Title: Using Consumer Sentiment to Enhance Service Quality in Fast Food Restaurants

Student Name:

Instructor Name:

Contents

1. Background/Introduction

The fast-food sector is very competitive, and the quality of service delivered to clients may be a critical distinction for firms (Hanaysha & Pech, 2018). In recent years, internet reviews have become a vital source of feedback for companies, since they give insights into consumer experiences and opinions. Yet, assessing vast volumes of client data may be complex and time-consuming. This research intends to examine the potential of sentiment analysis as a tool for enhancing service quality in fast food restaurants, with McDonald's reviews acting as a case study.

Customers today are much more likely to write reviews on sites like Yelp, Google Reviews, and social media, which has led to a rise in the significance of these types of online feedback (Rout et al., 2016). Customer feedback is a quick and easy way to assess interest in a company's goods and services. Success in the fast-food industry often depends on happy customers.

Sentiment analysis is a method for assessing how people feel and what they think based on what they review in texts. In this method, machine learning algorithms are used to determine whether a review from customer is positive, negative, or neutral (Kumar et al., 2021). Sentiment analysis can be used to extract interesting insights from fast food restaurants' online reviews and can assist the food industry to identify potential problem in their goods and services.

Analyzing customer evaluations of fast-food chains using sentiment analysis revealed that companies that responded to negative reviews had a greater impact on customer satisfaction than those that did not (Chen et al., 2020). This suggests that sentiment analysis not only assists businesses in identifying consumer issues but also affords them the opportunity to engage with customers and address their concerns.

2. Objectives/ Goals of the Project

The primary goal of this project is to use sentiment analysis to improve service quality in fast food restaurants, with McDonald's reviews serving as a case study. The findings of this study have the potential to help not just McDonald's, but also other fast-food companies looking to improve their service quality and customer happiness. The following are specific objectives for this project:

i. Using the McDonald's reviews dataset, create a sentiment analysis model able to correctly classify customer ratings as positive, negative, or neutral.

ii. To assess the sentiment of the McDonald's reviews dataset in order to identify the most common problems and areas for improvement, as well as to give insights into customer preferences and expectations.

iii. Based on the sentiment analysis results, provide suggestions for improving service quality at fast food outlets that fit to the particular needs and preferences of McDonald's customers.

3. Data Collection

The McDonald's reviews dataset is available on www.https://data.world.com website comprises reviews provided by customers of McDonald's restaurants located in a variety of cities. Each review is accompanied with the name of the city where the restaurant is situated, a list of the regulations that were broken, and a confidence rating in each of the policies that were violated. The reviews are written and take the form of text. They contain information regarding the customer's experience, such as the quality of the foods, the speed with which they were served, and the conduct of the employees. The reviews, numbering over a thousand, that were compiled for the dataset were gathered in February of 2015.

4. Data Visualization

Data visualization is crucial to the success of this project because it allows us to quickly and easily see patterns, trends, and insights within the data (Raschka, 2015). Customer sentiment may be visualized in a variety of ways, including bar charts, word clouds, and heat maps, which help to identify the interesting insights.

