**IOD Capstone Project**

****

**Pizza Hut USA Sentiment Analysis Report**

By

Vernon Chun

19 May 2023

**Table of Contents**

[1. Problem Statement 3](#_Toc135341749)

[2. Industry/Domain 4](#_Toc135341750)

[3. Stakeholders 4](#_Toc135341751)

[4. Business Questions 4](#_Toc135341752)

[5. Data Questions 5](#_Toc135341753)

[6. Data 5](#_Toc135341754)

[7. Data Science Process 7](#_Toc135341755)

[7.1. Data Pre-Processing 7](#_Toc135341756)

[7.2. Exploratory Data Analysis (EDA) 7](#_Toc135341757)

[7.3. Text Pre-Processing 14](#_Toc135341758)

[8. Modelling 19](#_Toc135341759)

[9. Modelling Results 21](#_Toc135341760)

[9.1. Machine Learning Results 22](#_Toc135341761)

[9.2. Ensemble Learning Results 23](#_Toc135341762)

[9.3. Deep Learning Results 24](#_Toc135341763)

[9.1. SVM Sample Predictions 25](#_Toc135341764)

[10. Data Answer 26](#_Toc135341765)

[11. Business Answer 26](#_Toc135341766)

[12. Response to Stakeholders 26](#_Toc135341767)

[13. Conclusions and Recommendations 28](#_Toc135341768)

[14. References 29](#_Toc135341769)

# Problem Statement

Pizza Hut in America is experiencing an overall decline in revenue for the last three years, as shown in the graph below from the financial statement information from its parent company Yum! Brands. Quarter 4 revenue in 2021 yielded the lowest value across the 3 years. The value was just more than a third of the revenue at the same point in time in 2020. Operating Margin is also trending downwards. The company needs to understand how to arrest this situation.

**Figure 1:** *Pizza Hut Earnings Release* by Yum! Brands Inc., 2019,2020,2021, 2022

**Figure 2:** *Pizza Hut Earnings Release* by Yum! Brands Inc., 2019,2020,2021, 2022

# Industry/Domain

Pizza Hut USA competes in the food service industry both in the United States and internationally. Although its speciality is pizza, it also offers pasta, chicken wings, salads and desserts. The industry is extremely competitive with numerous other significant players such as Domino’s Pizza, MacDonalds and KFC. Characteristics of the industry include extreme price competition, fast-changing consumer preferences and a focus on speed and convenience.

As a consequence of the recent COVID-19 pandemic, the entire industry has had to adapt their operating models from the traditional in-dining restaurants to takeout and delivery models, in order to meet the ever-changing needs of its consumers. Consequently, the ability to order online using mobile apps is key. In a similar vein social media are becoming more vital in driving revenue and customer engagement.

Given the heightened focus to better understanding customers. Companies that better understand customer emotions regarding products and services are more likely to enhance customer satisfaction, boost brand reputation, and ultimately drive revenue growth (Tan, Lee & Lim 2023).

The focus of this project is to ascertain Pizza Hut USA customer sentiment from their reviews on an established crowd-sourced business review site such as Yelp. This will equip Pizza Hut with the information to make data-driven decisions.

# Stakeholders

Given the magnitude of the problem, the entire Pizza Hut USA executive management team are involved in this critical project. These include:

* Chief Executive Officer
* Chief Operating Officer
* Chief Financial Officer
* Chief Marketing Officer
* Chief Digital & Technology Officer

# Business Questions

* 1. What is the cause of our declining revenue and operating margin from 2019 to 2022?
  2. Are there any particular states that can explain the decline?
  3. What strategies can we implement to better understand the market situation and what actions are required in order to return revenue and profitability to 2021 levels?

# Data Questions

Using a combination of the financial data from Pizza Hut’s parent company Yum! Brands in conjunction with publicly available data from Yelp, the following data questions can be answered:

* 1. What trends are evident from the financial data?
  2. What insights are shown from customer sentiment data?
  3. What is the best data-driven approach to predict customer sentiment for Pizza Hut?

# Data

The data was sourced from [*https://www.yelp.com/dataset*](https://www.yelp.com/dataset)*.* Yelp is one of the most established and best-known online review sites and frequently sits at the top of the search rankings.

The Yelp datasets comprise of 6 990 280 reviews for 150 346 businesses.The original data was in the form of two json files:

* **Review.json**

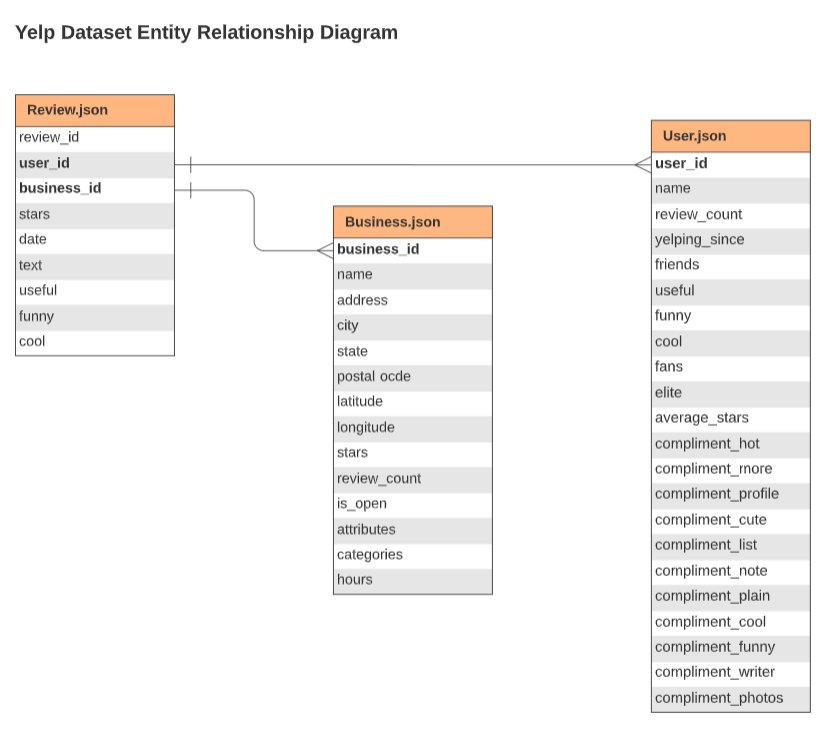
This file contains full review text data including the user\_id that wrote the review and the business\_id the review is written for.

* **Business.json**

This file contains business data including location data, attributes, and categories.

* **User.json**

This file contains user data including the user's friend mapping and all the metadata associated with the user.



**Figure 3:** *Yelp Dataset Entity Relationship Diagram*

A merged dataset was first generated by linking **review.json** and **business.json** using **business.id** as the key. The merged dataset was then linked to **user.json** using **user\_id**. The purpose of linking **user.json** was to extract “yelping\_since” feature so as to ascertain the user’s years of Yelp experience.

In order to process the data more effectively, a two-stage filtering process was adopted:

1. A new interim **pizza\_hut\_df** was created and its unique categories determined.
2. These categories were then used as a mechanism to create a new **reviews\_df** dataset of a more manageable 70069 rows.

# Data Science Process

## Data Pre-Processing

The following data pre-processing steps were performed on the **Pizza\_hut\_df** and **Reviews\_df** datasets:

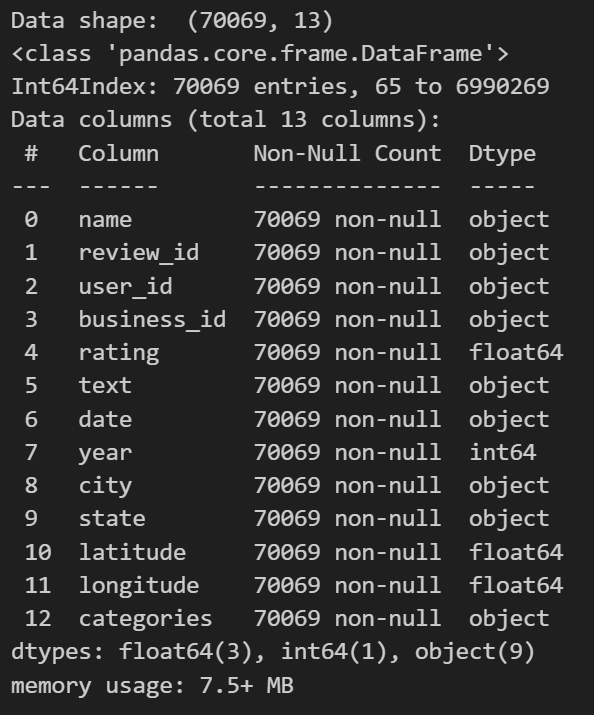
1. Drop unnecessary columns.
2. Rename “stars\_x” to “rating”.
3. Reformat “date” column by removing the timestamp*.*
4. Created a new “year” column and added it before the “date” column.
5. Moved “name” column to the beginning of the dataframe.
6. Moved “rating” column to the end of the dataframe.
7. Wrote the datasets to csv for the purpose of faster access in future.

