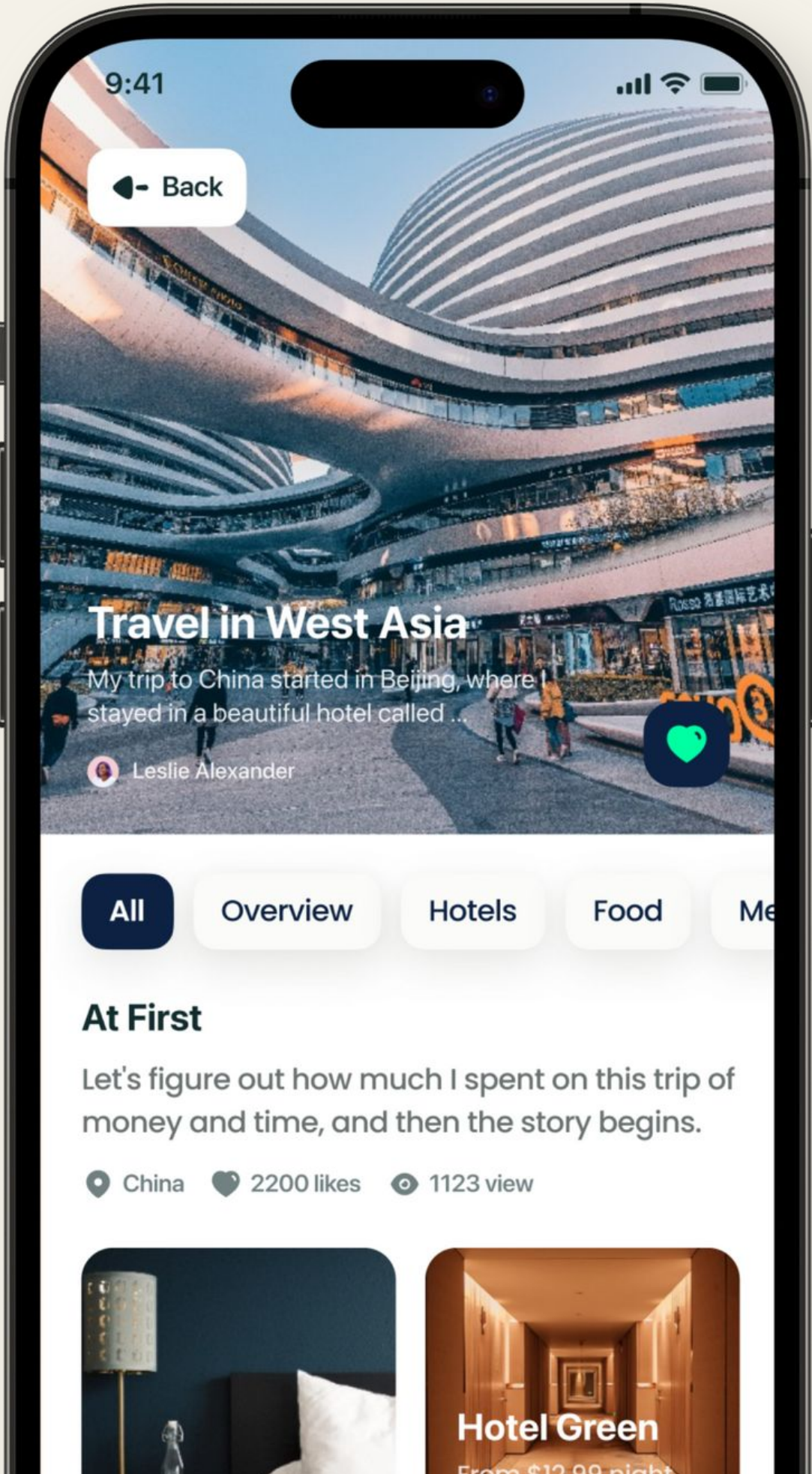


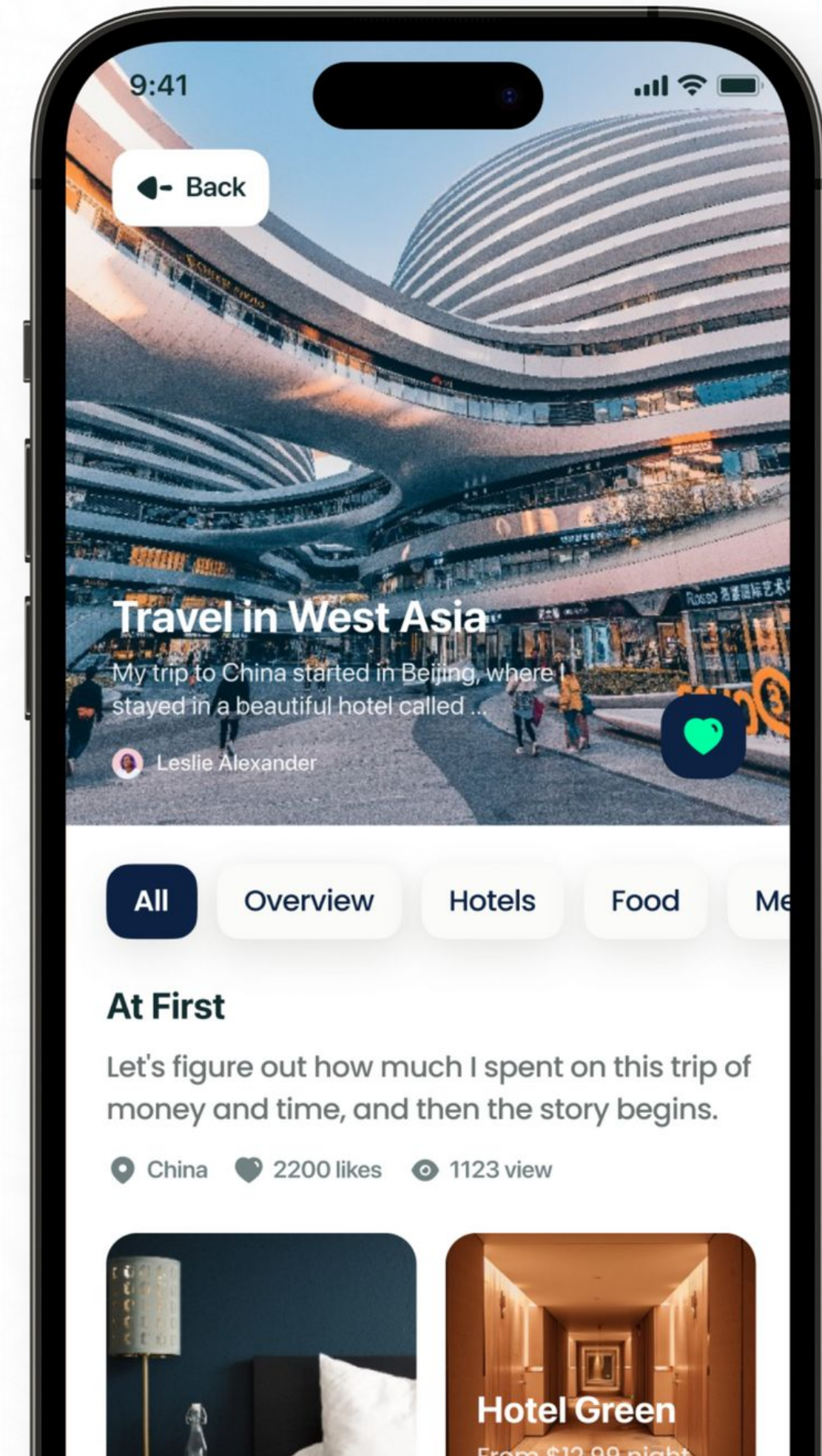
WE PLAN  
YOUR DREAM  
HOLIDAY





# Company Background

"We plan your dream Holiday ".  
Planner.Pro+ is a leading provider of  
personalized holiday itineraries,  
catering to travelers worldwide who  
seek exceptional and bespoke  
experiences.



