Journal of Applied Research and Multidisciplinary Studies (JARMS)
Volume 4 Issue 1, Spring 2023
ISSN(P): 2707-5087 ISSN(E): 2707-5095
Homepage: https://journals.umt.edu.pk/index.php/jarms

Article QR

Title:	**Employing Sentiment Analysis to Enhance Customer Relationships for Mobile Phone Operators Working in Pakistan**
Author (s):	Asad Jamil[1], Mira Aamir Mehmood[1], Imran Saleem[2], Raja Asif Wagan[1]
Affiliation (s):	[1]Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan [2]University of Management and Technology, Lahore, Pakistan
DOI:	https://doi.org/10.32350/jarms.41.09
History:	Received: March 12, 2023, Revised: April 15, 2023, Accepted: June 15, 2023, Published: June 30, 2023
Citation:	Jamil, A., Mehmood, M. A., Saleem, I., & Wagan, R. A. (2023). Employing sentiment analysis to enhance customer relationships for mobile phone operators working in Pakistan. *Journal of Applied Research and Multidisciplinary Studies*, *4*(1), 00–00. https://doi.org/10.32350/jarms.41.09
Copyright:	
Licensing:	
Conflict of Interest:	Author(s) declared no conflict of interest

A publication of
School of Professional Advancements
University of Management and Technology, Lahore, Pakistan

Employing Sentiment Analysis to Enhance Customer Relationships for Mobile Phone Operators Working in Pakistan

Asad Jamil [1], Mira Aamir Mehmood[1], Imran Saleem [2*], Raja Asif Wagan [1]

[1]Faculty of Information & Communication Technology, Balochistan University of Information Technology, Engineering, and Management Sciences, Quetta, Pakistan

[2]Dr Hasan Murad School of Management, University of Management and Technology, Lahore, Pakistan

Abstract

Customer Relationship Management (CRM) is a process through which a company or an organization manages its contacts and coordinates with its customers, usually by analyzing vast volumes of data. CRM systems collect information from a variety of sources, such as a company's website, phone, email, live chat, marketing materials, and, more recently, social media. Active coordination with the customers is the key to improve the quality of service being provided by a business. Social media has been proven a great tool to spread awareness regarding a particular topic and opinion forming which also gives limelight to public opinion. Thus, CRM enables the organizations to gain a better understanding of their target audiences and how to best respond to their demands, resulting in client retention and sales growth. Sentiment analysis can be characterized as a qualitative approach to data mining that recognizes and separates the subjective data as a source. Moreover, it also helps an organization to understand public opinion regarding the service and products it offers. The current study attempted to present the current state-of-the art in employing sentiment analysis in CRM.

***Keywords*:** customer relationship management (CRM), data mining, data science, natural language processing, sentiment analysis, social media analytics

Introduction

Technological advancements emerged in different fields of computing, such as the Internet of Things (IoT), Cloud Computing, and big data. This advancement in technology caused vast information generation on the internet, particularly on social media platforms. Such platforms offer a simple and convenient medium for their users to share their thoughts and experiences in the form of tweets, comments, blog posts, updates, and reviews. Moreover, it also enabled the customers to share their feedback regarding the services or products over the internet. Such experienced feedbacks about a product can highly influence the decision-making of other expected clients. The reviews in the form of tweets, blog posts, and so on, have also been referred to as free text which is available free of cost for people to read and is unstructured. This data has been used for numerous tasks, such as stock and financial market prediction, healthcare, and sentiment analysis (Sajid et al., 2020).

