X (Twitter) Sentiment Analysis Based on Hybrid Approach: An Application for Online Food Ordering

Araştırma Makalesi/Research Article

Yıldırım GÜNEŞ[1], Murat ARIKAN[2]

[1]Gazi University, Graduate School of Natural and Applied Sciences, Department of Supply Chain and Logistcs Management, Ankara, Türkiye
[2]Gazi University, Engineering Faculty, Department of Industrial Engineering, Ankara, Türkiye
yildirimgunes1973@gmail.com, marikan@gazi.edu.tr
(Geliş/Received:09.01.2025; Kabul/Accepted:27.02.2025)
DOI: 10.17671/gazibtd.1616709

***Abstract*—** For sentiment analysis of user opinions on online platforms such as X (formerly known as Twitter), dictionary-based approaches and machine learning methods are generally used. Recent studies emphasize that hybridizing these approaches improves model performance. In this study, we propose a hybrid classification model for sentiment analysis of texts on food ordering. In addition, we suggest a feature selection method based on aggregating words for the high-dimensionality problem of text classification. The main problems in that domain are low number of words with distinctive features, complexity of interpretation of food ordering field, domain dependency of text classification. The use of classification algorithms and a domain lexicon-based approach will contribute to overcoming these difficulties. For this purpose, two domain-specific lexicons are developed using data from online users' opinions, one for sentiment analysis and the other for product-service systems classification, referred to as basic lexicons. Basic lexicons have been transformed into new lexicons with fewer words, referred to as boosted lexicons, by grouping the words in basic lexicons and representing the groups with a single word in boosted lexicons. 144 models of combinations of six classification algorithms, three term weighting methods, and the lexicons are created in a hybrid approach for sentiment analysis. The study used two datasets of 21 039 and 14 389 tweets obtained from X between January 1 and December 31, 2020. The models were trained, tested on the first dataset, and the best models were selected. The second dataset is analyzed with the selected models, we present proposals for the industry.

***Keywords*—** X (twitter) sentiment analysis, lexicon-based classification, online food order, natural language process, feature selection

Hibrit Yaklaşıma Dayalı X (Twitter) Duygu Analizi: Çevrimiçi Yemek Siparişi Üzerine Bir Uygulama

***Özet*—** X (eski adıyla Twitter) gibi çevrimiçi platformlardaki kullanıcı görüşlerinin duygu analizi için, genellikle sözlük tabanlı yaklaşımlar ve makine öğrenmesi yöntemleri kullanılır. Son çalışmalar, bu yaklaşımların hibrit kullanımının model performansını iyileştirdiğini vurgulamaktadır. Bu çalışmada, yemek siparişi ile ilgili metinlerin duygu analizi için hibrit bir sınıflandırma modeli öneriyoruz. Ayrıca, metin sınıflandırmanın yüksek boyutluluk problemi için kelimeleri topluluştırmaya dayalı bir özellik seçim yöntemi öneriyoruz. Bu alandaki temel sorunlar, ayırt edici özelliklere sahip kelime sayısının düşük olması, yemek siparişi ile ilgili cümlelerin yorumlanmasının karmaşıklığı, metin sınıflandırmanın alan bağımlılığıdır. Sınıflandırma algoritmalarının ve alan sözlüğü tabanlı bir yaklaşımın birlikte kullanılması, bu zorlukların üstesinden gelinmesine katkıda bulunacaktır. Bu amaçla, çevrimiçi kullanıcıların görüşlerinden elde edilen veriler kullanılarak, biri duygu analizi için diğeri ise temel sözlükler olarak adlandırılan ürün-hizmet sistemleri sınıflandırması için olmak üzere iki alana özgü sözlük geliştirilmiştir. Temel sözlükler, bu sözlüklerdeki kelimelerin gruplandırılması ve sözkonusu gruplardan grubu temsil edecek bir kelimenin seçilmesiyle, daha az sayıda kelime içeren ve güçlendirilmiş sözlük olarak adlandırılan yeni sözlüklere dönüştürülmüştür. Duygu analizi için hibrit yaklaşımla, altı sınıflandırma algoritması, üç terim ağırlıklandırma yöntemi ve sözlüklerin kombinasyonlarından oluşan 144 model oluşturulmuştur. Çalışmada, 1 Ocak - 31 Aralık 2020 tarih aralığında X'ten paylaşılmış, 21 039 ve 14 389 tweetten oluşan iki veri seti kullanılmıştır. Modeller eğitilmiş, ilk veri seti üzerinde test edilmiş ve bunların arasından en iyi model seçimi yapılmıştır. İkinci veri seti seçilen modellerle analiz edilmiş ve sektör için öneriler sunulmuştur.

***Anahtar kelimeler*—** X (twitter) duygu analizi, sözlük tabanlı sınıflandırma, çevrimiçi yemek siparişi, doğal dil işleme, özellik seçimi

1. INTRODUCTION

X (formerly known as Twitter) is a platform where users can share their opinions on any topic; new ideas are added to the shared ideas at any time; and millions of tweets are outdated with the newly added tweets. Considering that as of 2022, there are 368.4 million monthly active users [1], it can be estimated how many and varied the number of tweets shared will be. These tweets, which contain user opinions for every sector and every field, are an essential source where business owners can collect positive and negative opinions about the sector. During the pandemic period, there has been an increase in people's tendency to eat at home or work, and these trends and habits, which have become permanent, have been reflected in X posts [2]. Providing timely feedback to the customer by evaluating the customer opinions to be obtained from social media tools that enable the rapid dissemination of customer opinions ensures customer satisfaction. These piles of textual data, which are continuously generated by users, are converted into usable data with sentiment analysis and text classification methods using automated methods.

X (formerly Twitter), user-generated platform, is a useful source of texts that enable customer insights through text classification including sentiment analysis. However, classifications of short texts on a domain to understand the customer emotions continues to be an difficulty due to the low number of words with distinctive features in the texts. The topic, food, is also difficult to interpret and domain dependent. The domain-based studies that make significant contributions to the correct understanding of emotion in a text are at a low level in languages other than English. It is seen that academic studies employ machine and deep learning algorithms, natural language process, domain-based lexicons for text classifications. On the other hand pre-trained language models are also run on textual data analysis. However domain dependency continues especially on sentiment analysis. [3-5].

Sentiment analysis and text classification are performed using natural language processing (NLP) on X texts containing customer opinions written in colloquial chats. For sentiment analysis, tweet contents can be classified as binary, positive and negative; ternary, positive, negative, and neutral; or multiple, with additional emotions such as anger, satisfaction, distrust, etc. [6]. These classifications are carried out using various methods, including artificial intelligence (AI) and NLP methodologies. Text classification is generally based on dictionary-based (as general dictionaries, domain specific lexicons and corpus-based lexicons) approaches, machine learning (ML) and deep learning (DL), transformative, and hybrid approaches [7,8].

Shinde et al [9] performed sentiment analysis on 25 000 tweets with an hybrid method using lexicon and machine learning approach. They employed SVM as classification algorithms, and unigram, bigram and trigram as feature selection method. They achieved in the range 57-62% performances. Vatambeti [10] et al proposed a hybrid model, Convolutional Bi-directional Long Short Term Memory, for tweet texts. The model performed ranging between 83-92 % in text classification. Trust and Minghim [11] studied the performance of seven large language models that are generally successful in text generation but not thoroughly studied in sentiment analysis, on text classification. In this study, it was found that the models performed ranging from 62-99 % in sentiment analysis tasks on five different datasets.

This study based on text classification with an hybrid method on food industry, which is an unstudied domain, will contribute to the solution for classification problem of short text and filling the gap in the domain of Turkish. In the study, we propose a hybrid classification model, deploying classification algorithms and domain lexicon-based approach, for sentiment analysis of texts on food ordering. Two domain-specific lexicons are developed for the model using data from online users' opinions. A classification of each tweet (document level) was made with the hybrid model. The classification models in the study provide performance at a level that can compete with state-of-the-art models.

The models created in the study are used to classify tweets' product and service features as positive and negative. Two models created in the study were used to classify the positive and negative emotions of product and service features of tweets about online food ordering. The results obtained from the model's rapid assessment of many customer opinions that cannot be evaluated manually, can contribute to the industry in two ways. The first is to present the opinions of previous customers to new customers with information such as "the level of customer satisfaction with the product and service," thus facilitating and guiding customers' decision-making. The second is to use the results obtained to improve the product and service by transforming them into tasks for the stakeholders involved in the supply chain. Thus, customer satisfaction and competitive advantage can be achieved.

In order to generate the dataset for the study, data was extracted from X with the Turkish keywords yemek sipariş (food order), yemeksepeti sipariş (basket of food order), döner sipariş (döner kebab order), lahmacun sipariş (thin Turkish pizza order), hamburger sipariş (hamburger order), and pide sipariş (pita order). The collected tweets were used in two separate datasets: 21 039 tweets and 14 389 tweets. The tweets belong to dates between 1st January and 31st December 0f 2020.

There are some problems encountered in text classification models. One of them is the domain dependency of the word and text [12]. The words used in a text have specific meanings related to the domain. Classifying with general dictionaries causes a decrease in classification performance since specific meanings are not taken into account [13,14]. Another challenge of text classification, such as short texts or tweet posts, is the small number of terms with distinctive features. If these distinctive terms in a document cannot be selected as features, the number of unclassified or misclassified documents increases. Another difficulty in classifying sentences on the topic of food ordering is the difficulty in interpreting conversations on this topic [3]. This study aims to develop models that contribute to solving these challenges through the hybrid use of ML and lexicon-based approaches. To this end, two domain-specific lexicons, called basic lexicons were developed using data from online users' opinions, one for sentiment and one for product-service system classification. Basic lexicons have been transformed into new lexicons with fewer words, referred to as lexicons, by grouping the words in the lexicon and representing the groups with a single word. The feature selection used in this transformation, based on the grouping of words, is a method that also contributes to solving the high-dimensionality problem of text classification.

As a result, we have contributed to the literature with two domain-specific lexicons, an approach that reduces the lexicon size by reducing the number of features for lexicon-based studies, and a model for sentiment analysis and classification of product-service systems.

The rest of the paper is organized as follows: In the second section, a literature review is conducted, focusing on sentiment analysis and text classification methods, as well as their applications in the food sector. The third section introduces the method used for developing the model, its stages, the dataset utilized, evaluation metrics, and the developed domain-specific basic and boosted lexicons. The fourth section evaluates and discusses the analysis, classification results, and the proposed model in detail. The final section presents the conclusions and outlines areas for future research.

2. LITERATUR REVIEW

The text contents in the posts of X users are observed to be unstructured, disorganized, ambiguous in meaning, suggestive, and varied in forms such as jargon and slang. The use of domain-specific emotion terms in such unformatted texts, differences in people's expressions of their emotions and thinking methods, spelling errors, implicit meaning, and ambiguities make sentiment analysis and text classification complicated [4]. The standard phases commonly used for analysis—understanding the task and data, data preparation, modeling, evaluation, and usage—also apply to text analyses [15]. The details of these analyses and phases, which can be used with different nomenclatures in different fields, may vary depending on the social media platform from which the data is drawn, the characteristics and content of the data set, and the analysis objective.

Machine learning and dictionary-based approaches, which are used as basic approaches for text classification and sentiment analysis, can be used in hybrid form as a third approach. Recent studies emphasize that hybrid approaches, which overcome the disadvantages of the basic approaches, improve classification model performance [7,8]. In addition, transformative approaches using advanced techniques such as deep learning also improve the performance of these basic approaches. Alongside the chosen methodologies, the characteristics of the dataset to be analyzed directly impact classification performance.

Text and sentiment classification is fundamentally a word-centric study, focusing on the characteristics of words. A text can be classified at three level—document, sentence, and aspect/feature level—using values derived from words. A commonly used method in these classifications is the aggregation of sentiment scores. These methods can be applied in various ways, such as combining the weight scores of words [16] or aggregating the classification results of different classifiers [17]. Mirtalaie et al. [18] aggregate sentiment polarities by considering the relationship between different features and the desired feature when determining the sentiment value of a specified target feature. In their study examining aggregation methods, Basiri et al. [19] determined the values of words and then performed aggregation at the sentence and general levels based on these values.

There are various challenges in determining the sentiment value of a word that is closest to its natural usage. To overcome these challenges, features such as the position and the specific meaning of the word can be analyzed separately [13,20,21]. The values obtained from these features determine the polarity score of the word. In a review of 47 studies that looked at problems with classifying emotions, Hussein [12] found problems that were seen in most of them. These problems included the fact that sentiment classification is domain-dependent and it can be hard to figure out whether negative sentences have explicit or implicit meanings. Despite the challenges in text classification, high-performance models are being developed for different sectors [22-24]. Recent studies indicate a growing trend toward hybridizing ML and lexicon-based classification methods. The lexicon-based method utilizes a sentiment lexicon to measure the power of emotions. There are two ways to prepare the sentiment dictionaries. The first is lexicon-based, using general dictionaries as a source, and the second is corpus-based, using the dataset as a source [25,26]. When a ready-made lexicon (the first one) is used, the classification may fail because words not included in the lexicon are not taken into account [27] or the specific meaning of the word is ignored [13,14]. When using a corpus-based approach, the problem of high dimensionality [13] is also encountered, as irrelevant words remain in the corpus even after text cleaning. In addition, the need to update lexicons due to the constant production of content with new and different structures on online platforms, shifts in word meanings, and the derivation of new words can also increase the failure of sentiment classification [28].

In many studies in the literature on dictionary-based classification and sentiment analysis in different languages, dictionaries such as Wordnet, Sentiwordnet, Bing, Afinn, Laughran, SentiStrength, NRC, Bing Liu Opinion Lexicon, and Textblob are utilized. Furthermore, domain-specific dictionaries generated from a limited number of seed word lists, dictionaries developed through automatic or manual translation methods, and specialized dictionaries that include idioms and proverbs are also used in text classification studies [29-45]. In classifying Turkish texts, dictionaries such as TS Corpus, Turkish National Corpus, Spoken Turkish Corpus, SentiTürkNet, and Turkish WordNet are also available as intuition dictionaries prepared with specific methods [46-59]. In languages lacking sufficient resources in terms of dictionaries and training data for text classification, building accurate models and improving model performance can be a significant challenge. Domain-specific studies and transfer learning models such as cross-lingual embeddings can contribute to solving this problem [60,61]. In their study, Kılıcer et al. [62] reported that in sentiment analysis for Turkish, Turkish classification dictionaries did better than translation dictionaries, and hybrid approaches did better than other approaches. However, there were not many hybrid studies at the time.

For ML, various methods can be employed, including supervised, unsupervised, semi-supervised, reinforcement, multi-task, ensemble, and instance-based learning, as well as neural networks [63,64]. The selection of an algorithm in ML depends on factors such as the type of problem, the number of variables, and the appropriate model type for the problem. The superiority of the algorithms can vary, and new methods and algorithms are continually developed to strengthen the weaknesses of previous approaches, leading to the introduction of new versions of algorithms [65-69].