## Exploratory Data Analysis (EDA)

The following EDA steps were performed on the **Reviews\_df** dataset:

* **Check data shape**

The dataset comprised of 70069 rows and 13 columns.



**Figure 4:** *Reviews\_df Data shape*

* **Missing value check**

There were no missing values in the dataset.

* **Check and plot Rating labels**

Five-star ratings formed the majority, whilst two stars the minority. This is shown in figure 5.

Chart, bar chart

Description automatically generated

**Figure 5:** *US Pizza Restaurant Star Ratings*

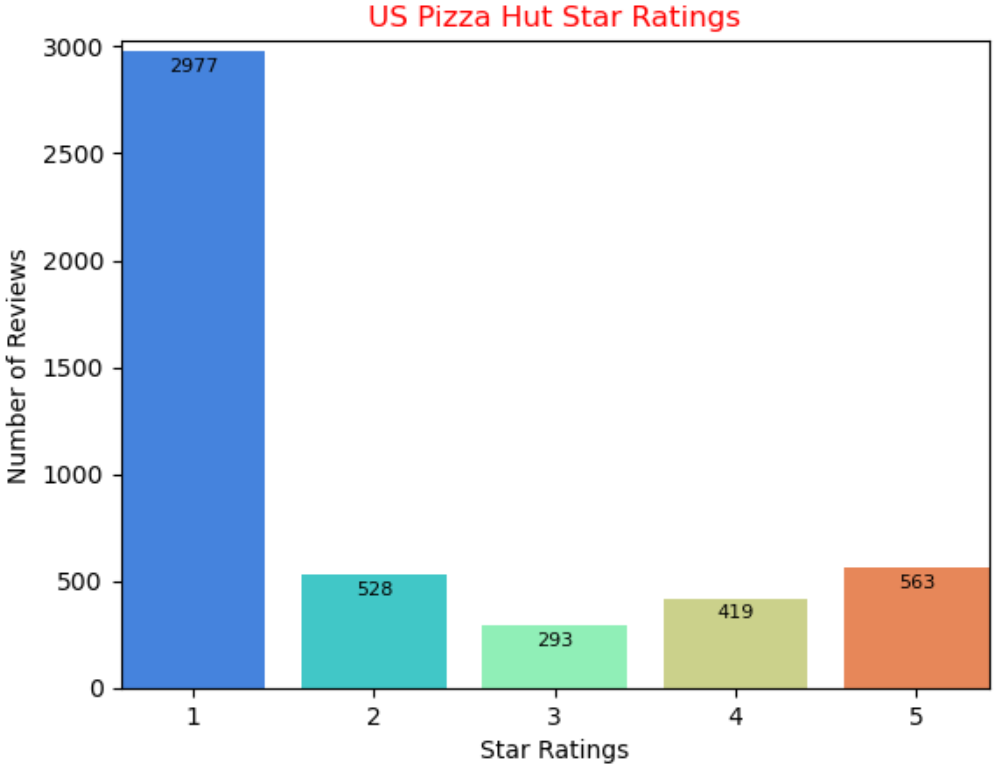
The proportions of the above ratings are reflected in the pie-chart in figure 6.

**Chart, pie chart

Description automatically generated**

**Figure 6:** *US Pizza Restaurant Star Ratings by percentage*

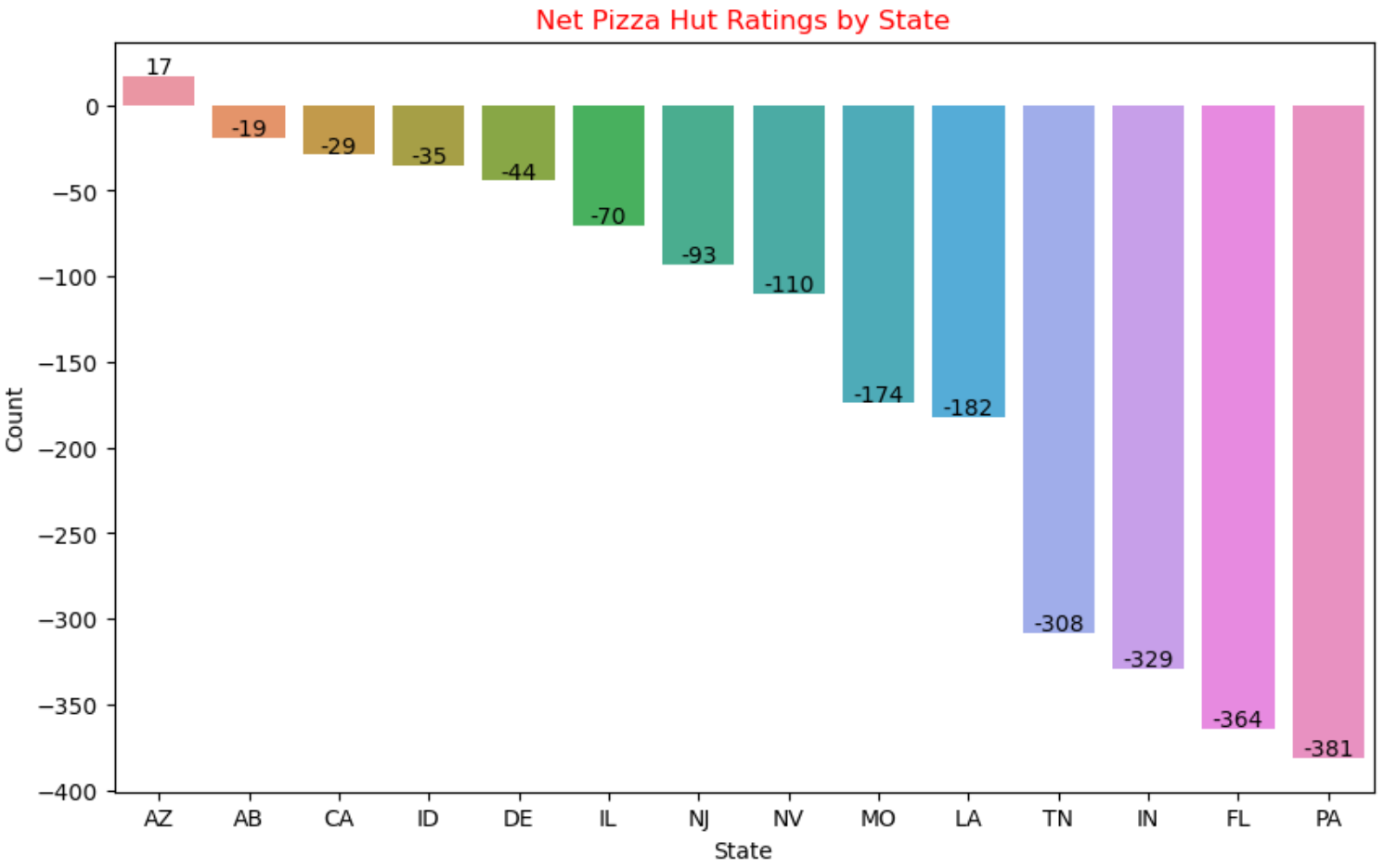
The high number of 1-star ratings for Pizza Hut in figure 7 is of concern, particularly given that the 5-star ratings across all pizza restaurants is in the majority.



**Figure 7:** *US Pizza Hut Star Ratings*

* **Plot plots of Net Pizza Hut Star Ratings by State**

To obtain a better picture of good ratings (3 stars and above) versus bad ratings (2 stars and below) by US State, counts of good and bad ratings were determined and the difference calculated and plotted. This is shown in figure 8.