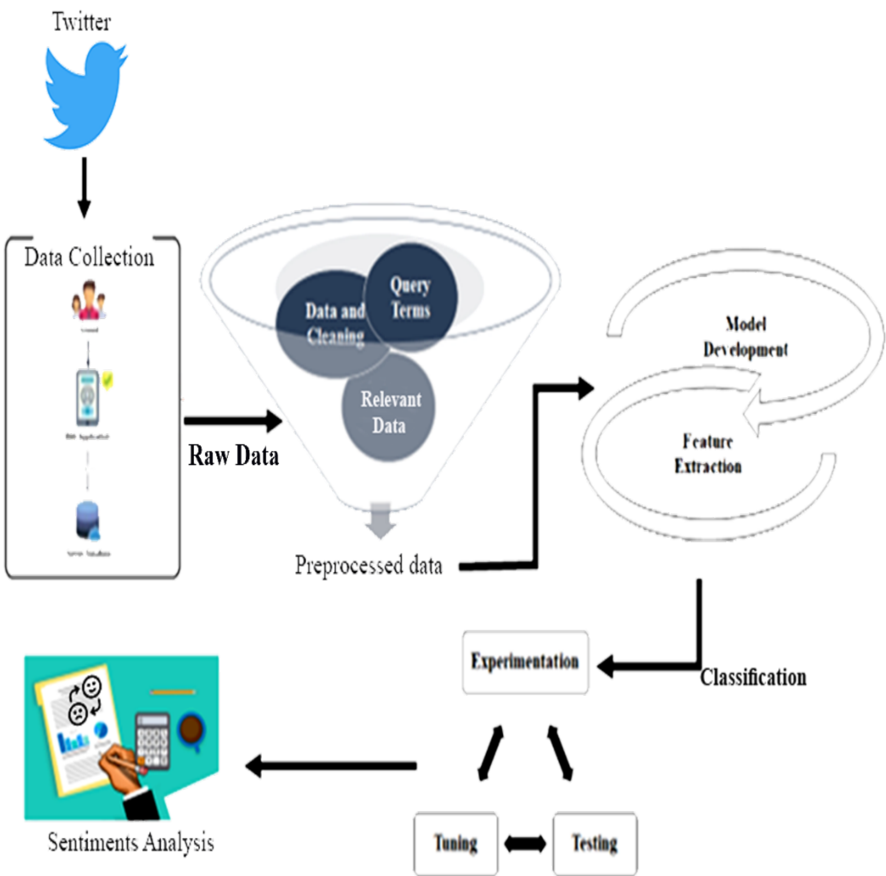
In today's connected world, it is unbearable to compromise on these services. The quality of services is often compromised due to poor Customer Relationship Management (CRM), scattered nature of the service, and a large number of customers are dispersed at different geographical locations which makes it difficult to reach and resolve the problems faced by the customers. Moreover, for telecom operators, inefficient CRM is a bottleneck (Ranjan et al., 2018). Hence, real-time sentiment analysis of customer feedback and CRM are essential to combat this issue (Ullah et al., 2019). The utilization of Twitter data, which is a freely available social media platform as a data source and performing sentiment analysis, can contribute to predict customer satisfaction.

Sentiment analysis is the technique which is used to analyze the public opinion shared on the internet, social media platforms, and any other source. By evaluating the data comprising of public opinion about a topic, product experience, and so on, one can find out the polarity of each data item. The basic concept of sentiment analysis is that it is used to understand the polarity of a text and for extracting subjective data from the source data to be mined. With the use of sentiment analysis for product reviews and such purposes, the concept of CRM emerged, where CRM is used by organizations for managing their contacts and customers (Abdullah et al., 2021).

For an organization, it is essential to evaluate the domains where better measures are required to improve the quality of service. As CRM enables the organizations to gain a better understanding of their target audiences by utilizing the available vast volume of data, helping the companies to better respond to the demands of their customers, results in client retention and sales growth (Khatoon, 2017). What parameters are to be addressed for the satisfaction of a customer? In the meantime, sentiment analysis also supports advertisers and business holders with new perceptions and insights, helping them to make efficient business decisions (Alamsyah & Bernatapi, 2019). The streaming data comes in the JSON form, a semi-structured format, which can be processed in real-time or stored and processed later on to identify the sentiments of the tweets (Hermansyah & Sarno, 2022).

The process of sentiment analysis involves the following generic steps, as depicted in Figure 1. These steps include data collection, pre-processing, model development, feature extraction, classification, experimentation, testing and tuning, and finally the analysis of sentiments generated. Figure 1 provides a visual representation of these sequential stages in sentiment analysis.

Figure 1
Twitter Sentiments Analysis



The current study focused to perform sentiment analysis of the customer feedback shared on the Twitter, that is, a social media platform. Moreover, it also identified the regions from where negative tweets are generated, measured the user satisfaction parameters against the negative tweets, and the most frequent problems faced by the end users. This study helped to understand the influential factors on customer satisfaction and its relationship with the quality of service. Twitter as the dataset, helped mobile phone operators working in Pakistan, to identify the public opinion regarding their services and company to develop competitive intelligence for the benefit of their businesses. This study also facilitated the mobile phone operators to develop better CRP and assist them in terms of improving the quality of service. Moreover, it would also be helpful to determine whether a tweet related to the company is positive or negative through the analysis which is an autonomous solution centered on sentiment analysis. For this study, Python packages, Tweepy, NLP Library, TextBlob, and other relevant packages were used. Python packages has been utilized on Anacondas for different fields of analytics as it provides extensive support to contemporary machine learning algorithms.