# Table of content

**Problem Statement**

Decline in Sales



**Solution**

Leverage the vast amount of feedback and inquiries



**Insights &  
Recommendations**

Trip curation and marketing teams



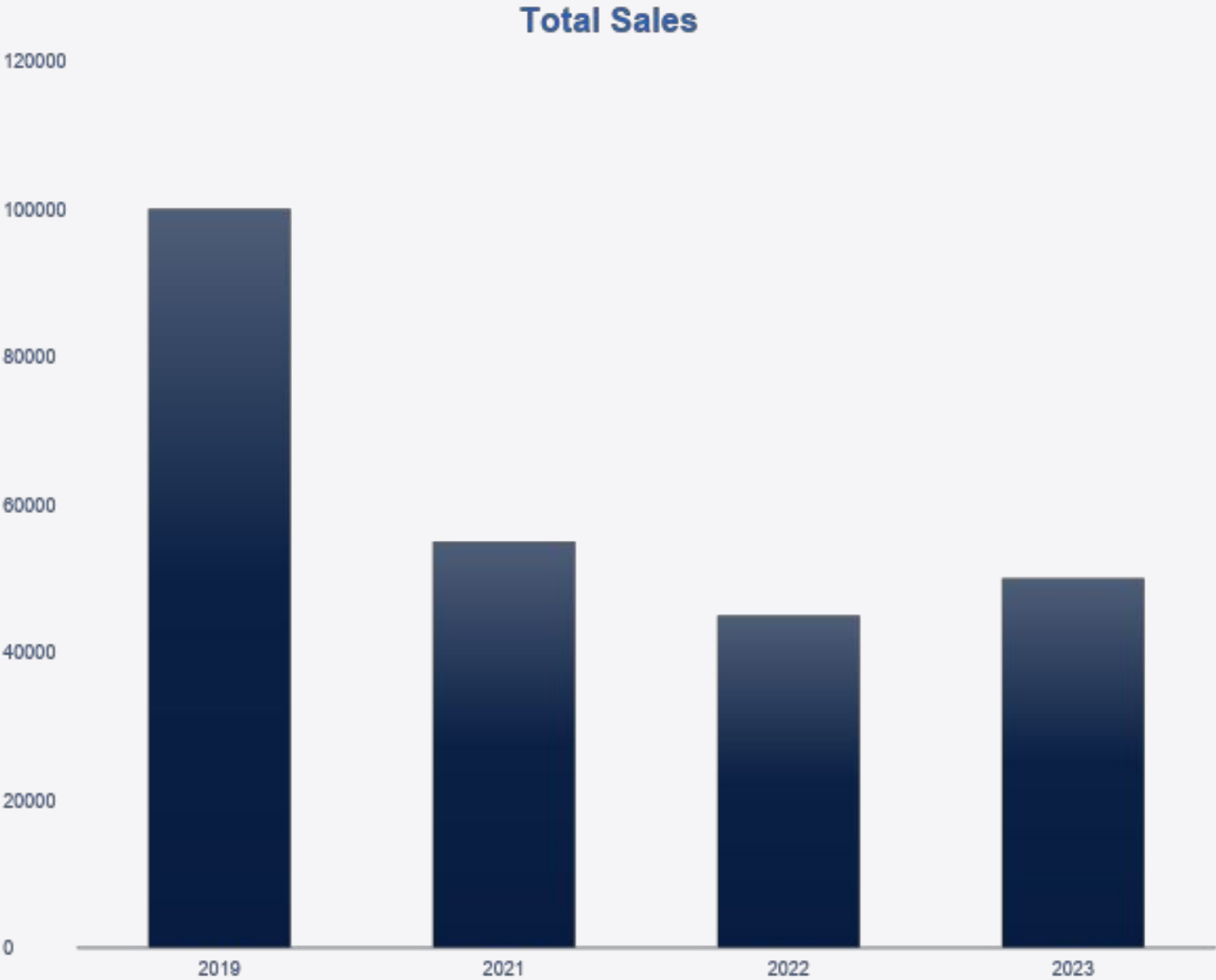
**Conclusion**

Key Takeaways





# Decline in Sales



Red Flag! Not personalised at all! Might as well as ChatGPT to plan my holiday!

Boring Trip !!!

17:57 54%

ChatGPT  
chat.openai.com

ChatGPT 3.5

**You**

Plan a 4 days 3 night family orientated trip to Hong Kong with Recommended Hotel

**ChatGPT**

Sure, here's a suggested itinerary for a family-oriented trip to Hong Kong:

**Day 1: Arrival and Exploring the City**

- Check into the Disneyland Hotel, which offers a magical experience for families and is conveniently located near Hong Kong Disneyland.
- Spend the day exploring Hong Kong Disneyland, enjoying the various themed lands, attractions, and shows.

# Solution

Leverage on the vast amount of feedback and inquiries received through multiple channels to generate valuable insights.