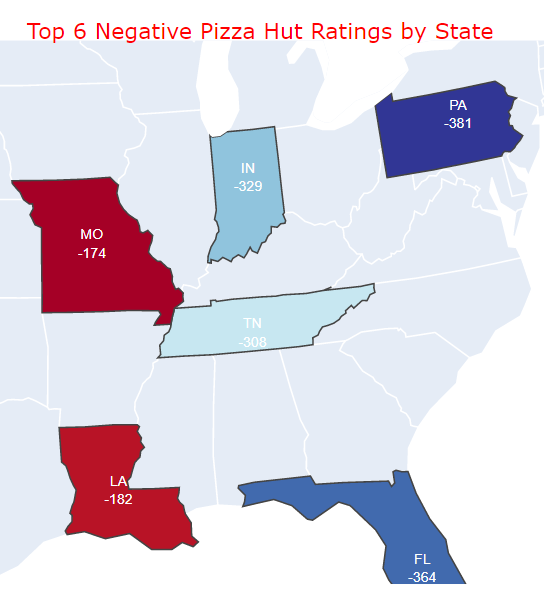


**Figure 8:** *Net Star Ratings by State*

It can be seen from the plot that Pennsylvania is the state which has more negative than positive reviews for Pizza Hut. Accordingly, it should be the first state followed by Florida where any rectification efforts are directed.

* **Plot map plots of 6 worst Negative Pizza Hut Star Ratings.**

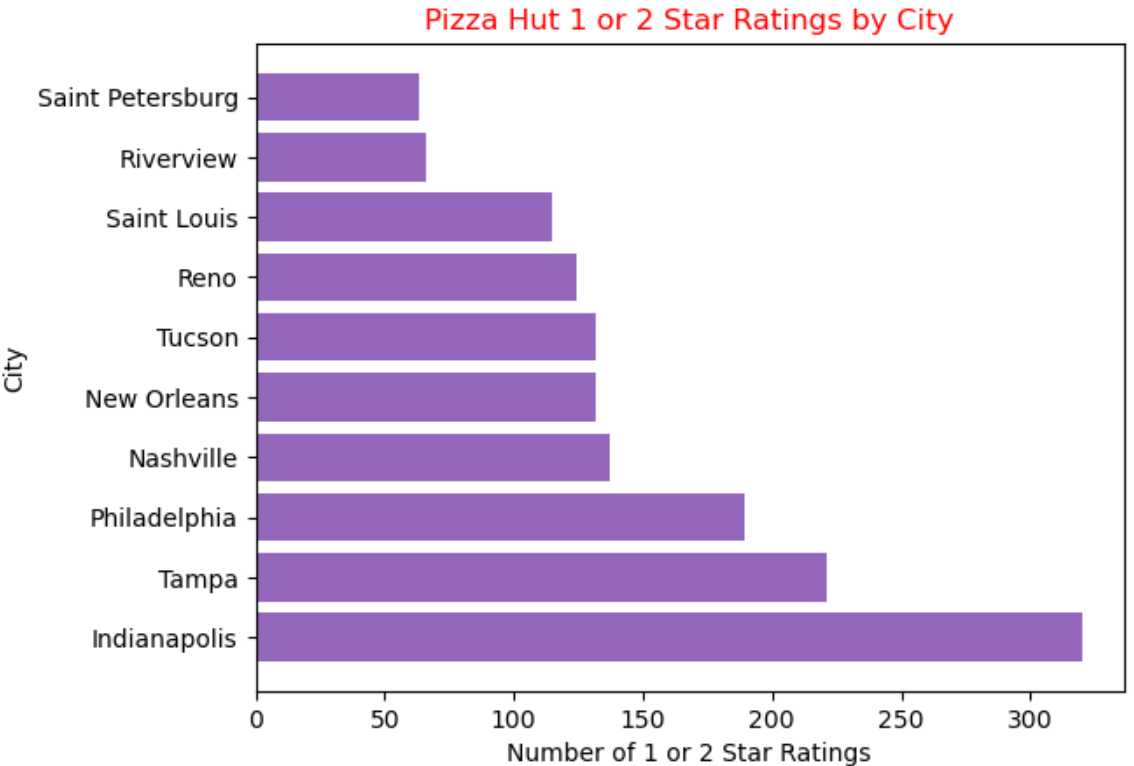
In order to determine a graphical representation of the 6 worst states for ratings, the net difference between good and bad ratings were plotted. As shown in figure 9, most of the poor ratings appear to be from the US East Coast States .

Timeline

Description automatically generated with low confidence

**Figure 9:** *Top 6 Negative Pizza Hut Ratings by State*

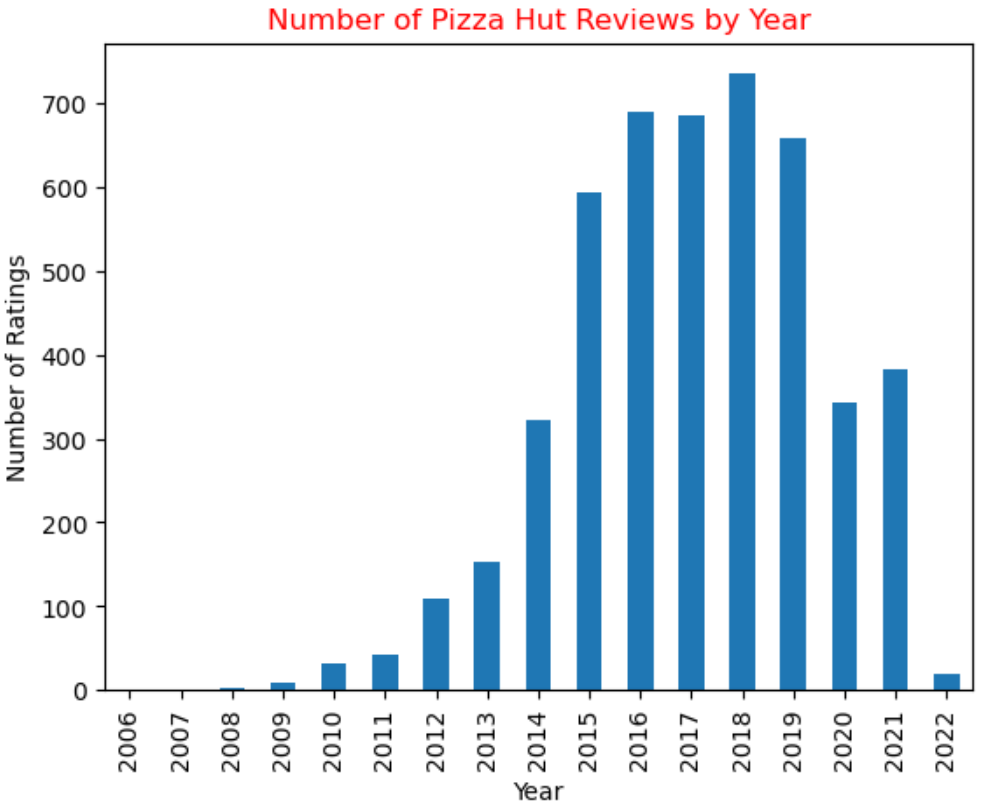
* **Plot histogram of Cities with 1- or 2-Star Ratings.**



**Figure 10:** *Top 10 Pizza Hut 1- or 2-Star Ratings by City*

As shown in figure 10, Indianapolis is the city with the majority of 1- or 2-star ratings for Pizza Hut. It is located in the third worst state for negative Pizza Hut ratings.