Literature Review

In this section, state-of-the art in sentiment analysis in general, and for Twitter, in particular, has been discussed. Sentiment analysis is a broad term that includes data ingestion, storage, data extraction, pre-processing and cleaning of the data, training models on the cleaned/pre-processed data, testing the model, and experimentation and research.

Sentiment analysis has been a growing trend in the domain of natural language processing and data science. Analysis of the sentiments refers to distinguish between the emotions being expressed, such as happiness or sadness. The data before sentiment analysis should go through the following 5 phases namely data gathering, data processing, model engineering, model tuning, and experimentation and results (Abdullah et al., 2021).

According to the findings of Ranjan et al. (2018), most of the research carried out in the domain of sentiment analysis either comprises of Lexicon (dictionary) based approach where a dictionary lookup database exists for classifying the linguistic or the machine learning based approach where the data processing and classification is done by the machine learning model which represents linguistics in a vector format. Therefore, there are two main approaches to test sentiment analysis, that is, machine learning and the lexicon based (dictionary based) approach. Additionally, few most common pre-processing steps used in literature are tokenization, stemming, lemmatization, and stop word removal.

Sentiment Analysis

The machine learning based method is one of the most used approaches in the literature on natural language processing and sentiment analysis. In machine learning based approach, the data is first required to be transformed in the form of vectors-a numerical representation, and then analyzed. There is a wide variety of machine learning algorithms, such as support vector machine (SVM), Naïve Bayes, and logistic regression. The algorithm is provided with the sentiment labeled, that is, positive, neutral, and negative, the model is trained over the data and with enough training examples the model could then make predictions on unseen data (Ullah et al., 2019).

Lexicon-Based Approach

The lexicon-based approach uses a lexical database to predict the sentiments, such as by using WordNet. WordNet is used to obtain a score for every single word that is available in the document. The polarity of the text is derived from set or words, where each of the words carry a weight in the dictionary. The lexicon database consists of opinionated words which are classified into negative and positive. Each word in the document is assigned a numeric value or score. The numeric score values, for the whole document, are summed up to assign the document a polarity (Abdullah et al., 2021).

Pre-Processing Steps

The most commonly used techniques involved in the processing of natural language are tokenization, stop word removal, stemming, and lemmatization. In the case of tweets, it also requires some additional processing, such as removing @mentions, #hashtags, and emoticons along with translating the tweets into a single language, such as English, since the Twitter data consists of multiple languages. Therefore, it can dominate the model if multiple languages are used in the dataset extracted against a topic (Khatoon, 2017).

After noise filtration, raw words are left behind which may contain the sentiment. Breaking down the sentences into words is called tokenization (Alamsyah & Bernatapi, 2019). The NLTK package Treebank word tokenizer was used in this study.

Some of the important preprocessing techniques used in sentiment analysis are stemming and lemmatization (Hermansyah & Sarno, 2022).

These approaches condense the words back to their base or normalized form. In stemming, a word is normalized by removing the suffix from that word, that is, connecting, connected, and connects would become connected. The Porter-Stemmer algorithm is used for performing stemming. It can normalize more complicated structures, such as knives would become knife and women would become woman.

The exclusion of unwanted symbols and words for refinement of free text is called stop word removal. It is the most commonly used pre-processing step for sentiment analysis (Al-Ansari, 2019). In this approach, the most frequently used words are removed, such as the, is, am, and so on. These words have no significance and emotional meaning and ultimately have no impact on sentiment score.

Critical Analysis of Contemporary Studies

The current research proposed a sentiment analysis solution for improving the quality of service by using a patients' feedback platform [PFP]. The authors performed free text analysis for enhancing the healthcare industry. The authors used an MLP (Multi-Layer Perceptron) as the model, claiming 88% accuracy in predicting whether a patient commends the healthcare service is received or not, based on which E-health care services could be improved (Sajid et al., 2020).

In the current study, the customer experience feedback has been studied pertaining to the services being shared on social media platform, that is, Twitter for the telecom brands in India where Twitter data and sentiment analysis are used to understand the polarity of the tweets. The study showed how sentiment analysis of tweets helps companies to identify and address user issues more efficiently in a particular geographical area. As more and more of the world is interconnecting by using technology, the number of telecom users is increasing and addressing the problems of customers through CRM is getting the primary need of the telecom sector, operating around the globe (Ranjan et al., 2018).