Figure 1: Bar chart showing the distribution of the sentiment

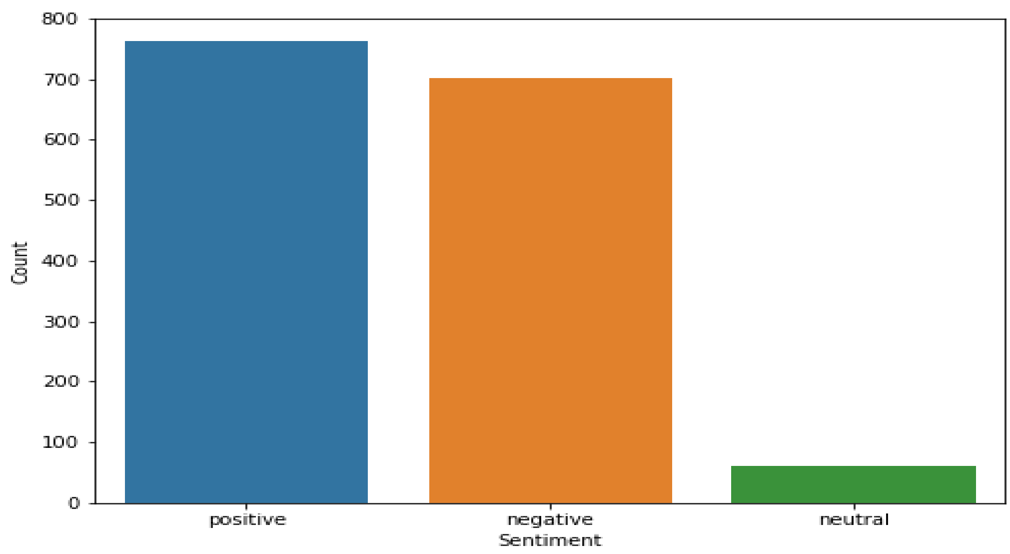

Figure 1 shows that there are more positive reviews than negative reviews, but both positive and negative reviews are fairly close in count. The number of neutral reviews is significantly lower than both positive and negative reviews.

Our analysis indicates that the text includes more positive than negative sentiment, giving it a sentiment score of 60. However, the mood is not generally positive, and there may even be some still negativity. It's important to remember that context, tone, and the precision of the analysis model may all affect the results of a sentiment analysis (Aljameel et al., 2020).

Figure 2: Stacked bar chart showing the distribution of the sentiment by city

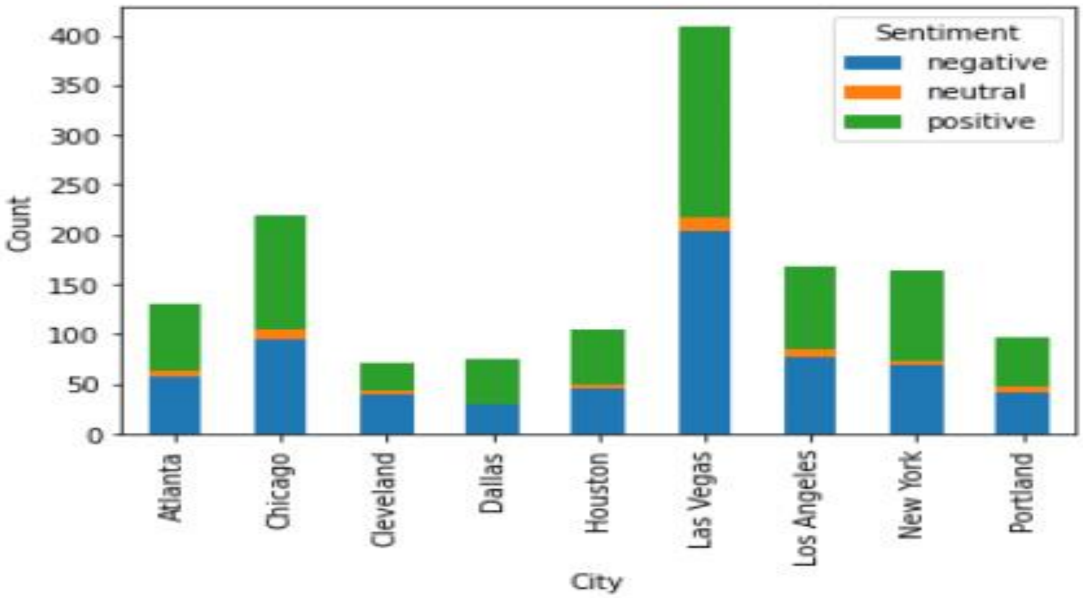

Figure 2 shows that the city with the highest positive reviews is Las Vegas. The city with the highest negative reviews is Chicago. The cities with negative reviews from highest to lowest are: Chicago, Los Angeles, Atlanta, New York, Houston, Portland, Cleveland, Dallas.

Figure 3: The word cloud for negative sentiment

Figure 3 indicates that the most often reported concerns about McDonald's are issues with waiting, meal quality, and staff. The words like "time," "want," "right," and "window" all associated with long wait times, incorrect orders, and issues with the drive-thru window. The words like "food" and "coffee" imply that customers were dissatisfied with the items or had a bad opinion of their quality. The words like "service," "order," and "place" all suggest that McDonald's consumers have had issues with the restaurant's service, maybe even with ordering or location. The words "location," "work," "people," "drive," and "day" all indicate that customers may have issues with the restaurant's real location, its personnel, or the overall experience on a given day.

