles by category

◈ Make judgments and inferences on associated regions and flavors

Classifier: **Methodology**

NLP Text Data

Machine Learning Model

Perform Logistic Regression

Word2Vec

Predict Accuracy

Neural Network Model

Perform LSTM

GloVe

Predict Accuracy
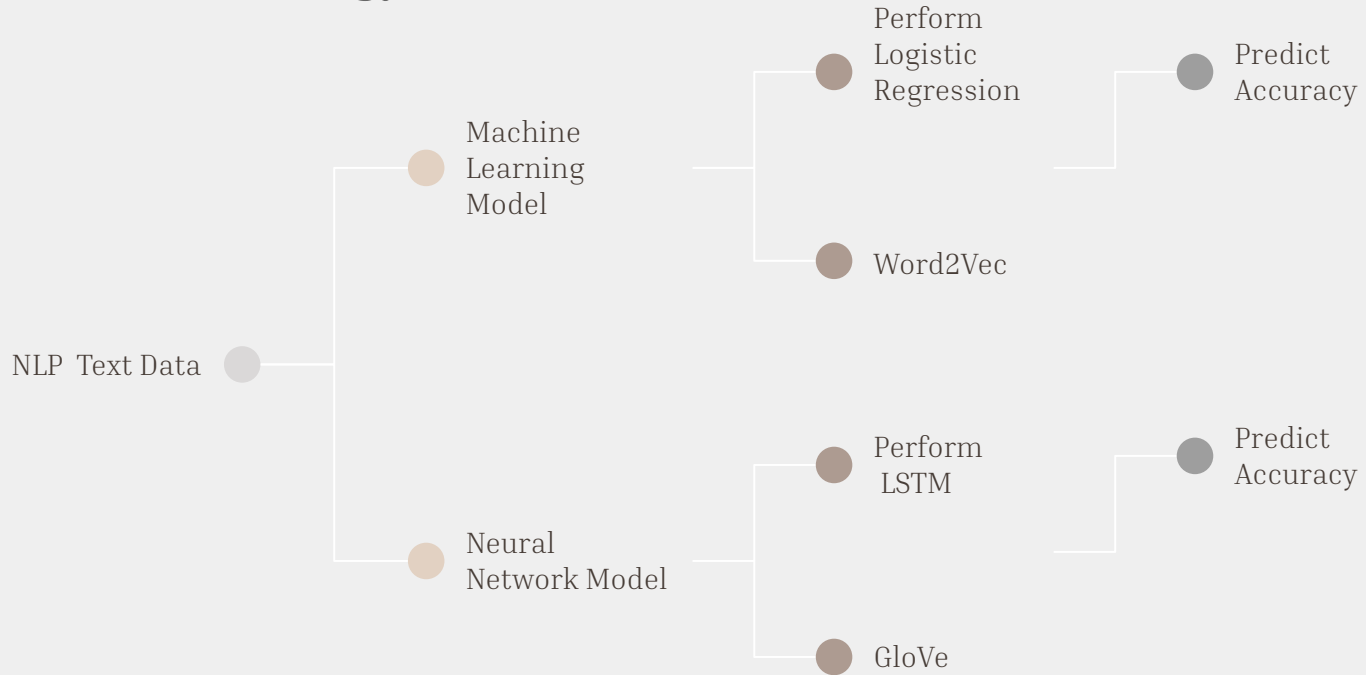
Topic Modeling

Topics in NMF model

Topic Modeling (BETTER)

Topics in NMF Model
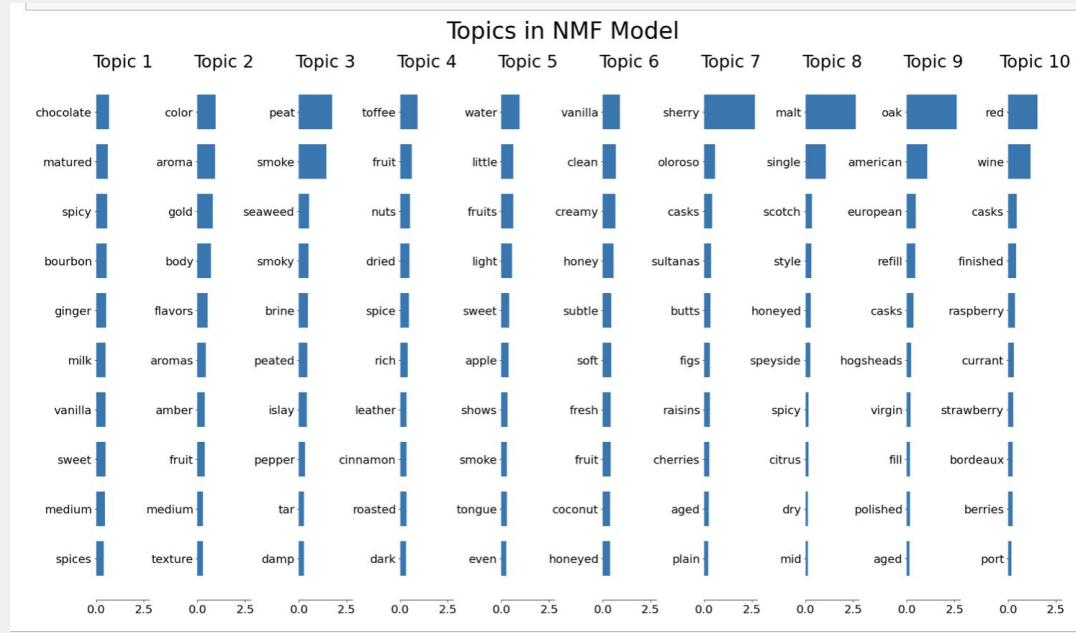
Regions & Flavors

1. **Islands**
   - Fruity, spicy, bold, sweet
2. **Speyside**
   - Honey, fruit, oak, light smoke
3. **Highlands**
   - Dry, spicy, honey, smokey
4. **Lowland**
   - Light, grassy, fruit, floral
5. **Campbeltown**
   - Brinky, smokey, fresh, frutiy
6. **Islay**
   - Bold, buttery, robust smoke

A map of Scotland's single malt whiskies

ISLANDS
FLAVOURS: big, bold, sweet
MEDIUM & FRUITY · MEDIUM & SPICY · ROBUST & COMPLEX

SPEYSIDE
FLAVOURS: honey, fruity, hint of spice, oak dryness and light smoke
MEDIUM & FRUITY · ROBUST & COMPLEX

HIGHLANDS
FLAVOURS: dry, spicy, honey
MEDIUM & FRUITY · ROBUST & COMPLEX · MEDIUM & SPICY · ROBUST & SMOKY

ISLAY
FLAVOURS: bold, buttery, very smoky
ROBUST & SMOKY

CAMPBELTOWN
FLAVOURS: briny, smoky
LIGHT & FRESH · MEDIUM & FRUITY

LOWLAND
FLAVOURS: light, grassy, mildly floral
MEDIUM & FRUITY · ROBUST & COMPLEX

Recommender

How The App Works:

◦─◇─◦

1. Choose the flavors (topic words) you want.

    (Enter your own description if you wish)

2. Pick the # of bottle recommendations you want

3. Wait…(for the system to take the pairwise distances using cosine similarity, to find you the most similar option results)…

    DONE!

Recommender In Action:

1. Choose between different topic words

Get A Recommendation!

Want a delicious Scotch recommendation!? From the list below, choose as many descriptive words (flavors) as you like, and the system will do the rest!

Choose an option ▾

soft

toffee

chocolate