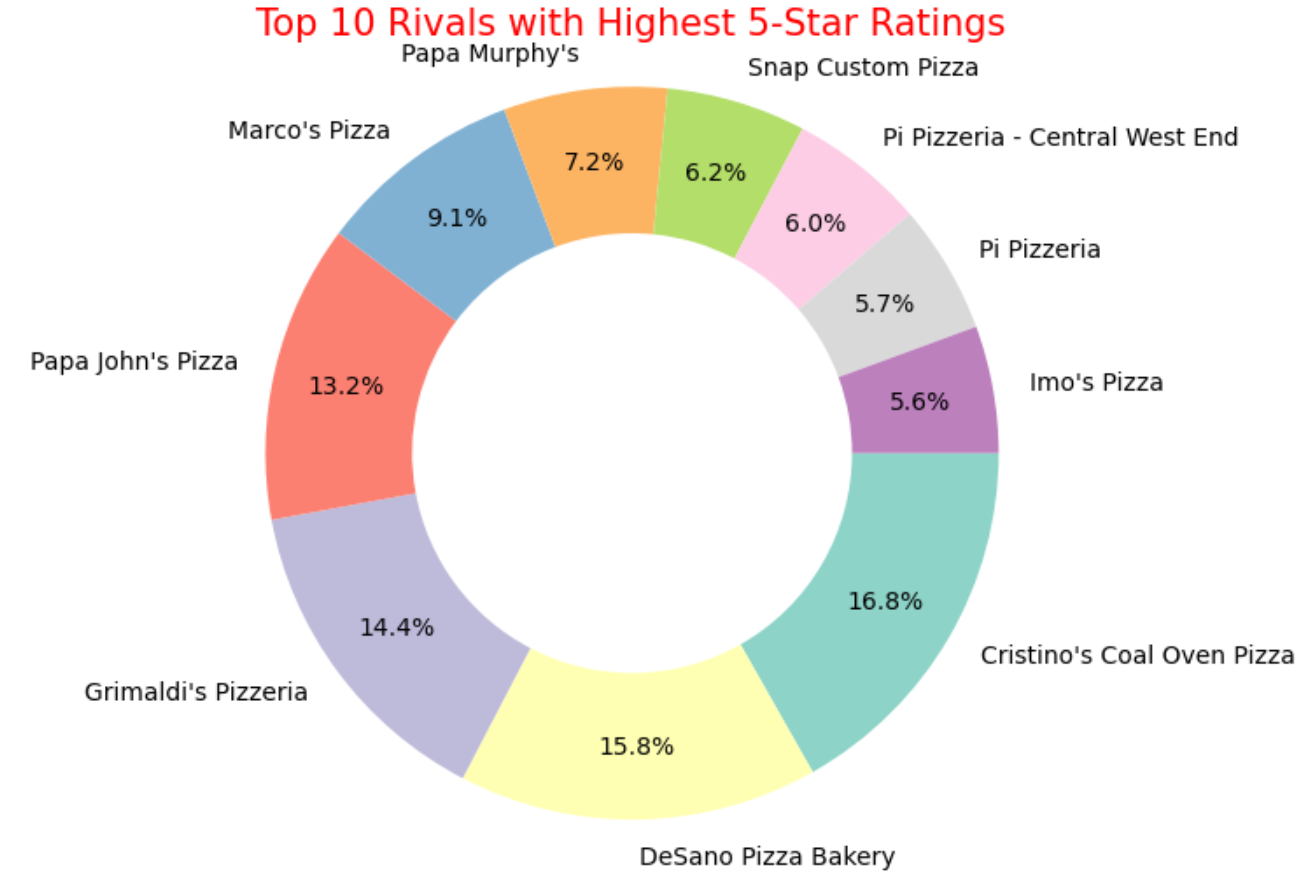
* **Plot histogram of Pizza Hut Reviews by Year**



**Figure 11:** *Number of Pizza Hut Reviews by Year*

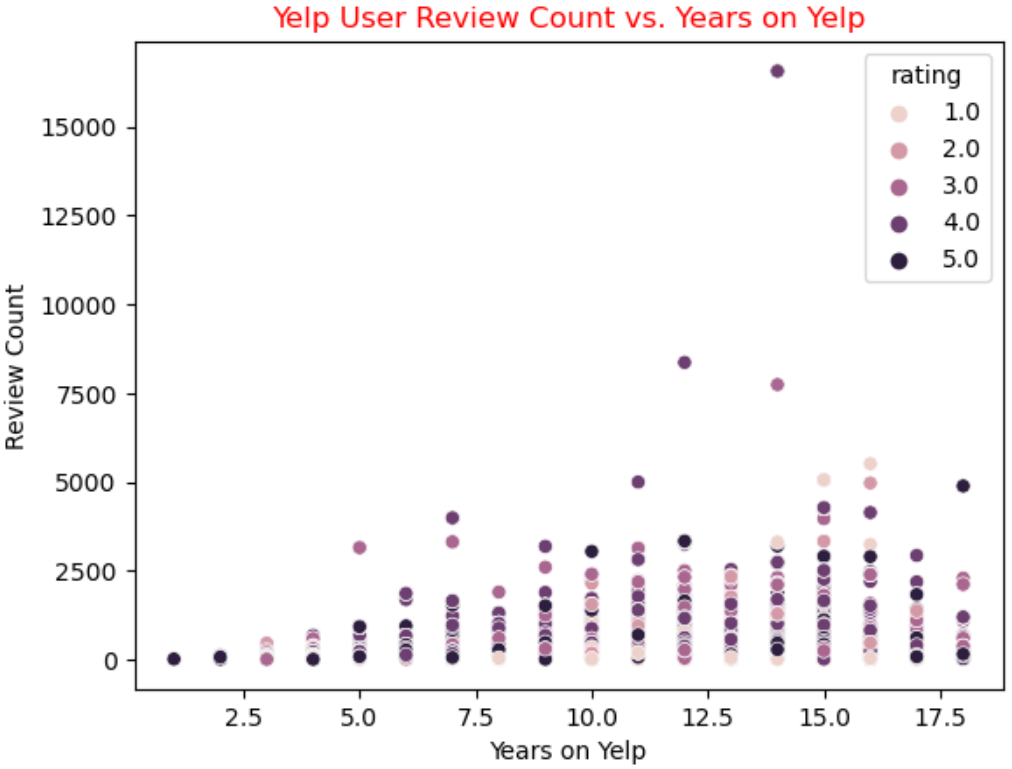
It can be seen from figure 11 that the lowest number of Pizza Hut reviews were in 2020 and 2021. This is aligned to when the COVID-19 pandemic was as its peak in the U.S.

* **Plot Donut Chart of Highest 5-Star Ratings**

**

**Figure 12:** *Top 10 Rivals with Highest 5-Star Ratings*

* **Plot Yelp User Review count versus Years on Yelp**



**Figure 13:** *Yelp User Review count versus Years on Yelp*

* **Plot count of User Years on Yelp**

A picture containing text, diagram, screenshot, plot

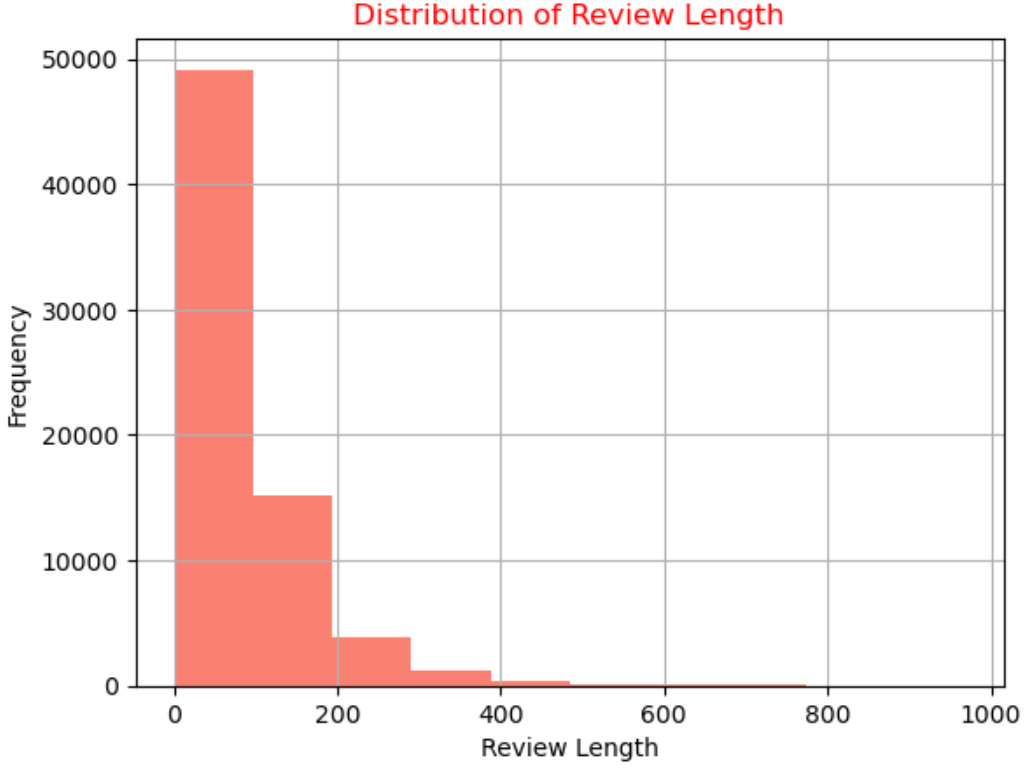
Description automatically generated

**Figure 14:** *Histogram of User Years on Yelp*

## Text Pre-Processing

In addition to “normal” pre-processing prior to Machine Learning modelling, additional pre-processing is required for text, in order for the data to be usable for making predictions. The following steps were performed on **reviews\_df.text**:

