

## The Impact of Mixue's Halal Announcement on Company's Brand Reputation: a Naïve Bayes Sentiment Analysis Approach

Gilbert Tandiawan<sup>1\*</sup>, Meditya Wasesa<sup>2</sup>

School of Business and Management, Institut Teknologi Bandung, Bandung, Indonesia<sup>1</sup>

Email: gilbert\_tandiawan@sbm-itb.ac.id

### Abstract

With the rise of beverage and ice cream market entry through franchising, Halal certification is more concerned than ever, especially in countries where Muslims, like Indonesia, make up a huge percentage of the population. The hindrance of business operation could potentially occur if customers felt dissatisfied and unwelcome toward the market entry that had not had Halal certification, ultimately affecting the brand reputation based on customer satisfaction above all. Hence, this research aims to determine whether Halal Announcement can turn a brand into a better brand reputation by analyzing comments retrieved through Instagram and evaluating the prediction model of Naïve Bayes using TF-IDF machine learning classification. The data for this research was separated into two posts, which have 3861 and 3128 unfiltered comments for occurrences before and after Halal Announcement, respectively. These comments are then processed to validate that the company rebounded the positive sentiments after the Halal Announcement, in which the hypothesis was accepted. Furthermore, the developed Naïve Bayes model is evaluated on the After Halal Announcement dataset, achieving an average accuracy of 65.99%, showing that the model can predict the sentiment fairly with several key takeaways noted.

**Keywords:** brand reputation; halal announcement; Naïve Bayes; sentiment analysis.

### A. INTRODUCTION

The beverage industry is on the verge of success, where part of lifestyle and well-being are growing to rely on drinks and refreshments. Many entrepreneurs are racing to build a prominent beverage business that can accommodate their market with what they need and value from their business. According to Statista 2021, Indonesia's beverage industry's prospect is expected to grow 13.54% in revenues from 2023 to 2027. No wonder beverage consumption, especially Sugar-Sweetened Beverages, has shown an increasing trend in developing countries, especially Indonesia.

As beverage industries in Indonesia continue to be progressively competitive, firms should be able to identify the strategy to capture what makes them "stand out" compared to other competitors. The build-up in raising competitiveness stake has become a buzzword for globalization (Bhawsar & Chattophaday, 2015), in which this action should also be considered by incorporating aspects of environmental uncertainty, mimetic behavior, and experiential learning (Townsend et al., 2009). Several firms pull the trigger in global expansion, especially in building brand credibility and reaching out to a new market yet to be uncovered. The expansion strategy can be either hit or miss, depending on how well the business can provide the service or products and how customers perceive the brand.

Mixue, a thriving franchised beverage industry established in 1997 in Zhengzhou, supported under Mixue Bingcheng Co., Ltd., has overcome the odds in serving high-quality and affordable products worldwide, which conveys its strategy to facilitate global market expansion, aligned with its mission. As one of the impactful expansions takes place in Indonesia, the business entered the market with quite a controversy concerning whether the product has been registered under BPOM and labeled to have used Halal products. With over 3 years of operations in Indonesia, Mixue has had over 1000 outlets, with the company providing anyone with authority to open a franchise at an affordable final expense that needs to

\* Corresponding author

be borne, yet 380 million IDR, under PT Zisheng Pacific Trading (CNBC Indonesia, 2022) while the opportunity of franchising is open wide, the development of franchising requires learning and relearning over the lifetime of the franchiser and not just a one-time event (Lindquist, 1996). Thus, franchising the brand requires external factors, such as people's sentiments and desire for the brand, as it is risky to conduct the ethics without any other consideration.

