

# The Travel Niche: Solo travel and Road trips

DSI Project 3 by: PREM HONGHIMAPHAN





# INTRODUCTION



# BOOKING

Destination

Arrival

Departure

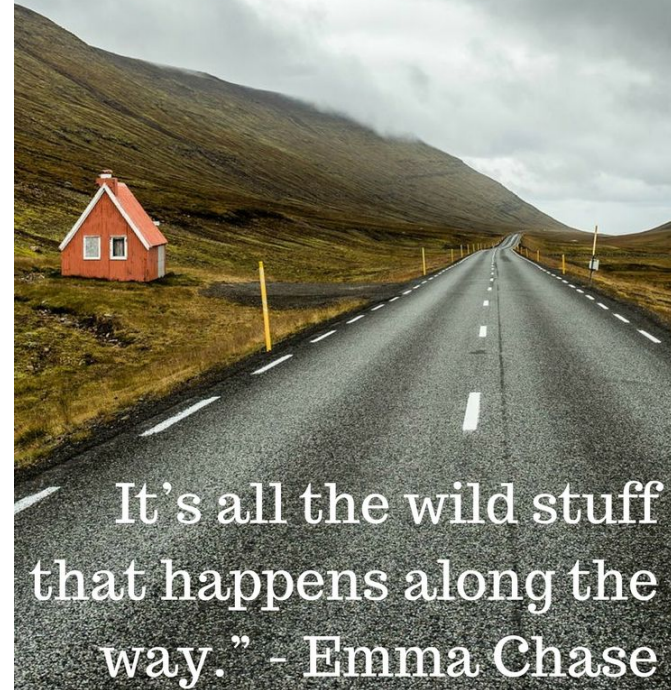
SEARCH

*"Solo travel not only pushes you out of your comfort zone, it also pushes you out of the zone of others' expectations."*

*- Suzy Strutner*



*"Because the greatest part of a road trip isn't arriving at your destination."*



*"It's all the wild stuff that happens along the way." - Emma Chase*

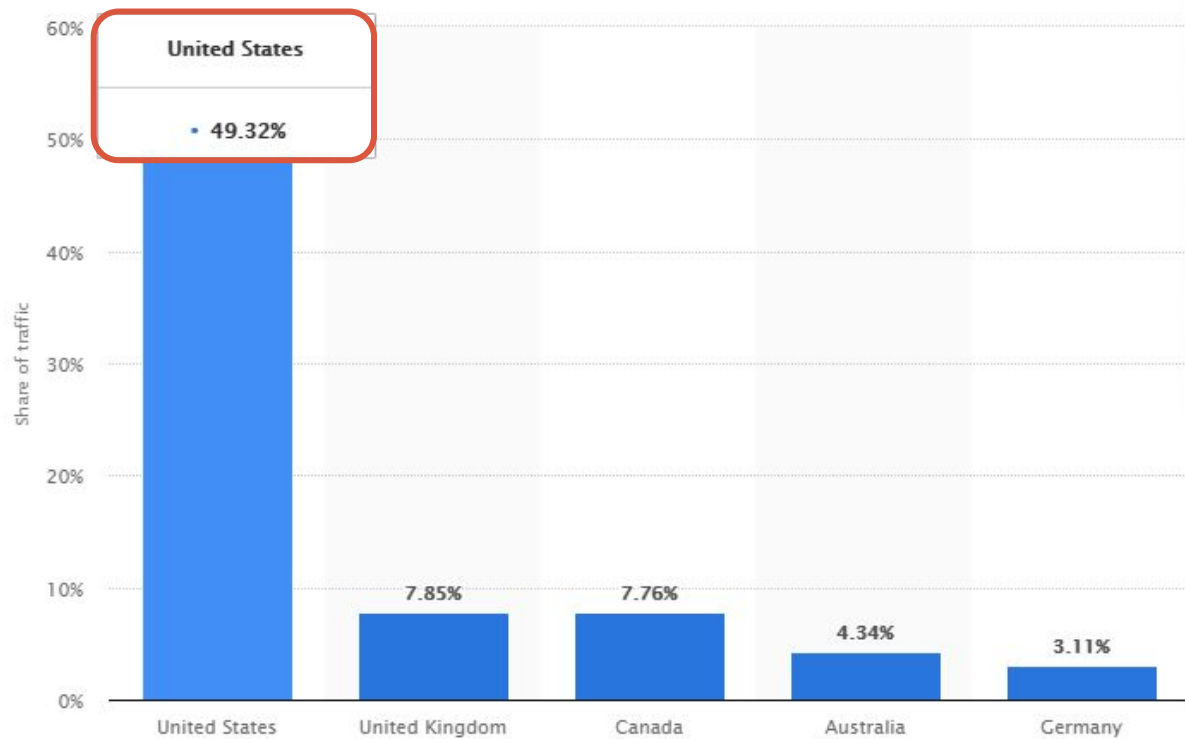




Target Destination: United States







© Statista 2021



### About Community

A place for all of those interested in solo travel to share their experiences and stories!