* **Application of text cleaning using regular expressions**
* **Checking of review length**



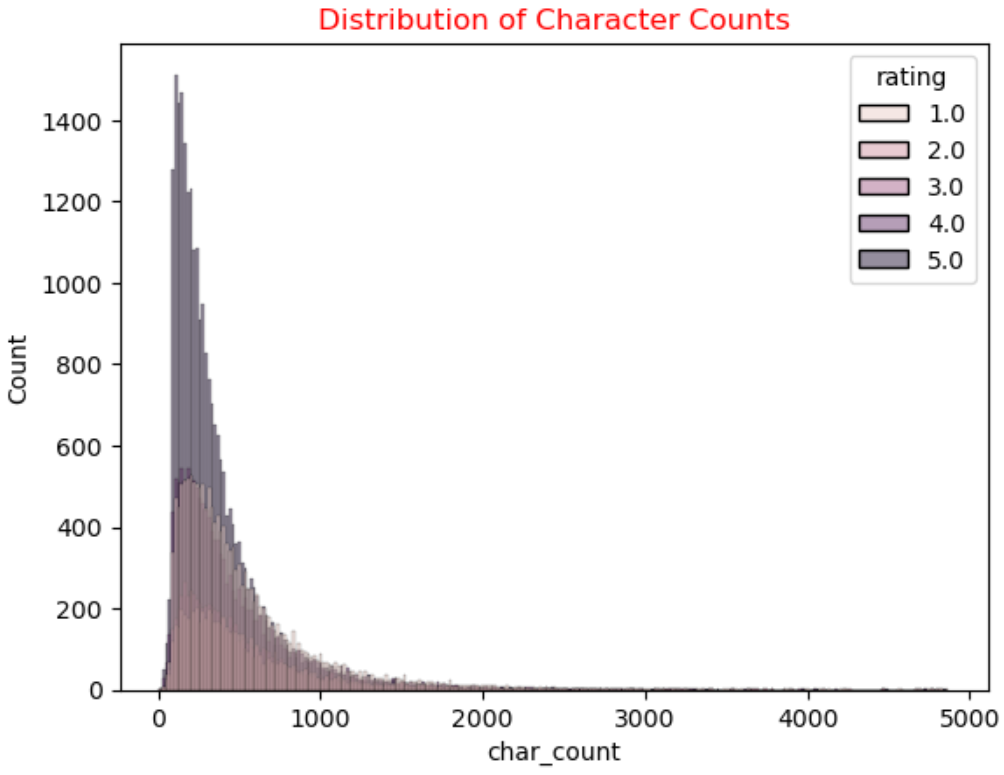
**Figure 15:** *Distribution of Review Length*

The mean length of a review is 86.96 characters. Whilst at the other end of the spectrum, the largest review is 967 characters.

* **Determine counts of review text**

A screenshot of a computer

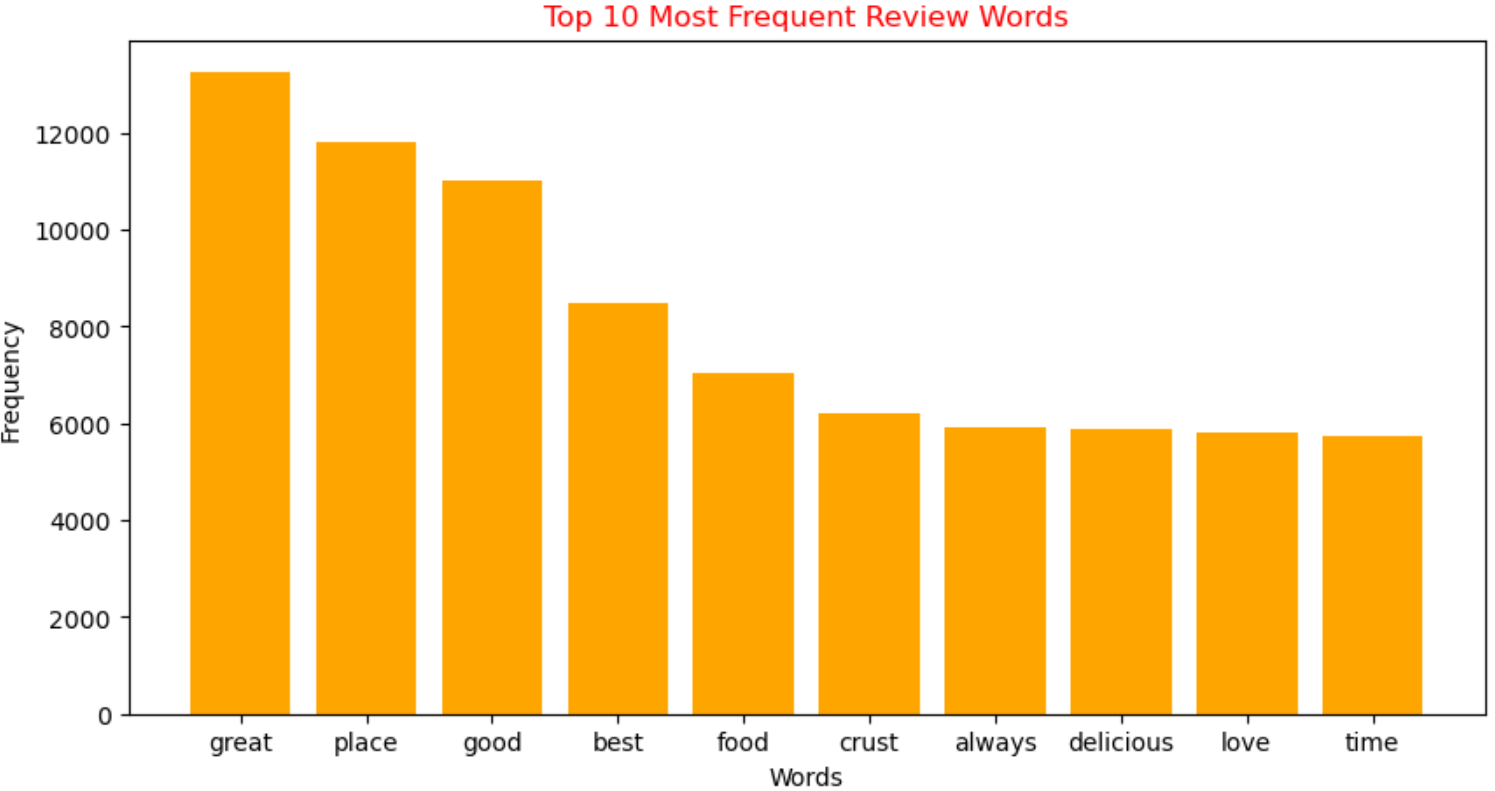
Description automatically generated with low confidence



**Figure 16:** *Histogram of Character Counts*

It is apparent from the histogram above that the higher star ratings tend to have longer reviews.

* **Plot the top 10 most frequent review words**



**Figure 17:** *Top 10 Most Frequent Review Words*

As seen above the top 10 words are associated with good reviews.

* **Plot Bi-Grams**

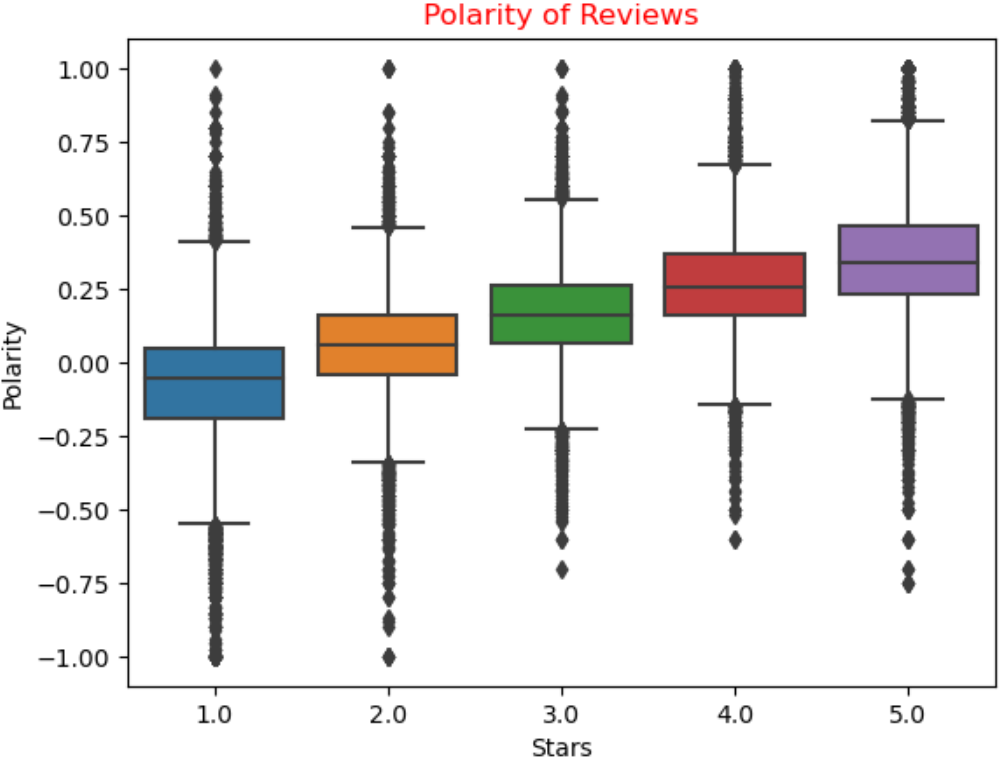
A picture containing text, screenshot, line, font