5. Data Analysis

I will use the frequency-inverse document frequency (tf-idf) approach to identify the most essential terms from customer evaluations in the data analysis section of this project (Mohd Nafis & Awang, 2021). Preprocessing the text data comprises converting all text to lowercase, tokenizing it, removing stop words, and stemming the words using the Porter stemmer method.

The TfidfVectorizer from the scikit-learn module will then be used to generate the tf-idf scores for each word in the preprocessed text (Hardeniya et al., 2016).

I will be developing logistic regression, random forest, support vector machine, and decision tree models to predict the sentiment of customer reviews in addition to tf-idf (Shafin et al., 2020). Each review's sentiment will be represented by a numerical score obtained from the sentiment labels in the dataset (-1 for negative, 0 for neutral, and 1 for positive). I will use the CountVectorizer to turn the text data into a numerical representation after preprocessing the text data and partitioning the dataset into training and testing sets. Then, using the training set, I will train a logistic regression model and evaluate its performance on the testing set by generating the accuracy score (Kasongo & Sun, 2020).

Table 1: The frequency of negative words

mcdonald	72.764727
order	68.956075
food	55.452807
drive	51.050202
time	48.005738
servic	45.550878
place	40.461686
locat	36.962073
wait	33.097516
fri	29.859012
good	29.260674
ï¾i	28.828419
ask	27.635806
peopl	27.242520
custom	26.216977
manag	26.216532
coffe	25.592168
minut	25.456518
work	25.445766
worst	25.349155

Table 1 shows that the name “Mcdonald” indicate that consumers are drawing connections between their poor experiences and the McDonald's brand as a whole. There may be problems with the ordering process or the quality of the meal being served if "order" and "food" are the second and third most often occurring terms, respectively.

Words like "drive" and "time" also seem to crop up often in customer complaints, pointing to problems with the drive-thru and long wait periods. Customers may be having issues with the service they get from employees and the convenience of the fast-food restaurants' locations, since "service" and "place" are both often used unfavorable words.

Table 2: The performance of the models

Model	**Accuracy**
Random Forest	74.10%
Logistic Regression	69.84%
Support Vector Machine	69.18%
Decision Tree	64.59%

Table 2 shows that Random Forest has the highest accuracy score of 0.7410. This implies that the model correctly classified 74.10% of the reviews as favorable, negative, or neutral. Logistic Regression received the second highest accuracy score of 69.84%, followed by SVM with an accuracy of 69.18% and Decision Tree with an accuracy of 64.59%. This means that Random Forest is the best method out of the four for the McDonald's reviews dataset. The accuracy of these models is not outstanding, and I will indicate what should be done to enhance them at the conclusion of the study (Song et al., 2022).

6. Conclusion

In conclusion, the insights I obtained from analyzing McDonald's reviews will help the company to better serve its customers and boost their satisfaction (Mcdonald et al., 2015). The positive reviews are more than negative, although there are still some problems. Positive reviews are more common for Las Vegas, while negative reviews are more common for Chicago. Customers have lots of problems with the wait time, the food, and the service.

McDonald's has a few options for tackling these issues, including tightening up on order accuracy, boosting meal quality, shortening drive-thru lines, upgrading customer service, concentrating on store management, lowering fry wait times, and strengthening management processes. These methods may assist the business boost client retention and loyalty. According to my findings, the Random Forest model performs best when categorizing reviews, with an accuracy rate of 74.10%. All the models aren't very accurate, however, so there's opportunity for improvement.

7. Recommendation

My analysis has shown actionable recommendations that McDonald's could apply to elevate the quality of its services and satisfy its customers.

The company may boost customer satisfaction and increase profits by concentrating on order precision, food quality, service, and location management. McDonald's could employ the suggestions that it receives from consumers to inform data-driven choices that will improve the firm and its products for everyone involved.

Using a bigger dataset, enhancing the quality of the data, and fine-tuning the models are all things that may be done to improve the models' accuracy and, their performance.