1.6m

Wanderers

425

Getting some rest



Created Jul 18, 2010

r/solotravel

### About Community

/r/roadtrip is your source for everything road trip related. Whether you enjoy traveling by motorcycle, car, or recreational vehicle this is your destination for everything related to road trips!

1.4m

Members

197

Online



Created Aug 5, 2009

r/roadtrip

# The Data

Number of Posts retrieved

Remove Duplicates

Replace Null Values

Solotravel

1,230

Solotravel

928

Road trip

1,240

Road trip

890

Total Data

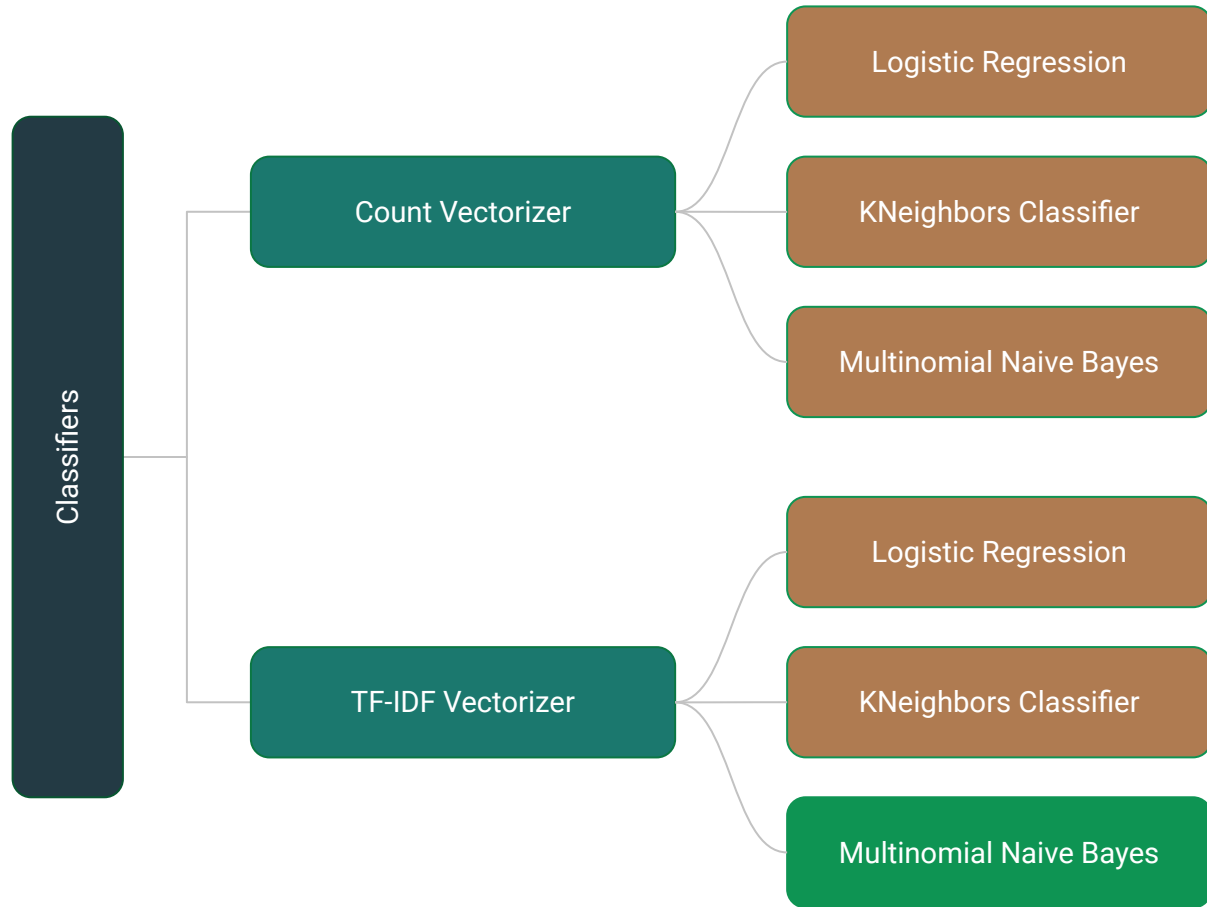
1,818

# About the Data

Feature	Type	Description
subreddit	object	Name of subreddit the posts belong to
selftext	object	Content of the post by users
title	object	Title of the post by users
alltext	object	Combination of selftext and title
clean_text	object	Cleaned text of alltext column
classifier	int	1=Solotravel, 0=Roadtrip



