UNIVERSITI TEKNOLOGI MARA

Cawangan Melaka

In partnership with

Tadulako University

i-JaMCSIIX 2023

EXTENDED ABSTRACT BOOK

Publication Date: 30 March 2024

ISBN: 978-967-15337-0-3

https://jamcsiix.uitm.edu.my

INTERNATIONAL JASIN MULTIMEDIA & COMPUTER SCIENCE INVENTION AND INNOVATION EXHIBITION (I-JaMCSIIX) 2023

EXTENDED ABSTRACT

ISBN: 978-967-15337-0-3
i-JaMCSIIX
Universiti Teknologi MARA Cawangan Melaka Kampus Jasin 77300, Merlimau, Melaka

Web: https://jamcsiix.uitm.edu.my

In partnership with
Tadulako University

ORGANIZING COMMITTEE

PATRON	PM DR ISMADI MD BADARUDIN
ADVISOR I	TS DR JAMALUDDIN HJ JASMIS
ADVISOR II	DATO' DR MOHD HAJAR HASROL JONO
PROGRAM DIRECTOR	**DR. NUR SUHAILAYANI SUHAIMI**
DEPUTY DIRECTOR	TS DR NURUL HIDAYAH BINTI MAT ZAIN
SECRETARY I	**ANIS SHOBIRIN ABDULLAH SANI**
SECRETARY II	FAIQAH HAFIDZAH HALIM
TREASURER I	**SITI AISYAH ABD KADIR**
TREASURER II	UMMU MARDHIAH JALIL
PUBLICATION	**NURBAITY BINTI SABRI** DR. SITI FEIRUSZ AHMAD FESOL DR. AHMAD FIRDAUS BIN AHMAD FADZIL SITI NURAMALINA BINTI JOHARI ROSNIZA ROSLAN
JURY	**Ts DR. ALYA GEOGIANA BUJA** NORBAHIYAH AWANG Ts. DR. NOR AFIRDAUS ZAINAL ABIDIN DR. RAIHAH AMINUDDIN NOOR AFNI DERAMAN
REGISTRATION	**SITI FAIRUS BINTI FUZI** BUSHRA BINTI ABDUL HALIM NORDIANAH BINTI JUSOH@HUSSAIN AINON SYAZANA BINTI AB HAMID SITI NURSYAHIRA BINTI ZAINUDIN
SYSTEM	**FADILAH EZLINA SHAHBUDIN** HAJAR IZZATI MOHD GHAZALLI FADHLINA IZZAH SAMAN NOR AZIDA MOHAMED NOH
INVITATION AND PROMOTION	**SHAHITUL BADARIAH SULAIMAN** IZNI SYAMSINA SAARI NOR ADILA KEDIN

	ADI HAKIM BIN TALIB MOHD AMIRUL BIN ATAN
MULTIMEDIA	**Ts. NURUL NAJWA ABDUL RAHID@ABDUL RASHID** NOOR ASHITAH ABU OTHMAN ANWAR FARHAN ZOLKEPLAY
AWARD	**ANITA BINTI MOHD YASIN** NURUL EMYZA ZAHIDI FATIMAH HASHIM SITI RAMIZAH JAMA DR NURUL HUDA NIK ZULKIFLI MARIATHY BINTI KARIM
CERTIFICATE	**KHAIRUL NURMAZIANNA ISMAIL** NUR NABILAH ABU MANGSHOR ZUHRI ARAFAH ZULKIFLI HAZRATI ZAINI
INTERNATIONAL RELATIONS	**Ts. DR. SITI RAHAYU ABDUL AZIZ** ALBIN LEMUEL KUSHAN SHAHADAN SAAD
LIAISON OFFICER	**SYAFNIDAR ABDUL HALIM** AJK WAKIL UNTAD
SPONSORSHIP	**ANIS AMILAH SHARI** MOHD RAHMAT MOHD NOORDIN DR YUZAIMI YUNUS DR SURYAEFIZA KARJANTO
SECRETARIAT & APPRECIATION BANQUET	**RAIHANA MD SAIDI** NUR SYUHADA BINTI MUHAMMAT PAZIL ANIS AFIQAH SHARIP SITI MAISARAH MD ZAIN HAZWA HANIM MOHAMED HAMZAH

UNTAD'S COMMITTEE FOR I-JAMCSIIX 2023:

PROF. IR. MARSETYO, M.AG., PH.D.