Training a ML Model

Generate Insights

Share with Departments



**Trip curation and  
marketing teams**

PLANNER.PRO+

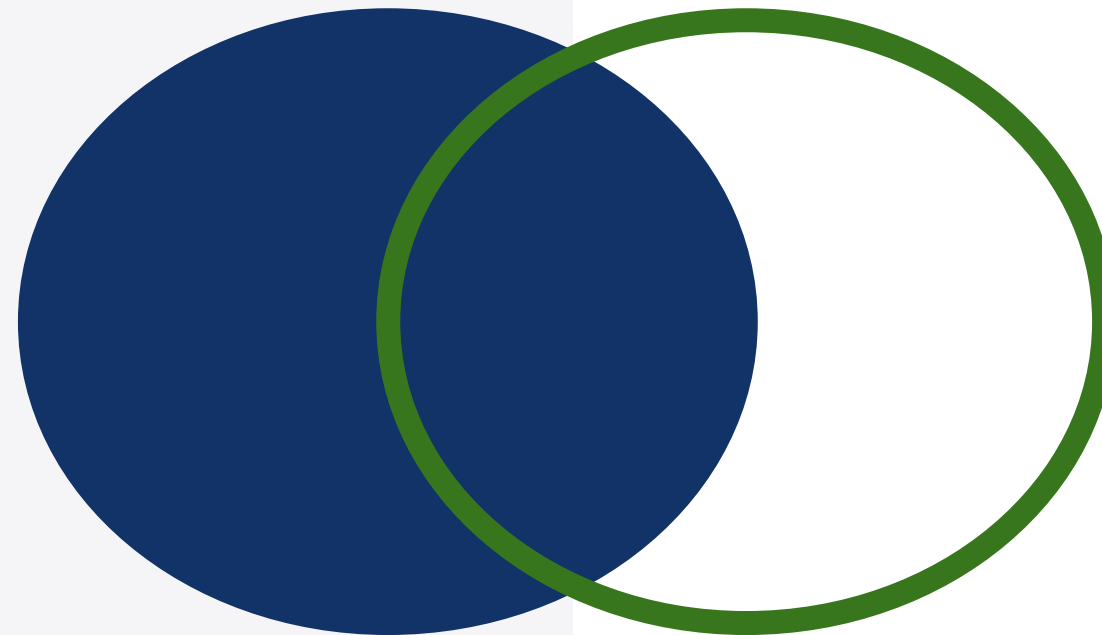
Total: Around 2000 posts

2 subreddits: Disney Parks & Universal Studio

| No. | Text Cleaning Methods | Examples  |
|-----|-----------------------|---|
| 1.  | Noise Reduction       | Remove Special Characters, HTML Tags, Punctuations, Emoji and Extra Space               |
| 2.  | Standardization       | Convert all text to lowercase   |
| 3.  | Expanded Contractions | Such as I'll to I will, she'd to she would  |
| 4.  | Stopword Removal      | common words like “the,” “and,” or “in”   |
| 5.  | Lemmatization         | Reduced words to their root forms – Eg Park and Parks                                   |
| 6.  | Handling Missing Data | Replaced image with a space instead   |
| 7.  | Tokenization          | Breaking text into individual words – Eg Went Disney Land to “Went” , “Disney” , “Land” |

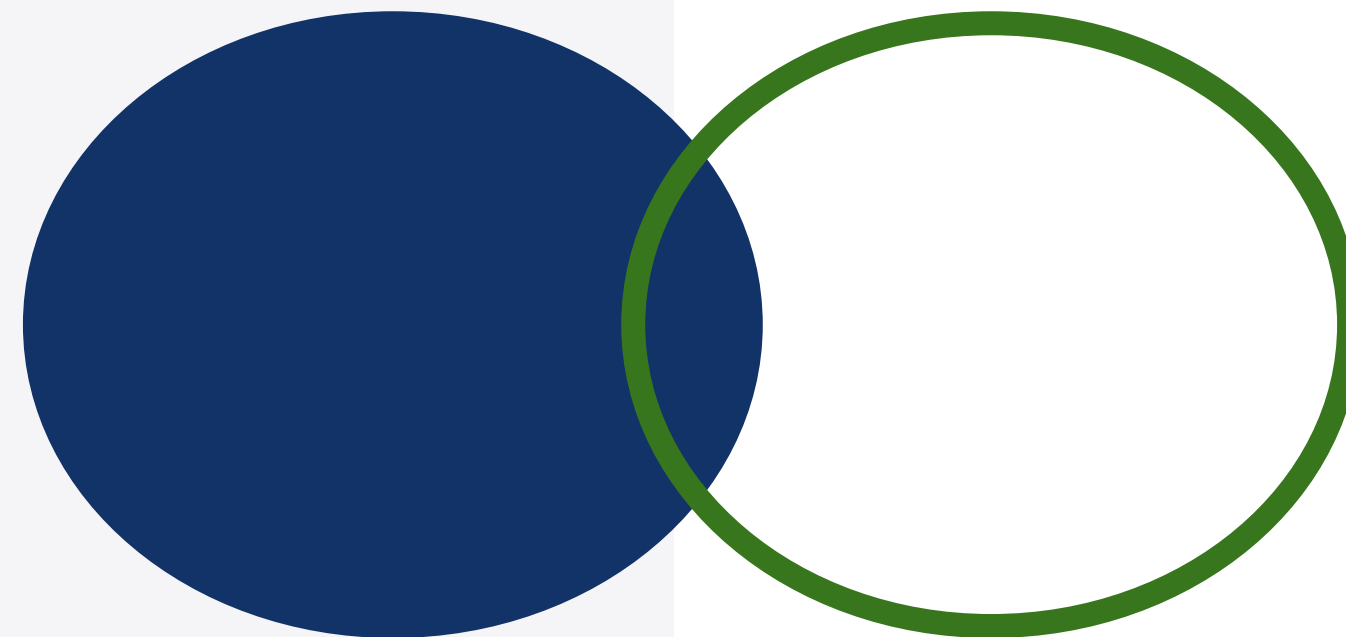
# Data Exploration : Common Words

Unigram



Eg “one”, “time”