It is observed that text classification algorithms exhibit different performances on various domains and datasets. Naive Bayes (NB), Decision Trees, Artificial Neural Networks, Support Vector, Instance Based and Statistical Language Model Based Classifiers are widely preferred in text classification applications [70]. In a literature review on sentiment analysis, Metha [71] stated that ML methods such as SVM, NB, and neural networks have the highest accuracy, considering them as fundamental learning methods. Numerous studies in the literature suggest the superior performance of ML, including DL, through proposed models and comparative analyses [8,10,13,23, 72-78]. Furthermore, it is often mentioned that combining multiple classifiers generally yields better experimental results than using a single classifier. However, in some cases, dictionary-based methods are also noted to be highly effective [79, 80]. When used together in a hybrid approach, ML and dictionary-based approaches can strengthen each other's weaknesses, allowing the development of higher-performing models. It is noted that hybrid models, with appropriate architecture and precise hyperparameter selection, can outperform all models [27,41,45,79,81-88].

Dey and Das [89], in a sentiment analysis study based on an approach proposing a modified TF-IDF term weighting method, achieved performance in the range of 62.1% to 89.2% on different datasets. Yoo and Nam [90] conducted a sentiment analysis study using machine learning algorithms and an electronic dictionary in Korean. In this study, they achieved a performance in the range of 76-80% with a hybrid approach on datasets of restaurants, computers, cinema, travel and clothing. Erşahin et al. [91] obtained performances of 73%, 86.32%, and 91.96% on three different datasets consisting of tweets about hotels and cinemas, using three different classification algorithms (SVM, NB, and J48) and the dictionary-based approach in their proposed hybrid models. There are comparative studies in the literature on dictionary-based, ML, and hybrid approaches used for sentiment analysis. In one study compiling 68 analyses, the highest performances were found to be 88.85% for dictionary-based studies, 98.29% for ML studies, and 91.96% for hybrid studies [62]. In another study, the performances of recent works using DL, ML, and hybrid approaches were reported to range from 74% to 91% [92]. Mahmood et al. [93] obtained performances of 86% and 90% in their hybrid study using Naïve Bayes and SVM as machine learning algorithms and Wordnet as the general dictionary.

The representation of emotion is considered one of the fundamental challenges in sentiment analysis, and it is noted that this area is still in its infancy [13]. In the classification of texts, weighted terms are used to determine the emotional direction of the text. Selecting features with high distinctiveness from weighted terms contributes to solving the high dimensionality problem in matrices created for analysis, enhancing model performance.

Due to the abundance of jargon meanings in tweets and the composition of very short sentences, feature selection becomes a critical process. The number of words and length of the text are elements that affect a document's score as determined by term weighting methods. It is observed that the score values of words in a dataset consisting of tweets are generally lower than the word scores in other datasets [12]. The high word count in customer reviews is used as a factor that increases the reliability of the review, assuming that a higher word count implies more information about the product [5]. In this regard, X differs from reviews containing evaluations directly related to a product or service. The sparsity of words with jargon meanings and short sentences in tweets can result in a lack of distinctive features in the text. In the literature on feature selection, basic techniques such as count-based methods such as bag-of-words, simple statistical values, term frequency (TF), inverse document frequency (IDF), co-occurrence of terms, n-gram statistics are widely used; PATricia (PAT) tree, SWN word groups, graph-based methods, the popular deep learning technique Bidirectional Encoder Representations from Transformers (BERT) architecture, and BERTweet built on top of it [13, 88, 94-99]. The count vector (CV), based on word frequency within sentences, allows for successful feature selection. Additionally, statistical methods such as the term frequency-information gain method (TF-IGM), which is suitable for multi-classification and considers class frequencies of terms, and the term frequency-inverse document frequency-inverse corpus frequency (TF-IDF-ICF), as well as the term frequency-inverse document frequency-inverse cluster size document frequency (TF-IDF-ICSDF), have been successfully employed for feature selection [100].

Alexandrovna et al. [101] stated that performance could be enhanced through careful and efficient feature weighting, and they achieved an improvement of 4%-5% in accuracy using a methodology that reduces the feature dimension. Bandhakavi [102] utilized a domain-specific dictionary and the unigram mixture model (UMM) to identify terms that best represent the text. Sarayna [103] employed the TF-IDF method on a dataset consisting of tweets, while Kaur [104] utilized the TF-IDF method in conjunction with n-grams. Alshehri and Algarni [105] utilized term frequency (TF) and term discrimination ability (TDA), which groups selected features based on their distinctiveness and weights them according to their contribution to each group. Sharma and Kumar [106] employed a multi-feature-based concept ranking algorithm that uses statistical, semantic, and scientifically named entity properties of terms.

Although there are few sentiment analysis studies in the food industry, it is observed that models and methods have been developed with satisfactory performance. Additionally, it is emphasized that there is a continued need for especially domain-specific studies in the field of sentiment analysis [22-24].

Hingle et al. [107] explored eating habits from X data related to the food industry, while Park et al. [108] investigated perceptions of Chinese, Japanese, Korean, and Thai restaurants. Mishra and Singh [109] focused on waste categories in the meat supply chain, and Singh et al. [110] examined dissatisfaction with beef products. El-Khchine et al. [111] conducted a study on the main areas of interest related to chicken products and proposed a model

Zahoor et al. [112] conducted studies on sentiment analysis and categorization of reviews about restaurants in Karachi, focusing on taste, ambiance, service, and value. Alamoudi and Alghamdi [113] performed sentiment classifications based on food, service, ambiance, and price as target features. In a study by Liapakis et al. [114], they analyzed customer reviews for the food and beverage industry for a one-month period in 2018. For the analysis, they identified five features: food quality, customer service, company image, price, and product quantity.

In their literature review examining sentiment analysis studies in the fast food sector, Adak et al. [3] noted that most studies in this field commonly employ lexicon-based and ML methods. They highlighted a limited number of studies applying DL techniques and mentioned that DL techniques exhibit better performance in other sectors. The authors also mentioned that 77% of models in this sector need help in interpretation because of their nature. Ahmed et al. [115] developed a model that considers implicit meanings to improve classification accuracy. The model, which utilizes domain-specific meanings and target features together, achieved a success rate of 89% when applied to a dataset related to restaurants. Aktaş et al. [116] achieved a performance of 86% in their ML study focused on the food and beverage sector.

This study developed classification models based on the methods and tools used for sentiment analysis and text classification. Methods for feature selection and dimensionality reduction were proposed to improve the performance of text classification models. In this context, lexicons called basic and boosted for domain-specific sentiment and product-service systems were prepared. For all tweets, the polarity scores of words in the lexicons were calculated according to three term weighting methods. These scores were placed in document-term matrix (DTM) cells, and the classes of tweets were determined using DTM with six classification algorithms. The classes determined by the algorithms and lexicons were compared with the actual classes of text to measure performance.

The proposed method of the study that reduces the feature dimension has been employed to transform the basic lexicons into the boosted lexicons. The transformation process that enables the aggregation of sentiment scores of words appearing as separate variables in the dataset and the lexicons is detailed in Section 3. The applied feature reduction method has resulted in an improvement in model performance.

3. METHOD AND MODEL PROPOSAL

Figure 1 depicts the stages of the strategy, which was developed by combining a lexicon-based approach with ML algorithms for sentiment and text classification.

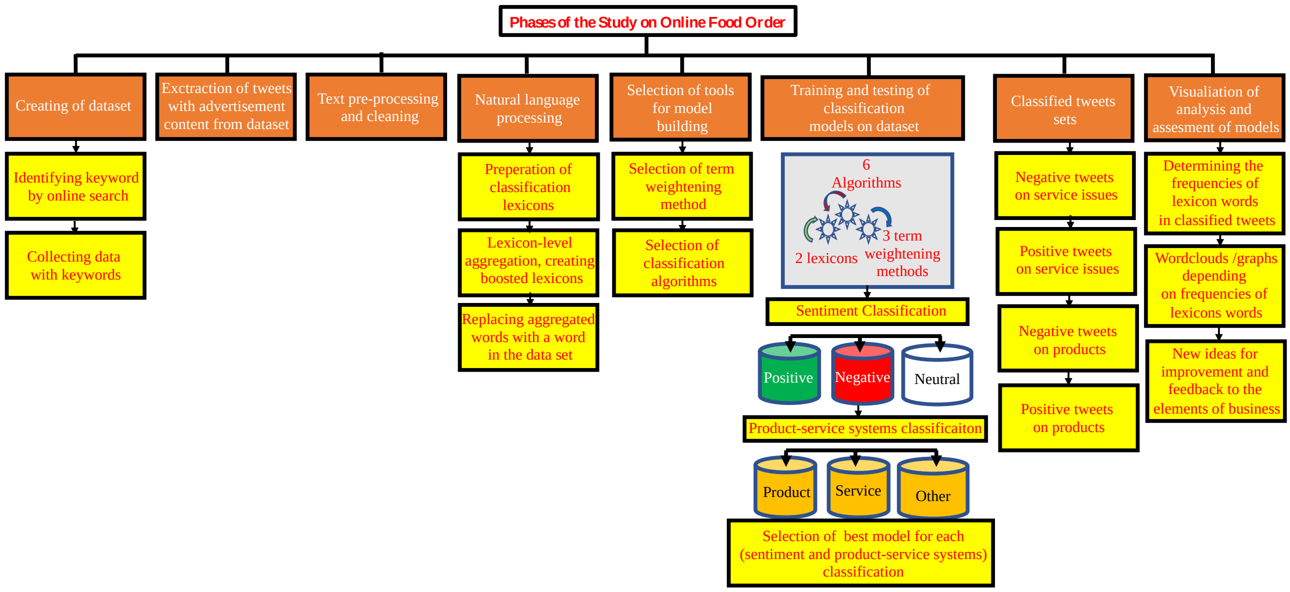

Figure 1. Phases of the study for classifications of tweets about online food ordering

The phases followed are described below.

3.1. The dataset and preprocessing

The data for the analysis was extracted from X tweets on food ordering using the paid Twitonomy application with the Turkish keywords yemek sipariş (food order), yemeksepeti sipariş, döner sipariş (döner kebab order), lahmacun sipariş (thin Turkish pizza order), hamburger sipariş (hamburger order), and pide sipariş (pita order). The tweets are from January 1 to December 31, 2020. The collected tweets are public. Using a specific advertising lexicon, advertising-related tweets were filtered out of the collected dataset. Duplicate tweets, as well as tweets consisting of two or fewer words, were removed. In the dataset, punctuation marks, shapes, and symbols were removed from the tweet texts as part of text preprocessing.

Additionally, the stop words from a list of 175 words prepared for the study were cleaned from the tweet texts (Annex-1). Finally, since the keywords used to collect tweets from X are present in all or most of the tweets in the dataset, which reduces their distinctiveness for the classifications, they were removed from the dataset. The numerical values, such as 8/10, 9/10, etc., found in tweet texts have yet to be cleaned from the text, considering X users' jargon and language usage in the domain. Instead, they have been included in the lexicons as variables (features).

In text classification, the performance measurement of the created models is determined by comparing the actual class of the text (the class label given by the evaluator) with the classification made by the model. The dataset is randomly divided into two groups: 21,039 tweets and 14,389 tweets. The first is used for model creation, training, and testing, and the classification model with the highest performance is identified. The results are presented using the model on the second dataset to show what kind of information can be accessed about the enterprises' fields of activity in the sector and to which areas they can direct their improvement and development activities.

Six evaluators labeled the dataset for the purpose of measuring the models' performance. The evaluators labeled the datasets as product, service, and other for target aspects and positive, negative, and neutral for sentiment classification. Manual labeling of tweets is time-consuming work and requires understanding the purpose of the analysis to do it correctly [117]. The evaluation of a tweet based on subjective evaluation by people with different ways of thinking [3, 8] can produce different labeling results. Due to the difficulties in interpreting sentences in the food sector and the vagueness and ambiguity of colloquial language, evaluators were informed about the labeling process verbally and through a briefing note. The briefing notes include the purpose of the study, the dataset characteristics, the target features of the classification, and Table 1[1]. Table 1 lists the subheadings extracted from food industry classification studies [107-114].

Tablo 1. Class labeling subheadings for product-service systems

Product and Product Quality	Service and Service Quality		
Taste and flavor	Terms of service	Attention, interest and helpfulness of the personnel	Advertising comments about the company/business
Healthy alternatives	Personnel and working conditions of them	Cleanness, hygiene	Online service
Product (menu) and product variety	Consistency	Industry-related advertisements	Discounts
Freshness	Packet	Service speed and duration, weather conditions (courier working conditions)	Promotion applications
Food safety	Price	Courier	Validity of meal cards
Recommended temperature of product	portion adequacy	Customer service staff	Speed of response to complaints, customer service support
Cooking aroma, food smell	foreign body presence in the food	Presentation of food	---
All topics not included in the Table have been labeled as “Other” category.			

The dataset of the study does not have an extreme imbalance [118]. Therefore, no balancing operation has been applied between the classes.To build the model, a first-group dataset consisting of 21,039 tweets was utilized, and classification lexicons and algorithms prepared within the scope of the study were employed for ternary (positive-negative-neutral and product-service-other) and binary (positive-negative and product-service) classifications. For the ternary classification model, all 21,039 tweets were used. In contrast, tweets with the "neutral" and "other" class labels were excluded from the binary classification model, leaving the remaining tweets for analysis.

3.2. The domain-based classification lexicons

Within the scope of the study, two domain-specific lexicons were prepared one for sentiment and the other for product-service systems. These lexicons are named the "basic sentiment lexicon" and the "basic product-service systems lexicon." As an example, he rankings of the first 20 words in the basic lexicon based on their frequencies in the dataset, along with class frequencies, are provided in Table 2.

Despite the progress in language models in recent years and the success of machine learning algorithms and dictionary-based models, these methodologies fail to capture the meanings of words accurately, and these meanings vary depending on the domain they belong to. In particular, the hybrid use of domain-specific dictionaries with one of these methodologies may provide a solution to this problem [11, 13].

A hybrid method of seed word list and corpus-based approach was employed in preparing the lexicons in the following steps: (i) A seed word list was created for the

[1] The details of evaluator briefing notes and labelling the tweet texts by evaluators are in the doctoral thesis which is in the ph.D. thesis of Y. Güneş which is supervised by M. Arıkan, “Twitter (X) Analytics for the Service Sector: An Application on Ordering Meal to Home and Offices”, PhD Thesis, Gazi University, Graduate School of Natural and Applied Sciences, 2024.

domain by examining various social media platforms and business web pages in the industry related to food orders (Annex-2). (ii) The seed word list was expanded with new words using synonym and antonym dictionaries. (iii) The jargon (such as biker, courier, basket maker), slang words, and commonly misused and misspelled words were added to the expanded word list, and the lexicons were formed. (iv) After determining the frequencies of lexicons' words in the dataset, words with zero frequency and the top two words ("food" and "order") with the highest frequency were excluded from the lexicons. However, the words "order" and "food" in a word group such as "order note" continue to be included in the lexicon. As a result, a basic sentiment lexicon consisting of 769 words and a basic product-service systems lexicon consisting of 684 words were obtained [17, 119-124].