PROF. I WAYAN SUDARSANA, S.SI., M.SI.

PROF. JUNAIDI, S.SI., M.SI., PH.D.

ELISA SESA, S.SI., M.SI., PH.D.

MUKRIM, M.ED., PH.D.

ZARKIANI HASYIM, S.PD., M.ED.

DR. HJ. ANI SUSANTI, M.SI.

DR. ISKANDAR, M.HUM.

DR. IR. ROIS., MP.

SYARIFUL ANAM, S.SI., M.SI., PH.D.

DR. NAHARUDDIN, S.PD, M.SI.

DR. DRG. ELLI YANE BANGKELE, M.KES.

HERMAN, SKM., M.MED.ED.

DR. IR. SAMLIOK NDOBE, M.SI.

DR. RAHMAT BAKRI, S.H., M.H.

DR. HAERUL ANAM, SE., M.SI.

DR. IR. BAKRI, S.T., PG. DIPL. ENG., M.PHIL.

DR. IR. MUHAMMAD YAZDI PUSADAN, S.KOM., M.ENG.

IR. SYAIFUL HENDRA, S.KOM., M.KOM.

RIZANA FAUZI S.T., M.T.

MOHAMMAD FAJRI, S.SI., M.SI.

NURUL FISKIA GAMAYANTI, S.SI., M.SI.

DR. NUR'ENI, S.SI., M.SI.

IMAN SETIAWAN, S.SI., M.SI.

FADJRIYANI, S.SI., M.SI.

LIST OF SPONSORS

External Company Sponsors

WIDAD FACILITY MANAGEMENT

Klinik Dr Jamaluddin
Klinik Mawar Jasin
Nasi Ayam Ala Cina Zul
ADS Oasis Enterprise
Noorys Enterprise
Che Ramli bin Che Ismail
Beria Maju Enterprise
Rintiz rezeki
H&K food cafe
HS Gerak Wawasan

Individual Sponsors

En. Muhammad Hanif bin Abdul Aziz
Nor Suhaida binti Karjanto

Table of Contents

i-JaMCSIIX 2023
UNIVERSITI TEKNOLOGI MARA Cawangan Melaka
International Jasin Multimedia & Computer Science Invention and Innovation Exhibition
In Partnership with: Tadulako University

Twitter Sentiment Analysis of Malaysian Fast Food Restaurant Chains: A Novel Approach to Understand Customer Perception using Naïve Bayes

Muhammad Hafeez Hakimi bin Muhd Zahidi Ridzuan[1], Khairul Nizam bin Abd Halim[2]

[1,2] College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Malaysia

hafeezhakimi01@gmail.com, khairulnizam@uitm.edu.my

***Abstract*— Social media has emerged as a prominent platform for users to share ideas, opinions, and thoughts, leading to more consumers expressing their product feedback through these channels rather than providing direct feedback to companies. Fast food has gained popularity recently due to its affordability, tastiness, and convenience. However, a lack of a dedicated platform for customers to access comprehensive reviews of fast-food restaurants in Malaysia results in time-consuming processes when trying to read online reviews. This study introduces a web-based system that uses Twitter sentiment analysis to visualise reviews of Malaysian fast-food restaurants. It employs the Naïve Bayes algorithm and Plotly library in Python to provide insights into customer perceptions, enhancing the fast-food brand experience in Malaysia. This system introduces a comprehensive solution to understand restaurant sentiments by employing a visualisation dashboard and conducting a comparative analysis between various companies. Moreover, it empowers users to analyse their Twitter data using a sentiment analyser, which predicts the sentiments associated with the provided textual data.**

Keywords—Twitter Sentiment Analysis, Classification, Naïve Bayes, Plotly, fast-food

I. INTRODUCTION

In the contemporary digital age, the fast-food industry has undergone a transformation of its own, becoming a culinary cornerstone in many societies. Fast food, characterised by its affordability, convenience, and delectable offerings, has resonated with urban consumers. The changing lifestyles of urban dwellers, marked by a preference for dining out and participating in social gatherings hosted at fast-food establishments, have fuelled this trend [1]. This shift has been further propelled by the demands of contemporary urban life, where work and leisure commitments leave limited time for home-cooked meals [2]. Consequently, the Malaysian fast-food industry has flourished, catering to the preferences of time-constrained consumers seeking quick and readily available meals.