Mixue's rapid expansion has sought people's interest with a high trend in the beverage industry and low franchising fees. As Indonesia is a developing country, populations are escalating with a high youth dependency rate of people below 15 (Atanda et al., 2012). Nonetheless, With Mixue's core strategy relies upon quantitative sales and penetration pricing, the firm can capture the presence of the global market by featuring strategic location, and full brand ownership is handed to the respective branch Given the supporting elements such as the cute mascot and catchy tracks, Mixue can create hype and generate sentiments through any communication and information media. Regardless of the success in gaining people's attention, there was still plenty of criticism concerning the Halal ingredients as the Halal symbol was not posted in each branch of Mixue's store, especially with the franchise expansion that keeps increasing over time.

The criticism of Halal concerns and Mixue's continuous expansion influences the gap in customer satisfaction and discomfort, especially with the certification not being released in 2022. With a business that can run for over 2 years and keep occupying new stores, these circumstances may generate people's worries in considering opening a franchise or purchasing the brand's products. Therefore, the researcher designs a sentiment analysis, which is a natural language processing of classifying subjective opinions (Luo et al., 2013), to analyze the gap and address the issue before and after the Halal Announcement of the company based on internet user's comments, yet utilizing a machine learning model of Naïve Bayes that may simplify the efforts for feedback filters on the brand.

## **B. RESEARCH METHOD**

The approach used in this study will highlight both quantitative and qualitative methods, in which comments and judgments generated will be adapted based on modeling criteria to perform the sentiment analysis. The researcher will filter irrelevant comments and spam before processing the data extracted from a specific post. Spam and Social bots have always tended to corrupt or pervert information surrounding social media as they may generate as much engagement or higher than humans (Bhowmick & Hazarika, 2016). Thus, consider the bias upon every sentiment provided using Natural Language Processing (NLP) outlook to perform sentiment analysis that heavily relies on data extraction.

On the quantitative side, the author designs a statistical framework to compare two Halal coincidences based on the research objectives. The framework will validate whether the hypothesis accepted that has been supported by several online sources and platforms on a qualitative side for an argument that halal announcement positively influences the company's brand reputation.

With the research flow starting from an exploratory approach, the author accumulates information to generate the hypothesis from a concrete literature review. Then, the problem arose when the author took two distinctive dates of the company's Instagram post, indicating before and after Halal Announcement. After vigorous data cleaning and modeling, the author analyses the gap of sentiment differences using a statistical approach, such as ANOVA, Box and Whisker Plot, TreeMap, and Net Reputation Score, to strengthen the hypothesis outcome as part of the evidence. The correlation between sentiment analysis and brand reputation was addressed by (Vidya et al., 2015), using a net reputation score from sentiment analysis based on customer satisfaction. Then, (Sudira et al., 2019) explored sentiment analysis to measure customer satisfaction with mobile providers. However, The data retrieved is selected in the time range of July 2022

to May 2023, as the Halal Announcement was posted on February 2023, where the author also evaluates the model of Naïve Bayes by taking before the Halal Announcement as part of the train data to implement and check upon the accuracy along with all posts that have been manually labeled.

### C. RESULTS AND ANALYSIS

#### Tree Map

From the frames below, it is intriguing that Mixue's Before Halal sample post was posted in July 2022, which is still supported by positive responses overall. Nevertheless, Mixue's Halal conversation started in 2023 with not-so-favorable sentiments. In January 2023, it was realized that the comments were generated more than in any other month without concerning any sentiment labels. It was presumably that people were waiting for clarity on Mixue's Halal certification, as Mixue might have pinned the posts with all the hype going on. Therefore, on 17th February, as Mixue announced that they finally received a Halal certification on Instagram, the sentiment shifted into positive with a huge margin compared to other months. Details are shown in Figures 1 and 2, which will also compare sentiment labeled quantity, and the comment traffics referring to each sentiment.

