Index N0: CT-05-40

Aspect Based Multi-Class Sentiment Dataset for Bilingual eWOM of Commercial Food Products

Binguni Edirisuriya (1st Author)
Department of Industrial Management
University of Kelaniya
Dalugama, Sri Lanka.
edirisur_im16018@stu.kln.ac.lk

Dinesh Asanka (2nd Author)
Department of Industrial Management
University of Kelaniya
Dalugama, Sri Lanka.
dasanka@kln.ac.lk

Janaka Wijayanayake (3rd Author)
Department of Industrial Management
University of Kelaniya
Dalugama, Sri Lanka.
janaka@kln.ac.lk

***Abstract* — Aspect-Based Sentiment Analysis for product review opinion analysis is commonly utilized by the commercial food products manufacturing businesses to drive decisions regarding products. However, the general consumers are not facilitated with decision-making ready end-user applications which generates insights to arrive at the purchase decision at the time of purchase due to the unavailability of products' attribute-wise analysis-ready data. Although Electronic Word of Mouth (eWOM) platforms are comprised of opinions with a diversity of languages and expression formats, themselves do not generate any value to make comparable decision making. Hence, there is an existing gap of impactful information retrieval by the consumer to aid the purchase. Therefore, creating a publicly available analysis-ready dataset for the commercial food product domain contributes significantly to the Sri Lankan consumers and Government organizations. Through our research work, a manually annotated bilingual eWOM opinion text dataset for selected commercial food products categories has been delivered in which the opinions expressed in the Sinhala language have been translated into English language and each opinion has been manually rated into five levels by two domain experts. Two product attributes, "Price of the product", "Safeness of product" have been considered as aspects to conduct the Aspect-Based Sentiment Analysis. This study describes the sub-tasks performed to conduct the Aspect-Based Sentiment Analysis on the dataset along with the basic statistical evaluation of the dataset. We have presented results on the performance of the dataset by utilizing an existing Long Short-Term Memory Model.**

Keywords — aspect-based sentiment analysis, commercial food domain

I. INTRODUCTION

Commercial food products play a vital role in fulfilling the consumption needs of individuals to manage food requirements with their busy lifestyles. In Sri Lanka, the commercial food market has been occupied by many local and multinational companies who are manufacturing the same varieties of food products with varied product quality attributes and products' prices. The production quality, packaging, and ingredients of the product reveal the "consumption-safeness of the product" which the consumer must focus on when making the purchase decision, as unsafe food product consumption may lead to health threats including Non-Communicable diseases. Consumer surveys reflect typical Sri Lankan consumer faces many challenges when making a purchase decision of commercial food products due to the lack of perceived knowledge on food hygiene and nutritional information analysis, being caught under influential, deceive marketing campaigns, promotions executed by business [1], inability to conduct a quick critical review on economical decision making when other alternative products are remained unobserved, unrevealing accurate information by the responsible government authorities and unavailability of a suitable unbiased supportive information resource which can be incorporated in analytical end-user applications which can aid the purchase decision at the time of purchase.

Nowadays, eWOM has become a prominent information resource that utilizes for business analytics purposes in many domains. Since eWOM comprises of diverse knowledge base, ideologies, expressions on experience with a high volume of user-base, with more than 6.40 million users in Sri Lanka, it can be recognized as a rich information resource.

A. Knowledge level on Commercial Food Products

The stakeholders associated with the commercial food product domain are the general consumers, commercial food manufacturers, Government Policy authorities like Consumer Affairs Authority, and Sri Lankan Medical Authorities. Consumer survey results and results of the interviews conducted in Sri Lanka in 2020, to assess the knowledge of food stakeholders [2],[3] depicted the fact, the lack of literacy exists in food stakeholders to recognize, distinguish, and select the most suitable food products to consume.

B. Text Mining in eWOM on Purchase Decision

The definition for Electronic Word of Mouth (eWOM) can be identified as "consumer-generated, consumption related communication that employs digital tools and is directed primarily to other consumers", which are in different forms as reviews, opinions, feedback where they become valuable, persuasive mode of information since most of the reviews are experienced-based, open-connected nature, active participation and affects the purchase decision making process of the consumer [4].

Text mining has been utilized for decision-making by the food products-related businesses, food science, and nutrition domain. Different data gathering mechanisms, data analysis methods have been discussed in the review paper analysis [5],[6] The evaluation of the review analysis was able to reveal that the key data source which can be used to identify consumers' opinions, food issues and the food consumption behavior is "social media".