Bigram

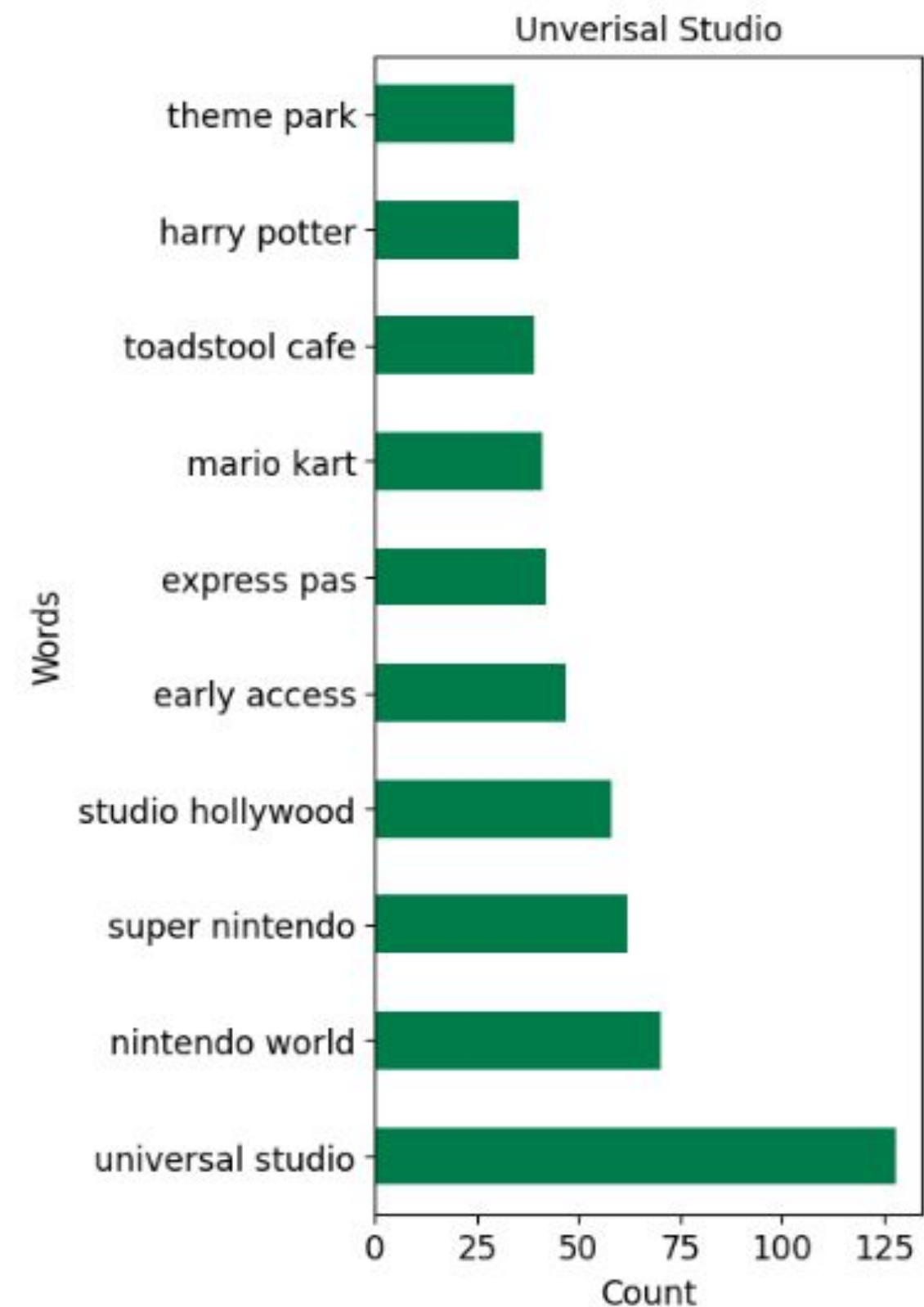
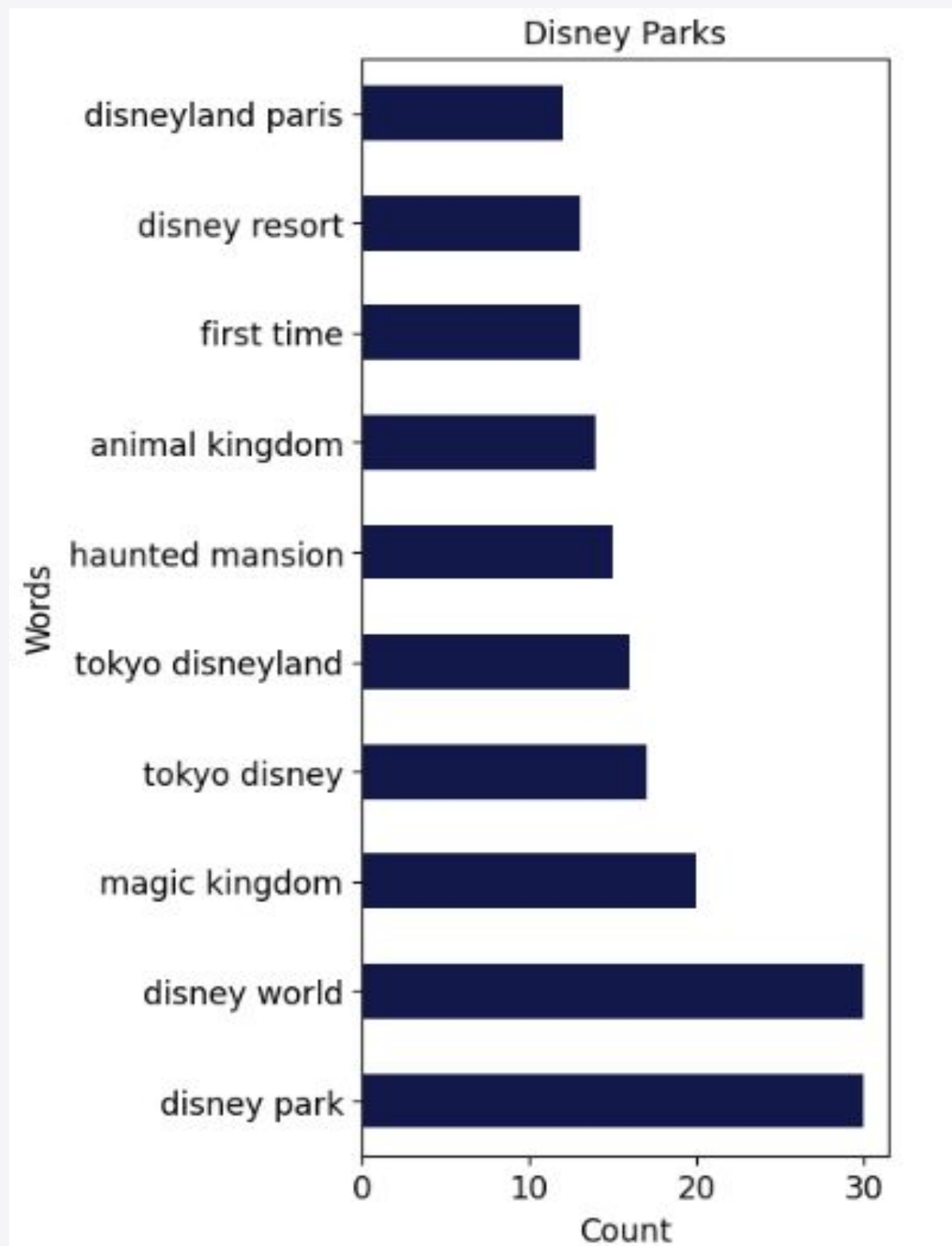


