IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 November 2023, accepted 27 December 2023, date of publication 5 February 2024, date of current version 10 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3362730

RESEARCH ARTICLE

Enhancing Minority Sentiment Classification in Gastronomy Tourism: A Hybrid Sentiment Analysis Framework With Data Augmentation, Feature Engineering and Business Intelligence

MOHD NORHISHAM RAZALI[1,2], SYAIFULNIZAM ABDUL MANAF[3], ROZITA BINTI HANAPI[1], MOHD RAFIZ SALJI[4], LEE WEN CHIAT[1], AND KASHIF NISAR[5]

[1]Faculty Business and Management, Universiti Teknologi MARA Sarawak, Kota Samarahan, Sarawak 94300, Malaysia
[2]Creative Advanced Machine Intelligence (CAMI) Research Centre, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Sabah 88400, Malaysia
[3]Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang 43400, Malaysia
[4]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA Sarawak, Kota Samarahan, Sarawak 94300, Malaysia
[5]Swinburne University of Technology, Sydney, Australia

Corresponding author: Mohd Norhisham Razali (hishamrazali@uitm.edu.my)

ABSTRACT The gastronomy tourism industry plays an important role in boosting local economies, enhancing the travel experience, and preserving culinary traditions unique to specific places. In this context, comprehending customer sentiments is of paramount importance for business decision-making, menu choice offerings, marketing strategies, and customer service improvements. Traditional sentiment analysis methods in gastronomy tourism tend to be time-consuming, prone to human error, and influenced by subjectivity. Furthermore, the absence of an effective visualization strategy hampers the reliability of sentiment analysis efforts. Compounding this, the data collected also often lacked balance across sentiment classes, making it challenging to predict minority sentiments accurately. To address these challenges, our research introduces a hybrid approach, combining various lexicon-based sentiment and emotional analysis algorithms, thereby enhancing the reliability of customer review analysis in the gastronomy tourism sector. Subsequently, we optimize machine learning sentiment classification by employing data augmentation in conjunction with feature engineering strategies, to improve the recognition of minority sentiment classes. Additionally, we present a comprehensive business intelligence and visualization solution that is personalized for the gastronomy tourism industry in Sarawak and offers real-time sentiment visualization. The optimization of sentiment classification, achieved through the integration of synonym augmentation and n-gram feature engineering in conjunction with kNN classifiers, has yielded impressive results. This approach attains optimal classification performance, boasting an accuracy rate of 0.98, a F1-score and a ROC-AUC score of 0.99. Notably, this methodology significantly enhances the recognition of minority sentiment classes within the dataset, addressing the main challenges in this research.

INDEX TERMS Business analytics, machine learning, sentiment analysis, tourism industry, data visualization.

I. INTRODUCTION

Gastronomy tourism has become increasingly known as a promotion strategy for economic and cultural development in a developed country. As a result, many destinations have invested in gastronomy tourism as a key component of their tourism strategy such as to ensure the quality of services provided by the eateries is up to a certain standard [5]. Sentiment analysis is an important task in gastronomy tourism as it helps businesses and destinations to better

The associate editor coordinating the review of this manuscript and approving it for publication was Walter Didimo.

understand their customers and improve the overall tourist experience ([8], [11], [25], [30]). By analyzing customer reviews and feedback in sentiment analysis, businesses can identify the important aspects of the tourist experience that should become the key strength of the businesses or require improvements in the future [24]. Other than that, sentiment analysis can help to sustain good reputations of the businesses. A vast number of studies have been conducted previously to improve the state-of-the-art sentiment analysis methods such as the improvement of Aspect Based Sentiment Analysis (ABSA) that incorporate linguistic process to overcome the semantic interference [21], stepwise multi-task learning model for holder extraction with RoBERTa and Bi-LSTM to address the issues of fine-grained sentiment analysis [23] and the concept of *Scope* introduced in [29], that highlights the structural text region related to a specific target in ABSA. However, in the local context of Sarawak, it rely on the traditional approaches in analyzing the customer's feedbacks [4]. Manual sentiment analysis involves a human reading and interpreting text to determine the writer's sentiment or emotion towards a particular topic. It has limitations that can impact the accuracy, speed, and cost-effectiveness of the analysis especially when dealing with large volumes of data. This can make it impractical for businesses that need to analyze customer sentiment in real-time. Manual sentiment analysis can also be subjective since different people may have different interpretations of the same text. This can lead to inconsistencies in the analysis results. Next, manual sentiment analysis has limited scope as a human can only analyze a limited amount of data within a specific timeframe, which can lead to a narrow scope of insights and lastly, a human can bring their own biases and opinions to the analysis, which can skew the results.

Another important challenges in machine learning for restaurant sentiment analysis is the difficulty faced by machine learning models in effectively distinguishing between positive, negative, and neutral sentiment classes, particularly when dealing with minority classes. Indeed, recent advancements in machine learning models, including LSTM, CNN, and hybrid CNN-LSTM, as well as the foundational models like SVM, Random Forest, and KNN, have demonstrated less convincing performance when evaluated in terms of F1-Score and ROC-AUC ([12], [16], [19], [28]). Moreover, a lack of studies has been found in evaluating feature engineering aspects to address the issues of class imbalance in restaurant sentiment analysis. Therefore, the objective of this research is to model a robust sentiment analysis and recognition in gastronomy tourism in which the contributions are highlighted as follows:

- We employed a hybrid approach that combines multiple lexicon-based sentiment and emotional analysis algorithms with machine learning techniques to enhance the accuracy and reliability of customer review analysis in the gastronomy tourism industry.
- We enhanced machine-learning-based sentiment analysis through data augmentation and feature engineering strategies to improve the recognition of minority sentiment classes in a local gastronomy reviews dataset.
- We conducted sentiment polarity analysis and generated sentiment visualizations to build a business intelligence dashboard tailored for the gastronomy tourism industry in Sarawak.

The remainder of this paper is organized as follows: section II reviews literature related to gastronomy tourism in Sarawak and sentiment analysis; section III presents the experimental setup and methodology of this study; section IV discusses the results, whereas section V provides the limitations and recommendations for future research.

II. RELATED WORKS

This section highlights the current state of gastronomy tourism development in Sarawak and the previous works of the sentiment analysis research in the context of gastronomy tourism.

A. GASTRONOMY TOURISM IN SARAWAK

Sarawak is a state on the island of Borneo in Malaysia that known for its unique cuisine and cultural heritage, making it a potential destination for gastronomy tourism [27]. The following statistics indicate that gastronomy tourism is an important component of Sarawak's tourism industry, with the state's unique cuisine and cultural heritage attracting both domestic and international tourists. In 2019 before the pandemic COVID-19 hit the world, Sarawak received a total of 4.66 million tourists, and out of these, 2.58 million were domestic tourists and 2.08 million were international tourists. (Ministry of Tourism, 2023). In 2022, Sarawak received a total of 2.03 million tourists after the border of the country has been reopened and expecting the number of tourists visiting Sarawak increased to 3 million in 2023 [14]. The contribution of food and beverage services to Sarawak's tourism industry was estimated at RM3.4 billion in 2018. (Department of Statistics, 2019). Sarawak is known for its diverse indigenous cuisine, including traditional dishes like manok pansoh (chicken cooked in bamboo) and umai (a raw fish salad). The state is also home to numerous local markets and street food stalls that offer a wide variety of snacks and delicacies. The annual Kuching Food Festival is a popular event that attracts both locals and tourists [26]. The festival features a variety of food stalls, cooking competitions, and cultural performances. In 2018, Sarawak Tourism Board launched the "Sarawak Culinary Heritage Trail" to promote the state's unique culinary traditions and cultural heritage. The trail includes a variety of local food experiences, from visiting local markets to trying traditional recipes. According to research conducted by [9], food is one of the reason that tourists visit Sarawak.

B. SENTIMENT ANALYSIS IN GASTRONOMY TOURISM

Gastronomy tourism is a rapidly growing industry that combines travel with culinary experiences. As gastronomy is a vital component of tourism, analyzing the sentiments associated with food and dining experiences can provide insights into the overall satisfaction of tourists. With the growth of social media platforms and online reviews, there has been a significant increase in sentiment analysis studies in the tourism industry. This section discusses various approaches for sentiment analysis in the restaurant industry using machine learning techniques. Reference [6] proposed a deep learning framework for aspect-level sentiment analysis using graph neural networks with a graph attention model to address the limitations of traditional sentiment analysis techniques. The proposed method captures context information and an extended structural attention model is developed to concentrate on a specific part of the graph. The approach was evaluated on SemEval 2014 dataset for laptop and restaurant reviews and achieved an accuracy of 72.88% and 78.75%, respectively. The findings demonstrate the effectiveness of the proposed technique in accurately identifying aspect-level sentiment polarity. Future works include the development of adaptive graph-based sentiment mechanisms.

One of the challenges in sentiment analysis is the language barrier. While most sentiment analysis research has focused on English language reviews, there is a growing need for sentiment analysis in other languages, including Arabic. Reference [5] developed a sentiment analysis approach using machine learning algorithms to identify aspect-level sentiment polarity of Arabic language reviews of restaurants in the Qassim region. The proposed approach utilized Support Vector Machine, Logistic regression, and Random Forest algorithms, achieving high accuracy, recall, and F-measure scores. The study used a dataset of 1,785 collected Arabic reviews, with findings of an F-measure of 0.93 (RF), accuracy of 0.89 (SVM), and recall of 0.92 (SVM). Future work includes the use of more data to further improve sentiment analysis accuracy. This study demonstrates the effectiveness of machine learning in accurately analyzing Arabic language restaurant reviews.

Reference [16] explored sentiment analysis in restaurant reviews using machine learning techniques. The study aimed to improve the accuracy of short text categorization, which has previously been challenging. The proposed technique used a deep neural network based on Long Short-Term Memory (LSTM) and Fuzzy logic model with incremental learning. The study used Yelp dataset and evaluated performance using F1-score, accuracy, precision, and recall. Results showed an accuracy of 81.04%, precision of 77.81%, recall of 80.63%, and F1-score of 75.44. The study demonstrated that the proposed technique can effectively classify sentiments from restaurant reviews with high accuracy.

In today's competitive market, understanding consumer preferences is crucial for businesses. However, existing approaches struggle to comprehensively analyze consumer choices for various attributes, hindering product or service providers. To address this, [31] propose a novel consumer preference analysis method based on reviews, including ratings and comments. Their approach uses multi-attribute, quadratic programming, aspect-based sentiment analysis, and recommendation decision making (MADM). Using the Ctrip.com dataset, the authors demonstrate the effectiveness of their approach by solving for weights and achieving better recommendation accuracy using mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE) metrics. This method offers businesses a way to better understand their customers and make informed decisions to improve their products/services.

Reference [19] proposed a deep learning approach for aspect-based sentiment analysis of restaurant reviews in Spanish. The purpose of the study was to process, analyze, and categorize large amounts of information generated from reviews. The proposed technique was a combination of two aspect-based convolutional neural networks (AB-CNN), evaluated on the SemEval dataset. Performance metrics included F1-measure and accuracy, with an accuracy score of 0.9333 and an F1-measure score of 0.6540. Future work may include incorporating dictionaries in the preprocessing stage to replace core words. This study demonstrates the potential of machine learning techniques for sentiment analysis in the restaurant industry, particularly in non-English languages.

Reference [17] aimed to perform text classification and clustering of consumer reviews using supervised and unsupervised machine learning algorithms. The study focused on two datasets, one on games and the other on restaurant reviews. They employed SVM and measured the performance of their proposed technique using accuracy, precision, recall, and F-score. Their findings show an accuracy of 0.82, precision of 0.82, recall of 0.821, and an F-score of 0.901. The study suggests the need for combining both supervised and unsupervised machine learning algorithms for sentiment analysis. The datasets used were obtained from Kaggle and a GitHub repository.

Reference [28] proposed a deep learning technique for aspect-level sentiment analysis (ABSA) based on BERT fusion multi-attention. The authors addressed the limitations of existing ABSA models that cannot effectively distinguish the importance of aspect words and words in the text and lack the utilization of the overall interaction between aspect words and text. The proposed approach captures the potential interaction between aspects and contexts and mines deep-level contextual internal semantic associations to improve aspect-level sentiment classification efficiency. The study was conducted on Semeval2014 data and the performance was evaluated using accuracy and F1-score. The results showed an accuracy of 84.3% and an F1-score of 74.8%. The study suggests that future research could explore other techniques to improve aspect-level sentiment analysis.

The study by [2] highlights the need for a literature review on deep learning (DL) methods and explainable artificial intelligence (XAI) techniques for predicting customer sentiments in the Food Delivery Service (FDS) domain.

The authors propose to use machine learning (ML) and DL models for sentiment analysis, with a focus on using XAI techniques to build trust in DL models. While the lexicon-based approach shows better performance than ML algorithms such as SVM, the DL methods of RNN, CNN, and LSTM show promising results. The study also recommends using topic categorization techniques such as LIME or SHAP for classifying negative sentiments into various topic categories.