C. *Aspect-Based Sentiment Analysis*

Aspect-Based Sentiment Analysis related research domain comprises many dedicated research studies targeting different domains. Publicly accessible improved datasets have been incorporated to evaluate the process and sub-tasks of Aspect-Based Sentiment Analysis. Using the publicly available restaurant review dataset, [7] study delivered annotated dataset considering six aspect categories and provide overall sentiment polarity and not focusing on aspect-wise sentiment polarity consideration. SemEval 2014 [8], [9] introduced laptops, restaurants datasets annotated with four fields as aspect term, aspect term polarity, aspect category, and aspect category polarity to be consumed in Aspect-Based Sentiment analysis. Driving a new approach SemEval 2016 [10] presented Multilingual datasets from 7 domains and 8 languages in which all were involved in Aspect-Based Sentiment Analysis tasks. The domains included in the datasets were hotels, laptops, mobile phones.

The research study conducted by [11] has focused on aspect-based sentiment analysis in the legal domain by fetching data from an available public dataset. They have taken 2000 legal statements and involved domain experts to annotate data manually by aspect identification through the legal statement and assign the sentiment level for each party by deciding the level of the beneficial party. They have focused on sub-tasks of sentiment analysis when annotating the dataset.

The research study conducted by the scholars, [12] has focused on analyzing the reviews given by university students regarding the learning experience at the University. The study has been driven considering the 22 different aspect levels of the different parties associated with the teaching and learning process. The reviews have been classified based on a weighted score for aspects to deliver unbiased results corresponding to each aspect.

This study would drive research attention on how the Aspect-Based Sentiment Analysis process can be incorporated to construct an analysis-ready user opinion dataset for the commercial food products domain. Parallel to that, methods for eWOM data extraction, data analysis, aspect-wise weight values assignment would be considered referring to the literature.

II. OBJECTIVES

This section discusses the main aim and objectives of constructing the dataset and its significance. As discussed in Section I, it is obvious that there is an existing gap for the retrieval of "decision-making" ready information by the Sri Lankan consumer, to make a beneficial commercial food product purchase for the wellbeing of the purchaser. To achieve that latency consumers should be supplied with an accurate public information retrieval source. Our main aim is to construct a publicly available dataset that can be utilized for opinion analysis purposes related to the food products domain. To achieve the main aim, several sub-objectives are executed to methodically gather dynamic, unstructured commercial food product related posts' data in Facebook eWOM platform, to perform sub-tasks of Aspect-Based Sentiment Analysis for aspect category identification and categorization of posts and aspect term identification with appropriate multi-class sentiment polarity assignment for each eWOM comment in the dataset, which would be utilized to conduct experiments.

Hence, this research would be significant, since it benefits the Sri Lankan community, with an analysis-ready data resource that can be incorporated as the analytical resource to develop analytical end-user applications to aid the purchase decision-making process. Thus, the researchers, business analytical officers, economists, health officials, and other government authorities such as Consumer Affairs Authority, Sri Lanka Standard Institute would gain the opportunity to conduct social-opinion-oriented surveys.

III. METHODOLOGY

A. *Data Source Selection*

This section describes the criterion considered for selecting the data sources for data collection. As stated in *StatCounter Global Stats*, 57.28% of social media users in Sri Lanka have been actively sharing opinions on the Facebook platform between September 2020 to September 2021. Two public Facebook discussion groups with members count over 100,000 and which focus on consumer experience on purchasing commercial food products were selected as the data collection sources.

B. *Product Aspects Selection*

This section describes how the products' aspects were elected to be considered in the research study. The *Consumer awareness Survey on Sustainable Consumption – 2018: Sri Lanka* has presented a consumer survey conducted targeting purchase decision making of food products where 423 individuals from five main densely populated districts in Sri Lanka were involved. The results of the survey revealed that the "price of the food product" was the main determinant for purchase decision with a weighted mean average of 3.78 while the "consumption safe quality of the product" was 3.56. In our study, the above values were taken into consideration as the aspect weight values to be assigned into opinions.