Eg “one time”

- Lesser shared words
- Capture more context
- Better distinguish between the 2 subreddits

Training the model with bigrams instead of unigrams

# Data Exploration : Bigram



Training the model with bigram instead of unigram



# Model Performance

2 subreddits: Disney Parks & Universal Studio

| No. | Models  | F1 Score on Test Dataset |        |
|-----|---|--------------------------|--------|
| 1.  | TF-IDF + SVC                                      | 0.930                    | Bigram |
| 2.  | TF-IDF + Logistics Regression + SVC + Naive Bayes | 0.930                    | Bigram |
| 3.  | TF-IDF + Logistics Regression                     | 0.927                    | Bigram |
| 4.  | TF-IDF + Naive Bayes                              | 0.926                    | Bigram |
| 5.  | TF-IDF + Random Forest                            | 0.891                    | Bigram |

# Model Performance

2 subreddits: Disney Parks & Universal Studio

| No. | Models       | F1 Score on Test Dataset |        |
|-----|--------------|--------------------------|--------|
| 1.  | Fast Text    | 0.932                    | Bigram |
| 2.  | TF-IDF + SVC | 0.930                    | Bigram |

PLANNER.PRO+.

*fast*Text

facebook

1



- useful for dealing with out-of-vocabulary words.

2

## Automatic Hyperparameter Optimization

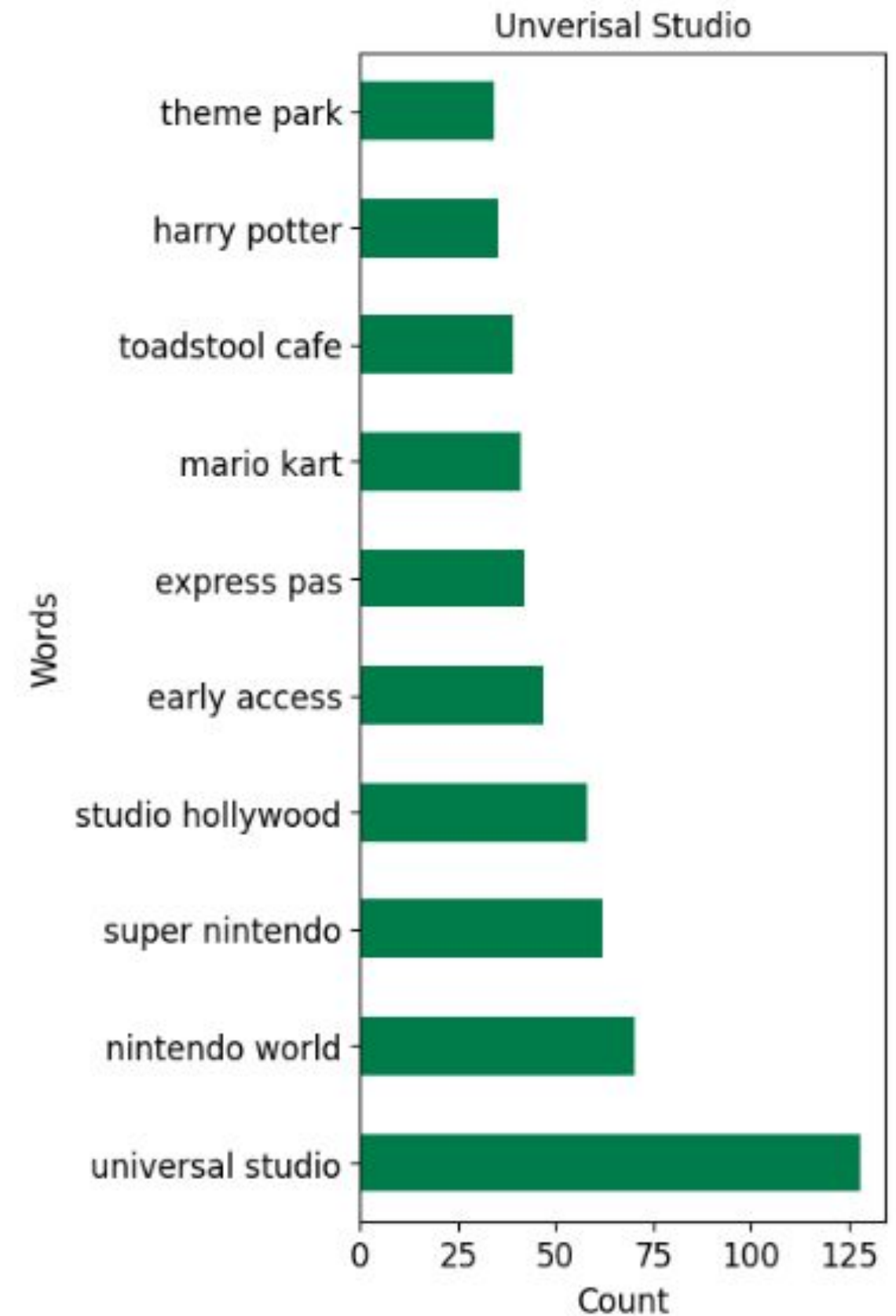
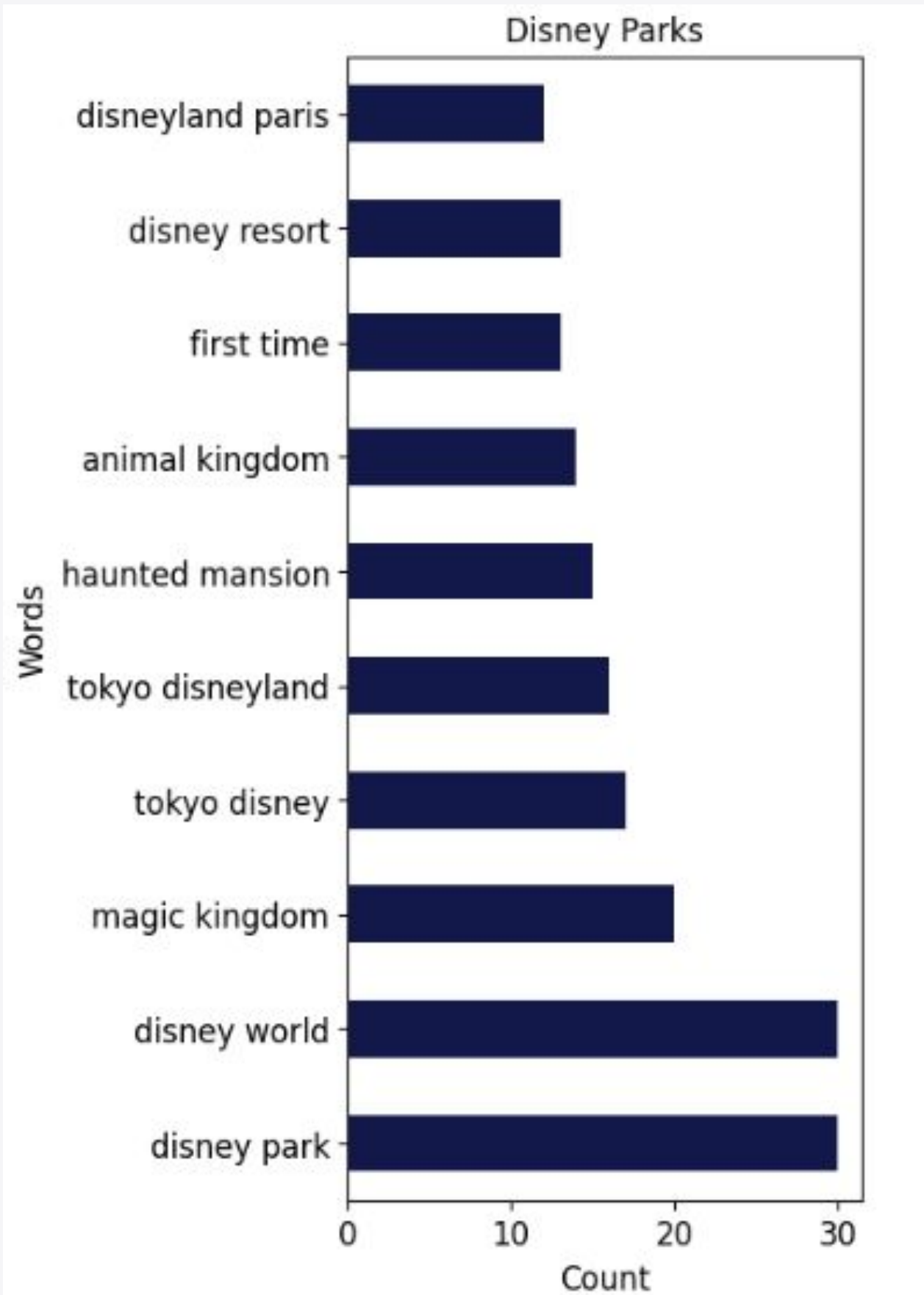
- Replace manual input in grid search