In any industry, customer feedback is invaluable for companies striving to improve their products and services. Measuring service quality and customer satisfaction can be challenging due to the unique characteristics of services: their intangible nature, variability, inseparability, and perishability [3]. Traditionally, companies have relied on manual methods, such as feedback forms, emails, or phone calls, to gather customer perceptions, complaints, and comments. However, a new avenue for collecting customer feedback has emerged with the advent of social media.

Although social media provides a platform for customers to express their opinions and experiences freely, manually sifting through vast amounts of data on platforms like Twitter can be time-consuming and daunting. As a result, there has been a need for a more convenient and automated approach to analysing customer feedback from social media platforms. One specific challenge in the context of Malaysia is the absence of a dedicated platform where customers can access comprehensive reviews of fast-food restaurants [4]. This absence leads to time-consuming processes when reading and analysing online reviews [5].

This paper presents a novel web-based system named MYFASTFOODVIBE, which leverages Artificial Intelligence (AI) through the Naïve Bayes algorithm to address these challenges. Unlike existing methods, MYFASTFOODVIBE can process text written in English and Bahasa Melayu, making it highly versatile for the Malaysian context. The classifier model used in MYFASTFOODVIBE is rigorously evaluated using confusion matrix to assess its accuracy. Additionally, the developed system undergoes usability testing to ensure an intuitive user experience. The paper is organised as follows: the subsequent sections detail the study methodology, followed by the results and discussion section, outlining the experimental findings and their interpretations. Finally, the conclusion summarises the study's results and their implications.

II. METHODOLOGY

A. Data Collection

In this study, secondary data in the form of tweets was collected online. The Malay dataset, sourced from the 'Malaya-Dataset' repository on GitHub curated by huseinzol05, included 57,913 negative, 44,670 positive, and 83,204 neutral sentiments. The English dataset from Kaggle comprised 85,249 negative and 85,457 positive sentiments. An additional dataset containing 61,990 rows of neutral sentiments was gathered from Kaggle to complete the sentiment categories. To capture real-world data, primary reviews of fast-food restaurants were scraped from Twitter, including tweets from @McDMalaysia, @KFCmalaysia, @pizzahutmsia, and others related to Malaysia.

B. Data Pre-Processing

Essential preprocessing steps were applied to the dataset to ensure its cleanliness and analysis readiness. Irrelevant information and columns were removed to maintain data quality. For the English dataset, tokenisation was performed using the NLTK library, followed by lemmatisation and excluding irrelevant and stop words. A parallel preprocessing approach was implemented for the Malay dataset, utilising the PySastrawi library due to linguistic similarities between the Malay and Indonesian languages.

C. Development and Evaluation of MYFASTFOODVIBE

The Waterfall model, a conventional software development approach, progresses linearly through requirements gathering, design, implementation, testing, and maintenance phases. One drawback is its inflexibility in accommodating changes post-phase completion [6]. In the development phase, the Iterative Waterfall model was applied. This model incorporates feedback loops between phases and enhances flexibility and adaptability to changing requirements [7], making it suitable for use in this study.

MYFASTFOODVIBE utilises a high-precision Naïve Bayes model for sentiment analysis, a critical aspect for understanding customer feedback. Before classification, the data undergoes transformation using TF-IDF and Bag of Words (BoW) techniques, converting textual inputs into numerical formats suitable for machine learning algorithms. This transformation reveals essential characteristics and keywords, enhancing the precision of sentiment analysis.

The Naïve Bayes algorithm plays a central role, undergoing training on labelled textual data, computing prior class probabilities and feature likelihoods, and ultimately selecting the class with the highest probability as the prediction using Bayes' theorem. To enhance accuracy, cross-validation is employed, dividing the training data into multiple folds to prevent overfitting and ensure adaptability to new data. Various hyperparameter configurations were tested using ten k-fold validations. All Naïve Bayes classifiers in this study were evaluated using a confusion matrix, and the results were recorded and compared to find the best Naïve Bayes classifier. The outperforming Naïve Bayes classifier was chosen and used for the MYFASTFOODVIBE application.