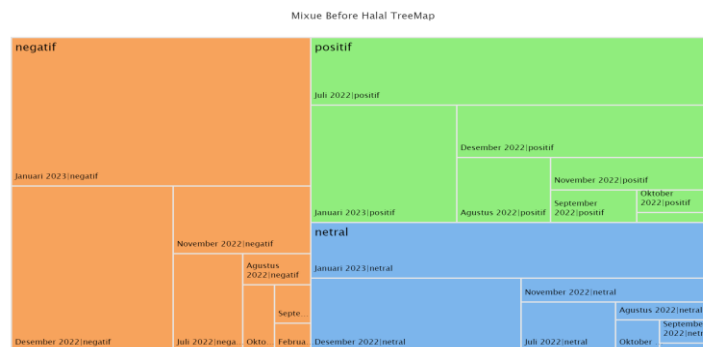


Figure 1. Tree Map Before Halal Announcement

Source: Research data, 2023

In Figure 2, The visualization of TreeMap will be based on the manually labeled data on positive, neutral, and negative, indicating the results are green, blue, and red, respectively, separated by quantities of sentiments from February 2023 to May 2023 as since there are slight differences between machine learning methodology and manual labeling in terms of its accuracy.



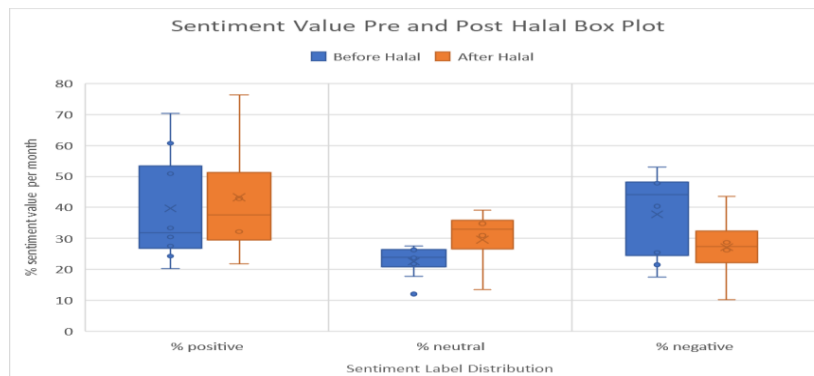
Figure 2. Tree Map After Halal Announcement (manual label)

Source: Research data, 2023

#### Box and Whisker Plot

The comparison between before and after Halal conditions is also shown through the box plot, shown in Figure 3. The distribution of the values is weighted from the percentage of sentiment per month,

which is scattered across the box and whisker. Furthermore, it is seen that positive sentiment after halal condition reaches 76.37% and is far away from the mean of 43.2875% in the x symbol as there are only 4 months counted yet from February to May.



**Figure 3: Sentiment Value Box and Whisker Plot**  
 Source: Research data, 2023

Meanwhile, as of neutral sentiment, the condition before halal seems to stay around 20%, with the highest and lowest being not very far away compared to after halal condition with one month far away from the data variance. However, we see more negative sentiments before halal conditions compared to after halal circumstances.

#### ANOVA

To strengthen the analysis, ANOVA compares the data variances between and within groups of positive sentiments. In this case, shown in Appendices A: One-Way Anova Steps, the researcher classified the analysis into 3 groups of month periods. The decision to group months is made to identify whether the mean when the train data was released differed from the data retrieved in 2023. As seen in Table 1, the separation between groups is referred to as follows: 1) July – Sept: Positive Sentiments are labeled more than Negative Sentiments; 2) Oct – Feb: There is a shift/campaign in which Negative Sentiments are labeled more than Positive Sentiment in the same post on July – September; and 3) Feb – April: The Halal Announcement Post turns the sentiment event around.

