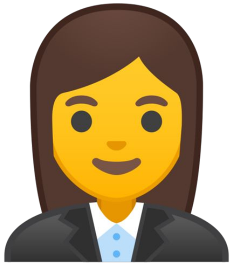
Aspect Based Sentiment Analysis of Luxury Hotel Online Reviews in Singapore

Bryan C

29 April 2021

Problem Statement



Big Picture Insights

Executive Team/GM

Online Reviews

- Multiple Sources

Processed by Marketing Staff/GM PA

(Hours spent on reading and 'annotating')

Unnecessary Skilled Labour Costs!

Department Heads

Insights Specific to their Dept

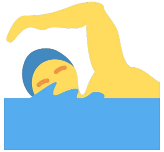
Online Reviews have predictable language usage and often talk about the same aspects for a given subject (in this case, across all hotels).

This makes processing online review data highly suitable for automation via a data analysis solution.

As such, the **objective** of the project was to create an Aspect-based model based on **Marina Bay Sands** review data and see if it could be applied to other hotels (in this case, the **Fullerton Hotel**).

What is Aspect-Based Sentiment Analysis?

To recognize different parts of a text which feature an aspect of the main subject and predicting sentiment with relation to that aspect.



Eg. I had the most wonderful time at the **pool!** However, the **breakfast** was objectively awful.

Methodology

01

Acquiring and Modifying Data

Review data of MBS and Fullerton scraped from TripAdvisor. Data is bootstrapped and Annotated Manually.

02

Building the Model

2-Step Multiclass classification problem for aspects (categories) and then sentiment with focus on accuracy.

03

Studying Insights from Model

The models reveal insights which may be useful for further work.

What is Bootstrapping?

Using a small set of annotated data in a model to predict (and thus annotate) even more data.
Both aspects and sentiment were bootstrapped

Annotated 1000 rows (4 hours) > TF-IDF Vectorizer and Random Forest Classifier

Annotated 1000 rows (2.5-3 hours) > TF-IDF Vectorizer and Random Forest Classifier

So on and so forth. Gets faster each time.

Data from 2019 and earlier

A look at the data

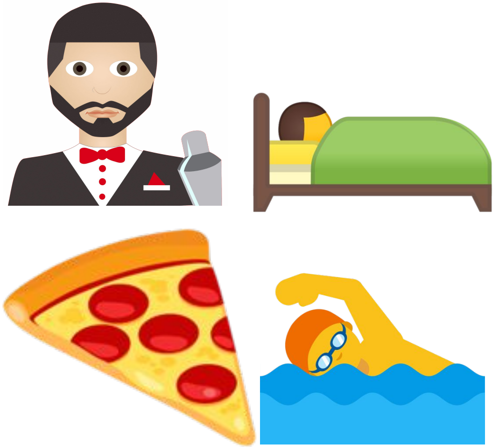
Marina Bay Sands

- Integrated Resort
- Booking.com: 9.0, TripAdvisor: 4.5

The Fullerton Hotel

- Colonial Luxury Hotel
- Booking.com: 9.1, TripAdvisor: 4.5

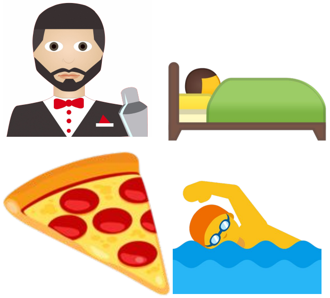
EDA: Aspects



Through Trial and Error, Decided Upon 6:

1. **Service** -- Anything involving staff, service
2. **Room** -- rooms, bathrooms, in-room amenities
3. **Pool** -- swimming pool, pool experience
4. **Food** -- breakfast, lunch, dinner, restaurants and bar
5. **Others** -- other features not in the rest (lifts, lobby, Gardens by the Bay)
6. **None*** -- Anecdotes, statements

Modeling (for Aspects)



Bag-of-Words Model. 3300+ rows of MBS data train-test split into 75-25% sets.