Reference [12] conducted sentiment analysis of local tourism in Thailand from YouTube comments using the Social Media Sensing framework (S-Sense) and Bidirectional Long Short-Term Memory (BiLSTM) methods. The study aimed to analyze the relationship between social media use and its effect on community-based tourism in Thailand. The study used lexicon and machine learning techniques to analyze comment data from 114 clips on YouTube. The proposed technique achieved an accuracy of 83.25%, precision of 87.01%, recall of 85.78%, and F1-score of 85.08%. Future works suggested collecting more comments and executing a parallel study on other languages.

Reference [18] addressed the challenge of sentiment classification of restaurant review images on social media using deep convolutional neural networks. Their purpose was to efficiently classify all types of restaurant images and improve the accuracy of the dataset. They proposed two models: the Dimensional Model and the Categorical Model, which utilized a convolutional neural network. Performance metrics including precision, recall, F1 score, and accuracy were used to evaluate the models, with impressive results. Precision was 0.96, recall was 0.90, F1 score was 0.924, and accuracy was 0.93. Future works include the consideration of further real-time images and videos streamed from social media for review.

The study by [15], addresses the lack of a restaurant-specific sentiment dictionary and explores the spatiotemporal trends of restaurant sentiments. The purpose is to develop a restaurant-domain-specific dictionary, visualize sentiment trends, and identify the main discrete emotions affecting customer ratings in a restaurant setting. The study uses multiple linear regression methods and Latent Dirichlet Allocation (LDA) to analyze restaurant review data in Chengdu, China. The findings suggest that positive and negative sentiment categories are not enough to fully capture the complexity of customer emotions in restaurant reviews. The authors suggest using longer time series, a larger range of data, and different clustering algorithms for future research.

Sentiment analysis has proven to be a useful tool for understanding the sentiments associated with gastronomy tourism. By analyzing the sentiment of online reviews and social media posts, tourism businesses can gain insights into the factors that influence customer satisfaction and improve their services accordingly. Furthermore, analyzing the sentiment of gastronomy events can provide insights into the overall success of such events and help to identify areas for improvement.

In summary, prior research has shown that several proposed methods have achieved F1-Scores slightly lower compared to their accuracy, with results such as 75.44% [16], 0.6540 [19], 74.8% [28], and 85.08 [12]. This indicates that these models struggle to effectively balance precision and recall, and encounter difficulties in distinguishing between positive and negative sentiment in the context of restaurant sentiment analysis. Notably, limited research has explored feature engineering techniques and their compatibility with machine learning algorithms to address class imbalance issues in this domain. Furthermore, it is suggested that a hybrid approach combining supervised and unsupervised machine learning methods could be considered for a more comprehensive and reliable sentiment analysis model, as proposed by [17].

III. EXPERIMENTAL SETUP

The data acquisition, experiments, data plotting and data visualization were carried out by using the programming environment, libraries and tools as depicted in Table 1 and Python libraries based on Spyder 4.2.2. Specifically, the feature extractions and classifications were performed by using Scikit-learn and Keras libraries.

A. DATA ACQUISITION

Web scrapping technique was used to collect and extract data in this study. Web scraping involves identifying the data to be collected, selecting the appropriate tools and techniques, writing the scraping script or program, testing and refining the script, and finally extracting and processing the data. Data acquired through web scraping was saved in CSV format before further analysed and processed using data analysis and visualization tools. In this study, the data related to restaurant's reviews were scrapped from Tripadvisor website by using Selenium library packages. In particular, 30 restaurants in Kuching city were randomly selected for web scrapping. The scrapped data including the review date, overall star rating, reviews, total reviews, restaurant name, address, cuisine types, food star rating, service rating and value rating. The pseudocode of data acquisition using Selenium is showed in Table 2. Table 3 depicted the descriptions and data type of each of attribute and Table 4 shows the total of reviews and scrapped reviews for each restaurants.

B. LEXICON-BASED SENTIMENT AND EMOTION ANALYSIS

Sentiment analysis was performed to determine the ratio positive to negative about the restaurant reviews scrapped from Tripadvisor website based on lexicon or rule-based approach. In particular, the polarity of reviews is calculated and return the decimal value in certain range. In this experiment, three types of lexicon based sentiment analysis were adopted which are AFINN, VADER and Textblob. AFINN [22] contains 3500 english words and each word has been manually label with the polarity between −5 to 5. The low polarity indicate more negativity of reviews.

TABLE 1. Experiment environment.

Item	Package
Integrated Development Environment	Sypder 4.1.5
Data Visualization	Tableau Desktop Professional Edition
Programming Language	Python 3.8
	afinn 0.1
	keras 2.4.3
	matplotlib 3.3.2
	NRCLex
	nltk 3.5
	spacy 3.0
	numpy 1.19.4
	opencv-python 4.5.1.48
	pandas 1.1.3
	scikit-image 0.17.2
	scikit-learn 1.0.2
	scipy 1.5.2
	selenium 3.141.0
Libraries	seaborn 0.11.0
	textblob 0.17.1
	webdriver-manager 3.5.2
	wordcloud 1.8.2.2

TABLE 2. Data acquisition pseudocode.

Algorithm 1: This algorithm automates the web browser to scrap the reviews of restaurant from Trip Advisor website.

Start
Import the necessary libraries: Selenium, webdriver, and pandas.

1. Set the default file format to CSV to store the scraped data.
2. Get the URL link of the restaurant from TripAdvisor.
3. Set the type of web browser to be used with the webdriver.
4. Start scraping by using a for loop to iterate through the pages:
    4.1 Find and assign the XPath for each element to be scraped from the webpage.
    4.2 Extract the data and store it in a pandas DataFrame.
    4.3 Write the scraped data to a CSV file.

End

Let's denote the AFINN sentiment score for a word as A(w), where "w" represents the word. A(w) is a real number representing the sentiment score of the word.

$$A(w) = \sum [\text{score}(wi)]/n$$

A(w): The AFINN sentiment score for the word "w."

score(wi): The sentiment score for each individual word "wi" in the word "w."

n: The total number of words in "w."

Next, VADER (Valence Aware Dictionary and Sentiment Reasoner) (Hutto & Gilbert, 2014) uses a lexicon of words and phrases that have been rated on a scale from −4 (most negative) to +4 (most positive) to determine the emotional sentiment of the reviews. VADER has capability to deal with the social media text, which can be more informal and contain more slang or misspellings than traditional written text.

Let S(w) be the VADER sentiment score for a given word "w." The VADER sentiment score for a text or sentence is calculated as the sum of the sentiment scores for each word in the text, normalized by the length of the text. Mathematically,

TABLE 3. Description of data attributes.

Attribute Name	Descriptions	Data Types
Review date	The date of reviews have been posted	Nominal
Overall star rating	The scale between 1 to 5 given by the customers to the restaurant as a whole.	Continuous
Review Title	Main highlight of reviews.	Nominal
Reviews	The evaluation written by the customers related to their experiences, quality, taste of foods or restaurants.	Nominal
Total reviews	The number of reviews received from the customer on particular restaurant.	Discrete
Restaurant Name	The name of business entity.	Nominal
Address	The specific location of the restaurants.	Nominal
Cuisine Types	The category of dishes.	Nominal
Food Star Rating	The scale between 1 to 5 given by the customers on foods.	Continuous
Service rating	The scale between 1 to 5 given by the customers on restaurant services.	Continuous
Value rating	The scale between 1 to 5 given by the customers on the value.	Continuous

the VADER sentiment score for a text "T" can be defined as:

$$S(T) = (\Sigma\ S(w))/\sqrt{}(\Sigma S(w)\text{^}2 + \alpha)$$

where:

S(T): The VADER sentiment score for the text "T."

S(w): The VADER sentiment score for each word "w" in the text.

Σ : The sum is taken over all the words in the text.

α : A constant used for normalization, typically a small positive value to prevent division by zero.

Meanwhile, Textblob is a python library that provides sentiment analysis features that used machine learning algorithm to determine the sentiment polarity score between −1 (most negative) and +1 (most positive). Textblob is built on top on Natural Language Tool Kit (NLTK). NLTK was also employed to tokenize the reviews data and to remove the stop words. All the punctuations comprising !"#$%&'()*+,-./:;?@[\]^_'{|}~ were also eliminated. Subsequently, n-grams analysis was performed to get some contextual meaning from the text. In particular, a contiguous sequence of N items from the text has been extracted. In this study, Bi-grams and Tri-grams have been extracted from the text. Bi-grams or 2-grams is a combination of two words appear in the text, whereby Trigrams or 3-grams is an N-gram containing up to three elements from the text. In this study, the top 50 for both bigrams and trigrams were analysed for all restaurants.

Afterwards, emotion analysis by using NRCLex python library has been conducted. The emotions score from the text is computed that measure the ten emotional affects which are fear, anger, anticipation, trust, surprise, positive, negative,

TABLE 4. Number of scrapped reviews for each restaurant.

Restaurant code	#reviews	# scrapped reviews
BAL	132	56
BEL	117	66
BIN	109	26
BLA	758	55
BLK	291	111
BOR	232	79
CEL	212	70
CHO	210	55
COM	118	24
HER	61	18
IZA	88	23
JAM	1,224	159
KOP	66	25
LEP	268	50
LIM	65	26
LAK	66	32
VIL	75	29
RAI	33	10
RAS	302	69
AYA	95	37
BIS	45	14
RUM	49	19
SAR	21	8
KAN	43	16
THE	589	195
DAK	298	67
LIN	183	44
SPO	134	24
TRI	40	10
NUS	64	27

sadness, disgust and joy. The affects dictionary in NRCLex contains approximately 27,000 words based on the National Research Council Canada (NRC) affect lexicon.

Let E(T) represent the emotion scores for a given text "T." Emotion analysis using NRCLex typically provides scores for multiple emotions, such as anger, fear, joy, sadness and so on.

Mathematically, the emotion analysis with NRCLex can be represented as a vector of emotion scores:

$$E(T) = [E_1, E_2, E_3, \ldots, E_n]$$

where:

E(T): The emotion scores for the text "T."

$E_1, E_2, E_3, \ldots, E_n$: The scores for individual emotions. The number of emotions and the specific emotions analyzed depend on the implementation of NRCLex.

C. SYNONYM AUGMENTATION

Data augmentation technique is employed in this study to address the problems of the current machine learning techniques such as LSTM, CNN and CNN-LSTM as well as in many traditional machine learning techniques that were less effective to recognize the minority sentiments such negative and neutral class in the dataset. In particular, synonym data augmentation technique was opted that generate new training data replacing words in existing sentences with their synonyms [1]. Synonym data augmentation is a simple but effective way to increase the size and diversity of a training dataset, which can lead to improved model performance.

The pseudocode shown in Table 6 describes a procedure for generating augmented reviews using synonym augmentation. The procedure takes as input a dictionary of restaurant reviews, where the keys are restaurant names and the values are lists of reviews. The procedure also takes as input the number of augmented reviews to generate for each review.

D. FEATURE ENGINEERING

This study was also explored on the feature engineering aspect in restaurant's sentiment analysis that is crucial to train machine learning model. Feature engineering in restaurant's sentiment analysis is a process to select, create and transform relevant features from the text or reviews. The evaluation of feature engineering is essential in this study as effective feature engineering should be tailored to the local context and to identify the most informative and optimal features. There were four feature engineering approaches being selected in this study which are Bag of Words (BoW), N-grams, Term Frequency-Inverse Document Frequency (TF-IDF) and Word2Vec [3], [13].

E. MACHINE LEARNING BASED SENTIMENT ANALYSIS

Numerous machine learning models were assessed to train the sentiment analysis model, encompassing a range of baseline models and state-of-the-art or domain-specific models. The baseline models include Linear Support Vector Machine (LSVM), Support Vector Machine (SVM), Gradient Boost Classification (GBC), Multilayer Perceptron (MLP), Random Forest (RF), and K-Nearest Neighbors (KNN). In contrast, the selection of domain-specific models features Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and CNN-LSTM. The LSTM has been employed by [16] to overcome the issues of short text classification in the restaurant sentiment analysis on Yelp dataset and [12] to analyze the relationship between social media use and its effect on community-based tourism in Thailand. On the other hand, CNN was used by [18] to deal with problem of review images sentiment classification for restaurants. As for training and testing data, the dataset was split into 80:20 in which 80% of the data used to train the machine learning model and 20% was used for testing.

F. TOP WORDS AND N-GRAMS ANALYSIS

The top words refer to the words that occur most frequently in the reviews. Analyzing the top words in this study can provide valuable insights into the underlying structure and meaning of a review. For instance, the key topics or theme written in the reviews could be identified to highlight the important issues that were being concerned by the customers towards the restaurants. Ultimately, the business entity would become more aware about what are the people talking about on the online platforms. On the other hand, N-grams analysis was also provided which consists of the contiguous sequences of n words from the review. By having N-grams analysis, it can help to capture the context in which words appear

TABLE 5. Pseudocode for sentiment and emotions analysis.