Aggregating the sentiment scores of words at different levels is a general method used in text classification problems. In this study, the aggregation method was used in the basic dictionaries to transform them into new classification dictionaries called the boosted dictionary, as described below. At this stage, MS Excel-365's synonyms and antonyms dictionary are deployed, and the aggregation process is performed by evaluating the dictionary information and context-semantic information together [125]. The jargon, slang, or low-value words that may be ineffective in classifications have increased the impact of the calculations through this aggregation at the lexicon level. (i) The words expressed in speech or incorrectly transcribed in the dataset have been grouped based on the correct spelling of the word. (ii) Synonyms, close meanings, or words that are considered to be used in the same sense in the text are groped together as Turkish words alan (field), bölge (region), etraf (around), konum (location), civar (vicinity, nearby), muhit (surroundings, environment), sokak (street), and semt (neighborhood). (iii) Words with the same root but different affixes (due to affixes, sound dropping, softening of hard letters, or vice versa) are grouped as a single word group. (iv) A word from each group is selected to represent the word group. The words other than the representative word in the word group are discarded from the basic lexicon. Simultaneously, the representative word is substituted for the other words in the group in the dataset's tweets. Thus, lexicons with fewer words, which are called boosted lexicons in the study, were obtained.

Tablo 2. Basic sentiment and basic product-service systems lexicon words with the frequencies-20 words

Basic Sentiment Lexicon			Basic Product-Service Systems Lexicon		
Lexicon Words	Dataset Frequencies	Class Dataset Frequencies	Lexicon Words	Dataset Frequencies	Class Dataset Frequencies
yok (unavailable)	1840	845	yorum (comment)	4711	1774
güven (trust)	900	11	saat (hour)	2159	1214
arkadaş (mate)	847	155	restoran (restaurant)	1436	1071
gelm (don't come)	846	611	lahmacun (meat filling)	1222	148
öner (suggest)	843	168	kurye (courier)	1148	995
sev (love)	696	264	burger (burger)	1078	196
istiyor (wants)	604	240	online (online)	879	749
iptal (cancel)	540	495	öner (suggest)	843	170
destek (support)	533	185	zaman (time)	827	250
yemek yok (no food)	475	44	telefon (telephone)	816	595
güzel (beautiful)	465	264	döner (döner kebap)	780	148
kara kara düşün (brood over)	432	15	firma (company)	772	545
zorunda kal (to be forced to)	431	65	paket (package)	768	413
verem (can't give)	358	281	hamburger (hamburger)	692	135
çıkar (self interest)	355	234	pizza (pizza)	672	139
isted (asked)	352	62	adres (address)	606	249
gerçek (real)	349	31	dk (minute)	601	321
getirm (not bring)	343	278	servis (service)	575	351
kazan (earn)	332	41	şimdi (now)	570	192
kalm (no left)	275	121	adam (men)	596	293
The red colors shows the negative class words, the others shows positive class words in the basic sentiment lexicon.			The blue colors shows the product class words, the others shows service class words in the basic product-service systems lexicon.		

As a result, the basic sentiment lexicon's word count decreased by 105, resulting in a boosted sentiment lexicon of 664 words (Annex-3 and Annex-4); the basic product-service systems lexicon's word count decreased by 93, resulting in a boosted product-service systems lexicon of 591 words (Annex-5 and Annex-6). One limitation of this dimension reduction method is that it introduces an additional process of grouping words before analysis and replacing the word representing the group in the texts.

The boosted lexicon structure aims to increase the frequency and weighted values of the words as variables and reduce the matrix size by lowering the number of words (the dimension reduction process). While high-frequency words are more successful for classification in classic feature selection approaches [88], low-frequency words have a negligible influence. The efficacy is increased by combining low-frequency terms with high-frequency words through the boosted lexicon structure.

Table 3 shows an example of how the proposed word representation in feature selection can lead to a rise in the number of times a word is used in term weighting formulas for boosted lexicons.

Tablo 3. Sample of word representation for boosted lexicon

Process order	Sample of Word Representation for Boosted Lexicon		
1.	Words in basic sentiment lexicon	Group of words (synonyms) characterized by a representative word	Representative word in boosted sentiment lexicon for the group of words
	Ödül (award, prize) Armağan (gift) Hediye (present)	ödül, armağan, hediye	hediye

2.	Tweets	Pre-processed Tweets	Replacement process of Tweet-1 for boosted sentiment lexicon
	Tweet-1	“dün gece etmeden aç aç yattığım kendimi *ödül* amaçlı kahvaltı pizza sipariş ettim”	“dün gece etmeden aç aç yattığım kendimi *hediye* amaçlı kahvaltı pizza sipariş ettim”
	Tweet-2	“kardeşim tatil *hediyesi* olarak etmiş aşko benzememeliydin”	“kardeşim tatil *hediyesi* olarak etmiş aşko benzememeliydin” (Because of the term of “hediye” is representing term for group of terms, the term in the Tweet-2 does not need to be replaced.)

3.	Frequency of terms in dataset based on basic sentiment lexicon	Frequency of terms in dataset based on boosted sentiment lexicon
	Ödül (Award, prize):47	---
	Armağan (gift):2	---
	Hediye (present):160	Hediye (present):209

3.3. Term weightening methods and classification algorithms

The formulas for three-term weighting methods can be found in Table 4. These are the count vector (CV), the term frequency (TF), and the term frequency-inverse document frequency-inverse class frequency (TF/IDF/ICF). Generally, tweet text score values are lower than in other datasets [12]. The method applied in the form of aggregating words involves transferring the weighted scores of words from the basic lexicon to the boosted lexicon structure by increasing them. The calculation of the weighted values used for the boosted lexicon structure of the words given in Table 3 is illustrated with an example in Table 4.

The Logistic Regression (LR), K-Nearest Neighbour (K-NN), Non-Linear Supportive Vector Machine (NL-SVM), Multi-Layer Perceptive Classification (MLPC), Gradient Boosting Machine (GBM), and eXtreme Gradient Boosting (XGB) algorithms were utilized for the creation of models.

When training the models, a 10-layer cross-validation method was applied, taking into account the amount of the data set in order to avoid the bias effect in the data set, and hyper-parameter tuning was performed to determine the best performance of the models [126].

Tablo 4. Term weighting formulas and application of them on basic and boosted lexicons' words with a sample

Term Weightening Methods	Words	Basic Lexicon	Boosted Lexicon
CV $W_{CV}(t_i) = TF(t_i, d_j)$	Hediye (present)	160	209
	Armağan (gift)	2	---
	Ödül (award, prize)	47	---
TF $W_{TF}(t_i) = TF(t_i, d_j)/T_j$	Hediye (present)	0.00052	0.00068
	Armağan (gift)	0.0000065	---
	Ödül (award, prize)	0.00015	---
TF-IDF-ICF $W_{TF-IDF-ICF}(t_i) = (TF(t_i, d_j)/T_j) * \{1 + \log(\frac{D}{d(t_i)})\} * \{1 + \log(\frac{C}{c(t_i)})\}$	Hediye (present)	0.0016	0.0022
	Armağan (gift)	0.000039	---
	Ödül (award, prize)	0.000552	---
$TF(t_i, d_j)$: the frequency of term i in document j.			
T_j: the total number of words in the collection/dataset.			
$d(t_i)$: the number of documents in which the term t_i occurs.			
$c(t_i)$: the number of classes in which the term t_i occurs.			
D: number of documents in the collection/dataset. C: the number of classes in the collection/dataset.			
Values for the terms used in the formula above: T_j=305 672; D=21 039; C=3. The values $d(t_i)$ =147 and $c(t_i)$ =3 for "hediye (present)"; $d(t_i)$ =2 and $c(t_i)$ =1 for "armağan (gift)"; $d(t_i)$ =44 and $c(t_i)$=3 for "ödül (award-prize)" are used in the formulas for basic lexicon. The values, $d(t_i)$ =193 and $c(t_i)$ =3 for "hediye (present)" are used in the formulas for boosted lexicon. This calculation is done for the ternary classification			

3.4. The evaluation of model performance

Measurement criteria such as precision, recall, accuracy, and F1 score values can be used for the performance testing of classification models. The calculations of these measurement criteria are based on a confusion matrix comparing the actual class of the data with the classes predicted by the models. The precision value indicates the overall success of the model in classification and is often a useful measure of the performance of datasets with a balanced class distribution. The F1 value, a more robust measurement tool in unbalanced datasets, can provide results by balancing precision and recall values.

However, the F1 calculation also does not consider the class's proportion of observations (samples). Therefore, the weighted average F1 value, which considers the distributional proportions of the classes, is the preferred method for imbalanced datasets. In the study, the weighted

F1 score was employed for performance comparisons, with class ratios used in computations, taking into account the modest imbalance (20% and below) in the positive and product classes in the dataset.

$$\text{(Weighted F1 score)} = \frac{Number\ of\ samples\ in\ the\ class}{Number\ of\ samples\ in\ the\ corpus} * \frac{2 * Precision * Recall}{Precision + Recall}$$

To create models for classifying sentiment and product-service systems, we looked at how well models built with basic and boosted lexicons, three-term weighting methods, six classification algorithms, and binary and ternary classifications worked. MS Excel's data analysis features, Python programming language, and its libraries were used for the application. The results of the analysis are presented in Section 4.

4. RESULTS AND DISCUSSION

For identifying the best classification model, 144 models were created, 72 for product-service system classification and 72 for sentiment classification. These models included ternary and binary classifications, basic and enhanced sentiment dictionaries, three-term weighting approaches, and six classification algorithms.

Considering the moderate imbalances in the dataset, the models established were initially evaluated based on the weighted average F1 scores to determine their performances. Fine-tuning, which allows the model to capture the best values of the parameters and learn better from the dataset, was used to improve the performance of the models. All comparisons between the models in the study were applied after fine-tuning [9].

4.1. Sentiment Classification

The performance of the models run for sentiment classification is shown in Figure 2. The findings related to the emotion models depicted in Figure 2 are as follows: (i) All binary classifications are more successful than ternary classifications. (ii) It has been observed that performance can be enhanced through hyperparameter tuning in the majority of models (performance before hyperparameter tuning is shown in black, and performance after tuning is shown in gray). (iii) Two-class models built with the boosted sentiment lexicon (average shown with a red dashed line) demonstrated, on average, a performance superiority of 1.36% over two-class models built with the basic lexicon (average shown with a green dashed line). This superiority was observed in ternary models at a rate of 0.26%. (iv) The model created using the boosted sentiment lexicon, binary classification, TF-IDF-ICF weighting method, and K-NN algorithm achieved the highest performance for sentiment classification with a rate of 85.217%.

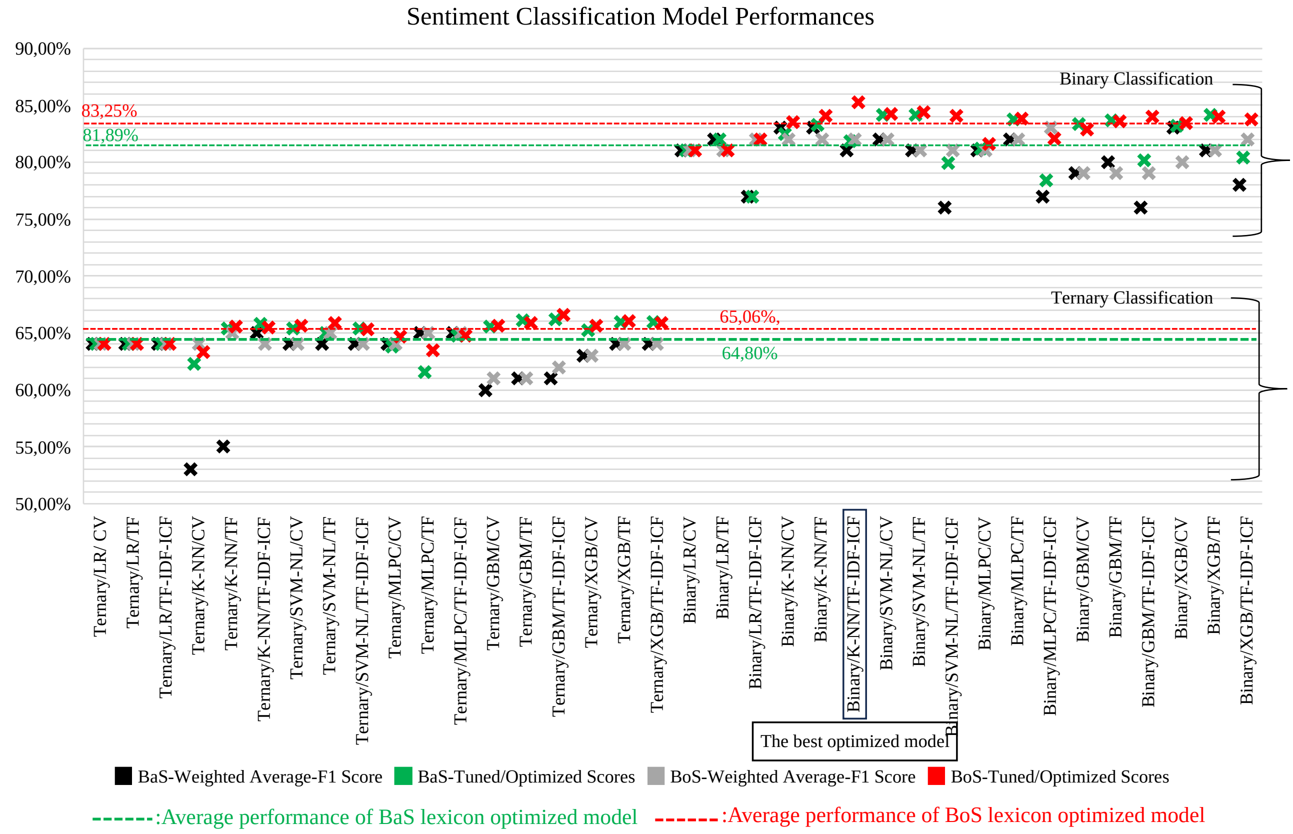

Figure 2. Sentiment classification model tuned performances of basic (BaS) and boosted (BoS) lexicons

The initial and optimized performances, along with the parameters of the proposed sentiment classification model, are presented in Table 5. During the fine-tuning process, various parameters—such as learning rate, maximum depth, number of estimators, minimum sample split, and number of neighbors in different algorithms—were evaluated using a random search approach. The optimal parameter combinations were then determined to enhance the performance of the optimized models [9].

Tablo 5. Proposed model final report for binary sentiment classification

Proposed Model for Binary Sentiment Classification			
Algorithm	Term Weightening Method	Lexicon for Classification	Initial Model
K-NN	TF-IDF-ICF	Boosted Sentiment Lexicon	make-pipeline (StandartScaler( ) KNeigbors Classifier())
Initial Performances			
Training Score	Test Score	10-K-Fold Score	
0.8673	0.8362	0.7667	
Final Performances			
	Precision	Recall	F1 Score
Macro-Avg	0.77	0.68	0.71
Weighted-Avg	0.82	0.84	0.82
Optimized Model Parameters and Performances			
Training Score	Parameters Tested in Initial Model for Optimisation	Final Model After Fitting the Parameters	Optimized Test Score
0.8416	{"n_neighbors": np.arange(1,50}	KNeighbors Classifier(11)	0.8522

4.2. Product-Service Systems Classification

The performance of the models run for product-service classification is shown in Figure 3. The findings related to the product-service systems models depicted in Figure 3 are as follows: (i) All binary classifications are more successful than ternary classifications. (ii) Similar to sentiment classification models, it has been observed that performance in the majority of product-service systems models can be enhanced through hyperparameter tuning. (iii) Two-class models constructed with the boosted product-service systems lexicon (average shown with a red dashed line) demonstrated, on average, a performance superiority of 1.12% over two-class models constructed with the basic lexicon (average shown with a green dashed line). This superiority is observed in ternary models at a rate of 3.95%. (iv) The model created using the boosted product-service systems lexicon, binary classification, CV weighting method, and GBM algorithm achieved the highest performance for product-service systems classification with a rate of 88%.