3

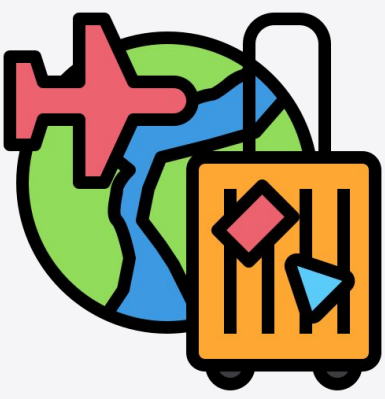
## More hyperparameters to Tune

- Learning rate
- Update learning rate
- Loss function
- Size of word vectors
- Epoch

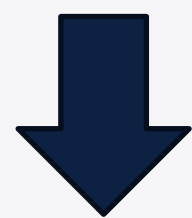




# Recommendations



| No. | Marketing Team   | Trip Curation Team  |
|-----|--|---|
| 1.  | Create content/blog that the public is interested in – Eg<br>“Do and don’t for your first time in Disneyland”, “Is it worth<br>staying at Disney Resort” and etc | Can be more detailed, such as what are the must play rides, route to<br>go to play most of the rides, where to eat inside the theme parks |



**Objective: Higher engagement and increase in website traffic**

- Improve the Search Engine Optimization(SEO) score
- Higher score = higher page ranking on the search page
  - leading to higher visibility and increase traffic



**Objective: More personalized trip**

- More satisfied customers
  - Higher return rate

# Conclusion

1. Best Model : **TF-IDF + SVC** & **Fast Text**

1. **Bigrams** vs Unigrams vs sub words :

- More meaningful content + lesser common words

1. Its property:

- work better with **non linear relationship** data
- **Less sensitive to noise** and **outliers**
- Effective with **high dimensional data** as it seeks to find the **hyperplane** that best separates the data points into different classes.
- **Non parametric model** - adaptive to different distribution
- More **hyperparameters to tune**



**WE PLAN  
YOUR DREAM  
HOLIDAY**

**WE PLAN  
YOUR DREAM  
HOLIDAY**