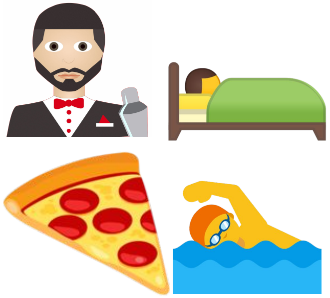
Text is from the Review Text Only (most Review titles too vague). **Words are sentencized using SpaCy** and then **Lemmatized with NLTK (root words are found).**

- TF-IDF Vectorizer used for all

Classifiers:

- Multinomial Naive Bayes
- Linear SVC
- Random Forest
- Ada Boost

Results



	F1 Score		Precision	Recall
	Micro (Acc.)	Macro	Macro	Macro
M-NBayes	0.74	0.72	0.77	0.68
Linear SVC	0.8	0.79	0.82	0.76
RForest	0.76	0.75	0.78	0.73
ABoost	0.75	0.7	0.77	0.68

- Optimized for **Micro F1 Score, analogous to accuracy.**
- Baseline to beat:

0.36

(predicting all as majority class of 'None')

Modeling (Sentiment Analysis)

Performed on both Lemmatized and Raw text.

Pre-trained Models:

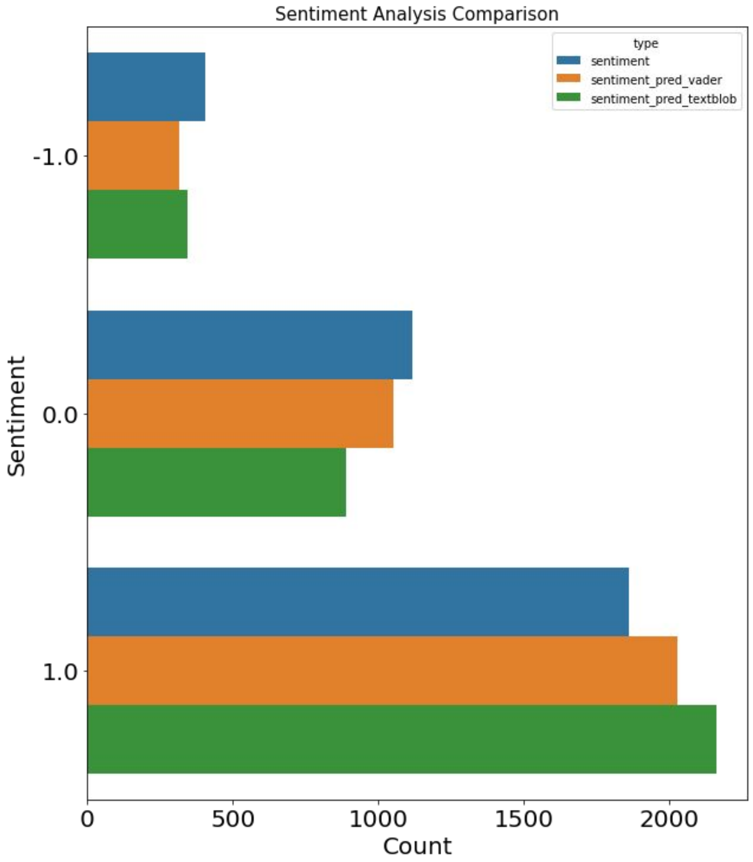
- NLTK VADER
- TextBlob

Benchmarked Against:

- Random Forest

Results

		Accuracy/F1
VADER	Unlemmatized	0.76
	Lemmatized	0.73
TextBlob	Unlemmatized	0.67
	Lemmatized	0.65
RForest		0.6



Will our models do on unseen data of the Fullerton Hotel?

EDA of 1100 of sentencized data shows that the Fullerton defers from MBS in several significant ways:

1. Much fewer rows talking about room stays
2. Many more about food, especially specific dishes
3. Many more about service personnel (MBS has a few, but not as many).