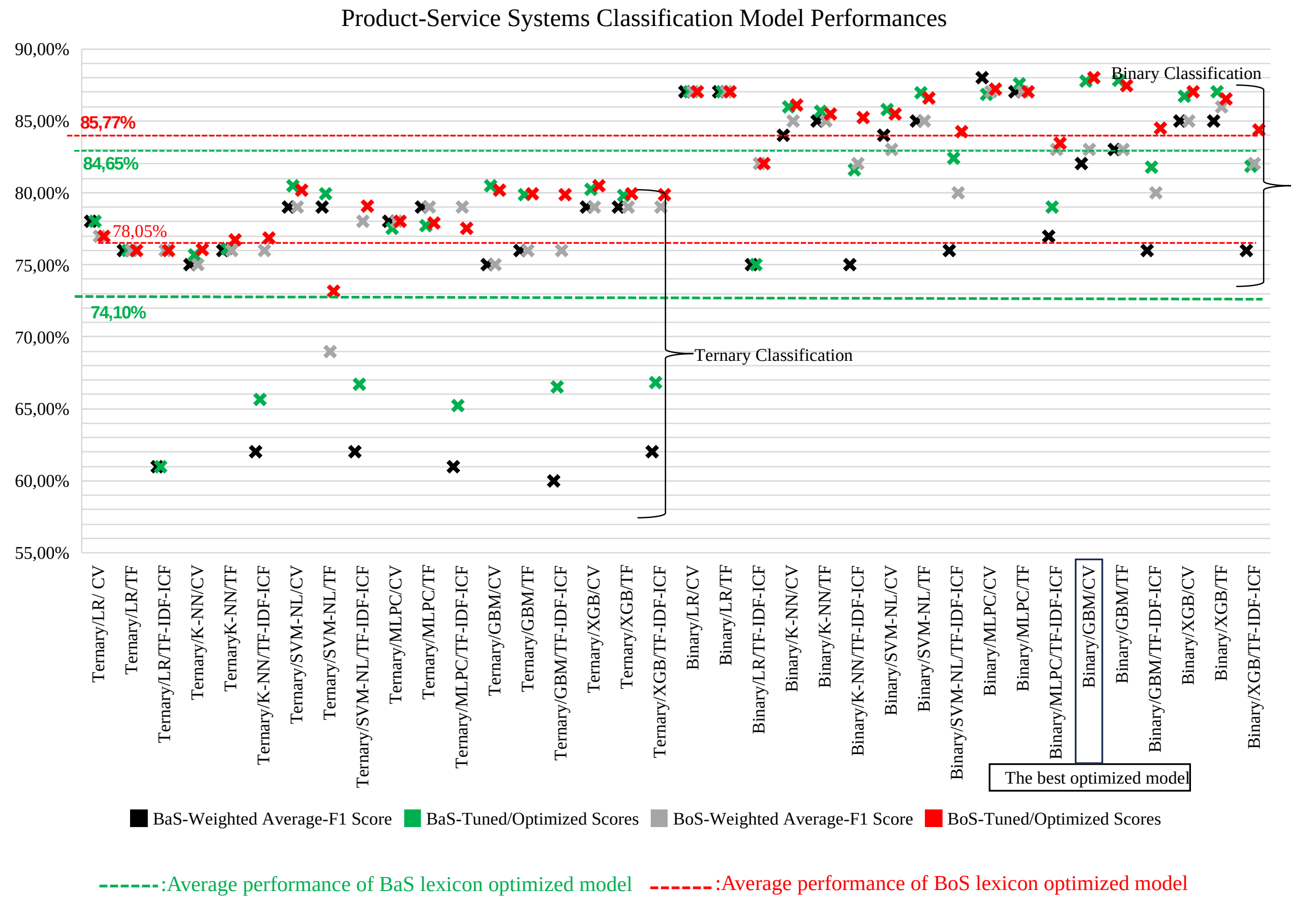

Figure 3. Product-service systems classification model tuned performances of basic (BaS) and boosted (BoS) lexicons

The initial and optimized performances and parameters of the proposed model for sentiment classification is shown in Table 6.

Tablo 6. Proposed model final report for binary product-service systems classification

Proposed Model for Binary Product-Service Systems Classification			
Algorithm	Term Weightening Method	Lexicon for Classification	Initial Model
GBM	CV	Boosted Product-Service Systems Lexicon	make_pipeline (Standart Scaler (), GradientBoostig Classifier())
Initial Performances			
Training Score	Test Score	10-K-Fold Score	
0.8769	0.8521	0.8477	
Final Performances			
	Precision	Recall	F1 Score
Macro-Avg	0.82	0.63	0.67
Weighted-Avg	0.84	0.85	0.82
Optimized Model Parameters and Performances			
Training Score	Parameters Tested in Initial Model for Optimisation	Final Model After Fitting the Parameters	Optimized Test Score
0.8829	{"max_depth":range(3,5),"n_estimators":[100,500,1000],"min_samples_split:[2,10]":}	GradientBoostingClassifier(max_depth=4,n_estimators=1000,min_samples_split=2)	0.8801

The best results from comparing 144 models created through the combination of different algorithms, term weighting methods, binary-ternary classifications, and the alignment of sentiment and product-service systems classifications are shown in Table 7. When the weighting methods and algorithms of the constructed models are examined in terms of average values across all classifications, (i) the best results for each of the three-term weighting methods were obtained with the GBM algorithm, and (ii) the TF method with the highest average of 79.3% was observed.

In addition to the model with the best performance highlighted in italic in Table 7, other binary classification models also exhibit satisfactory performance and can be used as classification models.

Table 7. Optimized/tuned classification results

Sentiment Classification Optimized/Tuned Test Scores			
Number of Classes	Term Weightening Method	Basic Sentiment Lexicon (769 words)	Boosted Sentiment Lexicon (664 words)
Three classes	CV	GBM: 65,53%	GBM:65,65%
	TF	GBM: 66,11%	XGB: 66,01%
	TF-IDF-ICF	GBM: 66,18%	GBM: 66,62%
Two classes	CV	NL-SVM: 84,15%	NL-SVM: 84,19%
	TF	NL-SVM/XGB: 84,15%	NL-SVM: 84,39%
	TF-IDF-ICF	KNN: 81,79%	*KNN: 85,22%*
Product-Service Systems Classification Optimized/Tuned Test Scores			
Number of Classes	Term Weightening Method	Basic Product-Service Systems Lexicon (684 words)	Boosted Product-Service Systems Lexicon (591 words)
Three classes	CV	GBM: 80,52%	XGB: 80,49%
	TF	NL-SVM: 79,92%	XGB: 79,94%
	TF-IDF-ICF	XGB: 66,82%	GBM/XGB: 79,9%
Two classes	CV	GBM: 87,77%	*GBM: 88%*
	TF	GBM: 87,82%	GBM: 87,43%
	TF-IDF-ICF	NL-SVM: 82,4%	KNN: 85,22%

It was seen that both the basic and boosted lexicons made for the study could be used for classifications. The boosted lexicon makes the model perform better than it did with the basic lexicon. In the literature, performance ranges in text classification using dictionary-based, machine learning, and hybrid approaches vary between 62% and 98% [58, 86-88]. Considering the unique challenges posed by tweet texts compared to other types of texts, the achieved performance of 85.22% and 88% in this study is considered superior and competitive compared to many studies in the literature.

4.3. Implementation of the Proposed Classification Models on the Second Group Dataset

After using the first group dataset, the suggested classification models are as follows: for sentiment classification, models made up of the boosted lexicon, binary classification, TF-IDF-ICF weighting method, and K-NN algorithm; and for product-service classification, models made up of the boosted lexicon, binary classification, CV weighting method, and GBM algorithm. The type of insights that can be derived using the recommended models in relation to the industry is demonstrated with the second group dataset, consisting of 14,389 tweets. For this purpose, the second dataset has been classified using the recommended two models, and the distributions resulting from the classification are shown in Figure 4. Based on the distribution ratios, it is observed that over 80% of the tweets in the second dataset consist of negative opinions.

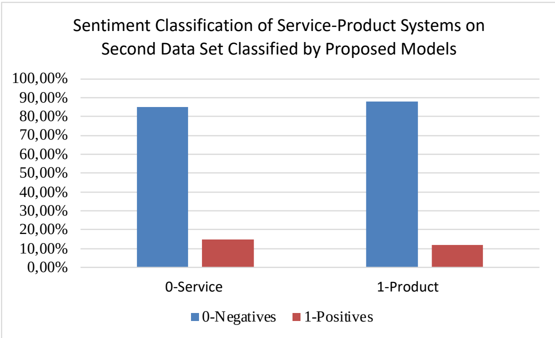

Figure 4. Proportion of second dataset class after classified by optimized models

At the end of the classification process of the second dataset, word frequencies in the tweets belonging to each class were determined, and word clouds were created based on these frequencies. The prominent topics in the word cloud can provide guiding results for businesses in the sector concerning improvement and development.

The prominent topics among users regarding the positive and negative aspects of the service class as a result of the classifications are shown in Figure 5. Accordingly, the most critical complaint topics that need improvement in the service aspect are observed to be restaurants (restoran), couriers (kurye), advertisements (reklam), cancellation processes (iptal), time-duration (zaman-süre), payment-related transactions (ödeme), delivery (teslim), and non-delivered (gelm). It is observed that users' positive opinions about the service are concentrated around winning (kazan), discounts (indirim), loyalty programs (joker), recommendations (öner), and friend (arkadaş), like (sev), plus (artı), want (istiyor), support (destek) topics.

Figure 5. Negative and positive highlights on service issues after classified second dataset

The prominent topics among users regarding the positive and negative aspects of the product class are shown in Figure 6. Accordingly, users express their negative opinions and complaints about the product mostly using words such as pizza (pizza), product (ürün), chicken (tavuk), cold (soğuk), spicy (acı), onion (soğan), salad (salata), missing (eksik), fatty (yağlı), dough (hamur), wrap (dürüm), minced meat (köfte), portion (porsiyon), stomach (mide), crispy (çıtır), awful (rezil). On the positive side, opinions about this matter are concentrated around words such as recommendation (öner), like (sev), chicken (tavuk), cold (soğuk), large size (büyük boy), delicious (lezzetli), instant (anında), product (ürün), sauce (sos), ketchup (ketçap), mayonnaise (mayonez), fried (kızartm), nutrition (besin), garnish (garnitür), vegan (vegan).

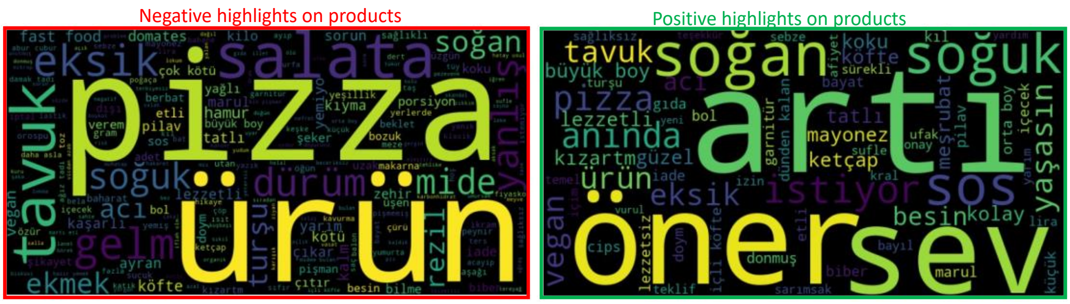

Figure 6. Negative and positive highlights on product after classified second dataset

There are many different models in text classification and sentiment analysis. These models continue to be developed, either on their own or in different combinations.One of the important problems in these models, including advanced language models, is the correct determination of the meaning of the word in the text in which it is used. Failure to correctly determine jargon, implied, and contextual meanings reduces model performance. Another problem of text classification for language models are high dimensionality, and extraction of keyword from the text efficiently [13]. In order to solve this problems, domain-based studies can provide important contributions to the domain. Domain-based lexicon helps to the model in focusing essential and meaningful words [127]. The words in these lexicons will be used only for classification related to the subject, they will not have high dimensionality as in general dictionaries, and words with low effectiveness in lexicons for classification can be excluded from the lexicons to decrease the dimensionalty. The use of a food-specific dictionary and classification algorithms together has achieved a level of success that can compete with state of the art models in this field. Further work in this area could form the basis for future advanced models in sentiment analysis and text classification.

4. CONCLUSION

In this study, an online food ordering classification model has been developed using a lexicon-based approach and classification algorithms in a hybrid method. A total of 144 model comparisons were conducted to form a model for sentiment and product-service system classification. The study will contribute to fill the gap in domain-based text classification and helps to industry to analyze with a robust text classification and sentiment analysis model in food domain. It also will encourage academicians to work on new classification models in other unstudied domains such as the clothing industry, cargo sector, which has potential for development in sentiment classification.

The study's contribution is the proposal of the boosted lexicons for use in sentiment and product-service classifications. The boosted lexicon structure not only yields better results compared to the basic lexicon but also reduces the complexity of the problem due to its smaller size. It has been observed that the applied method improves performance in both sentiment and product-service system classifications. The suggested approach and classification models obtained classification performance of 85% or higher, surpassing several studies on sentiment analysis and text classification found in the existing literature.

Within the scope of the study, four dictionaries were prepared specifically for the food ordering domain, including one basic and one boosted for both sentiment and product-service system classifications. It was observed that the boosted lexicon outperformed the basic lexicon, binary classifications performed better than ternary classifications, and product-service system classifications were better than sentiment classifications. Among the term weighting methods used in the models, TF was found to have the best performance average. Among the algorithms, GBM exhibited the highest performance. The recommended classification models, developed domain lexicons, and sentiment analysis conducted on customer feedback in the context of online food orders enable the measurement of customer satisfaction based on product and service target features. The results provide an opportunity to identify areas needing improvement that can potentially shape the industry.

In the following periods, the domain lexicons developed within the scope of the study can be developed and used in new studies specific to the field of food. The boosted lexicon structure, proposed as a solution to the dimensionality reduction problem, which is a significant issue in text classification problems, can be applied to classifications of other text types with a higher word count in text compared to tweets. Thus, the problem of high dimensionality in text classification issues is addressed, and performance comparisons with other models and methods can be made.

Limitations of the Study

The boosted lexicons created with the dimensionality reduction method have improved the classification performance. However, the recommended method also has some limitations. The method requires additional processes before the analysis operations. Words within word groups in the dataset should be replaced with representative words. The grouping of words for the method has been done considering synonyms and meaning similarities arising from jargon, domain-specific uses, and figurative expressions. Executing this method manually can be a time-consuming process. However, it contributes to producing useful domain lexicons for text classifications, considering the natural usage of language.

REFERENCES

[1] S. Dixon, Number of Twitter Users Worldwide from 2019 to 2024, https://www.statista.com/statistics/303681/twitter-users-world wi de/. 25 .04.2023.