Table 1. Product Aspects of Commercial Food Products

Product Aspect	Opinion Consideration by eWOM users
Safety of Food Product	Ingredient analysis, Packaging, Nutrition levels, Harmful effects, Experience on consumption, Health issues
Price of Food Product	Price among alternatives, Economical benefits, Inflation, Beneficial Purchase Decision

C. *Facebook Post data Crawling*

This section discusses the process of collecting data from Facebook posts with the mechanism used to crawl data. Observing the eWOM discussions in the public Facebook discussion groups selected food product categories were “Diary Products” and “Beverages” concerning timely discussions that occurred during the covid19 pandemic period. To conduct the Facebook data crawling, a python script was developed using Python 3 programming language and *Facebook scraper* library available in python language. About 1574 comments were collected along with the post details.

Algorithm 1: Facebook post details extraction

```
Input : page_id
Input : no_of_posts_crawl
Function scrape(page_id, no_of_posts_crawl):
    page = get_page (page_id, cookies = “cookie_file_name.txt”)
    all_posts = []
    posts = get_posts (page_id, no_of_posts_to_scrape)
    i = 0
    Loop For post in posts:
        all_posts [i] [‘post’] . add(post.id AND post.content)
        all_posts [i] [‘comments’] . add(post.comments)
        Do increase i by 1
    End Loop
End Function
```

Fig. 1. Algorithm: Extraction of Facebook Post Details & Comments

D. *Dataset Construction*

This section describes how the dataset construction was done with the extracted Facebook data. The post details were retrieved in nested JavaScript Object Notation (JSON) format and must be transformed into a data frame with post ID, post text, product category (Diary product or Beverage), product ID where the product’s brand name was masked to preserve the unbiasedness of the comments’ annotation process by the annotators, post aspect (price or safety of the product) and the comment.

E. *Comments Preprocessing Steps*

This section discusses the preprocessing steps conducted for bilingual comments pre-processing. Facebook eWOM comments are ambiguous with different linguistics, different variety of opinions expression modes with media types (images, videos, pictorial emojis), hashtags (#), punctuations, URL s, mentioning of an entity with @ sign and pictorial emojis. Therefore, preprocessing steps were conducted with Pandas python library to eliminate the above figures from comments.

.

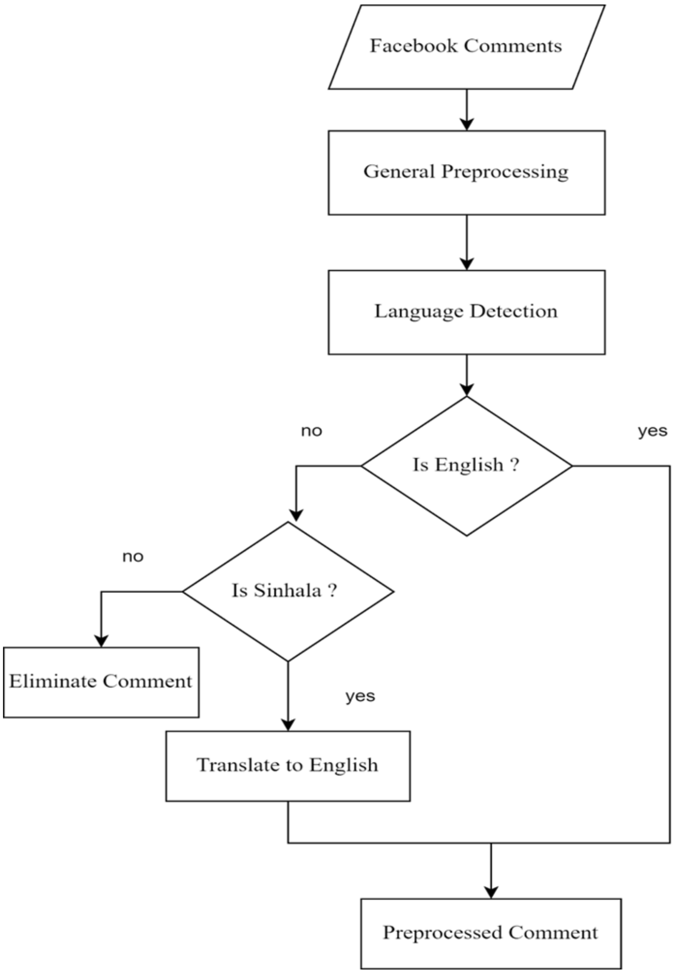

Fig. 2. Flowchart : Language Detection & Translation

The results elaborated that the comments which are expressed in “Singlish” format could not be recognized as one categorization. Currently, there’s no publicly available tool to translate the “Singlish” comments to the English language. Therefore, it reveals that the semantic orientation of Singlish opinion expression could not be captured considering the whole sentence level. Hence, for this level of the study comments which have been expressed in neither English nor Sinhala were eliminated from the consideration. The comments which were detected as “Sinhala” language were translated into the English language, using Google Translator API since sentiment analysis studies have incorporated it for language translation purposes [15].