Algorithm 2: This algorithm preprocess the data, compute polarity score and determine sentiment by using Textblob, VADER and AFINN and get the emotions.

```
Start
# Step 1
dataframe = read_csv_file("scrapped_data.csv")

# Step 2
for review in dataframe:
review_without_stopwords =   remove_stopwords(review)
review_without_punctuations =
remove_punctuations(review_without_stopwords)
nouns = tokenize(review_without_punctuations, "nouns")

# Step 3
for noun in nouns:
textblob_score = get_polarity_score(noun, "Textblob")
vader_score = get_polarity_score(noun, "VADER")
afinn_score = get_polarity_score(noun, "AFINN")

# Step 4
review_title = review["title"]
review_text = review["text"]
textblob_polarity = get_polarity_score(review_title + " " + review_text,
"Textblob")
vader_polarity = get_polarity_score(review_title + " " + review_text,
"VADER")
afinn_polarity = get_polarity_score(review_title + " " + review_text,
"AFINN")
if textblob_polarity < 0:
        textblob_sentiment = "Negative"
elif textblob_polarity == 0:
        textblob_sentiment = "Neutral"
else:
        textblob_sentiment = "Positive"
if vader_polarity <= 0.05:
        vader_sentiment = "Negative"
elif vader_polarity == 0:
        vader_sentiment = "Neutral"
else:
        vader_sentiment = "Positive"

if afinn_polarity <= 0:
        afinn_sentiment = "Negative"
elif afinn_polarity == 0:
        afinn_sentiment = "Neutral"
else:
        afinn_sentiment = "Positive"

write_sentiment_to_csv(textblob_polarity, vader_polarity,
afinn_polarity, textblob_sentiment, vader_sentiment, afinn_sentiment)

# Step 6
for review in dataframe:
review_title = review["title"]
review_text = review["text"]

emotion_scores = get_emotion_scores(review_title + " " + review_text)

write_emotion_scores_to_csv(emotion_scores)

End
```

in the reviews and to identify patterns that are not apparent by looking at the single words. In this study, bi-gram and tri-gram analysis were performed on reviews. In bi-gram and tri-gram analysis, the frequency of two-word and

TABLE 6. Pseudocode for synonym augmentation.

Algorithm 3: This algorithm preprocess the data, compute polarity score and determine sentiment by using Textblob, VADER and AFINN and get the emotions.

```
Input: TRIPADVISOR_SENTIMENTS_ALL FOR DATA AUG.csv
Output: augmented_reviews dictionary
    1.  Read the CSV file into a Pandas DataFrame
    2.  Separate reviews by restaurant into a restaurant_reviews
        dictionary
    3.  Initialize an augmented_reviews dictionary
    4.  For each restaurant in restaurant_reviews:
            o   Generate augmented reviews using the
                synonym_augmentation function
            o   Add augmented reviews to the
                augmented_reviews dictionary
    5.  Return the augmented_reviews dictionary
Define synonym_augmentation function:
Input: text, num_augmentations Output: augmented_texts list
1.  Tokenize the text
2.  Initialize an augmented_texts list
3.  For each token in the text:
4.  Get the synsets for the token
        o   If there are synsets:
        o   Get the synonyms for the synsets
        o   If there are synonyms:
                ▪   Randomly choose a synonym and add it to
                    the augmented_texts list
            Else:
                ▪   Add the token to the augmented_texts list
        o   Else:
                ▪   Add the token to the augmented_texts list
5.  Return the augmented_texts list
```

three-words sequences that appears consecutively have been extracted.

G. BUSINESS INTELLIGENCE DASHBOARD

All the data obtained through web scrapping, the outcomes of sentiment and emotional analysis, top-n and n-grams analysis and recommendation were compiled a csv file. The next process is to visualize the data into graphical forms to make the appealing, interactive and understandable data representation for the businesses. Subsequently, a business intelligence dashboard for Sarawak gastronomy tourism was created that provides a quantitative overview of the public perceptions towards Sarawak gastronomy industry. By having such a dashboard, the gastronomy and tourisms related industry may help in term of decision making, identify trends and take the countermeasures to rectify any issues as well as to improve the business performance. In this study, Tableau software was used to visualize the data and to create the business intelligence dashboard. Tableau provides user-friendly interface, powerful visualization capabilities, fast and efficient analysis as well as collaboration and sharing capabilities.

IV. FINDINGS AND EXPERIMENTAL RESULTS

This section presents the results of multiple analysis, including overall star rating, emotional analysis, n-grams analysis, sentiment analysis using Textblob, VADER and AFINN, sentiment word clouds, and business intelligence dashboards.

TABLE 7. Pseudocode of feature extraction using BoW.

```
Algorithm 4: Feature extraction using BoW
# Import necessary libraries
# Load the dataset
mydataset = read_csv('your_data_file.csv', encoding='latin-1')

# Extract features and labels from the dataset
features = mydataset.iloc[:, 1:2].astype(str)
labels = mydataset.iloc[:, 47:50].values.tolist()

# Define functions for text cleaning
def clean_text(text):
    # Replace symbols with spaces
    text = replace_symbols_with_spaces(text)
    # Remove non-alphanumeric characters
    text = remove_non_alphanumeric(text)
    # Remove numbers
    text = remove_numbers(text)
    # Remove stopwords
    text = remove_stopwords(text)
    return text

# Preprocess the text data
cleaned_features = features.apply(clean_text)

# Tokenize the cleaned text data
model = Tokenizer()
model.fit_on_texts(cleaned_features)
rep = model.texts_to_matrix(cleaned_features, mode='count')

# Save the Bag of Words representation to a CSV file
df_rep = DataFrame(rep)
df_rep.to_csv("review_titleBOW.csv")
```

The experiment results on the effect of synonym augmentation, feature engineering method and machine learning classifiers will be also presented.

A. STAR RATING

Customers on the Tripadvisor website can give specific star ratings ranging from 0 to 5 to the overall experience, as well as to the quality of food, service, and value. As illustrated in Figure 1, the average ratings for overall experience, food quality, and value were all 4 stars. However, the rating for service quality was slightly lower than the other dimensions. This data implies the need for further investigation into the service aspect to identify the underlying issues before proposing a viable solution. The specific star rating for each restaurant was also provided in the created business intelligence dashboard.

B. EMOTIONAL ANALYSIS

Next, the results from the emotional analysis by using aNRCLex as can be seen in Figure 2. aNRCLex is a powerful tool for extracting valuable insights from text data by quantifying the emotions expressed in the text and providing a detailed understanding of the emotional tone of the text. aNRCLex used dictionary-based approach where a predefined set of emotional categories and associated keywords are used to label and quantify the emotions expressed in the reviews. These emotions are categorized as anger, anticipation, surprise, trust, fear, positive, sadness, negative, disgust and joy. Every word in the lexicon is assigned a score indicating the strength of its association with a particular emotion.

TABLE 8. Pseudocode of feature extraction N-grams.

```
Algorithm 5: Feature extraction using N-grams
# Import necessary libraries
import pandas as pd
from sklearn.feature_extraction.text import CountVectorizer

# Define the path to dataset
dataset_path = 'path/dataset.csv'

# Load the dataset into a DataFrame
df = read_csv(dataset_path)

# Extract the 'Reviews' column from the DataFrame
reviews = df['Reviews']

# Initialize CountVectorizer for n-grams
ngram_vectorizer = CountVectorizer(ngram_range=(1, 2),
max_features=10000)
# You can adjust ngram_range and max_features as needed

# Fit and transform the 'Reviews' data into n-gram features using the
CountVectorizer
ngram_features = ngram_vectorizer.fit_transform(reviews)

# Convert the n-gram features to a dense array
ngram_features_array = ngram_features.toarray()

# Create a DataFrame with the n-gram features
ngram_df = DataFrame(ngram_features_array,
columns=ngram_vectorizer.get_feature_names_out())

# Optionally, add other relevant columns to the n-gram DataFrame
#ngram_df['Restaurant_Name'] = df['Restaurant_Name']
ngram_df['TB_Sentiment'] = df['TB_Sentiment']
ngram_df['Vader_Sentiment'] = df['Vader_Sentiment']
ngram_df['AF_Sentiment'] = df['AF_Sentiment']

# Define the path for saving the n-gram DataFrame as a CSV file
ngram_csv_path = 'path/to/save/ngram_features.csv'

# Save the n-gram DataFrame to the specified CSV file without including
the index
ngram_df.to_csv(ngram_csv_path, index=False)
```

The graph shown in Figure 3 revealed that most of the lexicons found in the reviews have obtained a good score in positive emotions (461.6), followed by joy (251.8), trust (243.7) and anticipation (133.6). Although the score of negative group emotions comprising anger, surprise, fear, sadness, negative and disgust were not that high individually, but if all of them is aggregated (291), the returned score is quite concerning as well. This has implied of some negative sentiments could not be denied and required appropriate attentions from both industry players and government. The emotional analysis shown in Figure 3 reflected the aggregated emotional scores for all restaurants' reviews. However, individual emotional analysis based on the restaurants are also could obtained dynamically by using the created business intelligence dashboard.

TABLE 9. Pseudocode of feature extraction using bow.

```
Algorithm 6: Feature extraction using TDF-IDF
# Import necessary libraries
import pandas as pd
from sklearn.feature_extraction.text import TfidfVectorizer

# Define the path to  dataset
dataset_path = 'path/dataset.csv'

# Load the dataset into a DataFrame
df = read_csv(dataset_path)

# Extract the 'Reviews' column from the DataFrame
reviews = df['Reviews']

# Initialize TfidfVectorizer
tfidf_vectorizer = TfidfVectorizer(max_features=10000)
# You can adjust max_features as needed

# Fit and transform the 'Reviews' data into TF-IDF features using the
TfidfVectorizer
tfidf_features = tfidf_vectorizer.fit_transform(reviews)

# Convert the TF-IDF features to a dense array
tfidf_features_array = tfidf_features.toarray()

# Create a DataFrame with the TF-IDF features
tfidf_df = DataFrame(tfidf_features_array,
columns=tfidf_vectorizer.get_feature_names_out())

# Optionally, add other relevant columns to the TF-IDF DataFrame
tfidf_df['Restaurant_Name'] = df['Restaurant_Name']
tfidf_df['TB_Sentiment'] = df['TB_Sentiment']
tfidf_df['Vader_Sentiment'] = df['Vader_Sentiment']
tfidf_df['AF_Sentiment'] = df['AF_Sentiment']

# Define the path for saving the TF-IDF DataFrame as a CSV file
tfidf_csv_path = 'path/to/save/tfidf_features.csv'

# Save the TF-IDF DataFrame to the specified CSV file without including
the index
tfidf_df.to_csv(tfidf_csv_path, index=False)
```

C. N-GRAMS ANALYSIS

Figure 4 and Figure 5 show the word cloud of bi-gram and tri-gram analysis respectively. By performing n-grams analysis on the reviews, the words and phrases that were appeared frequently could be identified. For instance, the positive words may represent strength or unique selling points of the business, while negative words may indicate a problem that need to be addressed by the business. The n-gram analysis provided in this section represented overall analysis for all restaurants.

Based on the bi-gram analysis depicted in Figure 4, the phrases were all positive words in which the phrases 'food good'(64), 'good food (48)', 'great food (39)', 'service good' (38), 'friendly staff' (38) were all the top 5 of frequent phrases mentioned in the reviews. It is also worth noting that the word 'laksa sarawak' was also among the frequent words (59) written in the reviews.