For the implementation of sentiment analysis from real-world data, data collection begins with tweets acquired through the Twitter API and Python-based scraping, stored for further analysis. The data was meticulously pre-processed to ensure quality and analysis readiness. After preprocessing, the previously identified outperforming Naïve Bayes classifier was used to predict sentiment in the tweets. MYFASTFOODVIBE generates comprehensive reports through the Naïve Bayes classifier's sentiment result and descriptive statistics for data analysis. This utilises the Plotly library, which provides interactive data visualisation capabilities. The MYFASTFOODVIBE application developed in this paper underwent usability testing, a user-centred design technique evaluating its intuitive design and user-friendliness. The testing ensured clarity and ease of presenting system workflow and information.

III. RESULTS AND DISCUSSION

A. Confusion Matrix Results

The confusion matrix results for both the training and test datasets provide insights into the accuracy and reliability of the Naïve Bayes classifier. In Fig. 1, the English classifier model demonstrates solid performance. During training, it achieves an 84% accuracy rate, indicating that it correctly predicts sentiment labels around 84% of the time. This translates to approximately 8 out of 10 accurate predictions for “positive”, “neutral”, or “negative” sentiments. The confusion matrix for training data offers further insights, revealing 55,329 correct predictions for “negative”, 45,590 for “neutral”, and 51,168 for “positive”.

The model maintains strong performance for testing data with a 79% accuracy rate, indicating that it correctly predicts sentiment labels on unseen data around 79% of the time. The corresponding confusion matrix showcases 12,789 correct predictions for “negative”, 11,335 for “neutral”, and 11,682 for “positive”.

Additionally, precision scores for different sentiment classes are provided. In training data, precision scores are 0.79 for “negative”, 0.97 for “neutral”, and 0.81 for “positive”. Similarly, for testing data, precision scores are 0.72 for “negative”, 0.96 for “neutral”, and 0.75 for “positive”.

```
Training Accuracy: 0.8408961479130612
Confusion matrix:
 [[55329   653 10731]
 [ 1265 45590  1458]
 [13750   919 51168]]
              precision    recall  f1-score   support

    Negative       0.79      0.83      0.81     66713
     Neutral       0.97      0.94      0.96     48313
    Positive       0.81      0.78      0.79     65837

    accuracy                           0.84    180863
   macro avg       0.85      0.85      0.85    180863
weighted avg       0.84      0.84      0.84    180863


Testing Accuracy: 0.7918878273177636
Confusion matrix:
 [[12789   219  3575]
 [  359 11335   385]
 [ 4580   292 11682]]
              precision    recall  f1-score   support

    Negative       0.72      0.77      0.75     16583
     Neutral       0.96      0.94      0.95     12079
    Positive       0.75      0.71      0.73     16554

    accuracy                           0.79     45216
   macro avg       0.81      0.81      0.81     45216
weighted avg       0.79      0.79      0.79     45216
```

Fig. 1. Result of Accuracy Testing for English Model.

```
Training Accuracy: 0.8430221743285803
Confusion matrix:
 [[53846  3138  1505]
 [ 7061 51025  2912]
 [ 5589  4268 26557]]
              precision    recall  f1-score   support

    Negative       0.81      0.92      0.86     58489
     Neutral       0.87      0.84      0.85     60998
    Positive       0.86      0.73      0.79     36414

    accuracy                           0.84    155901
   macro avg       0.85      0.83      0.83    155901
weighted avg       0.85      0.84      0.84    155901


Testing Accuracy: 0.7698070607553367
Confusion matrix:
 [[12770  1247   559]
 [ 2520 11577  1237]
 [ 1794  1615  5657]]
              precision    recall  f1-score   support

    Negative       0.75      0.88      0.81     14576
     Neutral       0.80      0.75      0.78     15334
    Positive       0.76      0.62      0.68      9066

    accuracy                           0.77     38976
   macro avg       0.77      0.75      0.76     38976
weighted avg       0.77      0.77      0.77     38976
```

Fig. 2. Result of Accuracy Testing for Malay Model.

In Fig. 2, the Malay classifier model demonstrates commendable accuracy. During training, it achieves an 84.30% accuracy rate, accurately predicting sentiment labels for the data it was trained on. The confusion matrix for training data reveals 53,846 correct predictions for “negative”," 51,025 for “neutral”, and 26,557 for “positive”.