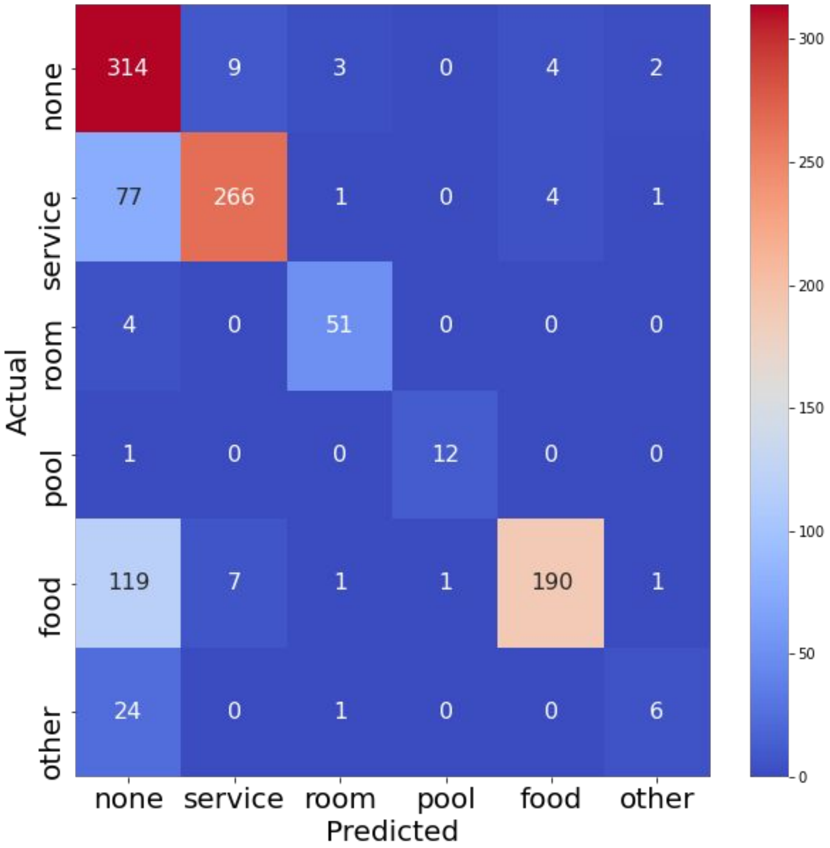
Can a Bag-of-words model handle word similarities?

Results (Fullerton)

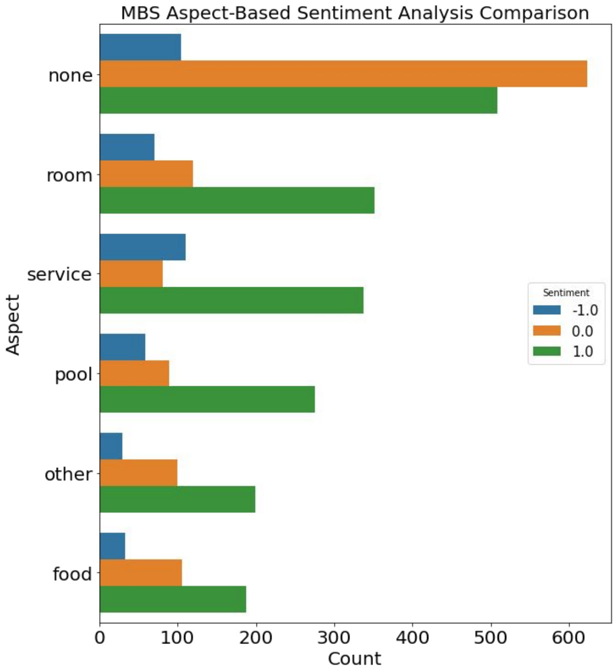
	F1 Score		Precision	Recall
	Micro (Acc.)	Macro	Macro	Macro
Linear SVC	0.76	0.74	0.82	0.72

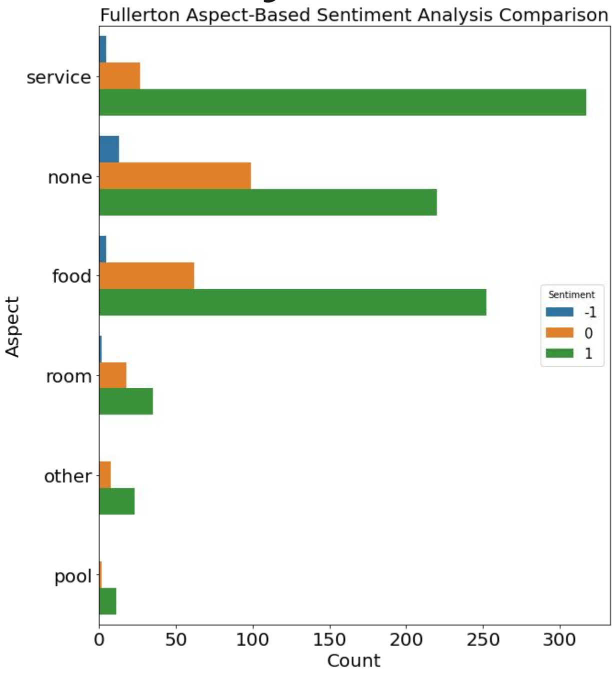
As expected, many ‘food’ aspect rows were mis-predicted.