[2] Y. Güneş and M. Arıkan, "Exploring Twitter dataset content by descriptive analysis: An application on online food ordering", *International Journal of Informatics Technologies (JIT)*, 16(2), 119-133, 2023.

[3] Adak A, Pradhan B and Shukla N. "Sentiment analysis of customer reviews of food delivery services using deep learning and explainable artificial intelligence", Systematic review. Foods, 11(10), 1500, 2022.

[4] L.R. Krosuri and R.S. Aravapalli, "Feature level fine grained sentiment analysis using boosted long short-term memory with improvised local search whale optimization", PeerJ Computer Science, (9)e, 1336, 2023.

[5] M.V. Gopalachari, S. Gupta and S. Rakesh, et al, "Aspect-based sentiment analysis on multi-domain reviews through word embedding", *Journal of Intelligent Systems*, 32(1), 2023.

[6] S. Atan and Y. Çınar, "Relations between financial news and market capitalizations of companies in BIST30 Index: Text mining and sentiment analysis methods", *Ankara University Journal of Social Sciences*, 74(1), 1-34, 2019.

[7] T. Shaik, X. Tao and Dann C, et al, "Sentiment analysis and opinion mining on educational data: A survey", *Natural Language Processing Journal*, 2, 1-11, 2023.

[8] W. Medhat, A. Hassan and H. Korashy, "Sentiment analysis algorithms and applications: A survey", *Ain Shams Engineering Journal*, (5), 1093–1113, 2014.

[9] G. K. Shinde. V. N. Lokhande, R. T. Kalyane, V. B. Gore and U. M. Raut. "Sentiment Analysis Using Hybrid Approach". *International Journal for Research in Applied Science and Engineering Technology*, (9), 282-285, 2021.

[10] R. Vatambeti, S.V. Mantena and K.V.D. Kiran, et al, "Twitter sentiment analysis on online food services based on elephant herd optimization with hybrid deep learning technique", Cluster Computing, 2023.

[11] Trust P, Minghim R. "A Study on Text Classification in the Age of Large Language Models". Machine Learning and Knowledge Extraction. 6(4), 2688-2721, 2024.

[12] DMEDM Hussein, "A survey on sentiment analysis challenges", *Journal of King Saud University, Engineering Sciences*, (30), 330-338, 2018.

[13] M. Bordoloi and SK Biswas, "Sentiment analysis: A survey on design framework, applications and future scopes", Artificial Intelligence Review, 2023.

[14] F. Sağlam, Automated sentiment lexicon generation and sentiment analysis of news. PhD Thesis, Hacettepe University, Computer Engineering, 2019.

[15] T. Keller, Mining the internet of things: Detection of False-Positive RFID Tag Reads using low-level reader data, PhD Thesis, University of St. Gallen, 2011.

[16] V. Singh, P. Rajesh and U. Ashraf et al, "Sentiment analysis of textual reviews; Evaluating machine learning, unsupervised and SentiWordNet approaches**", Institute of Electrical and Electronics Engineers 5th International Conference on Knowledge and Smart Technology-KST**, Chonburi, Thailand, 2013.

[17] O. Appel, F. Chiclana and J. Carter J et al, "IOWA & Cross-ratio Uninorm Operators as Aggregation Tools in Sentiment Analysis and Ensemble Methods.", **Institute of Electrical and Electronics Engineers International Conference on Fuzzy Systems, Naples**, Italy, 2017.

[18] M.A. Mirtalaie, O.K. Hussain, "Sentiment aggregation of targeted features by capturing their dependencies: Making sense from customer reviews", *International Journal of Information Management*, (53), 2020.

[19] M.E. Basiri, A. Kabiri and M. Abdar et al, "The effect of aggregation methods on sentiment classification in Persian reviews", Enterprise Information Systems, (14), 1394-142, 2020.

[20] Y. Li, Q. Pan and T. Yang et al, "Learning word representations for sentiment analysis", Cognitive Computation, 9(6), 843–851, 2017.

[21] H. Benghuzzi, M.M. Elsheh, "An investigation of keywords extraction from textual documents using word2vec and decision tree", *International Journal of Computer Science and Information Security*, 18(5), 13-18, 2020.

[22] S. Suneetha and V. Row, "Aspect-based sentiment analysis: A comprehensive survey of techniques and applications", *Journal of Data Acqusition and Processing-JCST*, 38 (3), 177-203, 2023.

[23] S. Ounacer, D. Mhamdi and S. Ardchir et al. "Customer sentiment analysis in hotel reviews through natural language processing techniques", *International Journal of Advanced Computer Science and Applications*, 14(1), 569-579, 2023.

[24] K. Du, F. Xing and E. Cambria, "Incorporating multiple knowledge sources for targeted aspect-based financial sentiment analysis", ACM Transactions on Management Information Systems, 4(3), 1-24. 2023.

[25] A.P. Rodrigues, R. Fernandes and A. Shetty et al, "Real-time Twitter spam detection and sentiment analysis using machine learning and deep learning techniques", Computational Intelligence and Neuroscience, 5211949, 14 pages, 2022.

[26] T. O'Keefe and I. Koprinska, "Feature selection and weighting methods in sentiment analysis", ADCS'14: Proceedings of the 14th Australasian Document Computing Symposium, Sydney, Australia, January, 2009.

[27] A. Giachanou F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods", ACM Computing Surveys (CSUR), 49(2), 1-41, 2016.

[28] S. Taj, A. F. Meghji and B. B. Shaikh, "Sentiment Analysis of News Articles: A Lexicon based Approach", **2nd International Conference on Computing Mathematics and Engineering Technologies (ICOMET),** United States, 30-31 Jan 2019.

[29] K. Tutar, M. O. Ünalır and L. Toker, "Development of a framework for ontology-based sentiment analysis on social media", *Pamukkale University Journal of Engineering Sciences*, 21(5), 194-202, 2015.

[30] G.A. Miller, “Beckwith R and Fellbaum C, et al. Introduction to WordNet: An online lexical database”, *International Journal of Lexicography*, 3(4), 235-244, 1990.

[31] S. Blair-Goldensohn, K. Hannan and R. McDonald et al, “Building a sentiment summarizer for local service reviews”, NLP Challenges in the Information Explosion Era (NLPIX), Beijing, China, April 22, 2008.

[32] N. Oliveira, P. Cortez and N. Areal, “Stock market sentiment lexicon acquisition using microblogging data and statistical measures”, Decision Support Systems, 85(6), 2-73, 2016.

[33] S. Baccianella, A. Esuli and F. Sebastiani, “Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining”, **The Seventh International Conference on Language Resources and Evaluation**, Valletta, Malta, 2010.

[34] G. Vural, B. B. Cambazoglu and P. Senkul et al, “A Framework for Sentiment Analysis in Turkish: Application to Polarity Detection of Movie Reviews in Turkish”, International Symposium on Computer and Information Sciences, London, UK, 29 October 2012.

[35] C. Türkmenoğlu and A.C. Tantuğ, “Sentiment analysis in Turkish media”, **International Conference on Machine Learning (ICML)**, Beijing, China, June 2014.

[36] K.S. Vishnu, T. Apoorva and D. Gupta, “Learning domain-specific and domain independent opinion oriented lexicons using multiple domain knowledge”, **Seventh International Conference on Contemporary Computing**, Noida, India, 2014.

[37] S. Akgül, C. Ertano and B. Diri, “Sentiment analysis with Twitter”, *Pamukkale University Journal of Engineering Sciences*, 22(2), 106-110, 2016.

[38] F.S. Çetin, G. Eryiğit, “Investigation of aspect based Turkish sentiment analysis subtasks – Identification of aspect term, aspect category and sentiment polarity”, *Journal of Information Technologies*, 11(1), 43-56, 2018.

[39] E. Bilgiç, A. Koçak. “Sentiment and opinion mining: Recent issues and an application for the third airport”, *Bingöl University Journal of Social Sciences Institute*, 9(17), 327-338, 2019.

[40] G. Arslantürk, “Anti-immigrant attitudes in social media: An examination within the frame of integrated threat theory”, Psychology Researches, 1(1), 06-16, 2021.

[41] L.S. Zaabar, M.R. Yaakub and M.I. Abu Latiffi, “Combination of lexicon based and machine learning techniques in the development of political tweet sentiment analysis model”, *International Journal of Synergy in Engineering and Technology (IJSET)*, 3(2), 72-83, 2022.

[42] A. Dey, M. Jenamani and J.J. Thakkar, “Senti-N-Gram: An n-gram lexicon for sentiment analysis”, Expert Systems with Applications, 2018.

[43] H.S. Hota, D.K. Sharma and N. Verma, “Lexicon-based sentiment analysis using Twitter data: a case of COVID-19 outbreak in India and abroad”, Data Science for COVID-19, 275–95, 2021.

[44] B. Muppuru, S. Vamsi and J. Jayashree et al, “Real time sentimental analysis of chat data using deep learning”, *Journal of Data Acquisition and Processing*, 38 (2), 4545-4556, 2023.

[45] B.S. Ainapure, R.N. Pise and P. Reddy, et al, “Sentiment analysis of COVID-19 Tweets ssing deep learning and lexicon-based approaches” Sustainability, 15, 2573, 2023.

[46] M. Hu, B. Liu, “Mining and Summarizing Customer Reviews”, 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, USA, 2004.

[47] R. Dehkharghani, Y. Saygin and B.A.Yanikoglu, et al. “SentiTurkNet: a Turkish polarity lexicon for sentiment analysis”, Language Resources and Evaluation, 50(3), 667-685, 2016.

[48] T. Sezer, What is TS Corpus Project, https://tscorpus.com. 05.04.2023.

[49] Turkish National Corpus (TNC), https://www.tnc.org.tr/tr/. 25.04.2023.

[50] Spoken Turkish Corpus, https://std.metu.edu.tr. 25.04.2023.

[51] U. Eroğul, Sentiment analysis in Turkish, PhD Thesis, Middle East Technical University, 2009.

[52] M. Kaya, G. Fidan and I.H. Toroslu, “Sentiment analysis of Turkish political news”, In Proceedings of WI-IAT-12 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology, Macau, China, 2012.

[53] E. Akbaş, Aspect based opinion mining on Turkish tweets, PhD Thesis, Bilkent University, 2012.

[54] M. Çetin, M.F. Amasyalı, “Supervised and traditional term weighting Methods for sentiment analysis” **Signal Processing and Communications Applications Conference (SIU)-21st**, (978):1-4, 2013.

[55] Ç. Aytekin, “An opinion mining task in Turkish language a model for assigning opinions in Turkish blogs to the polarities”, *Journalism and Mass Communication,* 3(3), 179-198, 2013.

[56] C. Özsert and A. Özgür, “Word polarity detection using a multilingual approach”, **Computational Linguistics and Intelligent Text Processing Conference Paper**, 75-82, 2013.

[57] M. Meral and B. Diri, “Sentiment analysis on Twitter”, IEEE 22nd **Signal Processing and Communications Applications Conference (SIU)**, Trabzon. Turkey, 2014.

[58] M. Baykara and U. Göktürk, “Classification of social media shares using sentiment analysis”, UBMK-17, **2nd International Conference on Computer Science and Engineering (UBMK),** Antalya, Turkey, 2017.

[59] Ö. Demir, A.I. Baban Chawai and B. Doğan, “Sentiment analysis using dictionary based approach in Turkish texts”, International Periodical of Recent Technologies in Applied Sciences, 1(2), 58-66, 2019.

[60] R. Ghasemi, S.A. Ashrafi Asli and S. Momtazi, “Deep Persian sentiment analysis: Cross-lingual training for low-resource languages”, *Journal of Information Science*, 48-1, 2020.

[61] R. Yang, “Machine learning and deep learning for sentiment analysis over students' reviews: An overview study”, Preprints.org, 2021.

[62] S. Kılıçer, R. Samlı, “Review of sentiment analysis studies on texts in different languages”, *Dicle University Engineering Faculty Journal of Engineering*, 13(3), 2022.

[63] D. Faggella, “What is artificial intelligence, an informed definition”, https://emerj.com/ai-glossary-terms/what-is-artificial-intelligence-an-informed-definition/.28.04.2023.

[64] B. Mahesh, "Machine learning algorithms-A review", International Journal of Science and Research (IJSR), 9(1), 381-386, 2018.

[65] H. Akpınar, 2000, "Knowledge discovery and data mining in databases", *Istanbul University Journal of the School of Business (IUJSB)*, 29(1), 1-22, 2000.

[66] M. Sağlam, "Key themes in brand reputation research: A bibliometric analysis with VOSviewer software", *Research Journal of Business and Management*, 9(1), 1-12, 2022.

[67] Ü.H. Atasever, "The use of boosting, support vector machines, random forest and regression tree methods in satellite images classification, PhD Thesis, Erciyes University, 2011.

[68] E. Akçetin, H. Turgut, "Data analyzing and data mining applications in incomplete data in the business database", *Route Educational and Social Science Journal*, 2(5), 181-199, 2015.

[69] H. Turgut, An application for the diagnosis of alzheimer's disease using data mining processing, PhD Thesis, Süleyman Demirel University, 2012.

[70] A.C. Tantuğ, "Text classification", *TBV Journal of Computer Science and Engineering*, 5(2), 2012.

[71] M. Muhammed, F. Rustam and F. Alasim, et al, "What people think about fast food: Opinions analysis and LDA modeling on fast food restaurants using unstructured tweets", PeerJ Computer Science, 2023.

[72] F. Benrouba, B. Rachid, "Emotional sentiment analysis of social media content for mental health safety", Research Square, 2023.

[73] H.J. Alantari, I.S. Currim and Y. Deng, et al, "An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews", *International Journal of Research in Marketing*, 39 (1), 1-19, 2021.

[74] P. Monika, C. Kulkarni and N. Harish Kumar, et al, "Machine learning approaches for sentiment analysis: A survey", *International Journal of Health Sciences*, 6(S4), 2022.

[75] S. Saifullah, R. Dreżewski and F.A. Dwiyanto, et al, "Sentiment analysis using machine learning approach based on feature extraction for anxiety detection", Computational Science-ICCS, (6), 2023.

[76] K. Gulati, S.S. Kumar and R.S. Kumar Boddu, et al, "Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic", Materials Today: Proceedings, 51(1):38-41, 2021.

[77] Y. Chen, Y. Rao and S. Chen, et al., "Semi-supervised sentiment classification and emotion distribution learning across domains", ACM Transactions on Knowledge Discovery from Data, 17 (5), 1-30, 2023.

[78] M. Shahzad, C. Freeman and M. Rahimi, et al, "Predicting Facebook sentiments towards research", *Natural Language Processing Journal*, 3 (100010), 2023.

[79] Y. Wang, J. Guo and C. Yuan, et al, "Sentiment analysis of Twitter data", Applied Sciences, (12) 22, 2022.

[80] Y. Al Amrani, M. Lazaar and K.E. El Kadiri, "Random forest and support vector machine based hybrid approach to sentiment analysis", Procedia Computer Science, (127), 511-520, 2018.

[81] M. Thelwall, K. Buckley and G. Paltoglou, et al, "Sentiment strength detection in short informal text*", Journal of the American Society for Information Science and Technology*, (61)12, 2544-2558, 2010.