Moving to testing data, the model maintains good accuracy, with a 76.98% accuracy rate, indicating accurate predictions for unseen data. The confusion matrix for testing data indicates 12,770 correct predictions for “negative”, 11,577 for “neutral”, and 5,657 for “positive”.

Furthermore, precision, recall, and F1-score metrics for each sentiment class are provided for testing data, offering insights into the model's performance. For the "Negative" class, precision is 0.75, recall is 0.88, and the F1-score is 0.81. For the “Neutral” class, precision is 0.80, recall is 0.75, and the F1-score is 0.78. For the “Positive” class, precision is 0.76, recall is 0.62, and the F1-score is 0.68. These metrics collectively reflect the model's effectiveness in classifying sentiments.

B. System Functionality Screenshots and Usability Testing Results

Screenshots of MYFASTFOODVIBE's user interface and functionality showcase how users can interact with the system and obtain valuable insights from social media data. The system's interface features a strategically designed landing page, as depicted in Fig. 3, guiding users to explore sentiment analysis results for fast food restaurants. As illustrated in Fig. 4, it offers an overview of sentiment distribution across categories, mentions of restaurants based on key aspects, and sentiment trends over time.

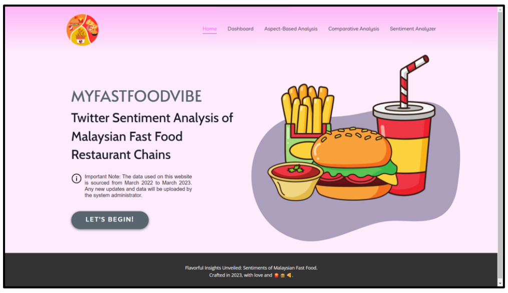

Fig. 3. Landing Page of the System.

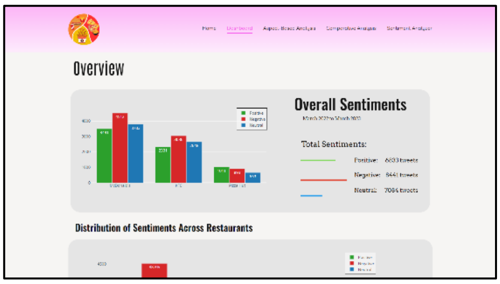

Fig. 4. Overview Page.

Users can explore detailed sentiment analysis dashboards for specific restaurants, exemplified by the McDonald's dashboard in Fig. 5, featuring summary metrics, sentiment breakdowns, and visualisations like bar charts, pie charts, and word clouds. The system provides a dedicated Comparative Analysis page, as shown in Fig. 6, for sentiment performance comparisons across selected companies. Interactive tools for single and multiple sentiment analysis are also available, demonstrated in Fig. 7 and Fig. 8, empowering users to analyse text for positive, negative, or neutral sentiment and visualise the results. These features collectively offer comprehensive insights into customer sentiments about fast-food establishments, facilitating informed decision-making. MYFASTFOODVIBE underwent usability testing involving feedback from 20 random individuals who used the System Usability Scale (SUS) questionnaire. The final average score on the SUS scale was 93.13%. This result demonstrates that the system's workflow and information were presented excellently and in an easy-to-use manner.

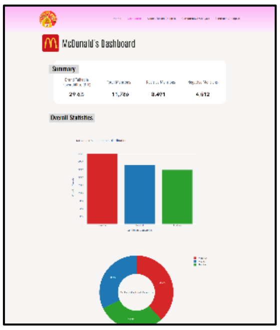

Fig. 5. Example of a fast food restaurant dashboard.

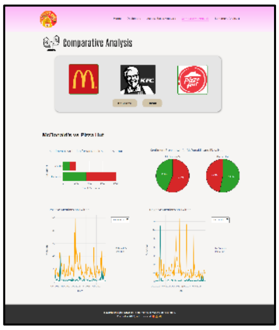

Fig. 6. Comparative Analysis page.

Fig. 7. Single Sentiment Analyzer page.