G. *Manual Data Annotation*

This section discusses the process driven by the two annotators to manually annotate the Facebook eWOM comment data. Reviewing the past literature, we were able to find that there is no publicly available tool to conduct bilingual eWOM text data annotation. As annotators, one graduate and one undergraduate from the University of Kelaniya were involved in the data annotation process. The annotator's task was to figure out the aspect categories expressed in each comment and assign a polarity rating for each aspect. To address the ambiguities in language and opinion expression, the need for a wider range of rating levels was obvious [16], [12]. Therefore, the annotators conducted the rating assignment adhering to the rating levels mentioned in Table 1.

Table 2. Rating levels of Annotation

Rating Level	Sentiment level assigned for rating
1	Very Negative
2	Negative
3	Neutral
4	Positive
5	Very Positive

Example 1

- *Comment 1.1 : D001milk powder contains toxic matter so , useless to buy at a high price.*

Table 3. Rating assignment for the comment by Annotator

Aspect: Sentence	Opinion Word	Rank Assignment				
		5	4	3	2	1
Price	Toxic					✓
Safety	High Price				✓	

The two human annotators assigned ratings for the Facebook eWOM comments separately, by identifying the aspect category expressed through the comment and assigning it with a rating value [11], [12], [17]. A single comment is assigned with two sentiment rating levels by a single annotator, since one comment may consist of opinions expressed concerning both aspects "product safety" and "product price". The comments which have been expressed an opinion only on a single aspect were assigned the rating level 3 (*Neutral*) for the aspect which is not included in the comment. The above example demonstrates how the comment annotation process was conducted by one single annotator. Consider the comment stated in Example 1, the annotator reads the comment and identifies the aspect categories expressed by the comment, next concerning each aspect annotator identifies the sentiment-bearing words and assigns two ratings per each comment under the two product aspects. Above Table 2 shows how the annotator assigns ratings for the above comment in Example 1.

After the two annotators have assigned ratings, the agreement level between the two annotators was assessed with "Inter- Annotator agreement level" to verify the interpretation assigned by the annotators. For that, "Cohen's Kappa" was calculated [18], which is a quantitative measure to measure the chance of how often that raters are agreeing on the same thing/opinion.

$$K = Pa - Pc / 1\text{-}Pc \quad (1)$$

The Cohen's kappa value was calculated separately for the two aspects "product price" and "product safety" and further the results are presented in Section IV.

H. Aspect-wise Class label Assignment for comments

This section discusses how the class labels were assigned for each comment concerning two product aspects considered in the study. Considering the results of Cohen's Kappa scores presented in Section IV, the final aspect rating for a particular comment was obtained by calculating the average rating score assigned by the two annotators since the Cohen's Kappa score revealed that the agreement level between two annotators is relatively same.

In the next step, the average rating score obtained for a certain comment was weighted with the weighted mean value considered for each aspect. The average rating score received for the "price aspect" was weighted with 3.78, while the average rating score received for the "safety aspect" was weighted with 3.56. Further, the weighted rating scores were subjected to Min-Max Scaling to derive a score between 0 and 1, to generalize the process of splitting the multiple-sentiment classes. Table 4 shows the counts of the comments received for each scaled value under the two aspects.

Table 4. Statistics: Comments for respective scaled values

Scaled Value	Count: Safety	Count: Price
0.0	4	4
0.1	0	0
0.2	61	122
0.3	1	9
0.4	92	139
0.5	2	4
0.6	478	670
0.7	11	3
0.8	255	120
0.9	46	12
1.0	249	116

The next step was to define the multiple sentiment classes considering the scaled value distribution. The scaled value range was sliced into five ranges with similar sizes and defined class labels corresponding to each range considering the lower score values to high score values as shown in Table 5. The weighted rating scores for each aspect of each comment were classified into sentiment classes using the "Binning Technique" in Pandas library.

Table 5. Statistics: Comments for respective scaled values

Range	Sentiment level of rating
0.0 <= score < 0.2	Very Negative
0.2 <= score < 0.4	Negative
0.4 <= score < 0.6	Neutral
0.6 <= score < 0.8	Positive
0.8 <= score < 1.0	Very Positive

Section IV presents the aspect-wise sentiment classification of the comments considered in the dataset.