Meanwhile, the word cloud of tri-gram analysis presented in Figure 5 shown the phrases 'good value money' (48), 'service good food' (39), 'great value money' (38), 'food

TABLE 10. Pseudocode of feature extraction using Word2Vec.

```
Algorithm 7: Feature extraction using Word2Vec
# Import necessary libraries
import pandas as pd
import numpy as np
from gensim.models import Word2Vec
from sklearn.preprocessing import LabelEncoder
from gensim.models import KeyedVectors

# Define the path to dataset
dataset_path = 'path/dataset.csv'

# Load the dataset into a DataFrame
df = read_csv(dataset_path)

# Define the path to the pre-trained Word2Vec model
word2vec_model_path = 'path/word2vec/model.bin'

# Load the pre-trained Word2Vec model
word2vec_model =
KeyedVectors.load_word2vec_format(word2vec_model_path,
binary=True)

# Define a function to average word vectors for a review
function average_word_vectors(words, model, vocabulary,
num_features):
  feature_vector = create_array_of_zeros(num_features, float32)
  nwords = 0
  for word in words:
    if word is in vocabulary:
      nwords += 1
      feature_vector = add_vectors(feature_vector, model[word])
  if nwords > 0:
    feature_vector = divide_vector(feature_vector, nwords)
  return feature_vector

# Prepare data for feature extraction
reviews = get_column(df, 'Reviews')
review_words = split_reviews_into_words(reviews)
num_features = get_vector_size(word2vec_model)
vocabulary = create_set(get_vocab_keys(word2vec_model)) # Use the
vocab attribute

# Extract features using Word2Vec embeddings
features = []
for review in review_words:
  feature = average_word_vectors(review, word2vec_model, vocabulary,
num_features)
  add_feature_to_list(features, feature)

# Convert features list to a NumPy array
features_array = convert_list_to_numpy_array(features)

# Encode sentiment labels
label_encoder = create_label_encoder()
labels = fit_transform_labels(label_encoder, get_column(df,
'AF_Sentiment')) # Assuming 'AF_Sentiment' is the sentiment column

# Add features and labels to a new DataFrame
feature_columns = generate_feature_columns(num_features)
feature_df = create_dataframe_with_features(features_array,
feature_columns)

# Add other relevant columns to the feature DataFrame if needed
#feature_df['Restaurant_Name'] = get_column(df, 'Restaurant_Name')
feature_df['TB_Sentiment'] = get_column(df, 'TB_Sentiment')
feature_df['Vader_Sentiment'] = get_column(df, 'Vader_Sentiment')
feature_df['AF_Sentiment'] = get_column(df, 'AF_Sentiment')

# Define the path for saving the feature DataFrame as a CSV file
features_csv_path =
'path/to/save/word_embedding_features_word2vec.csv'

# Save the feature DataFrame to the specified CSV file without including
the index
save_dataframe_to_csv(feature_df, features_csv_path)
```

TABLE 11. Pseudocode for top words, n-grams analysis and sentiment word clouds.

Algorithm 8: This algorithm analyse and clean the sentiment keywords and the reviews to determine the top words in the reviews and, the bi-grams and tri-gram for the sentiment keywords.
Start **1.** Read the CSV file containing the reviews and sentiment keywords and store the data in a dataframe. 2. Create a function named get_top_n_words: 2.1 Create a vocabulary vector. 2.2 Create a bag of words to calculate the frequency of each vocabulary. 2.3 Return the top 20 vocabulary frequencies. 3. Call the function get_top_n_words to get the top 20 vocabulary frequencies. 4. Create a function named get_top_n_bigram: 4.1 Define ngram_range=(2,2) as vocabulary. 4.2 Create a bag of words to calculate the frequency of each vocabulary. 4.3 Return the top 20 vocabulary frequencies. 5. Call the function get_top_n_bigram to get the top 20 vocabulary frequencies. 6. For all restaurants: 6.1 Read the CSV file containing the cleaned sentiment keywords from AFINN and Textblob. 6.2 Call the WordCloud function to plot the word cloud. 6.3 Save the word cloud as a PNG format. 7. Create a function named clean_text: 7.1 Remove stop words. 7.2 Remove symbols. 8. Apply the function clean_text on the reviews. 9. Visualize the results of bi-gram and tri-gram analysis in a word cloud. End

reasonable price' and 'best laksa sarawak' (36) were the top five words mentioned in the reviews. This has indicated that the food prices offered in the restaurants are reasonable and worthwhile. Other than 'laksa Sarawak' and 'mi kolok' which are the well-known and popular dish in Sarawak, it was also noticed that 'midin fried rice' was also highlighted more often in the reviews which could become the new trending dish in Sarawak.

D. LEXICON-BASED SENTIMENT ANALYSIS

In this section, the outcomes from the sentiment analysis by using Textblob, VADER and AFINN are presented. The frequency of sentiment words or lexicon for negative, positive and neutral is presented in the word cloud shown in Figure 6. The most frequent lexicon in the review are 'good'(1128), 'great'(730), 'nice' (648), friendly (379), 'best' (348) and 'delicious' (327).

E. CATEGORY OF SENTIMENTS

The percentages of sentiments analyzed using TextBlob, VADER and AFINN for all restaurants shown in Figure 7, Figure 8 and Figure 9 indicate that the majority of the reviews could be categorized as positive reviews. On the other hand, 12.12%, 8.80% and 7.13% of the reviews were classified as negative reviews.

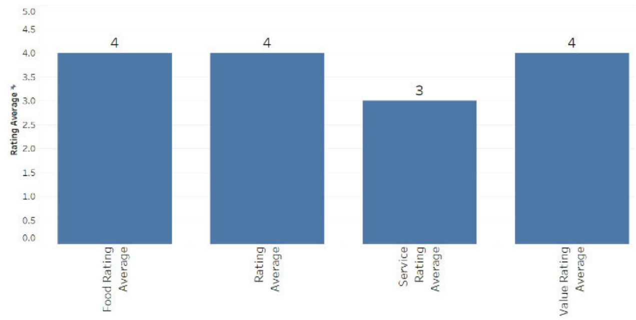

FIGURE 1. Overall star rating of the restaurants.

F. SENTIMENTS POLARITY SCORE

The degree of positive, negative and neutral sentiments category of sentiments of the reviews was determined based on the calculations of sentiments polarity. In Textblob, a sentiment polarity score greater than 0 indicates a positive phrase, while a score less than 0 indicates a negative phrase. A score of exactly 0 signifies a neutral sentiment. The histogram of sentiment polarity by using Textblob is depicted in Figure 10. According to this histogram, approximately 76.66% of the reviews have received a polarity score between 0.1 and 0.5. Interestingly, many negative reviews have a score between 0 and −0.2, indicating that they may not severe or intense.

The polarity score yielded by using AFINN as shown in Figure 11 demonstrated identical pattern with the Textblob. As explained in Section III-B, the polarity score less than 0 indicate negativity and the polarity score greater than 0 indicate positivity. The maximum value of polarity score is 25, while the minimum value of polarity score is −11.

Figure 12 shown the histogram of polarity score for VADER. This finding implies many reviews obtained high polarity score in between 0.6 and 0.9 that indicate the great level of reviews. However, there quite numbers of reviews obtained very low polarity score between −0.7 and 1.0 which indicate the very poor customer experience and required further attentions.

G. SENTIMENTS CLASSIFICATION BASED ON RESTAURANTS

Figures 13, 13, and 14 show comparisons of sentiment analysis results for all restaurants using the three sentiment analysis techniques. It was discovered that the restaurants AYA, BIN, CHO, IZA, NUS, and VIL had clearly gotten more unfavourable reviews than the other restaurants on Textblob as shown in Figure 13. Nonetheless, there are no issues about the restaurants SAR, SPO, and TRI.

Following this, it can be noticed from the sentiment analysis carried out using AFINN as shown in Figure 14 that restaurants HER, NUS, and VIL have received more negative reviews. In contrast, there are no detrimental reviews for the restaurants BAL, RAI, and RUM.

Last but not least, the sentiment analysis performed using VADER in Figure 15 revealed that the reviews for SAR and

FIGURE 2. Emotions word cloud from emotional analysis in gastronomy tourism dataset.

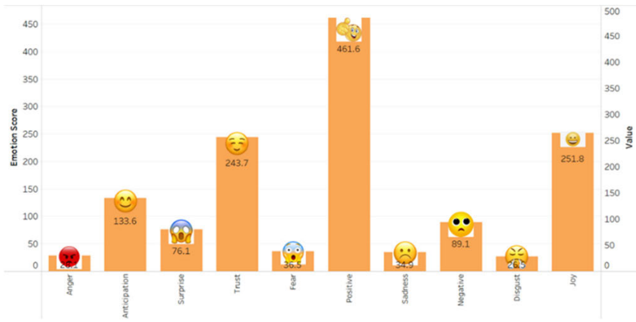

FIGURE 3. The distribution of the types of emotions towards the restaurants.

SPO were flawlessly perfect. Nonetheless, the restaurants HER, IZA, JAM, KOP, NUS, and VIL have received lot of undesirable reviews.

In a short, at least two sentiment analysis approaches consistently found that the restaurants IZA, NUS, VIL, and HER received a lot of unpleasant reviews. Meanwhile, the restaurants NUS and VIL were received many unfavourable reviews on all sentiment analysis approaches.

H. AVERAGE OF SENTIMENTS POLARITY SCORE FOR EACH RESTAURANTS

The average of sentiments polarity score by using Textblob, AFINN and VADER for each restaurant were derived and ranked to get some comparisons on overall customer's feeling towards the restaurant as illustrated in Figure 16, 17 and 18. Figure 16 shows the average of sentiment polarity using Textblob for all restaurants. The finding shows the RAI, COM, RAS, BLA and LAK are the top 5 restaurants with an average polarity of 0.4812, 0.4418, 0.4356, 0.4157 and 0.4002. Otherwise, the restaurants VIL, NUS, RUM, AYA, BLK and SPO are the bottom 5 with an average polarity of 0.1958, 0.2119, 0.2578, 0.2559 and 0.2843. Overall, every restaurant has an average rating of more than 0, which indicates the substantial meritorious reviews that inundate the adverse reviews.

The sentiment polarity scores on AFFIN indicate that the restaurants RAI, RAS, LIM, BLA, and COM have a high average score of 10.60, 9.57, 8.65, 8.35, and 8.00, respectively, as shown in Figure 17. In contrast, the restaurants HER,

FIGURE 4. Word cloud of Bi-grams analysis.

FIGURE 5. Word cloud of Tri-gram analysis.

SAR, CHO, NUS, and KOP received the lowest 5 scores, with 4.72, 4.50, 4.26, 3.85, and 3.84, respectively.

Figure 18 shows that the top five restaurants with the highest sentiment polarity scores for VADER are RAI (0.83), RAS (0.81), COM (0.79), LIM (0.77), and KAN (0.77). On the other hand, the bottom five restaurants with the lowest sentiment polarity scores are SPO (0.55), CHO (0.50), VIL (0.49), KOP (0.45), and NUS (0.44).

To conclude, the restaurants RAI, COM, RAS, LIM, BLA could be assumed to experience amiable perceptions from the previous customers due to persistent and best sentiment polarity score across Textblob, AFINN and VADER. However, there are also few restaurants need to have some factors in order to provide better dining experiences for the customers, especially the restaurants NUS, VIL and SPO.

I. SENTIMENT ANALYSIS BASED ON CUISINE TYPES

This analysis enables us to make meaningful comparisons of customers' sentiments across various cuisines, as shown in Figure 19, 20 and 21. Based on the sentiment analysis conducted by using Textblob in Figure 19, the Asian fusion, International fusion, Italian, Pizza and seafood cuisines have less negative reviews as compared to the other type cuisines. This finding is also consistent with the sentiment analysis conducted by using AFINN as shown in Figure 20. Meanwhile, the Indonesia and Japanese cuisines were received quite number of mischievous reviews from the customers.

V. MACHINE LEARNING AND FEATURE ENGINEERING EVALUATION FOR GASTRONOMY SENTIMENT ANALYSIS

The first level of sentiment analysis utilized the lexicon-based sentiment analysis using AFFIN, Vader and Textblob to determine the sentiment category from each of the restaurants'

FIGURE 6. Word cloud of sentiment lexicon.

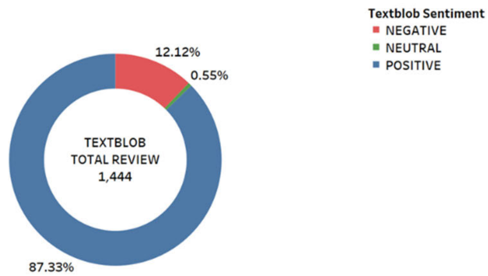

FIGURE 7. Sentiment analysis using Textblob.

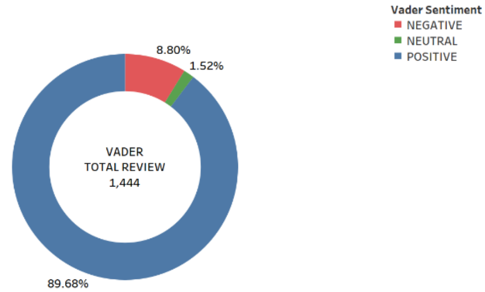

FIGURE 8. Sentiment analysis using VADER.

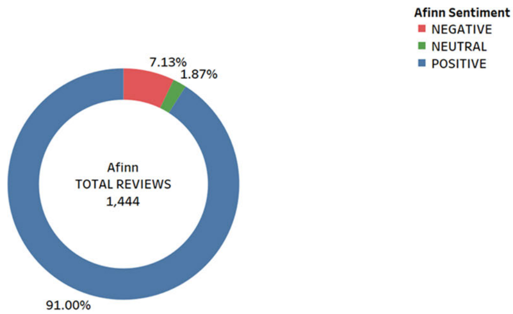

FIGURE 9. Sentiment analysis using AFINN.

reviews. In the second level of sentiment analysis, several machine learning algorithms were employed to train the sentiment classification model and the results pertaining classification performance are discussed in this section. The classification models were trained by using deep learning-based algorithms which are Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) and CNN-LSTM, as well as traditional machine learning algorithms comprising Linear Support Vector Machine (LSVM), Support Vector Machine (SVM), Gradient Boosting Classifier (GBC), Stochastic Gradient Descent (SGD), Multilayer Perceptron (MLP), Random Forest (RF) and k-Nearest Neighbour (kNN). The classification performance are measured based on accuracy, ROC-AUC and F1-Score.