Description automatically generated

**Figure 18:** *Top 10 Most Frequent Bigrams*

It can be inferred from the Bigrams above that key attributes of a good pizza experience are: location, customer service and a thin crust.

* **Plot Polarity of Reviews**



**Figure 19:** *Polarity of Reviews*

It is evident from the boxplots above that higher star ratings are positively correlated to increasing positive polarity of the users.

* **Perform Topic Modelling using LatentDirichletAllocation (LDA) & Plot**

A picture containing text, screenshot, plot, diagram

Description automatically generated

**Figure 20:** *Topic Distribution*

A black screen with white text

Description automatically generated with low confidence

It is evident from the topics above that place, crust and friendly service are what reviewers are looking for as key criteria for a satisfying pizza experience.

* **Generate Word Clouds**

A picture containing text, font, screenshot, number

Description automatically generated**s**

**Figure 21:** *Positive Word Cloud*

**A picture containing text, font, graphics, screenshot

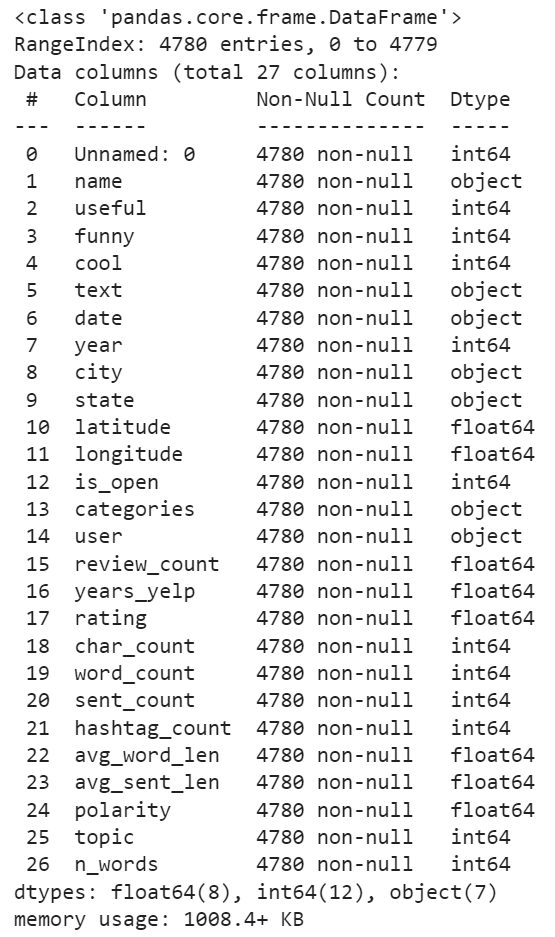
Description automatically generated**

**Figure 22:** *Negative Word Cloud*

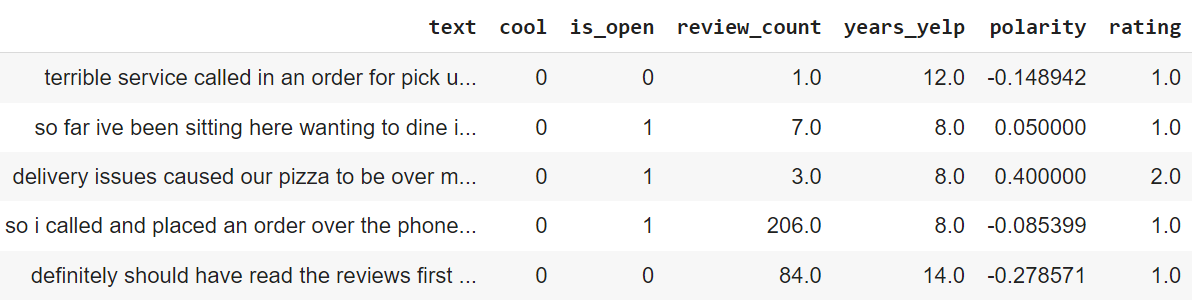
Both positive and negative word clouds have “place” and “time” words. The implication is that the ambiance of a restaurant can either make or break the customer experience, whilst time taken to process or bring the food are also important. Place was also a common thread in both the topics and bigrams suggesting its importance. Cheese is also a key ingredient and has a direct impact on customer satisfaction.

# Modelling

The dataset used for the modelling is based on the pizza\_hut2.csv dataset. This was generated from the reviews\_df. The details of the dataset are as follows:



The features used for the purpose of model other than “text” and “rating” are as follows:



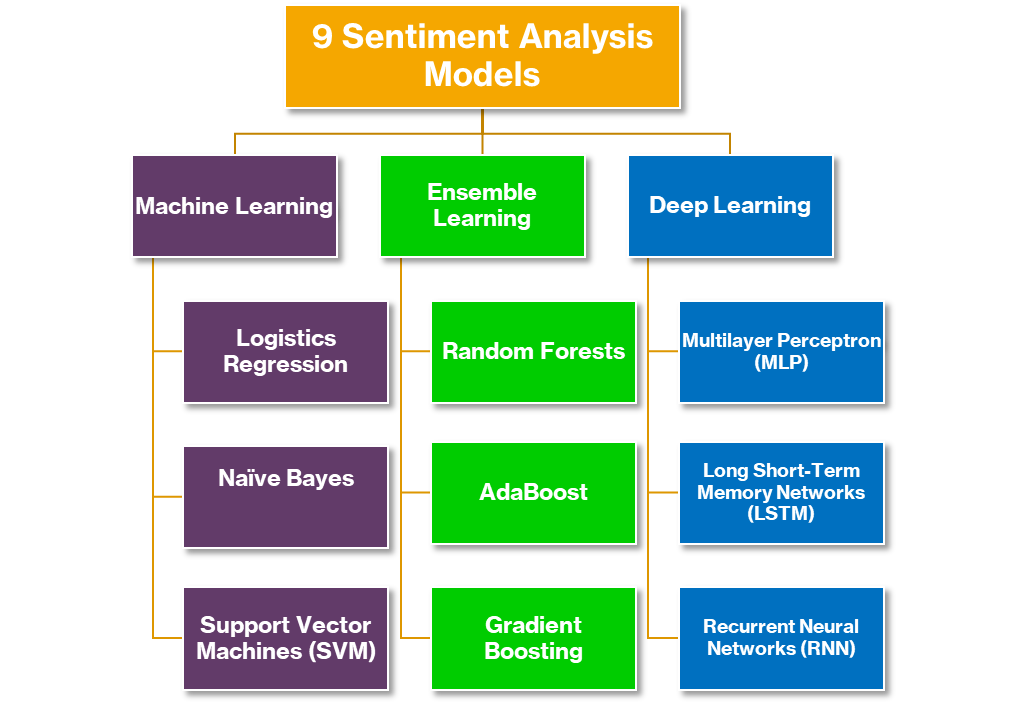
Given the iterative nature of the modelling process, the following steps were undertaken:

**Figure 23:** *Modelling Process*

1. The pizza\_hut2.csv file was ingested. The text and non-text components were defined for the purpose of modelling.
2. SMOTE was applied given the unbalanced nature of the df.rating for Pizza Hut.
3. Pipelines were defined for the purpose of vectorizing df.text using Count and TF-IDF vectorizers and the non-text columns were scaled.
4. Each model was then trained, tested and evaluated using Accuracy as a metric.
5. Hyper-parameter tuning using RandomizedSearchCV was then applied to the machine learning models for the purpose of minimising the likelihood of overfitting and optimising the model accuracy predictions.
6. The Deep Learning models were tuned by selecting the best combination of text and non-text features.