IV. RESULTS AND DISCUSSION

This section has presented the statistics of the dataset as a baseline evaluation, aspect-wise consideration of the statistics with some experiments driven to evaluate the performance of the manually annotated dataset.
Aspect-Based Sentiment Analysis is associated with several sub-tasks and in this study, we consider several sub-tasks as Aspect category identification, Aspect term identification, Aspect category polarity of the sentence. The dataset comprised of eWOM text with translated opinions that were expressed in the Sinhala language before. Hence, it adds complexity to the sub-tasks due to the ambiguity of eWOM text.

A. Aspect Category Identification

Initially, the extracted posts were categorized into the aspects “safety” and “price” considering the content of the post. One single comment may contain opinions for both aspects and some may not belong to any of the considered aspects. Annotators identify the aspects manually and decide on assigning a rate for the comment.

B. Aspect Term Identification

For the given eWOM comment, there can be several terms, a combination of terms that imply the polarity of the comment concerning the considered aspects. Dataset presents the terms which were concerned in the polarity assignment for the comment.

C. Aspect Polarity of the Sentence

Each comment in the dataset comprises of two rating levels between 1-5 scale, for each aspect “safety” and “price” corresponding to the polarity level of the comment. The two rating levels assigned have been utilized to conduct the experiments in the dataset.

D. Dataset Statistics

This section presents preliminary results of statistical analysis done on the collected data. A total of 1574 comments were collected initially and only 1197 comments were concerned for the data annotation process, since 377 comments which were expressed in “Singlish” format and other media formats, were eliminated from the consideration. Table 6 represents the post-aspect-wise statistics of the dataset, in which the respective word counts have been considered after eliminating “Stop-Words” using the NLTK library.

Table 6. Categorization of items as an example

Post Aspect	Comments Count	Word Count
Safety of Food Product	651	3832
Price of Food Product	546	2557

To minimize the effect of bias in the interpretation by a single annotator, two annotators were involved in the rating process, and an agreement analysis was conducted using “Cohen’s Kappa”. The objective of the analysis was to evaluate the reliability of the annotation process. Table 7 represents the Cohen’s Kappa scores obtained respectively to each aspect of eWOM opinion. As the Kappa measure for both aspects was above 0.80 [19], the agreement level between the annotators revealed a high level of confidence in the annotation process[18].

Table 7. Cohen’s Kappa Score for Annotator rating

Inter-Annotator Agreement	Sentiment level of rating
Cohen’s Kappa: Safety Aspect	0.925
Cohen’s Kappa: Price Aspect	0.956

Following the process described in Section H in Methodology, aspect-wise sentiment classification was conducted based on the average weighted score received for each eWOM opinion. Table 8 and Table 9 represent the statistics of the classification, respective to each product aspect, “Product Price” and “Product Safety”.

Table 8. Class Label Assignment : Product Price Aspect

Sentiment Class Label	Overall Polarity Score	Comments Count
Very Positive	+2	128
Positive	+1	793
Neutral	0	4
Negative	-1	148
Very Negative	-2	122

Table 9. Class Label Assignment : Product Safety Aspect

Sentiment Class Label	Overall Polarity Score	Comments Count
Very Positive	+2	295
Positive	+1	266
Neutral	0	480
Negative	-1	93
Very Negative	-2	61

Generally, comments/opinions expressed by the users in eWOM platforms are shorter in length since the eWOM user tends to express ideas relevant to the specific Facebook post content. In our study, the average word frequency received for comment was 5.3 approximately. Figure 2 and Figure 3 represent the frequency of the words present in comments of the considered dataset relative to each aspect, after “Stop Word Removal” using the NLTK library.