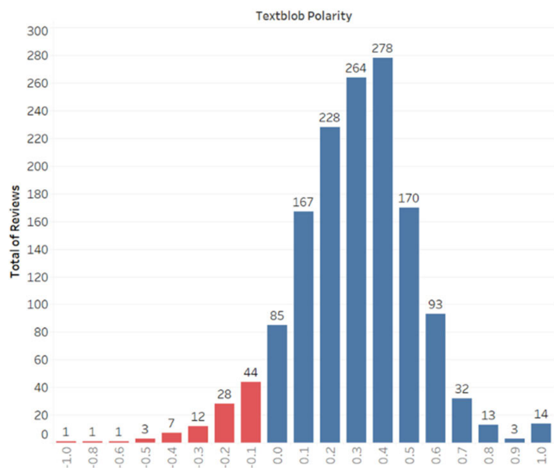

FIGURE 10. Histogram of Textblob polarity.

A. MACHINE LEARNING PERFORMANCE VERSUS LEXICON SENTIMENT ANALYSIS

1) AFINN

In the case of AFINN sentiment analysis, the Random Forest (RF) model exhibits the highest accuracy (0.9446) among all machine learning algorithms. In a flipside, its F1-Score (0.5244) is relatively lower, suggesting that it might struggle with class imbalance or may have precision-recall trade-offs. This means that while RF is good at overall classification, it might not perform as well in distinguishing between positive and negative sentiments. The CNN model achieves a remarkably high F1-Score of 0.9347, indicating a strong balance between precision and recall. K-Nearest Neighbors (KNN) also achieves a relatively high accuracy of 0.9376, while Linear Support Vector Machine (LSVM) and Support Vector Machine (SVM) perform consistently well across metrics. Gradient Boosting (GBC) and Stochastic Gradient Descent (SGD) models show varying performance, with GBC achieving higher accuracy but lower F1-Score.

2) TEXTBLOB

TextBlob sentiment analysis models demonstrate solid performance across the board. Gradient Boosting (GBC) leads with a relatively high Accuracy of 0.8885 and a strong Roc-Auc of 0.8308. Linear Support Vector Machine (LSVM) also achieves a balanced accuracy and Roc-Auc. Notably, the CNN model stands out with the highest F1-Score of 0.8508,

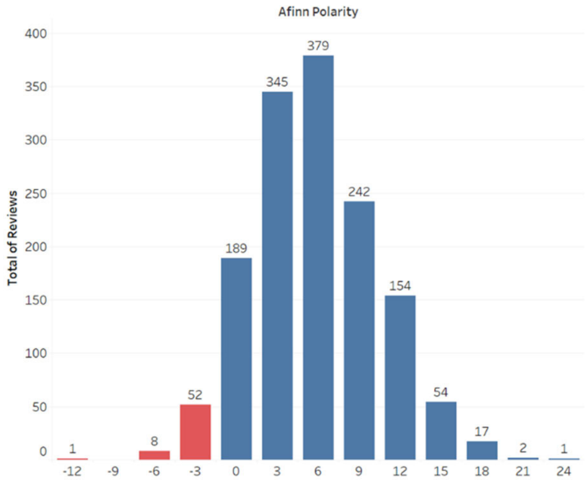

FIGURE 11. Histogram of AFINN polarity.

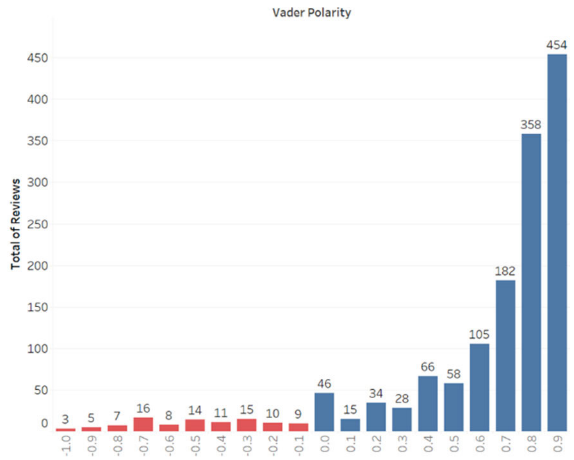

FIGURE 12. Histogram of VADER polarity.

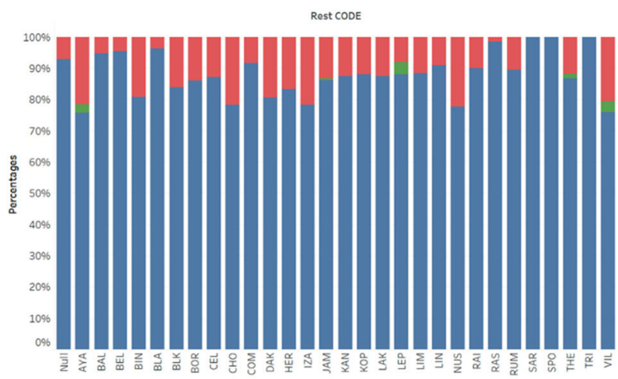

FIGURE 13. Sentiment analysis for all restaurants using Textblob.

indicating its ability to effectively classify restaurant reviews based on sentiment.

3) VADER

VADER sentiment analysis models showed some variability in performance, generally exhibit good accuracy in classifying restaurant reviews. The Multilayer Perceptron (MLP) model leads with an accuracy of 0.9261, a high Roc-Auc of 0.9192, and a competitive F1-Score of 0.5339. Linear Support Vector Machine (LSVM) also achieves solid results across all metrics. The CNN model, while scoring high in F1-Score (0.8996), demonstrates lower accuracy and Roc-Auc, indicating some limitations in broader sentiment analysis tasks.

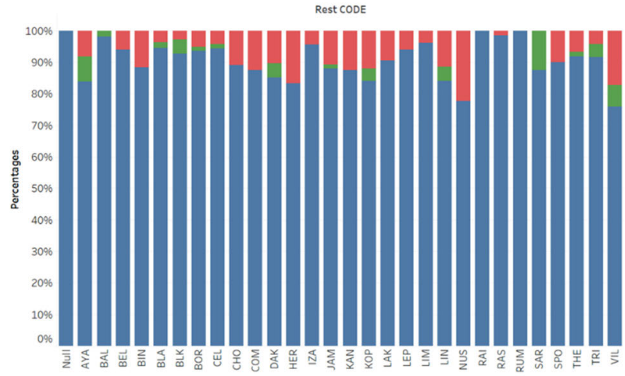

FIGURE 14. Sentiment analysis for all restaurants using AFINN.

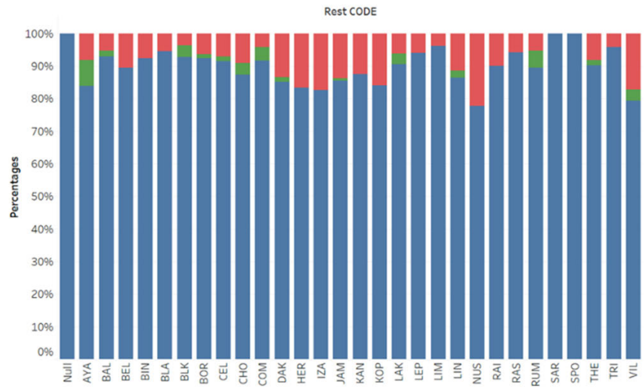

FIGURE 15. Sentiment analysis for all restaurants using VADER.

4) DISCUSSIONS

Based on the results shown in Table 12, the choice of lexicon-based sentiment analysis method (AFINN, TextBlob, or VADER) had a significant impact on model performance. Random Forest (RF) consistently performed well across all lexicon methods, achieving high accuracies and ROC-AUC scores in most cases. Linear Support Vector Machine (LSVM) and Support Vector Machine (SVM) also consistently performed well. Deep learning models, such as Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM), showed competitive results, with CNN often having the highest F1-Score. Gradient Boosting Classifier (GBC) achieved high accuracy in TextBlob-based sentiment analysis but had a lower ROC-AUC score.

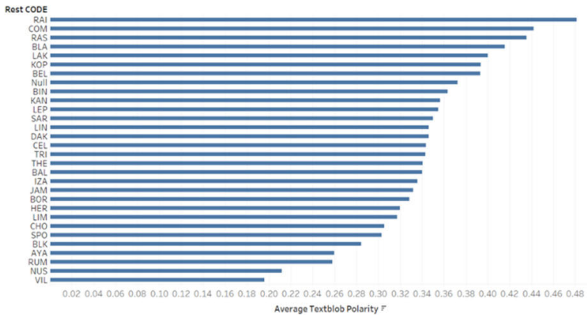

FIGURE 16. Average of sentiment polarity using Textblob for all restaurants.

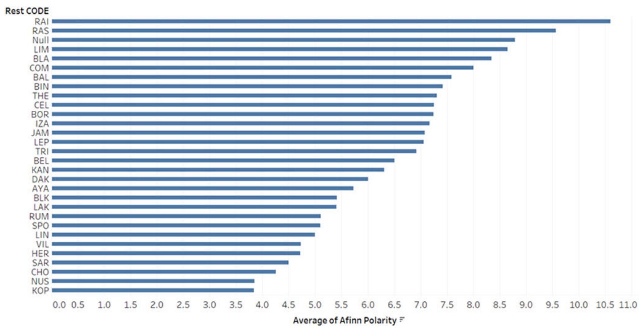

FIGURE 17. Average of sentiment polarity using AFINN for all restaurants.

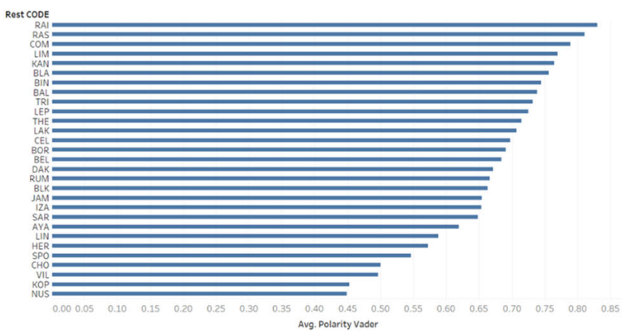

FIGURE 18. Average of sentiment polarity using VADER for all restaurants.

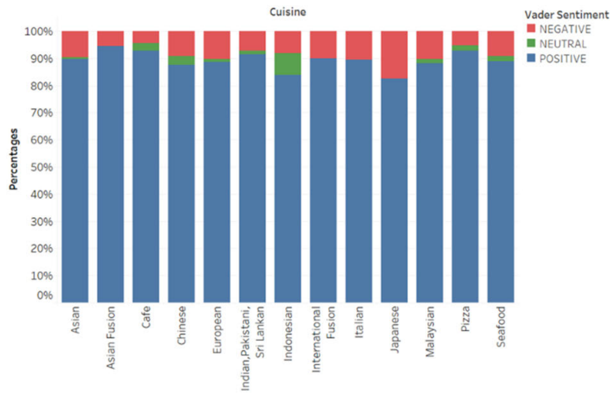

FIGURE 19. Sentiment analysis based on cuisine types using Textblob.

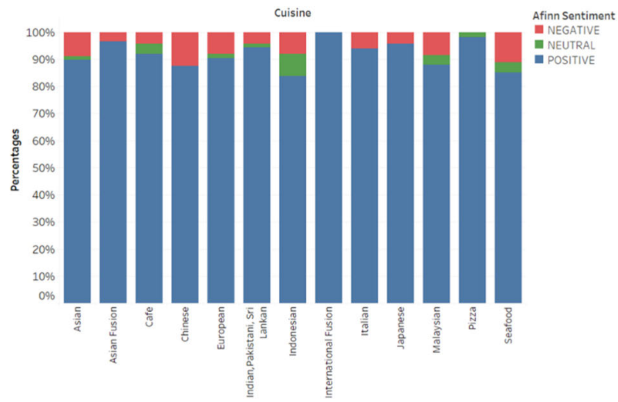

FIGURE 20. Sentiment analysis based on cuisine types using AFINN.

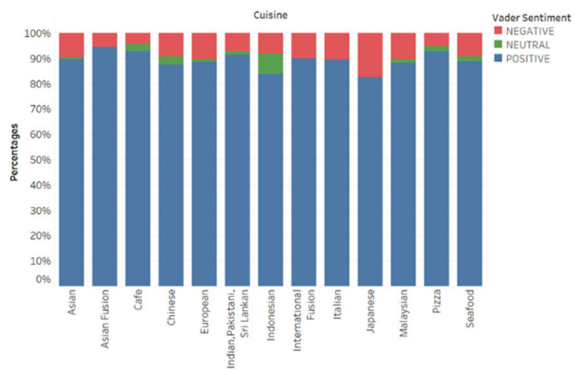

FIGURE 21. Sentiment analysis based on cuisine types using VADER.