[82] X. Bai, "Predicting consumer sentiments from online text", Decision Support Systems, (50)4, 732-742, 2011.

[83] C. Kaur, M.S.A. Boush and S.M. Hassen, et al, "Incorporating sentimental analysis into development of a hybrid classification model", *International Journal of Health Sciences (IJHS)*, 6 (S1), 1709–1720, 2022.

[84] D. Tiwari, B. Nagpal and B.S. Bhati, et al, "A systematic review of social network sentiment analysis with comparative study of ensemble-based techniques", Artificial Intelligence Review, 2023.

[85] S. Khaleghparast, M. Maleki and G. Hajianfar, et al, "Development of a patients' satisfaction analysis system using machine learning and lexicon-based methods", BMC Health Services Research, (23), 280., 2023.

[86] M. Wankhade, A.C.S. Rao and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges", Artificial Intelligence Review, (55), 5731–5780. 2022.

[87] M. Kristina, M. Marian and H. Miroslava, "Classification of special web reviewers based on various regression methods", Acta Polytechnica Hungarica, (17), 229-248, 2020.

[88] J.R. Chang, H.Y. Liang and L.S. Chen, "Novel feature selection approaches for improving the performance of sentiment classification", *Journal of Ambient Intelligence and Humanized Computing*, 2020.

[89] R.K. Dey, A.K. Das, "Modified term frequency-inverse document frequency based deep hybrid framework for sentiment analysis", Multimedia Tools Applications, (82), 32967–32990, 2023.

[90] G. Yoo and J. Nam, "A Hybrid Approach to Sentiment Analysis Enhanced by Sentiment Lexicons and Polarity Shifting Devices", The 13th Workshop on Asian Language Resources, Kiyoaki Shirai, Miyazaki, Japan, May 2018.

[91] B. Erşahin, Ö. Aktaş and D. Kılınç, et al. "A hybrid sentiment analysis method for Turkish", *Turkish Journal of Electrical Engineering and Computer Sciences*, 27(3), 1780 -1793, 2019.

[92] Y. Madani, M. Erritali and B. Bouikhalene, "A new sentiment analysis method to detect and analysis sentiments of Covid-19 moroccan tweets using a recommender approach", Multimedia Tools and Applications, 82, 27819–27838, 2023.

[93] A. T. Mahmood, S. S. Kamaruddin, Raed Kamil Naser, "A Combination of Lexicon and Machine Learning Approaches for Sentiment Analysis on Facebook", *Journal of System and Management Sciences*, 10(3), 140-150, 2020.

[94] S.K. Bharti, K.S. Babu, "Automatic keyword extraction for text summarization: A survey", arXiv preprint, 2017.

[95] P-I. Chen, S-J. Lin, "Automatic keyword prediction using google similarity distance", Expert Syst Appl; 37(3), 1928-1938, 2010.

[96] M.R.R. Rana, S.U. Rehman and A. Nawaz, et al, "Aspect-based sentiment analysis for social multimedia: A hybrid computational framework", Computer Systems and Engineering, 46(2), 2415-2428, 2023.

[97] T. Sun, L. Jing and Y. Wei, et al, "Dual consistency-enhanced semi-supervised sentiment analysis towards COVID-19 tweets", IEEE Transactions on Knowledge and Data Engineering, 1-13, 2023.

[98] S. Barreto, R. Moura and J. Carvalho, et al, "Sentiment analysis in tweets: an assessment study from classical to modern word representation models", Data Mining and Knowledge Discovery, (37), 318–380, 2023.

[99] D. Yulia, E.A. Bilashova, "Learning analytics of MOOCs based on natural language processing: CEUR", 3017(15), 187-197, 2021.

[100] T. Doğan, A.K. Uysal, "Comparing the Performances of Term Weighting Methods on Medical Document Classification", 6th International Symposium on Innovative Technologies in Engineering and Science (ISITES201), Antalya, Türkiye, 2018.

[101] M.M. Alexandrovna, A.L.M. Ali Mohsin, "Tweet sentiment analysis with CNN and XG-BOOST", Sustainability, 21(4), 1-9, 2023.

[102] A.S. Bandhakavi, W. Nirmalie and P. Deepak, "Lexicon based feature extraction for emotion text classification", Pattern Recognition Letters, 2016.

[103] S. Saranya, G. Usha, "A Machine learning-based technique with IntelligentWordNet lemmatize for Twitter sentiment analysis", Intelligent Automation and Soft Computing, 36(1), 339- 352, 2023.

[104] G. Kaur, A. Sharma, "A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis", *Journal of Big Data*, 10 (5), 2023.

[105] A. Alshehri and A. Algarni, "TF-TDA: A novel supervised term weighting scheme for sentiment analysis", Electronics, 12(7):1632, 2023.

[106] A. Sharma, S. Kumar, "Machine learning and ontology-based novel semantic document indexing for information retrieval", Computers and Industrial Engineering, 176, 108940, 2023.

[107] M. Hingle, D. Yoon and J. Fowler, et al, "Collection and visuliation of dietary behavior and reasons for eating using Twitter", *Journal of Medical Internet Research*, 15(6), e125, 2013.

[108] B.S. Park, J. Jang and C. Ok, "Analyzing Twitter to explore perception of Asian restaurants", *Journal of Hospitality and Tourism Technology*, 7(4), 405-422, 2016.

[109] N. Mishra and A.A. Singh, "Use of Twitter data for waste minimisation in beef supply chain", Annual Operations Research, (270), 337-359, 2018.

[110] A. Singh, N. Shukla and N. Mishra, "Social media data analytics to improve supply chain management in food industry", Transportion Research Part E: Logistics and Transportation Review, 114(C), 398-415, 2018.

[111] R. El-Khchine, A. Amar and Z.E. Guennoun, et al, "Machine learning for supply chain's big data: State of the art and application to social network's data", **MATEC Web of Conferences**, 2018.

[112] K. Zahoor, N.Z. Bawany and S. Hamid, "Sentiment analysis and classification of restaurant reviews using machine learning", **21st International Arab Conference on Information Technology (ACIT)**, 1-6 Giza, Egypt, 2020.

[113] E.S. Alamoudi, N.S. Alghamdi, "Sentiment classification and aspect-based sentiment analysis on yelp reviews using deep learning and word embeddings", *Journal of Decision Systems*, 30(2-3), 259-281, 2021.

[114] A. Liapakis, T. Tsiligiridis and C. Yialouris, et al, "A Sentiment lexicon-based analysis for food and beverage industry reviews, the Greek Language paradigm", Natural Language Engineering, (9), 21-42, 2020.

[115] K. Ahmed, M.I. Nadeem and Z. Zheng, et al, "Breaking down linguistic complexities: A structured approach to aspect-based sentiment analysis*", Journal of King Saud University - Computer and Information Sciences*, 35 (8), 1-21, 2023.

[116] Ö. Aktaş, B. Coskuner and I. Soner, "Turkish Sentiment Analysis Using Machine Learning Methods: Application on Online Food Order Site Reviews", *Journal of Artificial Intelligence and Data Science*, 1(1), 1-10, 2021.

[117] Z. Liu, C. Yang and S. Rüdian, et al, "Temporal emotion- aspect modeling for discovering what students are concerned about in online course forums", Interactive Learning Environment, 27 (5–6), 598–627, 2019.

[118] Imbalanced Data. https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data. 25. 04. 2023.

[119] V. A. Pitogo and C. D. L. Ramos, "Social media enabled e-participation: a lexicon-based sentiment analysis using unsupervised machine learning", **In: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, ACM,** 518–528, 2020.

[120] R. Purbasari, E. Munajat and F. Fauzan, "Digital innovation ecosystem on digital entrepreneur: Social network analysis approach", *International Journal of E-Enterpreneurship and Innovation*, 13(1), 21, 2023.

[121] Y. Kim, T. Y. Choi, T. Yan and K. Dooley, "Structural investigation of supply networks: A social network analysis approach", *Journal of Operations Management*, 29, 194-211, 2011.

[122] Zargan Translation Dictionary. https://www.zargan.com/tr/q/demeaning-ceviri-nedir, 2021.

[123] Tuareg Translation Dictionary. https://www.dictionary.com/browse/tuareg, 2021.

[124] F. Å. Nielsen, "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages 718 in CEUR Workshop Proceedings 93-98. 2011.

[125] Excel-365 Synonyms and Antonyms Dictionary. Microsoft Office-365, 2021.

[126] O. C. Atalaya, J. A. Tuesta, D. B. Mares, A. G. Pacheco, O. M. León, M. Q. Silvestre, G. T. Quispe and R. S. Bazalar, "K-Fold Cross-Validation through Identification of the Opinion Classification Algorithm for the Satisfaction of University Students", *International Journal of Online and Biomedical Engineering (iJOE)*. 19(11), 140-158, 2023.

[127] A. Srivastava. "A Review on Sentiment Analysis of Twitter Data Using Machine Learning Techniques". *International Journal of Engineering and Management Research*, 14(1), 2024.

ANNEXES

Annex-1- Turkish Stopword List

Turkish Stopwords Prepared fort the Study
acaba, akabinde, akebinde, altı, altına, altında, altta, ama, ancak, aralarında, arasında, arasından, arkada, arkasında, artık, asla, aslında, aşağı, aşağıdan, aşağısı, aşağıya, aynen, aynı, ayrıca, az, azıcık, bana, başka, bazen, bazı, bazıları, bazısı, belki, ben, beni, benim, benzer, benzeri, beş, bi, bide, bile, binaenaleyh, bir, bir defa, bir hayli, bir kere, bir kere daha, bir kerecik, bir kimse, bir miktar, bir şey, bir şeyi, bir takım, bir vakitler, bir zamanlar, biraz, biraz önce, birbirinden, birçoğu, birçok, birçokları, birde, biri, birisi, birkaç, birkaçı, birşey, birşeyi, bişey, bişi, biz, bizatihi, bize, bizi, bizim, bizimki, bizzat, bizzat ben, bizzat kendileri, bizzat kendimiz, bizzat kendisi, boyunca, böyle, böylece, böylelikle, böylesine, bu, bu gibi, bu kadar, bu noktada, bu suretle, bu şekilde, bu türlü, bugünlerde, buna, buna benzer, bundan, bundan başka, bunlar, bunu, bunun, bunun gibi, bununla birlikte, burada, buraya, bütün, civarında, çevresinde, çoğu, çoğuna, çoğunu, çok, çok az, çünkü, da, daha, daha çok, daha evvel, daha fazla, daha önce, daha ziyade, dahi, de, dedi, dedik, dediler, dedim, dedin, dediniz, değin, demek, demek ki, demi, dışarda, dışarı, dışarıda, dışarıya, diğer, diğeri, diğerleri, dimi, diye, diyor, dokuz, dolayı, dört, ediyor, eğer, ek olarak, elbette, en, epeyce, eski, eskiden, esnasında, etmek, etrafında, evvelce, evvelki, fakat, falan, felan, filan, gene, gibi, hala, halbuki, halinde, hangi, hangisi, hangisini, hani, hatta, hayır, hem, hemen, hemen sonra, henüz, hep, hepsi, hepsine, hepsini, her, her biri, her ikisi, her ikisini, her ne, her ne kadar, her tarafa, herbiri, herhangi bir, herhangibir, herkes, herkese, herkesi, hiç, hiç birine, hiç birini, hiç kimse, hiçbirine, hiçbirini, içerisinde, içerisine, içi, için, içinde, içinden, içine, iken, iki, ikisi, ikisini, ilaveten, ile, ileri, ileride, ileriye, ilk, ise, ise de, işbu, işte, itibarıyla, itibarıyle, itibariyla, itibariyle, iyi, kaç, kadar, karşı, kendi, kendi kendine, kendi kendini, kendi kendinize, kendi kendisine, kendi kendisini, kendilerinde, kendilerine, kendilerini, kendiliğinden, kendim, kendin, kendine, kendini, kendinin, kendiniz, kendinizde, kendinize, kendisi, kendisinin, keza, ki, kim, kime, kimi, kimin, kimisi, kimse, lakin, madem, mı, mıdır, mısın, mısınız, mıydı, mıyım, mi, midir, misin, misiniz, miydi, miyim, mu, mudur, musun, musunuz, muydu, muyum, mü, müddetince, müdür, müsün, müsünüz, müydü, müyüm, nasıl, ne, ne kadar, ne sebeple, ne vakit, ne zaman, neden, nedeniyle, nedir, nerede, nereden, neredeyse, nereye, nesi, netice olarak, neyse, niçin, niye, o anda, o halde, o hususta, o kadar, o noktada, o türlü, o vakit, o yer, o yere, o zaman, o zamanın, o zamanki, oldukça, olmak, olmakla beraber, olur olmaz, on, ona, ondan, ondan sonra, onlar, onlara, onlardan, onları, onların, onlarınki, onu, onun, onunki, ora, orada, oradaki, orası, orasında, oraya, oysa, oysaki, öbür, öbürü, ön, önce, önceden, önceki, önünde, ötede, öteki, öteye, ötürü, öyle, öyle ise, öylesine, özellikle, pek çok, rağmen, sadece, sana, sanki, sebebiyle, sebep, sekiz, sen, senden, seni, senin, seninki, sırf, siz, sizden, size, sizi, sizin, son, son derece, sonra, sonuç olarak, sözcük, süresince, şahıs, şahsı, şayet, şey, şeyden, şeye, şeyi, şeyler, şimdi, şöyle, şu, şu anda, şu halde, şu kadar, şu sırada, şuna, şunda, şundan, şunlar, şunu, şunun, şurada, şuraya, ta kendisi, ta ki, takdirde, takriben, tam, tamamen, tamamı, tamı tamına, tastamam, tekrar, tıpkı, tıpkısı, tüm, tümü, üç, üstelik, üstü, üstünde, üstüne, üzere, üzerinde, vaktiyle, var, vasıtasıyla, vb, ve, veya, veyahut, vs, ya, ya da, yada, yahu, yahut, yakınında, yaklaşık, yalnız, yanında, yani, yalnız, yapar, yapıyor, yapmak, yeniden, yerine, yıl, yine, yoksa, yukarı, yukarısı, yukarısında, yukarıya, yüzünden, zarfında, zaten, zira, ziyade