Abdullah et al. (2019) studied reputation measurement for the Saudi based telecom companies by using a hybrid sentiment analysis approach. The Twitter was used to conduct this study. The total number of 5000 tweets were extracted and tweets containing other languages, such as English were removed. The authors only used texts, retweets, and favorite attributes of tweets to conduct the study. The Decision Tree (DT), Naïve Bayes, and

SVM classifiers were trained on the pre-processed data. The models achieved Naïve Bayes 87%, Decision Tree 68%, and the SVM achieved the highest score of 95% among the classifiers.

Abdullah et al. (2021) performed the churn prediction for the telecom sector companies by using machine learning approach. The authors studied customer satisfaction and factors behind the customer loss in telecom sector by using CRM's existing data. The authors concluded that identifying the customer satisfaction aspects and employing CRM can help in customer retention.

Khatoon (2017) performed real-time analysis of Twitter data of telecom companies to develop competitive intelligence and improve CRM aspect. A dataset of 3000 tweets against 3 popular networks of Saudi was collected.

In their study, Alamsyah and Bernatapi (2019) examined the evolution of Customer Experience Management (CEM) for internet service provider companies by using customer analytics. Naïve Bayes Classifier was used over 7249 tweets collected from Twitter by using Tweepy API. The model achieved 82% accuracy. The study concluded that the variables of customer concern were service quality, customer support, and billing cost based on the results after topic modelling.

In their study, Hermansyah and Sarno (2022), performed a quality aspect based sentiment analysis to improve the services of Indonesian telecom companies. Twitter data was used to conduct this study. The authors used Term Frequency Inverse Cluster Frequency (TF-ICF) to find the hidden topics from the tweets and expand the product quality dimensions. Moreover, the SVM, Senti WordNet, Senti Circle, and Random Forest were used for the classification of the tweets. The results showed that the SVM model attained the highest accuracy of 96.3%. The researcher discovered that the service aspect (serviceability) had the highest negative sentiments.

The authors performed social media sentiment analysis to choreograph the emotions by enhancing the CRM for telecom sector of Oman. A total number of 83,981 tweets were extracted from Twitter. A major part of these tweets were in Arabic, therefore the authors used Google translate and Google spreadsheet to translate the dataset into English. The NVivo, a software for qualitative data analysis, was used for the data analysis and sentiment generation (Al-Ansari, 2019).

CRM data mining and impact of knowledge management were examined in a study. Moreover, Collaborative CRM (concerns association control with resource channels and partners), ii. Operational CRM (contains automation of customer interaction), and iii. Analytical CRM (concerns customer behavior analysis) were also discussed in the conducted study. Additionally, questions, such as why they are essential for business organizations' growth and survival, were also addressed in it. Classification, regression, forecasting, clustering, association analysis, and visualization being the main data mining techniques were performed on customer loyalty data (Srivastava et al., 2019).

The current study introduced the usage of a CRM system as a new way of information processing. The study primarily emphasized the application and research of data mining technology in CRM, terming it as the new way of information processing. The results clustered the data into positive and negative categories with the highest accuracy of 86% by the multilayer perceptron. The study utilized the information available on social media platforms and websites. This work has been optimal in sentiment analysis and clustering them into categories (Song & Liang, 2021).

A study attempted to explore the impact of customer engagement on Twitter by using data mining. It also predicted the connection between customer engagement along with identifying the social networks generated by customers. The authors used K-Nearest Neighbor (KNN) classifier, Decision Tree C4.5 (C4.5), Support Vector Machine (SVM), and Naïve Bayes (NB). By using the tweets about IKEA as the dataset and applying the results on 4000 tweets, the patterns of electronic word-of-mouth (eWOM) were explored. The results found three types of eWOM as subjective statements, knowledge sharing, and objective statements. Moreover, based on the nature of tweets it was concluded that satisfied customers mostly tweet knowledge sharing eWOM and tend to share their experiences with others more (Díaz-Martín et al., 2015).

A study attempted to introduce a multiclass approach in order to predict the user behavior by utilizing a deep learning framework over Twitter data of general elections in India. The study performed experiments with datasets in two sizes, that is, 25K and ii. 50k tweets. The study concluded that the traditional supervised models including Random Forest, Decision Tree, NB, and SVM performed less efficiently with large-sized data, while the deep learning model improved accuracy up to 99% when the size increased (Mohbey, 2020).