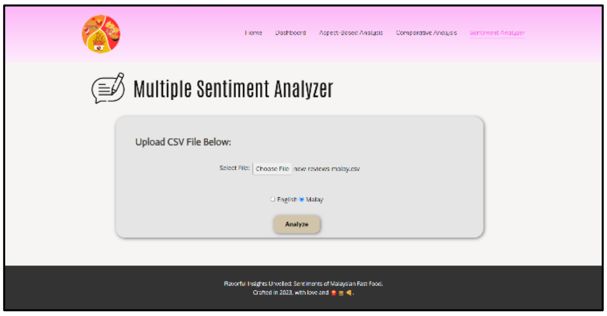

Fig. 8. Multiple Sentiment Analyzer page.

IV. CONCLUSIONS

MYFASTFOODVIBE is a system that has achieved notable success in the realm of fast-food sentiment analysis. It leverages a high-accuracy Naïve Bayes classifier to effectively address the challenges of analysing customer feedback on social media platforms. This accomplishment underscores its significance as a tool that benefits fast-food managers by providing valuable insights into customer perceptions of their brands and empowers customers to make informed choices based on Twitter reviews. Moreover, the system's AI-driven capabilities enable continuous knowledge updates and offer an interactive platform for effortless analysis, enhancing Malaysia's overall fast-food brand experience.

Furthermore, MYFASTFOODVIBE's versatility is a key advantage. It is not limited to fast-food chains like McDonald's, KFC, or Pizza Hut. Instead, its adaptable framework can be extended to encompass a broader range of restaurants, thereby expanding the scope of this study and the potential benefits it can bring. This adaptability ensures that MYFASTFOODVIBE remains a valuable and relevant tool for the dynamic fast-food industry in Malaysia.

ACKNOWLEDGMENT

Alhamdulillah, praises and thanks to Allah for His Almighty and His utmost blessings, I could finish this study within the given timeframe. I extend my heartfelt gratitude to my dedicated supervisor, Mr. Khairul Nizam Bin Abd Halim, whose unwavering support, invaluable guidance, and belief in my abilities were instrumental throughout this research journey. His patience, constructive feedback, and tireless efforts supervising me over two semesters have been truly appreciated. I also express my sincere thanks to my diligent examiner, Dr. Nurulhuda binti Zainuddin, whose insightful evaluation and valuable comments greatly contributed to the enhancement of this research project. Their expertise and mentorship have been indispensable, and I am profoundly grateful for their unwavering commitment to excellence. A special thank you also goes to my beloved parents and siblings for all the endless prayers and support, especially for always believing in me. Last but not least, I would like to thank all my friends who have assisted in many ways upon completing this project.

REFERENCES

[1] Kwek Choon Ling, Yeong Wai Mun, & Hiew Mee Ling. (2011). Exploring factors that influence customer loyalty among Generation Y for the fast food industry in Malaysia. African Journal of Business Management, 5(12), 4813–4823.

[2] Habib, F. Q., Abu Dardak, R., & Zakaria, S. (2011). Consumers' preference and consumption towards fast food: Evidences from Malaysia. Business and Management Quarterly Review (BMQR), 2(1), 14-27.

[3] Naik, C. K., Gantasala, S. B., & Prabhakar, G. V. (2010). Service quality (SERVQUAL) and its effect on customer satisfaction in retailing. European journal of social sciences, 16(2), 231-243.

[4] Shamsudin, M. F., Yazid, M. F. M., Hasan, S., & Kadir, B. (2020). Review on service quality in fast food restaurant. Customer satisfaction as mediator. Journal of Critical Reviews 7(16), 997-1006.

[5] Pecar, S. (2018). Towards opinion summarization of customer reviews. Proceedings of ACL 2018, Student Research Workshop.

[6] Adenowo, A. A., & Adenowo, B. A. (2013). Software engineering methodologies: a review of the waterfall model and object-oriented approach. International Journal of Scientific & Engineering Research, 4(7), 427-434.

[7] Goel, P. (n.d.). What is Waterfall Model in Software Engineering? || Modified Waterfall Model. The Study Genius. https://www.thestudygenius.com/what-is-waterfall-model/.

i-JaMCSIIX 2023

PUBLISHED BY:
i-JaMCSIIX
Universiti Teknologi MARA Cawangan Melaka
Kampus Jasin
77300 Merlimau, Melaka

Tel: 062645000
Email: jamcsiix@uitm.edu.my
Web: https://jamcsiix.uitm.edu.my/