Annex-2- Seed Word List

Seed Words for Preparing the Classification Lexicons
abartma, abla, acı, acılı, adalet, adres, alakasız, açık paket, aç kaldık, açım, adet, adres, ağbi, ağız tadı, aksaklık, akşam, alakasız, alış veriş, alışveriş, amatör, anne, aracılık, artık yeter, asla, aşırı, aşırı yağlı, aynı hata, ayran, az, az yağlı.
baba, bafra pidesi, bahane, bakmak, bayat, bayat ürün, bekle, berbat, besleyici, bıçak, bıkmak, bilet, bimutlulukgetir, bir daha asla, boş, boş köfte, bozuk, bozuk salata, bölge, bulantı, burger, buz, buz gibi, büfe, büyük boy, büzüşmüş.
cafe, canım çekti, canlı yardım, cevap.
çatal, çeşit azlığı, çiğ köfte, çocuk, çok, çok az, çok iyi, çok kızarmış, çok kötü, çöp, çözüm.
dağılmış, dakika, değer, deli olmak, deneyim, dışarıdan sipariş, dikkatsiz, dip, diş kovuğu, diyet, doğru bir nokta, domates, double, doymadım, doymayan, döner, duble, dükkan, dürüm, düzeltme.
eğitim, eğitimsiz personel, ekmek, ekmek arası, eksik, eksik geldi, eksik ürün, en beğendiğim, en güzel, en güzel yanı, en kötü, en kötü yemek, et et döner, etraf, extra, ekstra, ev, evde yemek, ev yemeği.
fark, fast food, fast food zinciri, fazla fiyat, fındık lahmacun, fıstık lahmacun, firma, fiyasko, fiyat, fiyat politikası.
gaflet, gıda, gram, gece, gece yarısı, gecikme, geç gelen, gönderi, gönderme, güzel.
haber ver, hak, hak etmiyor, hamur, hamur gibi, hamburger, hata, hayal kırıklığı, hazır yemek, hediye, helal etmiyorum, helal olsun, hesap, hızlı dönüş, hızlı yemek, hijyen, hizmet, hizmet kalitesi, hizmet sıfır, hizmet verme, homeburger.
ıslak, ıslak hamburger.
iade, iade talebi, içecek, içi boş, iftar, iletişim, iletişim sorunu, ilgisiz, indirim, indirim kodu, insaf, insan sağlığı, internet sitesi, internetten yemek siparişi, ishal, itham, itibar, iptal, işte yemek, İtalyanpizza, İtalyan lezzetleri, iyi fikir.
joker, joker indirim, kaba, kampanya, kardeş, karadeniz pidesi, karışık pide, kart, kasa, kaşar, kaşarlı, kaşık, kavurmalı, kavurmalı kaşarlı, kayış gibi, kazanma, ketçap, keyif, kıkırdak, kıl, kıral gibi, kıralsın, kıymalı, kıymalı pide, kızarmış, kızartma, king burger, koku, kola, köfte, köpek gibi açım, kötü, kötü puan, kredi kartı, kurumsal, kurye, kusma, kuşbaşılı, kuşbaşılı kaşarlı, kuş tüyü, küçük, küçük boy,.
lahmacun, latte, lanet, lanet olsun, lanet ediyorum, leziz, lezzetli, lezzetsiz, limit
mağdur, mağduriyet, mahal, mail adresi, malzeme, malzeme eksikliği, manipülasyon, mayhoş, mayonez, memnuniyet, memnuniyetsizlik, menü, mönü, merkez, mesaj, mide, minimum, minimum tutar, mis, mis gibi, mobil uygulama, muamma, muhatap, multinet, mutfak, mükemmel, mükemmel ürün, müşteri, müşteri hizmetleri, müşteri memnuniyeti, müşteri memnuniyetsizliği.
nakit, ne yesem, niyet, numara.
objektif, olumlu, olumlu yorum, olumsuz, olumsuz yorum, onay, online, online ödeme, online sipariş, orta boy, otomatik ulaşmak, oyala.
ödeme, ödeme yöntemi, öğle, öğrenci, öğün, öncelik, öneri, özen, özensiz, özür.
paket, paketleme, paket servisi, paket sipariş, para, para iadesi, patates, personel, personel ilgisizliği, peynirli, peynirli pide, pide, pilav üstü döner, pişmanlık, pipet, pişmiş, pizza, pizzasipariş, pizza siparişi, poşet, problem, promosyon, puan.
resmi tatil, restoran, restoran zinciri, restorant, restorant zinciri, restaurant, rezalet.
saat, saç, saçmalık, sağlıklı, sağlıksız, salata, Samsun pidesi, Samsun pidecisi, sanki, sayfa, saygısızlık, servis, servis elemanı, servis sıfır, set card, setcard, setkart, severek, seviliyorsun, seviliosun, seviyorum, sıkıntı, sinek, sipariş, sipariş hattı, sipariş iptali, sipariş notu, sipariş onayı, sipariş öncesi, sipariş sonrası, sistem, sodexo, soğuk, soğumuş, son sipariş, sonuç, sorumlu, sorumsuzluk, sorun, sos, sübjektif, suç, suçlama, sufle, süre, şımartmak, şikayet, şube, şüpheli.
taahhüt, takip, talep, talep etmek, tam zamanında, tat, tavuk, tavuk burger, tavuk döner, tavuk hamburger, tecrübe, tehlike, telafi, telefon, temiz, terleme, teslim, teslim etmek, teslimat, teslimat süreci, teslimat süresi, teşekkür, ticket, tövbe, trend, tutar, tuz, Türk mutfağı, tüy.
uğraşıyorum, ulaşamamak, umursamaz, umursuz, unutmuş, urfa, uygulama, uygun lokasyon, uzak.
ücret, ücret iadesi, ürün, üstü boş, üşenmek, üye.
vıcık, vıcık vıcık, vurdum duymaz.
Web sitesi
yağlı, yalan, yanlış, yalnış, yanlış sipariş, yanlış sipariş, yanmış, yapışmış, yaptırım, yardım, yarım, yaşasın, yazık, yemek, yemek arası, yemeksepeti, yemek sepeti, yemek yapma, yemek yok, yeter artık, yetersiz, yetki, yetkili, yettim, yoğunluk, yol, yorum, yumurtalı, yüzde.
zaman, zamanında, zehirlendim, zehirlenme, zevk veren.

Annex-3- Grouped Words for Boosted Sentiment Lexicon

Grouped Words (Aggregation of Words) for Boosted Sentiment Lexicon
{"acele":["acele", "acil"], "ağzımın su":["ağzımın su", "azımın su"], "aksiyon":["aksiyon","ekşın"], "aralıksız":["aralıksız", "durmadan"], "hediye":["armağan", "ödül", "hediye"], " aşık":["asko", "aşko", "aşkito", " aşık"], "beğeni":["beğeni","beğend"], "canı çek":["canım çek", "canı çek"], "çık hayat":["çık aklı", "çık hayat"], "eline sağlık":["eline sağlık", "ellerinize sağlık"]," esas ":[" elit"," esas "], "gönül":["gönül", "gönlü"], "hakikaten":["hakikaten","hakkaten"], "hapır hupur": ["hapır hupur", "hapur hupur"], "hayırlısı":["hayır ol","hayırlısı"], "hastası":["hastası", "hasta ol"], "hemen":["hemen", "hızla", "hızlı", "ivedi", "çabuk"], "inşallah":["inşallah", "inş "], " istiyor": [" istiyor"," ister"," istey"," istemiş"," isted"], "iyilik":["iyice", "iyilik"], "kalp": ["kalp", "kalb"], "kardeşim":["kardeşim", "kardeşlerim"], "keyif":["keyf","keyif"], "lezzet":["leziz", "lezzetli", "nefis", "lezzet"], "müsaade":["müsade","müsaade"], "otomatik":["otomasyon", "otomatik"], "özenli":["özendir", "özenli"], "saygılı":["saygılı", "saygın"], "şükür":["şükür", "şükran"], "taktir":["taktir", "takdir"],"uygun":["uygun"," uyum"],"usta":["usta"," piri "], "teşvik":["teşfik","teşvik"], "yakışıklı":["yakışıklı", "yakışır"], "abuk sabuk":["abidik gubidik", "abuk sabuk", "antin kuntin"], "acayip":["acayip", "absürt"],"acemi":["acemi","toy"], "açlık": ["aç kald", " açım"," açız", "açlıktan bayıl"], "ağlıcam":["ağlıcam","ağlicam"], "sövmek":[" amk","amq","aq"," mk ", "skm","skt"," söv"], "aşırı yağlı":["aşırı yağ", "çok yağlı"], "bağlanmak": ["bağlanam","bağlanm"], "bomboş":["bomboş"," boş "], "yasak":["banla", "yasak"], "bela": ["belanı versin", "belanızı versin"], "bıkmak":["bıkkın", "bıktım"],"boşuna":["boş yere", "boşuna"], "bozuk":["bozuk", "bozulmuş"], "donmuş":[" buz ","donmuş"],"pişmemiş":[" çiğ ", "pişmemiş"], "çökmüş":["çökmüş", "çöktü"], "dağınık":["dağınık", "dağıl"], "dağ başı":["dağ baş", "dağbaş","dağın baş"], "deli ol":["deli ol", "delir"], "duygu sömür":["duyar kas", "duygu sömür"], "gecik":["gecik", "geç gelen"], "gergin":["gergin", "geril"], "getirmek": ["getirem", "getirm"], " uzak": [" ırak "," uzak"], " istemiyor":[" isteme"," istemiyor"],"kafayı ye": ["kafayı ye", "kafayı yi"], "kahrolsun":["kahretsin","kahrolsun"], "kaldık":["kaldık","kaldım"], "kirli": ["kirlenmiş","kirli"], "lanet":["lanet","nalet"],"negatif":["negatif"," eksi "], "orospu":["or*spu çocuğu","or*spuçocuğu","orospu"], "rezil":["rezalet", "rezil"], "soğuk":["soğuk", "soğumuş"], "trip":["trib","trip"], "ulaşam":["ulaşam","ulaşmam", "ulaşm"], "umrunda değil":["umrumda değil", "umrunda değil"], "üzgün":["üzgün","üzül", "üzüyo", "üzücü"], "yanık":["yanık", "yanmış"], "yeter artık":["yeter artık", "yeter yahu"], "yetersiz":["yetersiz", "yok"], "zarar ":["zarar ","zararlı"], "zoraki":["zor bela", "zoraki", "zorla", "zorunda bırak", "zorunda kal"]}

Annex-4- Boosted Sentiment Lexicon

Boosted Sentiment Lexicon Words
["10 puan","number one", "10/10", "6/10", "8/10", "acar", "acele", "adalet", "aferin", "afiyet", "ağzımın su", "ak pak", "aksiyon", "alakalı", " alfa", "anında", "anlamlı", "aralıksız", "arkadaş", " artı", "arzu"," asıl "," aşık", "aşer"," aşkın", "avantajlı", "bağımlı", "bahşiş", "başarılı", "başlıca", "bayıl", "bayram", " baz ", "bebeğim", "beğeni", "beklenti", "beyefendi", "bilerek", "bilgili", "bilinçli", "boğazımdan geçm"," bol ","bravo", "buruk", "canı çek", "cansın", "centilmen", "ciddi", "çare", "çekiliş", "çeşni", "çık hayat", "çılgın", "çok iyi", "çözüm", "daima", "dayanışma", "değerli", "demlenmiş", "dengeli", "derle", "derman", "destek", "devamlı", "dikkatli", "dilek", "doğal", "doğru","dolu","doyam","dua", "duyarlı", "dürüst", "düşünceli", "düzelt", "düzenli", "efendi", "eğitim", "eksiksiz", "ekstra", "ev yemeği", "eline sağlık"," esas ", " esen "," eşit"," etkili", "faydalı", "fazilet", "gayet", "gerçek ","gerekli"," göm", "gönül", "görev", "görgülü", "gurur", "güven", "güzel", "hakikaten ", "hakkıyla", "halis", "hapır hupur", "hasret kal", "hassas", "hastası", "havalı", "hayati", "hayırlısı", "hayran", "hediye", "helal", "hemen", "heves", "hijyen"," hoş ","hukuk", "huzur", "içim", "ikram", "ilave"," ilgili", "iltifat", "incelik", "indirim", "insan sev", "insani", "inşallah"," istiyor", "işbirli", "iştahlı", "itibar", "iyi fikir", "iyi puan", "iyi yemek", "iyilik"," izin", "jest", "joker", "kabul", "kalp", "kalıcı", "kaliteli", "kanka", "kardeşim", "karlı", "kazan", "kefil", "kesintisiz", "keyif", "kızarmış", "kibar", "kolay", "kral", "kurban", "latif", "layık", "lokum", "makbul", "mantıklı", "medeni", "memnuniyet", "merhamet", "mesut", "meşhur", "minimum tutar", "minnet", " mis ", "motivasyon", "muhteşem", "mutlu", "mükemmel", "müsaade", "müsait ","nazik", "nitelikli", "nizam", "olağanüstü", "olumlu", "onay", "onur", "optimum", "otomatik", "öncelik", "önem", "öner", "özel", "özenli", "patla", "pişkin", "pişmiş"," pls ", "pozitif", "prestij", "profesyonel", "promosyon", "rahat", "rica", "rüzgar", "safa", "sağlam", "sağlıklı", "sakin", "salim", "saygılı", "sefa"," seri", "sev", "sıcacık", "sipariş onay","sistem","sorumlu","stalk","sürekli", "sürpriz", "şahane", "şerefine", "şevk", "şımart", "şükür", "taahhüt", "tadı güzel", "takip", "taktir", "talep", "tarafsız", "tatlı", "tavla", "taze", "tecrübe", "teklif", "telafi", "temel", "temiz", "terbiyeli", "teşekkür", "teşvik"," tez ","titiz", "tolerans", "toparla", "toplu", "tutul", "ucuz", "umut", "usta", "uyanık"," uygun", "ücretsiz", "üstlen", "üstün", "vaktinde", "verimli", "vurul", "yakın", "yakışıklı", "yarar", "yardım", "yaşasın", "yemek video", "yeni ","yerinde", "yeterli", "yoğunlaş", "yöntem", "yüce", "zahmet olmazsa", "zevkli", "abart", "abes", "abuk sabuk", "acayip", "acemi"," acı ","acitasyon", " açık ", "açlık", " adi ", " ağır","ağlıcam", "ahlaksız", " aksak"," aksi ","alakasız", "alt tarafı", "sövmek", "andaval", "anksiyete", "anlamsız", "arıza", "asılsız", "aşağı", "aşınmış", "aşırı yağlı", "ayar ol"," ayıp"," ayı "," azal", "azar ye", " azarla", "bağlanmak", "bahane", "balık hafıza", "balon", "yasak", "basit", "başarısız"," bat", "bayağı", "bayat", "baygın", "beceriksiz", "beddua", "beğenm", "beklet", "bela", "bencil", "bitmiştir", "berbat", "bereket", "beter", "beyhude", "beyinsiz", "bez", "bıkmak", "bilgisiz", "bilinçsiz"," bilme", "bin pişman", "daha asla", "bitkin", "blokl", " bok ", "bomboş", "boşuna", "botla", "boykot", "bozuk", "böcek"," bulan", "donmuş", "büyütme", "cahil", "camış","cansız", "cenabet", "cennet", "cereme", "cesaretim yok", "ceza"," cılk", "crash","çakal","çamur","çekin","çelişki","çıkar","çıkış","çıldır","pişmemiş","çirkin", "çok kötü", "çökmüş"," çöp ","çürü", "dağ başı", "dağınık", "dalgın", "dandik"," dar ", "dayanma sınır", "dedikodu", "değersiz", "deli ol", "dengesiz", "dert"," dışı", "diken üst", "doym", "dökülmüş", "dönek", "dram", "duygu sömür", "duyarsız", "düğüm", "düşman", "düşük", "düşüncesiz", "düzenbaz", "düzensiz", "edepsiz"," eksik ", "eleştiri", "enayi"," engel", "erimiş"," ertele"," eşek ","eyvah"," ezik","fake","fani", "fasa fiso", "faydasız", "fazla", "felaket", "feleğim şaş","fena","fırsatç","fiyasko", "gaflet", "gazabına uğra", "geber", "gecik", "geçici", "gelm", "gereksiz", "gergin", "getirmek", "gıcık", "gına gel", "görgüsüz", "gözü karart", "gudubet", "hadsiz", "hak etmiyor", "hakaret", "haksız", "halt", "hamur", "haram", "hatalı", "hayal kırık", "haysiyetsiz", "hayvan", "hazetm","hikaye"," uzak", "ıslanm","ıssız","iade","ibne","iflah olm", "iflas", "iğren", "ihmal", "ihtar", " ilgisiz"," illet", "insaf", "iptal", "israf"," istemiyor", "istismar", "isyan", "işkence", "işsiz", "iştahsız", " itici", "iyi değil", "kaba ","kafayı ye", "kahrolsun", "kalas", "kaldık", " kalm","kan emici", "kanser", "kapısının önü", "kara kara düşün","kara liste", "karışık", "kaybet", "kayış", "kazık", "kendimi tut", "keşke"," kını", "kırıl", "kıtlık", "kifayetsiz", "kilo ", "kirli", "kitle", "kokan", "kokm", "kopya", "korkunç", "köle", "kömür", "kötü ","köylü"," kuru ", "kusm", "kusur", "küçük", "küflü", "küstah"," laf ", "lakayıt","lanet", "lastik", "leke"," leş ", "lezzetsiz", "lüzumsuz", "mağdur", "mahsur", "mantıksız", "manyak", "maraz doğ", "mesele", "midem bulan", "mikrop","minik","muhatap","mutsuz","nefret","negatif","olacaksa ol", "odun"," oha ", "olumsuz", "orospu", "ortadan kaldır", "oyala", "öküz"," ölü ","özensiz", "özür", "pahalı", "panik", "paranoyak", "pezevenk"," pis ", "pişman", "problem", "psikopat", "rasgele", "reddet", "rezil", "risk", "rötar", "ruh hasta", "saçma"," saç ", "sağlıksız", "sahte", "sakat", "sakınca", "salak", "salla", "saman", "sası", "savsak", "saygısız", "sebepsiz", "serseri", "sert", "ses seda yok", "sıfır", "sıkıl", "sıkıntı", "sıradan", "sızlan", "sinek", "sipariş hata", "sipariş iptal", "sitem", "sivrisinek", "skandal", "soğukta", "soğuk", "sorm", "sorumsuz", "sorun", "sömür", "sözde"," suç", "sürün", "şaka", "şans gülm", "şerefsiz", "şeytan", "şımarık", "şikayet", "şişir", "şişko patates", "şopar", "şüpheli", "taciz", "tadı yok", "takat", "talan et", "talihsiz", "tasa ", "taş ","tatava yap", "tatsız", "tehlike", "tekel", "telaş", "tembel", "terbiyesiz", "ters", "tırs", "toz", "trip", "tuzsuz","tüketme","tükür","tüy","uçurum","ufak","uğraş","ukala"," ulan", "ulaşam", "umrunda değil", "umursam", "unutmuş", "unutul"," usan"," utan", "uyar", "uydur", "ürkütücü"," üşen", "üzgün", "vahim", "vasat", "vazgeç"," verem", "verimsiz", "vıcık","vicdan azab","vicdansız", "vizyonsuz", "yağı don", "yağsız", "yalaka", "yalan", "yanlış", "yamuk", "yanık", "yanıl", "yapışmış", "yaram", "yaratık", "yaşlan", "yavaş", "yavşak", "yazık", "yemek yok","yemiyo","yerlerde", "yeter artık", "yetersiz", "yorgun", "zarar ","zehir", "zevksiz", "zıkkım", "ziyan", "zoraki"]