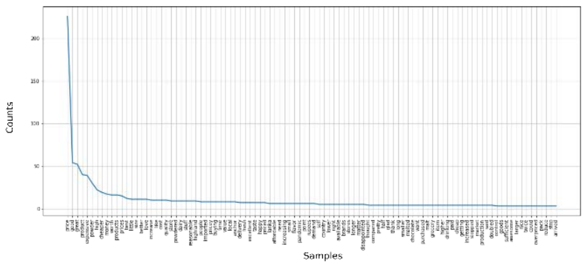

Fig. 3. Word Frequency Plot (100 samples): Price Aspect

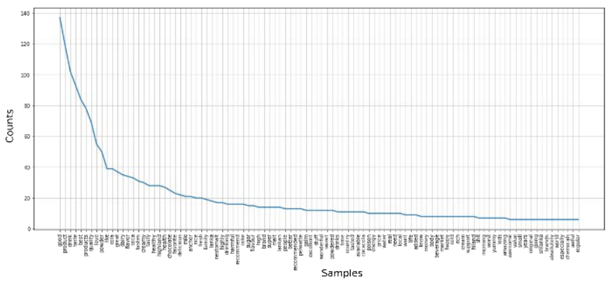

Fig. 4. Word Frequency Plot (100 samples): Safety Aspect

E. *Experiments*

In this section, the results of the basic experiment conducted with an existing Deep Learning-based Bi-Directional Long Short-Term Memory (Bi-LSTM) to evaluate the performance of the dataset have been presented.

As the initial step, the dataset was subjected to preprocessing. Sequences were generated to be utilized as features for the model. The 80% of the dataset randomly was split in to train split and 20% of the dataset was split as the test set to conduct the experiments with the model. One Bi-Directional Layer was considered with a total of 240 neurons and finally classified the polarity labels by feeding to a softmax activation function, while total trainable parameters for the embedding layer, Bi-Directional layer, and dense layer was 164280, 231360, and 329929 respectively. The model training was conducted considering up to 200 training epochs on the training dataset for both aspects separately. Figure 5 and Figure 6 respectively represent the accuracy score plot obtained for both aspects.

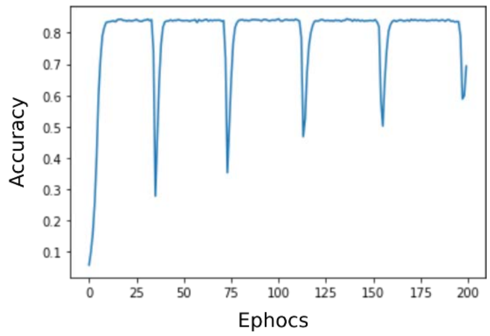

Fig. 5. Accuracy Plot: Safety Aspect

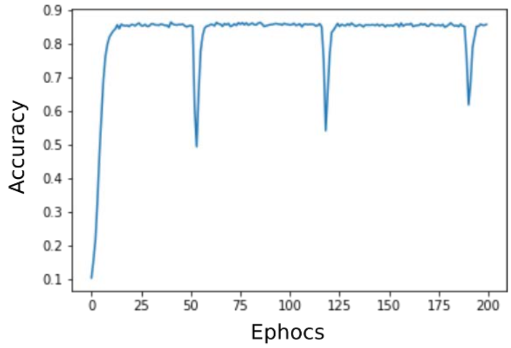

Fig. 6. Accuracy Plot: Price Aspect

At this level of study, a manually annotated eWOM dataset consisting of 1197 opinions was considered for the above experiment. Using the above-explained Bi-LSTM model we were able to obtain an average accuracy score of 0.83 (83%) for the price aspect and 0.85 (85%) for the safety aspect. This dataset can be further improved to utilize for analysis purposes by the parties stated in Section II.

V. CONCLUSION

This study has been driven with the objective to construct a human-annotated dataset for Aspect-Based Sentiment Analysis in the commercial food products domain with the utilization of the Facebook eWOM entity which is comprised of diverse opinions. The current version of the study presented three sub-tasks of Aspect-Based Sentiment Analysis as aspect category extraction, aspect term recognition, aspect category-wise polarity detection for each comment. Since this version of the dataset performed with a satisfiable accuracy for both “product price” and “product safety” aspects, it can be justified that, the study achieved the main aim of constructing an analysis-ready resource for the commercial food products domain.

REFERENCES

[1] Shivany, S. and Jayasinghe, S., Critical Discourse Analysis of Sinhala Cultural Identity in Celebrity Endorsement in Advertisements. *Journal of Business Studies*, 3(2), p.56, 2018.

[2] Hettiarachchi, C., Food Act of Sri Lanka: A Systematic Review on Awareness and Practice. *World Nutrition*, *11*(4), 56-65. doi: 10.26596/wn.202011456-65, 2020.

[3] Talagala, I., & Arambepola, C., Use of food labels by adolescents to make healthier choices on snacks: a cross-sectional study from Sri Lanka. *BMC Public Health*, *16*(1). doi: 10.1186/s12889-016-3422-1, 2016.