# Modelling Results

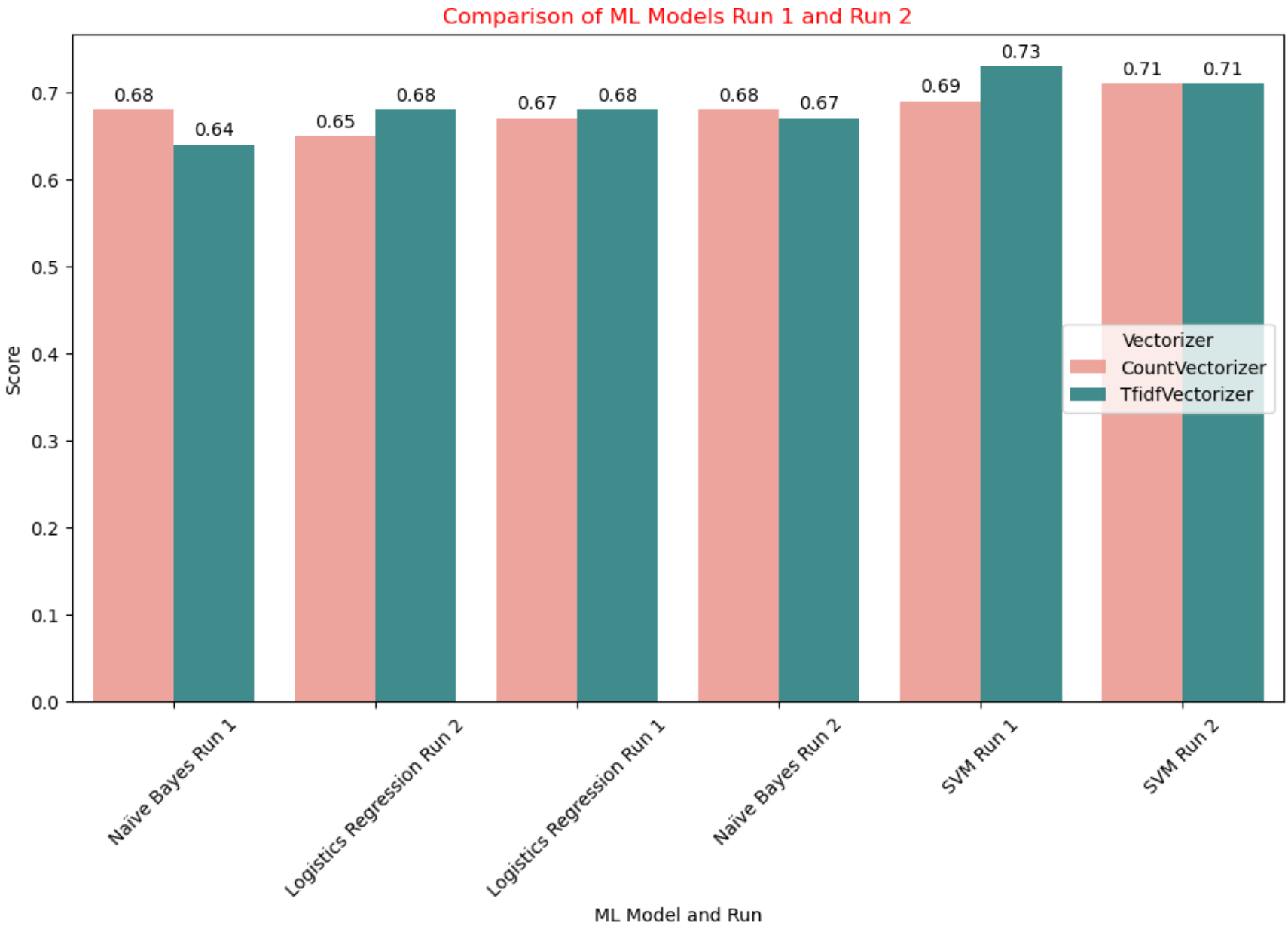
The Sentiment Analysis modelling process covered the following:



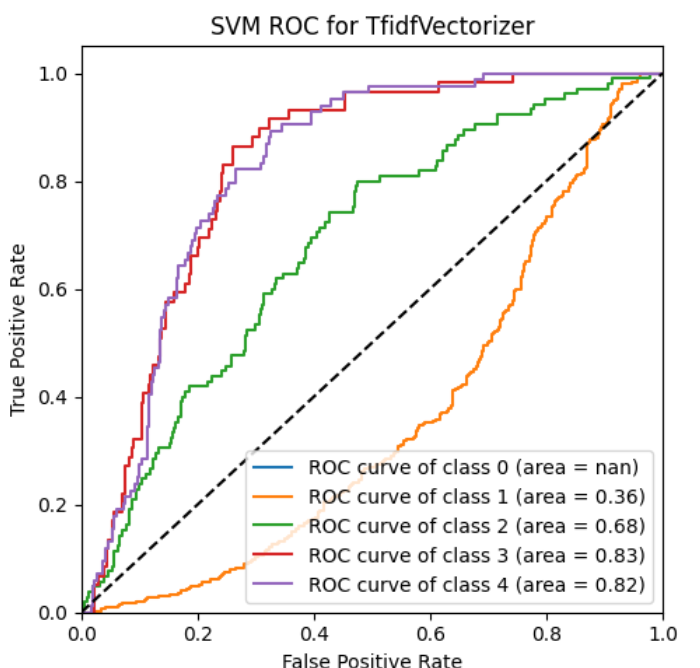
**Figure 24:** *Sentiment Analysis Models*

As show above, three models from Machine Learning, Ensemble Learning and Deep Learning were chosen for modelling purposes. Run 1 represents the pre-tuning whilst Run 2 represents the modelling result post RandomSearchCV. Each model was also processed using Count and TF-IDF Vectorizers and standardised. The results from the various runs are presented below.

## Machine Learning Results



**Figure 25:** *Comparison of ML Models Run 1 and Run 2*

 A picture containing text, screenshot, font, number

Description automatically generated

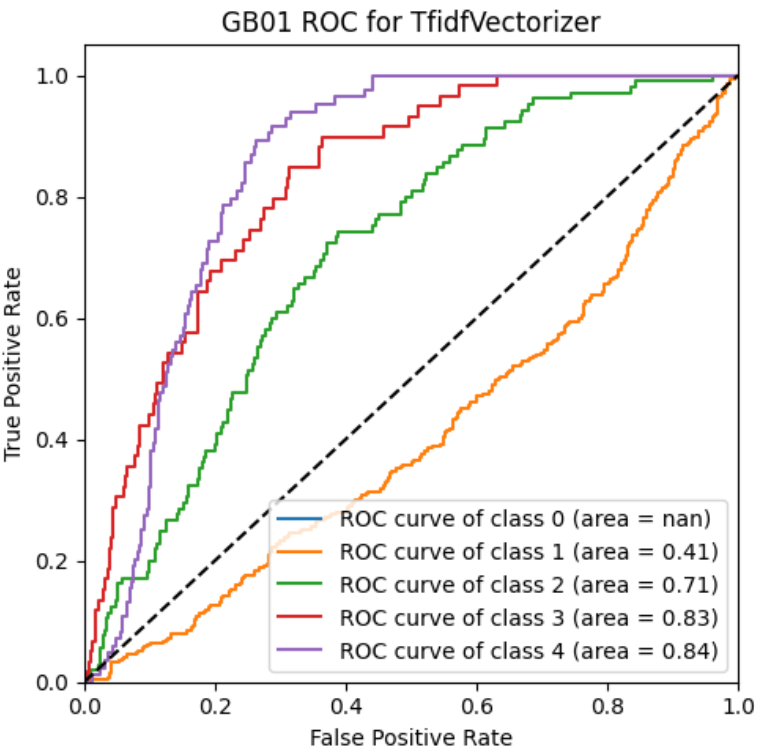
As shown from the results above the best accuracy achieved by was that of the SVM and the TF-IDF Vectorizer on its first run independent of RandomSearchCV. It achieved the highest accuracy of any algorithm with a value of 0.73.