Annex-5- Grouped Words for Boosted Product-Service Systems Lexicon

Grouped Words (Aggregation of Words) for Boosted Product-Service Systems Lexicon
{"adet":["adet", "tane"],"çok yağlı":["aşırı yağ", "çok yağlı"],"bozuk":["bozuk", "bozulmuş"], "donmuş":[" buz ","donmuş"], "canı çek":["canım çek", "canı çek"],"pişmemiş":[" çiğ ", "pişmemiş"], "damak tadı":["damak tadı", "damak zevk"], "duble":["double","duble"], "doyurucu":["doyum", "doyurucu"], "ekşimiş":["ekşim"," ekşi"], "ev yapımı":["el yapım", "ev yapım ", "ev yemeği"],"fast food":["fast food","fastfood"],"lezzetli":["leziz", "lezzetli", "nefis"], "soğuk":["soğuk", "soğumuş"], "vejetaryan":["vejetaryan", "vejeteryan"], "lira":["₺","lira", "nakit", " para ", " tl"], "acemi":["acemi", "toy"], "konum":[" alan ","adres", "bölge", "civar"," etraf", "konum", "mahal", "mevki", "muhit", "semt", "sokak"], "alışveriş":["alış veriş", "alışveriş"], "alt sınır":["alt limit", "alt sınır"], "anasayfa":["anasayfa","ana sayfa"], "aplikasyon": ["aplikasyon", " app","mobil uygulama"], "boşuna":["boş yere", "boşuna"], "callcenter": ["callcenter", "call center","canlı destek", "canlı yardım"], "çökmüş":["çökmüş", "çöktü"], "dağın baş":["dağ baş", "dağbaş","dağın baş"], "dağınık":["dağınık", "dağıtıcı", "dağıtım ağı"], "dakika": ["dakika","dk","saat"], "davranış":["davranış","davranm"], "debit":["debit","setcard","sodexo", "multinet"], "dışardan sipariş":["dışardan sipariş", "dışarıdan sipariş"], "entegrasyon": ["entegrasyon","entegre"], "fiyat":["fiat","fiyat"], "gecik":["gecik", "geç gelen"], "gel al":["gel al", "gel-al"], "gelen abi":["gelen abi", "gelen arkadaş"], "gönderi": ["gönderi","gönderm"], "hemen": ["hemen", "hızlı"], "hes cod": ["hes cod", "hes kod"], " uzak":[" ırak ", " uzak"], "iletm":[" ileti", "iletm"], "kapalı":["kapalı", "kapanmış", "kapanış", "kapatmış", "kapatmak"], "kara liste":["kara liste","karaliste"], "kokorecci":["kokoreççi", "kokorecci"], "kurye":["kuriye", "kurye"], "posta":["mail", "posta"], "min tutar":["min paket tutar", "min sipariş tutar", "minimum sipariş tutar", "minimum tutar"], "motokurye":["motokurye", "motosiklet", "motor"],"nerde kal":["nerde kal", "nerden gel"], "otomasyon":["otomasyon", "otomatik"], "özenli":["özenli", "özen göster"], "paket servis":["paket servis", "paket sipariş"], "rasgele": ["rasgele", "rastgele"], "saygılı":["saygılı", "saygın"], "sipariş iptal":["sipariş hata", "sipariş iptal"], "soğuk hava": ["soğukta", "soğuk hava"], "telefon":["telefon","mesaj","sms"], "yemekçi": ["sepetçi", "yemekçi"],"trip":["trib","trip"], "ulaşmam":["ulaşam", "ulaşmam", "ulaşm", "ulaşım"]}

Annex-6- Boosted Product-Service Systems Lexicon

Boosted Product-Service Systems Lexicon Words
["abur cubur", " acı ", "adet", "ağız tadı", "altın günü yiyecek", "ana yemek", "aperatif", "çok yağlı", "ayran", "azıcık", "baharat", " bal ", "bayat", "besin", "biber", "bisküvi", " bol ", "bozuk", "böcek", "cips", "börek", "burger", "buruk", "donmuş", "büyük boy", "canı çek", "çeşni", "çevirm", "çıtır", "pişmemiş", "çoban", " çöp ", "çörek", "dağıl", "damak tadı", "dolma", "domates", "duble", "doym", "doyurucu", "döner", "dünden kalan", "dürüm"," ekle ","ekşimiş", " ekmek ", " eksik ", "ev yapımı", " erzak", "gıda", "eşantiyon", " etli", " etsiz", "fast food", "futuristik", "garnitür", "gevrek", "gluten", "gram", "gurme", "hafif yemek", "hamburger", "hamur", "haram", "hastası", "hatay usul", "havyar", "hazır yemek", "helal"," hoş "," ısıt", "ıspanak", "iade", "içecek"," içki ", "içli köfte", "ikram", "kadınbudu", "kalın hamur", "kayış", "karbonhidrat", "karışık", "kaşarlı", "katık", "kavurma", "köpek yem", "kestane", "ketçap", " kıl ", "kıtır"," kıyma", "kızarmış", "kızartm","kilo ", "kişi başı", "klasik", "kokan", "kokm", "koku", "konserve", "köfte", "kömür", "kurt", " kuru ", "kusm", "kuşbaşı", "küçük", "küflü", "lahmacun","lastik", "latif","latte", "lüfer", "lezzetli", "lezzetsiz", "lokma", "lokum", "makarna", "martı eti", "marul", " maya", "mayonez", "meyve", "meşrubat", "meze ","mide", "minik"," mis ","nane", "organik", "orta boy", "öğün", "ölçü", "patates", "patlıcan", "peynir", "pide", "pilav", "pişmiş", "pizza", "poğaça", "porsiyon", "pörsüm", "reçel", " saç ", "sağlıklı", "sağlıksız", "salata", "salça", "saman", "sarımsak", "sası", "sebze", "sıcacık", "sinek", "soğan", "soğuk", "son kullanma tarih"," sos ", "sucuk", "sufle", "sulu", "şalgam", "şeker", "şerbet", "taraftar", "taş ", "tatlı", "tatsız", "tavuk", "taze", "tereyağ", "turşu", "tuzlu", "tuzsuz", "ufak", "urfa", "ürün", "vegan", "vejetaryan", "vıcık", "yağı don", "yağlı", "yağsız", "yanık", "yanmış", "yapışmış", "yarım", "yemiş", "yeşillik", "yöresel", "yudum", "yumurta", "zehir","zevkli", "lira","3d secure", "abone", "acele", "acemi"," açık "," açıl", "adalet"," adam", "adisyon"," adi ", "ahlaksız", "ak pak", " aksi ", "aktarm", "alakalı", "alakasız", "konum", "alçal", " algı", "alışveriş", "alt sınır", "altyap", "anasayfa", "anında", "aplikasyon", "anlamsız","anlaşılm", "anlayış", "aracı", " asıl ", "asistan", "aşağı"," ayı "," azarla", "bağlantı", "bahşiş", "bakım", "basit", "başlıca", "bayağı", " baz ", "beceriksiz", "bedel", "beklet", "belgeli", "bencil", "beyhude", "beyinsiz", "bıçak", "bildiri", "bilet", "bilgisiz"," anda ", "blokl", "boşuna", "bölüm", "buton", "cahil", "callcenter", "cevap", "crash", "cüzdan", "çaba", "ivedi", "çağrı", "çakal", "çalışan", "çalışm", "çatal", "çelişki", "çevre", "çıkış", "çiçek", "çirkin", "çökmüş", "dağın baş", "dağınık", "dakika", "dalga", "dalgın", "danışma", "davranış", "debit", "değerli", "değersiz", "deney", "denk", "dert", "destek", "devre", "dezenfekte", "dış kapı", "dışardan sipariş", "dikkate alm", "diyet", "doğru", "dönem", "duyarsız", "dükkan", "dürüst", "düşük", "düşüncesiz", "düşünceli", "düzensiz", "edepsiz", " ederi"," eğitim", "eksiği", "eleman", "emekçi", "entegrasyon", " esas ", " esnaf "," eşek ", " eşit", "fatura", "faydasız", "fazilet", "fazla", "fiyat", "filtre", "firma", " fiş", "gamsız", "gayret", "gece yarısı", "gecik", "gel al", "gelen abi", "gelir", "gerçek ", "gereksiz", "getiren", "getirtm", "gönderi", "görev", "görgülü", "görgüsüz", "görmem", "görüş", "götürm", "haberleş", "hack", "hain", "hakaret", "hata ver", "havalı", "hayvan", "hediye", "hemen", "hes cod", "hesap", "hizmet", "hukuk", " uzak", "ısmarlam", " icra", "ihmal", "ihtar", "iletm", " ilgili", "ilgisiz", " ilişki", " ilkel", "indirim", "influencer", "insan sağlığı", "internet site", "internetten yemek siparişi", " işçi", "işlem", "işletme", "işyeri", " izin", "joker", "kaba ", "kaide", "kalas", "kampanya", "kapalı", "kapıda öde", "kara liste", "kargocu", "karşılı", "kasiyer", "kaşık", "katır", "kaytar", "kazan", "kazık", "kdv", "kebapçı", "kısır", "kıymetli", "kifayetsiz", "kokorecci", "komisyon", "konsept", "kota", "kart", "kullanıcı dost", "kupon", "kural", "kurye", "kurumsal", "küçümseme", "küstah", "lakayıt", "legal", "posta", "maliye", "malzeme", "mekan", "memur", "mendil", "menü", "merkez", "mesafe", "mesele", "meslek", "mevsim", "mezun", "miktar", "min tutar", "misli", "motokurye", "muamele", "muhatap", "mukabil", "mutfak", "mücadele", "müdavim", "müddet", "müşteri", "naçiz", "nazik", "negatif", "nerde kal", "nizam", "nöbet", "numara", "odun", "online", "operatör", "ortak", "otomasyon", "ödeme", "ödev", "öküz", "öner", "örgüt", "özenli", "özensiz", "pahalı", "paket servis", "paket", " pay ", "peçete", "perhiz", "personel", "pipet", "platform", "portör", "pos cihazı", "poşet", "pratik", "promosyon", "prosedür", "puan", "range", "rasgele", "rejim", "reklam", "restoran", "rozet", "rötar", "ruhsat", "rut dışı", "sağanak", "salah", "sapa", "satış", "savsak", "saygılı", "saygısız", "yemekçi", " seri", "server", "servis", "sezon", "sıradan", "sırılsıklam", "sipariş hattı", "sipariş iptal", "sipariş not", "sipariş onay", "sipariş önce", "sipariş sonra", "sistem", "sitem", "soğuk hava", "sorumlu", "sorumsuz", "sorun", "stil", "story", "sunucu", " süre", "şimdi", "şirket", "şube", "taciz", "tahsil", "tahsis", "taksit", "tarih", "tarz", "taşı", "tatbik", "tayfa", "tecrübe", "tehlike", "tek kişi", "teklif", "teknoloji", "telefon", "temas", "temel", "temiz", "temsilci", "terbiyesiz", "ters", "teslim", " tez ","ticket", "titiz", "token", "tolerans", "toplam", "trafik", "trip", "tutar", "tutum", "tüketici", "tükür", "türe", "tüy", "uğraş", "ukala", "ulaşmam", "umursam", "unutmuş", "usta"," usul", "uyar", "uygula", " uyum", "ücret ", "üyeli", "üzücü", "vakit"," vale", "vasıta", "verimli", "verimsiz", "viral", "virüs", "web site", "webchat", "yağmur", "yakışıklı", "yalın", "yanlış", "yanıt", "yaptırım", "yarar", "yasal", "yatkın", "yavşak", "yazılım", "yerel", "yetersiz", "yetkili", "yoğunluk", "yollam", " yol", " yorum", "yoz", "yönetici", "yöntem", "yürütm", "zaman", " zam ", "zararlı", "zihniyet"]