**Table 1. ANOVA Test Result**

SUMMARY					
Groups	Count	Sum	Average	Variance	
July – Sep	54	406	7.51851	618.103	
Oct - Feb (before Halal)	82	393	4.79268	39.4256	
Feb - May (After Halal)	58	1742	30.0344	24853.6	

ANOVA						
Source of Variation	SS	Df	MS	F	P-value	F crit
Between Groups	23974.25	2	11987.1	1.57615	0.20944	3.0432139
Within Groups	1452612	191	7605.30			
Total	1476587	193				

Source: Research data, 2023

The Halal Announcement in February generated the most positive sentiments, even with the previous data combined, affecting a high variance in the following months after February. Eventually, since the Fstat is lower than the F critical value, the null hypothesis failed to be rejected, meaning that there is still no significant evidence that the mean between groups is not the same, with 20% of a null hypothesis considered correct.

### Net Reputation Score

In order to measure the brand reputation from sentiment analysis, Net Reputation is used to track the superior sentiments between positive and negative, as neutral will not be counted as part of the calculations. The outcome of the net reputation analysis is presented below.

**Table 2. Net Reputation Score**

Before Halal Announcement Post	After Halal Announcement Post
-11,71%	70,11%

Source: Research data, 2023

As seen in Table 2, there is a massive rise in net reputation score before Halal and After the Halal Announcement, indicating the post impacts toward more on positive sentiment rather than negative, shown from a shift of negative to positive.

### Model Evaluation

To recognize whether the Naïve Bayes model is enough to predict the sentiment through machine learning, the sentiment performance analysis is proposed through the confusion matrix, shown in Table 3. The designed confusion matrix will determine whether the predicted classification belongs to its actual event. As the matrix is a 3x3, the outcome is split into 6 categories of True Positive (TP), True Neutral (TNeu), and True Negative (TNeg). False Positive (FP), False Neutral (FNeu), and False Negative (FNeg).

Known that Naïve Bayes predicts the text to be labeled as 1595 positive, 403 neutrals, and 440 negative sentiments, it is obtained from the confusion matrix in Table 3 that the true positive, neutral, and negative are 1333, 131, and 145, respectively. Given the true values, the Naïve Bayes also identifies the text classification incorrectly in the false values cell matrix.

**Table 3. Naïve Bayes Confusion Matrix Result**

Predicted Class	Actual Event		
	Positive Class	Neutral Class	Negative Class
Positive Class	1333	166	96
Neutral Class	207	131	65
Negative Class	202	93	145

Source: Research data, 2023

Based on Table 3, Naïve Bayes classifies the positive class better than the other two. As the predictions for neutral and negative classes are not as accurate as positive classes, the Naïve Bayes model interprets the event in favor of positive classification rather than the actual event itself. Therefore, to analyze Naïve Bayes' performance, Table 4 and Table 5 were designed to evaluate the accuracy, precision, recall as well as the F1 score of the model.

**Table 4. Naïve Bayes Classifier's Accuracy**

Indicator	Accuracy
Result	65.99%

Source: Research data, 2023

**Table 5. Naïve Bayes Model Evaluation**

Class	Precision	Recall	F1 score
Positive Class	0.8357	0.7652	0.7989
Neutral Class	0.3251	0.3359	0.3304
Negative Class	0.3295	0.4738	0.3887

Source: Research data, 2023

The overall accuracy of the Naïve Bayes model is 65,99%, so the classification is further evaluated with various precision, recall, and F1 scores. Knowing that the F1 score is aimed to determine the model performance while the values are still categorized based on the class, the calculation to receive a weighted F1 score is necessary and provided in Table 6. along with the micro and macro F1 Average.

**Table 6. Naïve Bayes Classifier Performance**

	Micro F1 Average	Macro F1 Average	Weighted F1 Score
Result	0.6599	0.506	0.6724

Source: Research data, 2023

The classification performance includes micro and macro F1 averages to notice whether the dataset reflects the importance of dataset samples. As seen in Table 6, the micro f1 average is equal to the accuracy, while it differs from the macro f1 in average. Thus, the weighted F1 score is used to represent the performance of Naïve Bayes in an imbalanced data set of sentiment labels, highlighting the positive sentiments after the Halal announcement as proposed.