Multi-Layer Perceptron (MLP) achieved the highest accuracy among VADER-based models but did not consistently have the highest F1-Score.

These insights indicate that different combinations of lexicon-based sentiment analysis and machine learning algorithms can yield varying results. The choice of the most suitable combination depends on the specific requirements of the sentiment analysis task, including the importance of precision, recall, accuracy, and ROC-AUC score. Further tuning and experimentation may help identify the best-performing model for a particular application. Determining the "best" sentiment analysis model depends on your specific use case and the prioritized performance metric. If accuracy is the main concern, Random Forest (RF) achieved the highest accuracy in the AFINN-based sentiment analysis. Among TextBlob-based models, Gradient Boosting Classifier (GBC) had the highest accuracy. For VADER-based sentiment analysis, Multi-Layer Perceptron (MLP) had the highest accuracy. If a balanced model that provides good precision and recall (F1-Score) is important, Convolutional Neural Network (CNN) performed well across different lexicon-based sentiment analyses. It achieved the highest F1-Score in both AFINN and TextBlob-based models. If the concerns are about the ability to distinguish between positive and negative sentiment, the ROC-AUC score should become the highlights. Linear Support Vector Machine (LSVM) achieved the

TABLE 12. Baseline performance of machine learning sentiment analysis.

Lexicon Sentiment Analysis	Machine Learning Algorithms	Accuracy	Roc-Auc	F1-Score
AFINN	RF	**0.9446**	0.8476	0.5244
	CNN	0.9413	0.6746	**0.9347**
	KNN	0.9376	0.6939	0.4766
	LSVM	0.9238	**0.8954**	0.5573
	SVM	0.9215	0.8545	0.3197
	SGD	0.9178	0.5692	0.3419
	CNN-LSTM	0.917	0.6185	0.8772
	MLP	0.9122	0.8538	0.4697
	GBC	0.9099	0.1721	0.3688
	LSTM	0.8866	0.5094	0.87
TEXTBLOB	GBC	**0.8885**	0.8308	0.4517
	LSVM	0.8845	**0.904**	0.477
	RF	0.8822	0.7343	0.4191
	CNN	0.8805	0.5699	**0.8508**
	KNN	0.8799	0.6008	0.4147
	MLP	0.8776	0.8591	0.4302
	SVM	0.8661	0.6864	0.3094
	SGD	0.8661	0.5	0.3094
	CNN-LSTM	0.8581	0.4603	0.7926
	LSTM	0.8454	0.5084	0.7976
VADER	MLP	**0.9261**	0.9192	0.5339
	LSVM	0.9238	**0.9341**	0.5206
	CNN	0.9203	0.6116	**0.8996**
	RF	0.9192	0.8242	0.4514
	KNN	0.9099	0.6709	0.4454
	GBC	0.8985	0.5702	0.3467
	LSTM	0.8969	0.4858	0.8366
	SVM	0.8915	0.82	0.3142
	SGD	0.8915	0.5391	0.3142
	CNN-LSTM	0.8858	0.5685	0.8322

highest ROC-AUC score in both AFINN and TextBlob-based models. Among VADER-based models, LSVM and MLP had strong ROC-AUC scores. In term of computational complexity, Deep learning models like CNN and Long Short-Term Memory (LSTM) may require more computational power and training data compared to traditional machine learning models like Random Forest (RF) or Support Vector Machines (SVM). Simpler models like Random Forest (RF) are often more interpretable compared to complex deep learning models.

B. EFFECT OF SYNONYM AUGMENTATION

The experiment have explored the effect of synonym data augmentation on gastronomy tourism sentiment analysis on AFINN, TextBlob, and VADER. The results are presented machine learning models under two conditions, "NO" (without data augmentation) and "YES" (with data augmentation) as shown in Table 13, Figure 22, Figure 23, and Figure 24.

TABLE 13. Effect of synonym augmentation on machine learning sentiment analysis.

Lexicon SA	ML	Data Augmentation F1-Score		Roc-Auc		Accuracy	
		NO	YES	NO	YES	NO	YES
AFINN	CNN	0.9347	**0.9975**	0.6746	0.8186	0.9413	0.9976
	CNN-LSTM	0.8772	0.8816	0.6185	0.8212	0.9170	0.9129
	GBC	0.3688	0.8519	0.1721	0.9012	0.9099	0.9673
	KNN	0.4766	0.9941	0.6939	**1.0000**	0.9376	**0.9977**
	LSTM	0.8700	0.8744	0.5094	0.4938	0.8866	0.9061
	LSVM	0.5573	0.9911	0.8954	0.9995	0.9238	0.9965
	MLP	0.4697	0.7848	0.8538	0.9763	0.9122	0.9638
	RF	0.5244	0.8368	0.8476	0.9924	0.9446	0.9700
	SGD	0.3419	0.9760	0.5692	0.9990	0.9178	0.9908
	SVM	0.3197	0.8896	0.8545	0.9999	0.9215	0.9761
TEXTBLOB	CNN	0.8508	0.9933	0.5699	0.8187	0.8805	0.9934
	CNN-LSTM	0.7926	0.8369	0.4603	0.7978	0.8581	0.8604
	GBC	0.4517	0.6784	0.8308	0.7774	0.8885	0.8746
	KNN	0.4147	**0.9965**	0.6008	0.9989	0.8799	**0.9977**
	LSTM	0.7976	0.8134	0.5084	0.5054	0.8454	0.8674
	LSVM	0.4770	0.9876	0.9040	**0.9994**	0.8845	0.9919
	MLP	0.4302	0.8990	0.8591	0.9913	0.8776	0.9704
	RF	0.4191	0.9134	0.7343	0.9979	0.8822	0.9519
	SGD	0.3094	0.9576	0.5000	0.9965	0.8661	0.9738
	SVM	0.3094	0.9187	0.6864	0.9972	0.8661	0.9627
VADER	CNN	0.8996	**0.9929**	0.6116	0.8107	0.9203	0.9930
	CNN-LSTM	0.8322	0.8641	0.5685	0.8460	0.8858	0.8984
	GBC	0.3467	0.0084	0.5702	0.5000	0.8985	0.0127
	KNN	0.4454	0.9914	0.6709	0.9946	0.9099	**0.9958**
	LSTM	0.8366	0.8516	0.4858	0.4929	0.8969	0.9001
	LSVM	0.5206	0.9807	0.9341	0.9973	0.9238	0.9908
	MLP	0.5339	0.8039	0.9192	0.9669	0.9261	0.9627
	RF	0.4514	0.9047	0.8242	0.9980	0.9192	0.9658
	SGD	0.3142	0.9412	0.5391	0.9973	0.8915	0.9792
	SVM	0.3142	0.9069	0.8200	**0.9995**	0.8915	0.9719

For the AFINN Lexicon sentiment analysis from Figure 22, CNN model achieved a high F1-Score of 0.9975 with data augmentation. This substantial increase in F1-Score indicates a significant boost in precision and recall balance, which is vital for sentiment classification. A similar trend is observed across various other models, such as CNN-LSTM, GBC, KNN, and LSVM. Roc-Auc that measuring the models' abil-

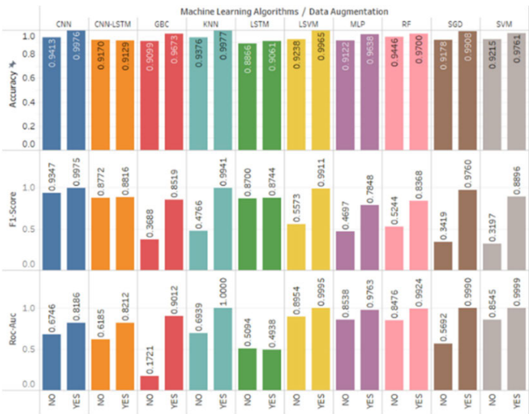

FIGURE 22. Effect of data augmentation (AFFIN).

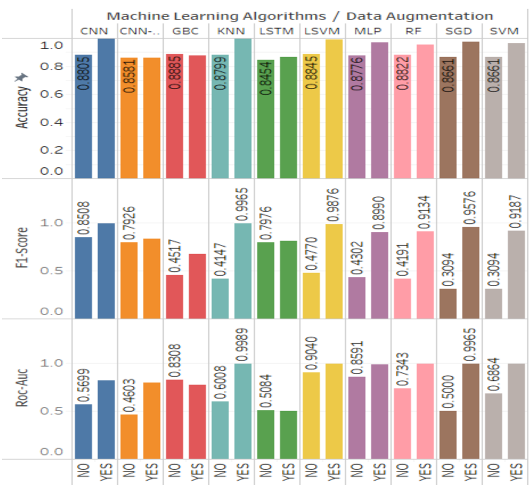

FIGURE 23. Effect of data augmentation (Textblob).

ity to discriminate between positive and negative reviews, also show a marked improvement with data augmentation. The accuracy metric has also improved, reflecting a reduction in misclassifications and an overall improvement in prediction accuracy with data augmentation. CNN, KNN and LSVM and may work well with AFINN.

In the case of TextBlob sentiment analysis as presented in Figure 23, data augmentation similarly enhances model performance across various machine learning models. F1-Scores, Roc-Auc values, and Accuracy metrics see notable improvements. The KNN model demonstrates the most substantial improvement in F1-Score when data augmentation is applied, increasing from 0.4147 to 0.9965. This signifies that data augmentation significantly enhances its precision and recall balance, making it highly effective in sentiment analysis. The KNN model stands out with a remarkable ROC-AUC of 0.9989 with data augmentation, showcasing its strong ability to discriminate between positive and negative sentiments. Linear Support Vector Machine (LSVM) also benefits greatly from data augmentation, with its F1-Score improving from 0.477 to 0.9876 and achieving a high ROC-AUC of

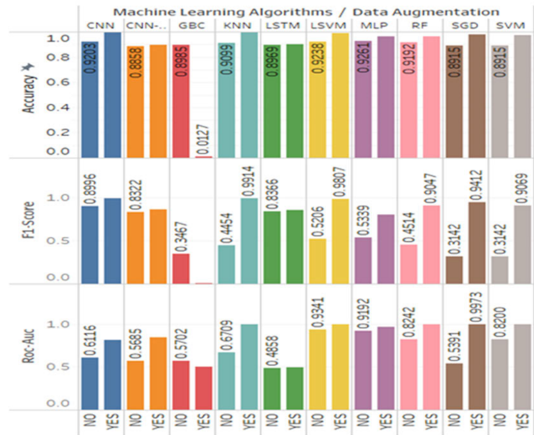

FIGURE 24. Effect of data augmentation (Vader).

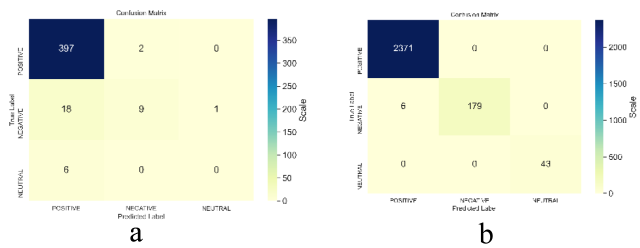

FIGURE 25. (a) No data augmentation (b) With data augmentation (KNN).

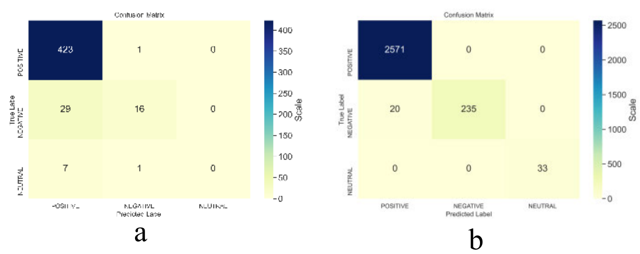

FIGURE 26. (a) No data (b) With data augmentation (CNN).

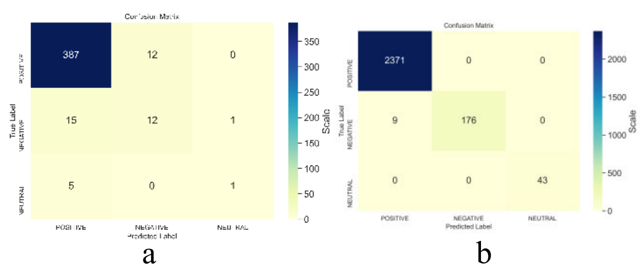

FIGURE 27. (a) No data augmentation (b) With data augmentation (LSVM).

0.9919 with data augmentation. Overall, KNN and LSVM demonstrated the best performance as it has precision and recall balance and discrimination capability.