# MODELING





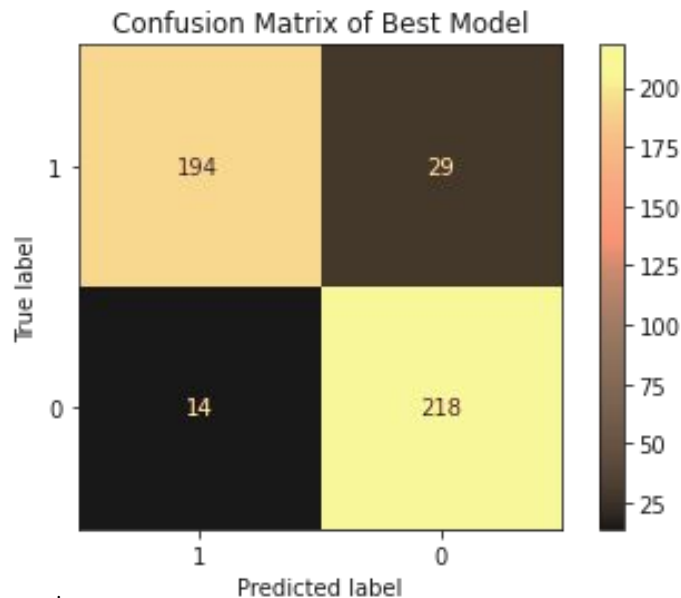
# Multinomial Naive Bayes with TF-IDF Vectorizer using Hyperparameters

- ❖ Training Score: 0.909
- ❖ Testing Score: 0.9055

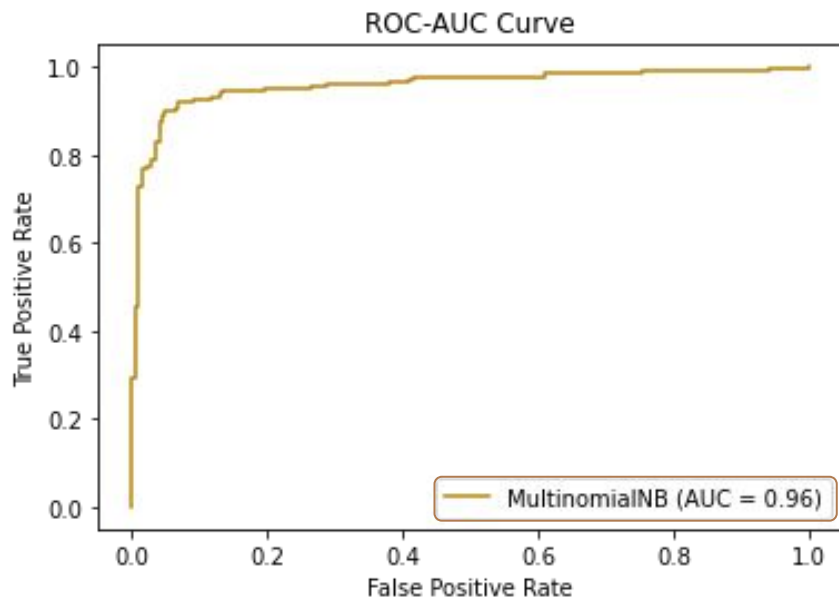
Best features retrieved from doing Gridsearch:

- `max_df = 0.9`
- `min_df = 3`
- `max_features = 2000`
- `ngram_range = (1,2)`
- `alpha = 2`

# Confusion Matrix



1 = Solotravel  
0 = Roadtrip





# FINDINGS

# Top 10 words that are correctly classified

## Solotravel

- Commun pleas
- Shit
- Complain
- Complain certain
- Share victori
- Constantly
- Love travel
- Content blog
- Contin
- Reminisc