This reveals the **limitation of bag-of-words models** -- unable to generalize words with close semantic meanings.



Final Aspect-Based Sentiment Analysis (Predicted)

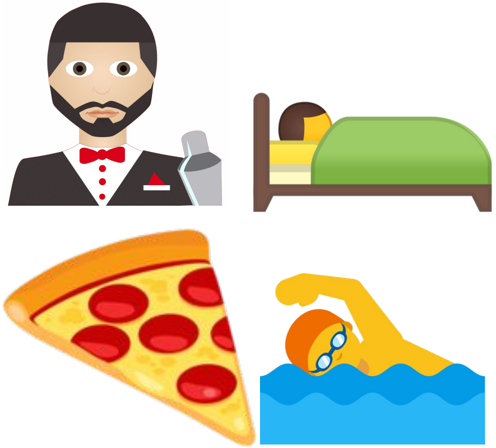




Conclusion

1. Aspect Prediction Model has F1 Micro Score of 0.80
2. Sentiment Prediction Model has a F1 Micro of 0.75
3. **ABSA model achieves purpose, but needs to be more sophisticated for production use.**

Things I attempted but could not complete



1. Using SpaCy to accurately **identify noun and adjective/verb pairs using dependency trees**
2. A way to **generalize certain types of nouns into special features and insert them back into the review text** (eg. names into something like {NAME} and features like Gardens By The Bay into {LANDMARK}

Other Future Work

1. Method to handle compounded lines ('the food and room were good but the service was not!'). An algorithm able to handle multiclass classification will have to be found.
2. **Beyond bag-of-word models:** Find a way to handle semantic meanings of words better. Eg., 'the cupcakes were good' could not be detected to be talking about food if 'cupcake' was not found in the train set and labeled to the category 'food'.
   a. How to incorporate algorithms like GloVe and Word2Vec into the model?

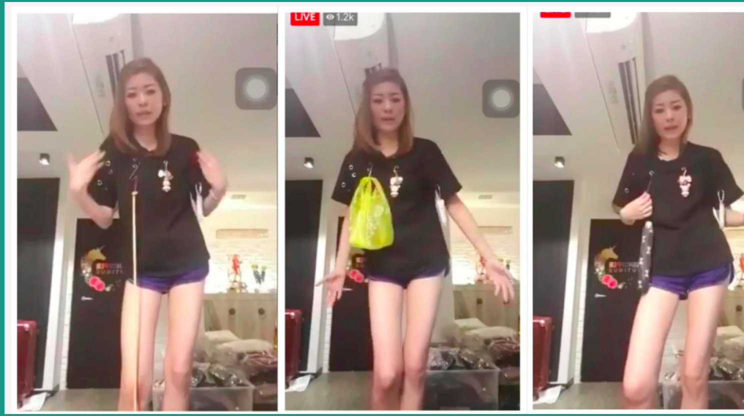
Now For Some Funny Ones

Someone likes costumed capers very much (MBS)

"Personnel is **costumer oriented.**"

Fashion Forward Russians (MBS)

"Oh and Russians.
Those **pesky Russians** and their desire to **wear the hotel provided robe and slippers from their hotel room on level 4 all the way to the pool on level 57.**
For **some reason the rest of the folk follow that trend.**"



Hooks and warm arses (Fuller)

"The bathroom was huge and we had a **ToTo toilet that warmed are arses** 😬 **and Hooks!!**

Believe it or not you get points when you have more than one hook"

Thank you.
Q&A!