VADER sentiment analysis, however, demonstrates different results as depicted in Figure 24. Several models, such as KNN, LSTM, LSVM, and RF, benefit from data augmentation by experiencing substantial improvements in F1-Score,

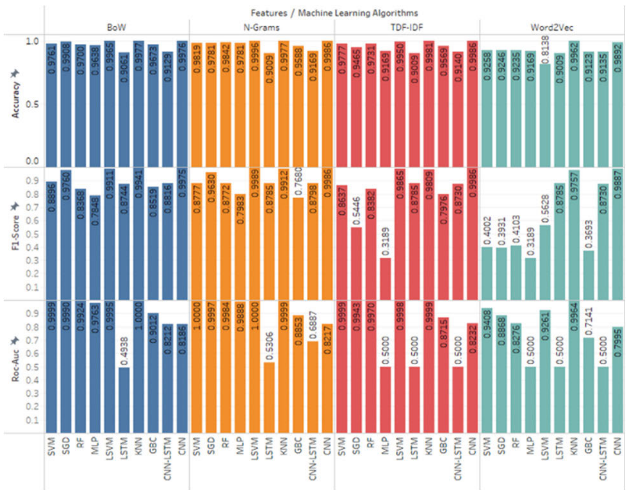

FIGURE 28. The model performance of different feature engineering techniques.

FIGURE 29. The model performance of different machine learning techniques.

ROC-AUC, and accuracy. On the other hand, models like GBC and SVM do not benefit significantly from data augmentation, with low F1-Scores and ROC-AUC values. KNN achieves the highest F1-Score of 0.9914, indicating that it effectively balances precision and recall. KNN also exhibits an exceptional ROC-AUC of 0.9946, demonstrating its ability to discriminate between positive and negative sentiments with high accuracy. Additionally, KNN maintains a high accuracy of 0.9958, making it the top performer across multiple metrics.

In summary, data augmentation consistently enhances sentiment analysis in AFINN and TextBlob sentiment analysis tools, leading to improved F1-Scores, higher Roc-Auc values, and increased Accuracy across various machine learning algorithms. However, the impact of data augmentation on VADER sentiment analysis appears to vary, with some models benefiting from it and others demonstrating more inconsistent results. The choice of the sentiment analysis tool and specific machine learning model is critical, as this can influence the effectiveness of data augmentation. Further exploration and fine-tuning may be required to optimize sentiment analysis for restaurant reviews using these tools.

C. EFFECT OF FEATURES ENGINEERING

The evaluation on the feature engineering techniques is conducted in order to assess how different feature engineering impact the performance of sentiment analysis models. Besides, we can determine set of features is the most effective for sentiment analysis task especially on the specific use case. Figure 28 shows the results of restaurant sentiment analysis experiments based on different feature engineering techniques (Bag of Words - BoW, N-Grams, TF-IDF, and Word2Vec) combined with various machine learning algorithms.

Based on the findings in Figure 28, it is evident that BoW, N-Grams, and TDF-IDF consistently demonstrated strong performance across all machine learning algorithms, with N-Grams in combination with LSVM achieving the highest accuracy (0.9996). Conversely, the accuracy of machine learning models utilizing Word2Vec features appears to exhibit a slight decrease. In another view, the performance of the machine learning models on all these features are fluctuate in terms of F1-Score and ROC-Auc. Several machine learning models applied to Word2Vec and TDF-IDF features have produced subpar F1-Scores. However, it's noteworthy that CNN and KNN stand out for consistently delivering high F1-Scores across all feature types. Additionally, BoW and N-grams features have also demonstrated commendable F1-Scores, with the highest observed in the case of N-Grams paired with LSVM (0.9989). The results for ROC-AUC across all models have been notably favorable for both BoW and N-Grams features. However, it's worth noting that LSTM, CNN, and CNN-LSTM have exhibited comparatively less impressive ROC-AUC results in all feature types. The most outstanding ROC-AUC performance was achieved by KNN when using BoW features, boasting a perfect score of 1.00. Similarly, LSVM in conjunction with N-Grams also demonstrated exceptional ROC-AUC results. BoW and N-Grams appear to be highly effective feature engineering techniques for restaurant sentiment analysis across a wide range of machine learning algorithms. TF-IDF's performance is more mixed and depends on the specific algorithm used. Word2Vec did not perform well in this context and may not be suitable for sentiment analysis of restaurant reviews. The choice of feature engineering technique and machine learning algorithm should be based on the specific requirements of the sentiment analysis task, the available data, and the desired performance metrics. In this experiment, BoW and N-Grams stand out as strong choices for feature engineering.

D. EFFECT OF MACHINE LEARNING ON FEATURES

Figure 29 presents the results of restaurant sentiment analysis experiments, specifically focusing on the impact of different machine learning classifiers on various feature engineering techniques (BoW, N-Grams, TF-IDF, and Word2Vec).

FIGURE 30. Sentiment word clouds for all restaurants.

The results shown in Figure 29 has revealed that the KNN, CNN, SVM and LSVM have robust performance on all feature types. On the contrary, the performance of CNN, CNN-LSTM, and LSTM appears to be consistently lower in terms of ROC-AUC across all feature types. Additionally, it's evident that SVM, RF, SGD, GBC, and MLP have underperformed, particularly in terms of F1-Score.

Sentiments Word Cloud and Review Samples

A word cloud is used to understand the most commonly used words, in this context sentiment lexicons in the reviews for individual restaurant as shown in Figure 30. Word cloud displays the most frequently occurring sentiment lexicons in a larger font size, making it easy to see which words appears most often. By analysing the words in a word cloud, the most common sentiment-related lexicons can be quickly identified and to get a sense of the overall sentiment expressed in the reviews. Table 14 and Table 15 provide ten samples of positive and negative review titles that were taken randomly.

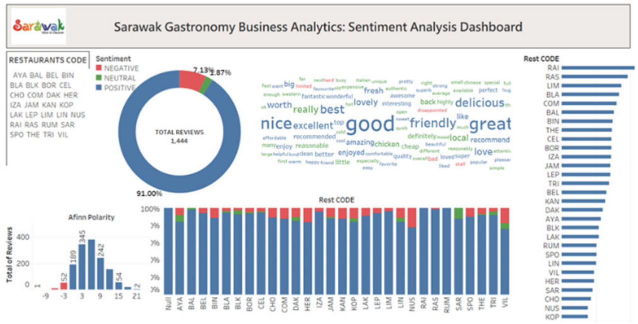

FIGURE 31. Sentiment analysis dashboard.

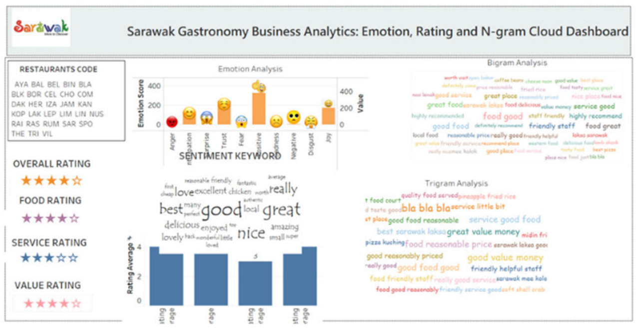

FIGURE 32. Emotion, rating and N-grams analysis dashboard.

TABLE 14. Positive reviews.

No	Samples of Positive Reviews Title
1.	Wonderful Pizza in town...
2.	Good Italian Restaurant in Kuching
3.	Great coffee
4.	Food very tasty in nice setting
5.	Great freshly cooked food and juices!
6.	Tasty Laksa
7.	Best Sarawak Laksa and more!!!
8.	Great Native Food of Sarawak
9.	Friendly atmosphere and vegan food options
10.	Delicious indigenous food

VI. SARAWAK GASTRONOMY BUSINESS INTELLIGENCE DASHBOARD

The present study demonstrates the development of a Tableau-based business intelligence (BI) dashboard that allows for visualizing important indicators related to sentiment analysis outcomes. The dashboard has been designed specifically for food businesses, to facilitate real-time monitoring of customer sentiment and feedback. By providing easy access to critical information, prompt identification of areas of concern, and efficient data comprehension, this

TABLE 15. Negative reviews.

No	Samples of Negative Reviews Title
1.	Disappointed with their attitude and services.
2.	Disappointed with some racist casher.
3.	quite expensive
4.	Good but expensive
5.	Not veggie friendly
6.	Sushi Unfresh
7.	Terrible service
8.	taste ok.but very poor service
9.	maybe a LITTLE overrated...
10.	Overprized, not very special

dashboard has significant potential to enhance the performance of food businesses. The dashboard comprises two types, which have been depicted in Figures 31 and Figure 32.

Figure 31 depicts a sophisticated business intelligence (BI) dashboard that enables the visualization of sentiment analysis outcomes for both individual restaurants and overall sentiment. The dashboard comprises several components, including restaurant codes, percentages of positive, negative, and neutral sentiments, a sentiment lexicon word cloud, a histogram of sentiment polarity, and a ranking of sentiment polarity. Users can interact with the dashboard by selecting a specific restaurant code to view sentiment information for that restaurant. The dashboard updates all other sentiment information dynamically based on the user's selection, providing a powerful means of analyzing and understanding sentiment data.

Figure 32 presents a BI dashboard that effectively showcases customers' star ratings across different categories, including overall rating, food rating, service rating, and value rating. Not only that, the dashboard also provides a wealth of additional information such as emotion analysis, sentiment lexicons, bi-gram and tri-gram analysis, making it a valuable tool for gaining deeper insights into customer preferences and opinions.

VII. CONCLUSION

Sentiment analysis is a powerful tool in gastronomy tourism, uncovering customer preferences, opinions, and feedback. These insights empower businesses and destinations to make data-driven decisions that improve the tourist experience, leading to happier customers, repeat business, and positive word-of-mouth. This research focuses on the evaluation of accuracy and reliability within the context of customer review sentiment analysis in the gastronomy tourism sector. To tackle these challenges, a hybrid sentiment analysis model has been meticulously crafted, incorporating diverse components such as lexicon-based sentiment analysis, emotional analysis, and machine learning-based sentiment analysis. Subsequently, the model underwent further enhancements through the integration of data augmentation alongside a feature engineering strategy. This augmentation and feature engineering were employed to boost the model's performance in recognizing and classifying the minority sentiment classes especially the F1-Score and ROC-Auc measures, a necessity arising from the inherently skewed nature of the data procured from web-crawling activities. Subsequent to these enhancements, a comprehensive sentiment analysis was undertaken, encompassing both sentiment polarity analysis and the formulation of a visualization strategy. This endeavor culminated in the creation of a business intelligence dashboard tailored specifically for the gastronomy tourism sector in Sarawak. Additional insights gleaned from this study's findings indicate that the outcomes of the sentiment analysis reveal a noteworthy trend. Specifically, all three lexicon-based sentiment analysis techniques consistently yielded positive sentiment scores exceeding 80% across all restaurants assessed. Furthermore, an in-depth examination of the average polarity scores produced by the three lexicons revealed that they consistently surpassed the threshold for positive sentiments across each restaurant under scrutiny. Moreover, our emotional analysis brought to light that a considerable portion of reviews could be categorized into positive emotions, encompassing sentiments of joy, trust, and anticipation. These findings collectively lead us to the conclusion that an abundance of positive sentiments frequently appears in restaurant reviews, a trend harmonious with the overall star ratings awarded by customers. Nonetheless, our results also cast light upon the imperative need for certain restaurants to make substantial improvements in their service quality, as reflected in the presence of negative reviews. For the refinement of our sentiment analysis algorithm, forthcoming studies should prioritize the collection of an expanded volume of review data obtained from diverse online platforms. This augmentation in data acquisition aims to provide a more accurate and holistic portrayal of public sentiment regarding specific businesses within a given locale. Furthermore, future research endeavors could concentrate on the development of an algorithm capable of discerning sentiments expressed in specific aspects of a business, such as service quality, and expanding the applicability of the algorithm to non-English language reviews.

REFERENCES

[1] H. Q. Abonizio, E. C. Paraiso, and S. Barbon, "Toward text data augmentation for sentiment analysis," *IEEE Trans. Artif. Intell.*, vol. 3, no. 5, pp. 657–668, Oct. 2022, doi: 10.1109/TAI.2021.3114390.

[2] A. Adak, B. Pradhan, and N. Shukla, "Sentiment analysis of customer reviews of food delivery services using deep learning and explainable artificial intelligence: Systematic review," *Foods*, vol. 11, no. 10, p. 1500, May 2022, doi: 10.3390/foods11101500.

[3] R. Ahuja, A. Chug, S. Kohli, S. Gupta, and P. Ahuja, "The impact of features extraction on the sentiment analysis," *Proc. Comput. Sci.*, vol. 152, pp. 341–348, Jan. 2019, doi: 10.1016/j.procs.2019.05.008.

[4] S. M. S. Alam, "Evaluation of destination image among foreign tourists: The case of Sarawak," Tech. Rep., 2019.

[5] L. M. Alharbi and A. M. Qamar, "Arabic sentiment analysis of eateries' reviews: Qassim region case study," in *Proc. Nat. Comput. Colleges Conf. (NCCC)*, Mar. 2021, pp. 1–6, doi: 10.1109/NCCC49330.2021.9428788.