[4] Hewage, W., & Madhuhansi, T., Impact of Social Media Characteristics on Consumer Buying Decision 47 Process in Western Province, Sri Lanka. *Global Journal Of Management And Business Research: E Marketing*, *19*, 2019.

[5] Tao, D., Yang, P., & Feng, H., Utilization of text mining as a big data analysis tool for food science and nutrition. *Comprehensive Reviews In Food Science And Food Safety*, *19*(2), 875-894. doi: 10.1111/1541-4337.12540, 2020.

[6] Chua, A., & Banerjee, S., Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality. *Computers In Human Behavior*, *54*, 547-554. doi: 10.1016/j.chb.2015.08.057, 2016.

[7] Ganu, G., Elhadad, N., & Martin, A., Beyond the stars: Improving rating predictions using review text content. *Webdb*, *9*, 1-6, 2009.

[8] Oepen, S., Kuhlmann, M., Miyao, Y., Zeman, D., Flickinger, D., & Hajic, J. et al., SemEval 2014 Task 8: Broad-Coverage Semantic Dependency Parsing. *Proceedings Of The 8Th International Workshop On Semantic Evaluation (Semeval 2014)*. doi: 10.3115/v1/s14-2008, 2014.

[9] Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., & Androutsopoulos, I., SemEval-2015 task 12: Aspect based Sentiment anaysis. *Proceedings Of The 9Th International Workshop On Semantic Evaluation (Semeval 2015)*, 2015.

[10] Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., & AL-Smadi, M. et al., SemEval-2016 Task 5: Aspect Based Sentiment Analysis. *Proceedings Of The 10Th International Workshop On Semantic Evaluation (Semeval-2016)*. doi: 10.18653/v1/s16-1002, 2016.

[11] Mudalige, C., Karunarathna, D., Rajapaksha, I., Silva, N., Ratnayaka, G., Perera, A., & Pathirana, R., SigmaLaw-ABSA: Dataset for Aspect-Based Sentiment Analysis in Legal Opinion Texts. *2020 IEEE 15Th International Conference On Industrial And Information Systems (ICIIS)*. doi: 10.1109/iciis51140.2020.9342650, 2020.

[12] Valakunde, N., & Patwardhan, M., Multi-aspect and Multi-class Based Document Sentiment Analysis of Educational Data Catering Accreditation Process. *2013 International Conference On Cloud & Ubiquitous Computing & Emerging Technologies*. doi: 10.1109/cube.2013.42, 2013.

[13] Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A., Gelbukh, A., & Zhou, Q., Multilingual Sentiment Analysis: State of the Art and Independent Comparison of Techniques. *Cognitive Computation*, *8*(4), 757-771. doi: 10.1007/s12559-016-9415-7, 2016.

[14] Asanka, D., & →, V., Language Detection in Azure Machine Learning with basic Text Analytics Techniques. Retrieved 2 November 2021, from https://www.sqlshack.com/language-detection-in-azure-machine-learning-with-basic-text-analytics-techniques/, 2021

[15] Fernando, A., & Wijayasiriwardhane, T., Identifying religious extremism-based threats in SriLanka using bilingual social media intelligence. *2020 International Research Conference On Smart Computing And Systems Engineering (SCSE)*. doi: 10.1109/scse49731.2020.9313010, 2020.

[16] Kaur, W., Balakrishnan, V., Rana, O., & Sinniah, A., Liking, sharing, commenting and reacting on Facebook: User behaviors' impact on sentiment intensity. *Telematics And Informatics*, *39*, 25-36. doi: 10.1016/j.tele.2018.12.005, 2019

[17] Rahman, M., & Kumar Dey, E., Datasets for Aspect-Based Sentiment Analysis in Bangla and Its Baseline Evaluation. *Data*, *3*(2), 15. doi: 10.3390/data3020015, 2018.

[18] Alqaryouti, O., Siyam, N., & Shaalan, K., A Sentiment Analysis Lexical Resource and Dataset for Government Smart Apps Domain. *Advances In Intelligent Systems And Computing*, 230-240. doi: 10.1007/978-3-319-99010-1_21, 2018.

[19] Takala, P., Malo, P., Sinha, A., & Ahlgren, O., Gold-standard for topic-specific sentiment analysis of economic text. *LREC*, 2152-2157, 2014.