8. References

Aljameel, S. S., Alabbad, D. A., Alzahrani, N. A., Alqarni, S. M., Alamoudi, F. A., Babili, L. M., Aljaafary, S. K., & Alshamrani, F. M. (2020). A Sentiment Analysis Approach to Predict an Individual's Awareness of the Precautionary Procedures to Prevent COVID-19 Outbreaks in Saudi Arabia. *International Journal of Environmental Research and Public Health*, *18*(1), 218. https://doi.org/10.3390/ijerph18010218

Chen, W.-K., Riantama, D., & Chen, L.-S. (2020). Using a Text Mining Approach to Hear Voices of Customers from Social Media toward the Fast-Food Restaurant Industry. *Sustainability*, *13*(1), 268. https://doi.org/10.3390/su13010268

Hanaysha, J. R., & Pech, R. J. (2018). Brand Prestige and the Mediating Role of Word of Mouth in the Fast Food Industry. *Global Business Review*, *19*(6), 1494–1514. https://doi.org/10.1177/0972150918794736

Hardeniya, N., Perkins, J., Chopra, D., Joshi, N., & Mathur, I. (2016). Natural Language Processing: Python and NLTK. In *Google Books*. Packt Publishing Ltd. https://books.google.co.ke/books?hl=en&lr=&id=0J_cDgAAQBAJ&oi=fnd&pg=PP1&dq=tokenizing+it

Kasongo, S. M., & Sun, Y. (2020). Performance Analysis of Intrusion Detection Systems Using a Feature Selection Method on the UNSW-NB15 Dataset. *Journal of Big Data*, *7*(1). https://doi.org/10.1186/s40537-020-00379-6

Kumar, V. V., Raghunath, K. M. K., Muthukumaran, V., Joseph, R. B., Beschi, I. S., & Uday, A. K. (2021). Aspect based sentiment analysis and smart classification in uncertain feedback pool. *International Journal of System Assurance Engineering and Management*, *13*(S1), 252–262. https://doi.org/10.1007/s13198-021-01379-2

Mcdonald, G., Jackson, D., Vickers, M. H., & Wilkes, L. (2015). Surviving workplace adversity: a qualitative study of nurses and midwives and their strategies to increase personal resilience. *Journal of Nursing Management*, *24*(1), 123–131. https://doi.org/10.1111/jonm.12293

Mohd Nafis, N. S., & Awang, S. (2021). An Enhanced Hybrid Feature Selection Technique Using Term Frequency-Inverse Document Frequency and Support Vector Machine-Recursive Feature Elimination for Sentiment Classification. *IEEE Access*, *9*(302), 52177–52192. https://doi.org/10.1109/access.2021.3069001

Raschka, S. (2015). Python Machine Learning. In *Google Books*. Packt Publishing Ltd. https://books.google.co.ke/books?hl=en&lr=&id=GOVOCwAAQBAJ&oi=fnd&pg=PP1&dq=Data+visualization+is+crucial+to+the+success+of+this+project+because+it+allows+us+to+quickly+and+easily+see+patterns

Rout, J. K., Singh, S., Jena, S. K., & Bakshi, S. (2016). Deceptive review detection using labeled and unlabeled data. *Multimedia Tools and Applications*, *76*(3), 3187–3211. https://doi.org/10.1007/s11042-016-3819-y

Shafin, M. A., Hasan, Md. M., Alam, Md. R., Mithu, M. A., Nur, A. U., & Faruk, Md. O. (2020, December 1). *Product Review Sentiment Analysis by Using NLP and Machine Learning in Bangla Language*. IEEE Xplore. https://doi.org/10.1109/ICCIT51783.2020.9392733

Song, J., Hobensack, M., Bowles, K. H., McDonald, M. V., Cato, K., Rossetti, S. C., Chae, S., Kennedy, E., Barrón, Y., Sridharan, S., & Topaz, M. (2022). Clinical notes: An untapped opportunity for improving risk prediction for hospitalization and emergency department visit during home health care. *Journal of Biomedical Informatics*, *128*(23), 104039. https://doi.org/10.1016/j.jbi.2022.104039