[6] J. Chen, H. Hou, Y. Ji, J. Gao, and T. Bai, "Graph-based attention networks for aspect level sentiment analysis," in *Proc. IEEE 31st Int. Conf. Tools Artif. Intell. (ICTAI)*, Nov. 2019, pp. 1188–1194, doi: 10.1109/ICTAI.2019.00164.

[7] Department of Statistics. (2019). *Regional Tourism Satellite Account Sarawak 2018*. [Online]. Available: https://www.dosm.gov.my/v1/index.php?r=column/pdfPrev&id=Qlc3cW5WRkE5cEtKYStkaUs1UzNJdz09

[8] D. Gupta, A. Roup, D. Gupta, and A. Ratre, "Food review analysis and sentiment prediction using machine learning models," in *Proc. 2nd Int. Conf. Adv. Electr., Comput., Commun. Sustain. Technol. (ICAECT)*, Apr. 2022, pp. 1–7, doi: 10.1109/ICAECT54875.2022.9807907.
[9] A. Z. H. Sapheri, A. M. Jainal, S. K. A. Bakar, and M. S. M. Zahari, "Factors influencing tourists visiting Kuching, Sarawak," *Int. J. Academic Res. Bus. Social Sci.*, vol. 8, no. 16, pp. 136–143, 2018, doi: 10.6007/IJARBSS/v8-i16/5123.
[10] C. J. Hutto and E. Gilbert, "VADER: A parsimonious rule-based model for sentiment analysis of social media text," in *Proc. 8th Int. Conf. Weblogs Social Media*, 2014, pp. 216–225, doi: 10.1609/icwsm.v8i1.14550.
[11] I. H. Junaid, F. Hossain, U. S. Upal, A. Tameem, A. Kashim, and A. Fahmin, "Bangla food review sentimental analysis using machine learning," in *Proc. IEEE 12th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2022, pp. 347–353, doi: 10.1109/CCWC54503.2022.9720761.
[12] S. Khruahong, O. Surinta, and S. C. Lam, "Sentiment analysis of local tourism in Thailand from YouTube comments using BiLSTM," in *Multi-Disciplinary Trends in Artificial Intelligence*. Berlin, Germany: Springer, 2022, doi: 10.1007/978-3-031-20992-5_15.
[13] X. J. Lee, T. Tzen, V. Yap, H. Ng, and V. T. Goh, "Proceedings of the international conference on computer, information technology and intelligent computing (CITIC 2022)," in *Proc. Int. Conf. Comput., Inf. Technol. Intell. Comput.*, vol. 1, 2022, pp. 488–502, doi: 10.2991/978-94-6463-094-7.
[14] S. Ling. (2023). *Three Million Visitor Target for 2023, Says Sarawak Tourism Minister*. [Online]. Available: https://www.thestar.com.my/news/nation/2023/02/21/three-million-visitor-target-for-2023-says-sarawak-tourism-minister
[15] J. Liu, Y. Yu, F. Mehraliyev, S. Hu, and J. Chen, "What affects the online ratings of restaurant consumers: A research perspective on text-mining big data analysis," *Int. J. Contemp. Hospitality Manag.*, vol. 34, no. 10, pp. 3607–3633, Aug. 2022, doi: 10.1108/ijchm-06-2021-0749.
[16] A. Londhe and P. V. R. D. P. Rao, "Dynamic classification of sentiments from restaurant reviews using novel fuzzy-encoded LSTM," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 10, no. 9, pp. 112–124, Sep. 2022, doi: 10.17762/ijritcc.v10i9.5714.
[17] K. Makkar, P. Kumar, M. Poriye, and S. Aggarwal, "A comparative study of supervised and unsupervised machine learning algorithms on consumer reviews," in *Proc. IEEE World Conf. Appl. Intell. Comput. (AIC)*, Jun. 2022, pp. 598–603, doi: 10.1109/AIC55036.2022.9848880.
[18] M. Mamatha, S. Shivakumar, J. Thriveni, and K. R. Venugopal, "Visual sentiment classification of restaurant review images using deep convolutional neural networks," in *Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol. (CONECCT)*, Jul. 2022, pp. 1–6, doi: 10.1109/CONECCT55679.2022.9865746.
[19] B. C. Martínez-Seis, O. Pichardo-Lagunas, S. Miranda, I. J. Perez-Cazares, and J. A. Rodriguez-Gonzalez, "Deep learning approach for aspect-based sentiment analysis of restaurants reviews in Spanish," *Computación y Sistemas*, vol. 26, no. 2, pp. 899–908, Jun. 2022, doi: 10.13053/cys-26-2-4258.
[20] Ministry of Tourism. (2023). *Visitor Arrivals Into Sarawak 2019*. [Online]. Available: https://mtcp.sarawak.gov.my/upload/file_folder/VIsitorsArrival/VisitorArrivalsintoSwak2019(Dec).pdf
[21] R. Mokhosi, C. Shikali, Z. Qin, and Q. Liu, "Maximal activation weighted memory for aspect based sentiment analysis," *Comput. Speech Lang.*, vol. 76, Nov. 2022, Art. no. 101402, doi: 10.1016/j.csl.2022.101402.
[22] F. Koto and G. Y. Rahmaningtyas, "Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs," in *Proc. Int. Conf. Asian Lang. Process. (IALP)*, Dec. 2017, pp. 93–98.
[23] H. M. Park and J. H. Kim, "Stepwise multi-task learning model for holder extraction in aspect-based sentiment analysis," *Appl. Sci.*, vol. 12, no. 13, p. 6777, 2022, doi: 10.3390/app12136777.
[24] M. N. Razali, E. G. Moung, F. Yahya, C. J. Hou, R. Hanapi, R. Mohamed, and I. A. T. Hashem, "Indigenous food recognition model based on various convolutional neural network architectures for gastronomic tourism business analytics," *Information*, vol. 12, no. 8, p. 322, Aug. 2021, doi: 10.3390/info12080322.
[25] P. Rita, C. Vong, F. Pinheiro, and J. Mimoso, "A sentiment analysis of michelin-starred restaurants," *Eur. J. Manag. Bus. Econ.*, vol. 32, no. 3, pp. 276–295, Jul. 2023, doi: 10.1108/ejmbe-11-2021-0295.
[26] M. Sarawak. (2022). [Online]. Available: https://www.sarawaktourism.com/StoryDetail.aspx?sId=1213 and https://www.sarawaktourism.com/StoryDetail.aspx?sId=1213
[27] D. J. T. Weng. (2021). *Sarawak As a Potential Gastronomy Tourism Destination. Campus & Beyond*. [Online]. Available: https://www.swinburne.edu.my/campus-beyond/sarawak-as-a-potential-gastronomy-tourism-destination.php
[28] J. Xiao and X. Luo, "Aspect-level sentiment analysis based on BERT fusion multi-attention," in *Proc. 14th Int. Conf. Intell. Human-Machine Syst. Cybern. (IHMSC)*, Aug. 2022, pp. 32–35, doi: 10.1109/IHMSC55436.2022.00016.
[29] L. Xu, X. Pang, J. Wu, M. Cai, and J. Peng, "Learn from structural scope: Improving aspect-level sentiment analysis with hybrid graph convolutional networks," *Neurocomputing*, vol. 518, pp. 373–383, Jan. 2023, doi: 10.1016/j.neucom.2022.10.071.
[30] S. Yarkareddy, T. Sasikala, and S. Santhanalakshmi, "Sentiment analysis of Amazon fine food reviews," in *Proc. 4th Int. Conf. Smart Syst. Inventive Technol. (ICSSIT)*, Jan. 2022, pp. 1242–1247, doi: 10.1109/ICSSIT53264.2022.9716410.
[31] B. Zhu, D. Guo, and L. Ren, "Consumer preference analysis based on text comments and ratings: A multi-attribute decision-making perspective," *Inf. Manag.*, vol. 59, no. 3, Apr. 2022, Art. no. 103626, doi: 10.1016/j.im.2022.103626.
[32] S. Azodolmolky, J. Perelló, M. Angelou, F. Agraz, L. Velasco, S. Spadaro, Y. Pointurier, A. Francescon, C. V. Saradhi, P. Kokkinos, and E. Varvarigos, "Experimental demonstration of an impairment aware network planning and operation tool for transparent/translucent optical networks," *J. Lightw. Technol.*, vol. 29, no. 4, pp. 439–448, Feb. 2011.

MOHD NORHISHAM RAZALI received the B.Sc. (Hons.) and M.Sc. degrees in information technology from Universiti Teknologi MARA (UiTM), Malaysia, and the Ph.D. degree in intelligent systems from Universiti Putra Malaysia. He is currently a Senior Lecturer with 14 years of teaching experience in computer science and business. In addition, he possesses professional qualifications certified by the Lembaga Teknologis Malaysia. He served at two institutions which are Universiti Malaysia Sabah and Universiti Teknologi MARA. He is an accomplished researcher and has published several articles in peer-reviewed journals. His notable research contributions are reflected in a series of impactful publications. These include indexed publications such as "Landmark Recognition Model for Smart Tourism Using Lightweight Deep Learning and Linear Discriminant Analysis" and "Indigenous Food Recognition Model Based on Various Convolutional Neural Network Architectures for Gastronomic Tourism Business Analytics." His research interest includes predictive models for student grading to the performance evaluation of masked face recognition for Covid-19 Standard Operating Procedure (SOP) compliance monitoring. Beyond their scholarly publications, he has also made significant contributions through exhibitions and presentations, exemplified by their work on "Mr. Penauk-a Campus Virtual Assistant Bot" and the development of a deep learning architecture-based local food recognition system for Sabah's gastronomic tourism. He received a number of awards and recognition for his contributions.

SYAIFULNIZAM ABDUL MANAF received the bachelor's degree in computer science from Universiti Teknologi Malaysia, the first master's degree in computer science from Universiti Kebangsaan Malaysia, and the second master's degree in remote sensing and GIS and the Ph.D. degree in intelligent systems from Universiti Putra Malaysia. He is currently a Senior Lecturer with the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia. In addition, he possesses professional qualifications certified by the Institution of Geospatial and Remote Sensing Malaysia (IGRSM). He is also an accomplished researcher and has published several articles in peer-reviewed journals.

ROZITA BINTI HANAPI received the Diploma degree in secretarial science, the bachelor's degree (Hons.) in office systems management, and the M.Sc. degree from Universiti Teknologi MARA (UiTM), Cawangan Shah Alam. Since December 2009, she has been a Senior Lecturer with the Faculty of Business Management, UiTM Sarawak, Kota Samarahan. She has a background in office system management and information technology. She specializes in information technology management and published her research in many journals and conferences.

MOHD RAFIZ SALJI received the Ph.D. degree in computer security from Universiti Putra Malaysia. Since then, he has been teaching and research with Universiti Teknologi MARA Sarawak. He is a Senior Researcher with more than ten years of experience in information management. He has published numerous articles in leading journals. His research interest includes computer security. His research contributions are significant, as he was actively involved in a national project aimed at developing a framework for health-related programs and activities in Malaysian public libraries. His research has also led to the publication of several indexed articles, covering a wide range of topics, including health-related programs, access control models, and privacy preservation. In addition to research, he plays a role in nurturing the next generation of researchers. As the main supervisor for a master's candidate, he guides and supports research on intelligent control systems for perioperative anesthesia information.

LEE WEN CHIAT received the bachelor's degree (Hons.) in economics (industrial economics) and the corporate master's degree in business administration from Universiti Malaysia Sarawak, Malaysia, and the Ph.D. degree from Universiti Utara Malaysia. He is currently a Senior Lecturer with the Faculty of Business and Management, Universiti Teknologi MARA Sarawak. He is an expert in fisheries and agricultural economics. His commitment to research is evident in his involvement in numerous national projects. Notably, he serves as the Principal Investigator for a project titled "Evaluating the Effectiveness and Performance of the Asajaya Area Fishermen Association In Sarawak," funded by Dana Dalam Negeri and spanning, from 2022 to 2023. His scholarly contributions are well-documented, with a substantial body of work published in indexed journals and proceedings. His research interests include the impact of demand factors on economic growth, technology adoption, labor force, capital formation in rice production, and fisheries conflict management. He also dedicates his expertise to postgraduate supervision, co-supervising research on the impacts of Chinese culture inheritance, and training on economic innovations. Beyond his research and academic commitments, he has been appointed to the Expert Panel for the Samarahan Tourism Development Plan.

KASHIF NISAR received the Ph.D. degree in computer network and information technology from Universiti Teknologi PETRONAS, Malaysia, with a fully funded scholarship. He was a Postdoctoral Researcher with the Auckland University of Technology, Auckland, New Zealand. He is currently a Lecturer with the Swinburne University of Technology, Sydney, Parramatta, NSW, Australia. He received solid training in research and development (R&D), writing funding proposals, journal publications, and as a consultant.