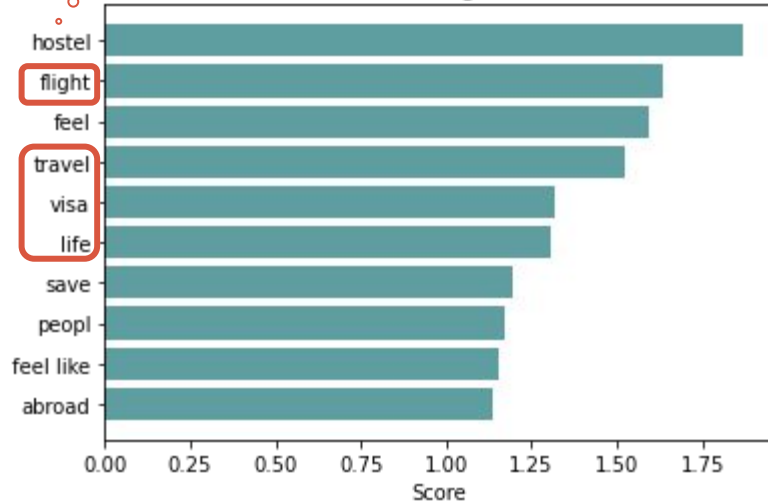
## Road Trip

- Savannah
- Geocach
- Oklahoma
- Dalla
- Bay area
- Bc
- Carolina
- North carolina
- Friend plan
- Awawego blogspot

# Top 10 words that appear in both

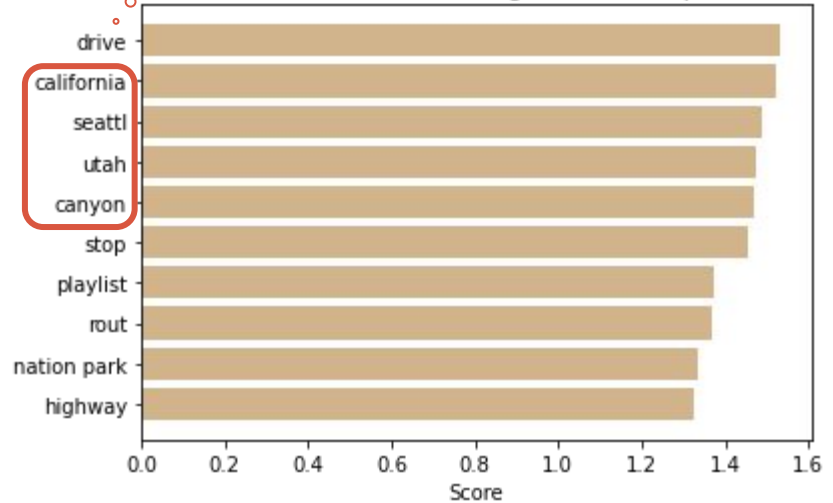
Hostel

Words with more weight for Solotravel



Drive

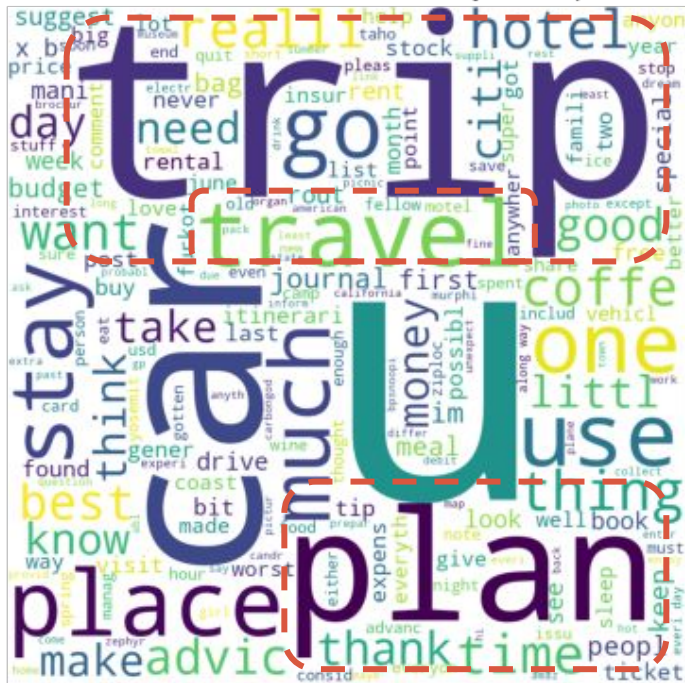
Words with more weight for Roadtrip





# Words causing misclassifications

Predicted Solotravel, but actually Roadtrip



## Predicted Roadtrip, but actually Solotravel





# CONCLUSION

## Solo Travel



'Visa', 'Flight', 'Hostel'

## Road Trip



'California', 'Grand Canyon'



'Plan', 'Time'



Travel is dependent on time of year and weather conditions



Other discussion forums should also be taken into account, global and domestic.



THANK  
YOU