## Ensemble Learning Results

A picture containing text, screenshot, diagram, line

Description automatically generated

**Figure 26:** *Comparison of Ensemble Models Run 1 and Run 2*

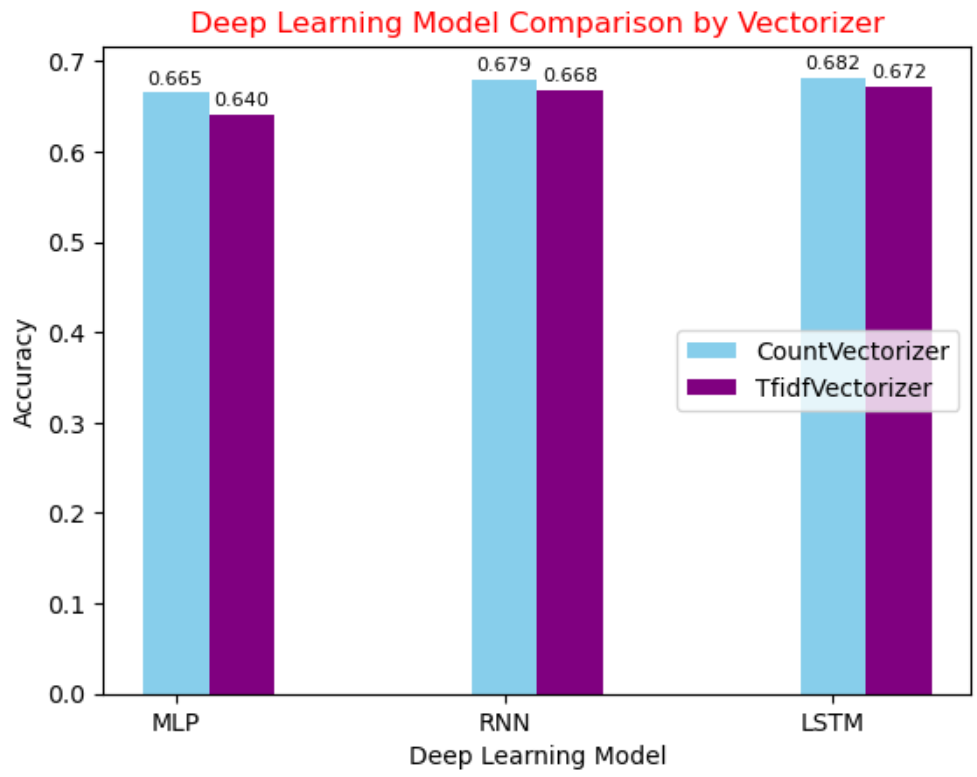


As can be seen from the Ensemble Models, they all had similar levels of performance with exception of Adaboost which was the worst performer. Gradient Boosting post RandomSearchCV + TF-IDF Vecotizer performed the best of the Ensemble models with an accuracy of 0.71.

A picture containing text, screenshot, font, number

Description automatically generated

## Deep Learning Results

****

**Figure 27:** *Comparison of Deep Learning Models*

**A picture containing text, screenshot, font, algebra

Description automatically generated**

**A picture containing text, diagram, plot, screenshot

Description automatically generated** **A picture containing text, line, plot, screenshot

Description automatically generated**

An interesting observation from the Deep Learning Models is that in each instance the best accuracy was achieved in combination with the Count Vectorizer. In addition, none of them could match the best of either the Machine Learning Models or the Ensemble Models. The best Deep Learning Model was that of LSTM in combination with the CountVectorizer. It achieved an accuracy of 0.682. The other interesting point is that accuracy was maximised at less than an Epoch, suggesting that there is insufficient data for the Deep Learning Models to achieve their full potential.

# SVM Sample Predictions

Despite the relatively low peak accuracy of SVM with TF-IDF Vectorizer, the results using the following three test sentences were reasonable:

A screenshot of a computer

Description automatically generated with medium confidence

**Figure 28:** *SVM Sample Predictions*

# Data Answer

The low representation of classes 0 (NAN) and 1 (0.36) in the ROC/AUC plots is indicative of insufficient representation of all 5 classes despite apply SMOTE to counter the unbalanced nature of the Target variable. Whilst an accuracy of 0.73 is reasonable, it can be improved through training the models on more data volumes, particularly for the Deep Learning models.

# Business Answer

The EDA process has partially answered the questions as to where the problem Cities and States from the perspective of low reviews. The other questions as to why the decline in revenue has occurred and what can be done going forward will be answered through the implementation of a SVM Sentiment Analysis model. This model will provide Pizza Hut with the capability to quickly gauge unfavourable and favourable sentiment allowing timely action to be taken. Consequently problems be resolved sooner with the net improvement in future ratings as an outcome.

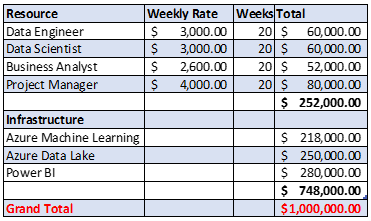
# Response to Stakeholders

Implementation of the SVM model will realise the following benefits:

A close-up of a graph

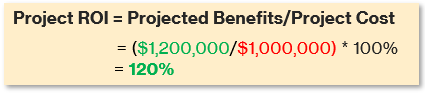
Description automatically generated with low confidence

The estimated project cost is:



**Figure 29:** *Estimated Project Cost*

The projected Return on Investment over the 2 years is:



The indicative project implementation timeline is:



Given the financial and non-financial benefits the project will bring to Pizza Hut, it is recommended that the project is approved for implementation.

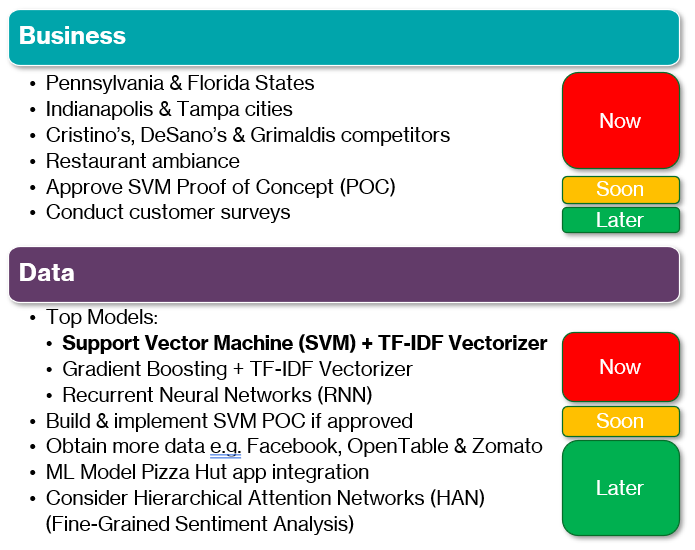
# Conclusions and Recommendations

The Modelling process has revealed that there is not much difference between the traditional Machine Learning Models, Ensemble Models and Deep Models for the purpose of Sentiment Analysis using NLP. Results marginally improved for Machine and Ensemble Learning when RandomSearchCV was applied.

Notwithstanding the closeness of the results, it is recommended that SVM in conjunction with the TF-IDF Vectorizer will provide the highest level of accuracy. Gradient Boosting was not far behind however it does require more memory and processing power than SVM.

Another take-away is that more data is required given that there were insufficient samples for a ROC/Area Under the Curve for class 0.

The following is a representation of the business and data tasks, along with an anticipated timeframe that are required to move the project forward:



**Figure 30:** *Summary of Business & Data next steps*

# References

* Elizabeth, E 2020, *11 Best Restaurant Review Sites for 2022 - Restaurant Clicks*, restaurantclicks.com, viewed 9 May 2023, <https://restaurantclicks.com/restaurant-review-sites/>.
* Ibisworld 2023, *IBISWorld - Industry Market Research, Reports, and Statistics*, www.ibisworld.com, viewed 6 May 2023, <https://www.ibisworld.com/united-states/market-research-reports/pizza-restaurant-franchises-industry/>.
* Lake, A 2022, *4 Major Fast-Food Chains With Declining Sales This Year — Eat This Not That*, *eatthis.com*, viewed 21 November 2022, <https://www.eatthis.com/news-major-restaurant-chains-declining-sales-2022/>.
* Tan, KL, Lee, CP & Lim, KM 2023, ‘A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research’, *Applied Sciences*, vol. 13, no. 7, p. 4550.
* yelp 2019, *Yelp Dataset*, Yelp.com, viewed 6 May 2023, <https://www.yelp.com/dataset>.
* Yum! 2023, *Financial Reports - Yum! Brands*, investors.yum.com, viewed 12 May 2023, <https://investors.yum.com/financial-information/financial-reports/>.