#### *Discussion*

Received from the weighted F1 score is 0.6724, the value indicates that the Naïve Bayes' prediction is a fair classification. Nevertheless, the model supports more in favor of positive sentiments rather than the other sentiment label, as many models can be entrusted and used for Mixue's available dataset. On the other hand, Mixue's Halal Announcement hypothesis positively influenced Brand Reputation and was accepted as several statistical approaches have been conducted.

#### **D. CONCLUSION**

According to this study's findings, Halal Announcement impacts sentiment upturn as it affects brand reputation per study literature. This research highlights the data taken from before Halal Announcement and after the post, as which positive sentiments are superior after the announcement, indicating that audiences are more satisfied with the current circumstances. Both analyzed posts were manually labeled to validate the hypothesis as this research also applies the model of Naïve Bayes and evaluates it based on the manually labeled data provided. Hence, with the model designed by Naïve Bayes having a flaw to consider, it is best to explore other machine learning tools that may benefit most on sentiment prediction in this case.

Moreover, the study suggests collecting more consistent data with fewer irrelevant comments on the post yet using more posts and variables within the comments. Future research is also expected to use various supporting prediction models to research the most effective analysis method. Consequently, implementing other classifying machine learning tools can be improved to predict the implication of sentiment analysis with Halal Announcement as its trigger.

#### **REFERENCES**

##### **Journal Article:**

- Atanda, Akinwande A. & Aminu, Salaudeen B. & Alimi, Olorunfemi Y., 2012. "The role of population on economic growth and development: evidence from developing countries," MPRA Paper 37966, University Library of Munich, Germany.
- Bhawsar, P., & Chattopadhyay, U. (2015). Competitiveness: Review, Reflections and Directions. *Global Business Review*, 16, 665–679. <https://doi.org/10.1177/0972150915581115>
- Bhowmick, A., & Hazarika, S. (2016). Machine Learning for E-mail Spam Filtering: Review, Techniques and Trends.
- Fladmoe-Lindquist, K. (1996). International franchising: Capabilities and development. *Journal of Business Venturing*, 11(5), 419–438. [https://doi.org/https://doi.org/10.1016/0883-9026\(96\)00056-0](https://doi.org/https://doi.org/10.1016/0883-9026(96)00056-0)

Luo, T., Chen, S., Xu, G., & Zhou, J. (2013). Trust-based Collective View Prediction. <https://doi.org/10.1007/978-1-4614-7202-5>

Townsend, J. D., Yenyurt, S., & Talay, M. B. (2009). Getting to Global: An Evolutionary Perspective of Brand Expansion in International Markets. *Journal of International Business Studies*, 40(4), 539–558. <http://www.jstor.org/stable/25483384>

Luo et al., Trust-Based Collective View Prediction, DOI: 10.1007/978-1-4614-7202-5\_4, Springer Science+Business Media New York 2013

Sudira, H., Diar, A. L., & Ruldeviyani, Y. (2019, October). Instagram sentiment analysis with naive bayes and KNN: exploring customer satisfaction of digital payment services in Indonesia. In 2019 International Workshop on Big Data and Information Security (IWBIS) (pp. 21-26). IEEE.

Vidya, N. A., Fanany, M. I., & Budi, I. (2015). Twitter sentiment to analyze net brand reputation of mobile phone providers. *Procedia Computer Science*, 72, 519-526.

**Internet Source:**

Statista. (2021). Beverages – Indonesia Statista.

Retrieved

<https://www.statista.com/outlook/dmo/ecommerce/beverages/indonesia#revenue>

from

Salsabilla, R. (2022). Mau Buka Gerai Mixue? Siapkan Duit Segini untuk Franchise! CNBC Indonesia. Retrieved from <https://www.cnbcindonesia.com/lifestyle/20221228123610-33-400838/mau-buka-gerai-mixue-siapkan-duit-segini-untuk-franchise>