The research reviewed the concept of assigning polarity to the processed tweets in order to find out whether an author expresses a positive or negative opinion. The concept of stop words removal, changing data structure, and emoticon removal was introduced for better cleansing of data. Moreover, a predefined algorithm, pointwise mutual information (PMI-IR 2) was used for the sentiment analysis of the data extracted from Twitter (Seki, 2016).

The impact of CRM was elaborated on a company, moreover the question, why businesses need to use sentiment analysis was also addressed. CRM is the latest marketing concept that is customer centric. The impact of CRM on a business can be explained through certain factors. These factors include value equity (VE) which is a combination of offers, the convenience of a product, brand equity (BR) which includes brand awareness, and brand reputation along with customer equity (CE) which is the classification of customers and also helps to predict the potential profitability concerning the projected loyalty. Authors analyzed the data collected from different sources, such as blogs, Twitter, forums, and news consisting of 1927623 total mentions and determined that 78.2% equivalent to 134000+ were positive, 21.8% equivalent to 34000+ were negative, where the net sentiment was positive and about 56%. The results demonstrated that the system achieved better results with a vast amount of data (Di et al., 2018).

A study presented three different approaches to perform sentiment analysis, namely 1. lexicon based model, 2. machine learning, and 3. psycholinguistic method. Moreover, a comparison was also performed between three approaches focusing on the psycholinguistic method. The data collection was carried out between service outage of Skype and nearly 10,000 tweets were collected per day by using the twitter4j API of Java. The lexicon method is a domain-free classification algorithm that does not require training data and extracts unigram (single word) out of a tweet by removing stop words and punctuation marks. The results determined that the machine learning and lexical model correlated better than the psycholinguistic method. While, psycholinguistics offers additional information in insight, such as emotional intensity which traditional approaches cannot provide (Griesser & Gupta, 2019).

An approach was proposed in a study to improve the lost/won classification of complex deals, such as high price tags, the risk factor is higher for the customer, and so on, of the CRM activity notes. The support vector machine (SVM) model called ε-insensitive classification (ε-SVM), extended by sentiment polarity features is used for opinion prediction, the R language, and its 'e1071' library, Python, and its NLTK (Natural Language Toolkit) library. The study utilized the data of complex sales over the period of 7 years, having 598 sales containing 28462 notes with a total number of 66258 sentences. The study revealed that combining different attributes of complex sales for closing a deal, acquired a high accuracy of (0.8933) in sentiment analysis in lost/won opportunities and high accuracy of (0.9337) in the F1 measure of sentiment analysis. The study also concluded that removing full stops and treating the text as a large blob decreases the accuracy of the model (Rotovei & Negru, 2017).

A study proposed a novel approach for the classification of text in Italian and English by using sentiment analysis for effective CRM by using NLP techniques. The proposed model classified the text and assigned it a confidence score along with a class. It also introduced word embedding (WEs) which captures semantic similarities among words. The study was conducted over 30,000 data items, collected from two different datasets, out of which 40% were from a CRM department of the Italian company and the remaining 60% were from public data. The results achieved an accuracy of 0.89 and 0.79 for Italian and English, respectively. The study predicted that such incremental models are efficient in sentiment accuracy as well as being utilized in multiple languages to target a bigger audience. The outcomes demonstrated that the hierarchical method was useful for the sentiment analysis and classification (Capuano et al., 2021).

By utilizing machine learning probabilistic classification and natural language processing, sentiment analysis was performed on an online customers' dataset to develop a business decision support system. The study covered sentiments in 2 aspects. Firstly, the overall semantic measurement of reviews of the dataset and the product-oriented semantics measures which cover the percentage of positive, negative, and neutral comments for each product. The results showed that the accuracy of NLP-based analysis for a single product was (55.40%), while the accuracy for multiple products by using multinomial Naïve Bayes was (64.06%). The results concluded that the model was effective in business decision- making (Al Asaad & Rotovei, 2020)

The NLP and AI-based methods were performed in the study to solve sentiment analysis. The research studied customers' sentiments shared on the official accounts of Coca-Cola and Pepsi on six different social media platforms. The R language libraries were used in the study. The results showed that Coca-Cola on average has 68.32% positive reviews among 6 platforms, the least preference percentage than Pepsi which has on average 75.67% positive reviews (Pavaloaia et al., 2019).

The research proposed a multi-agent expert system for sentiment analysis of the product aspect level in relation with supply chain management. The NLP-based bag-of-words algorithm was used for sentiment analysis. The study utilized JADE (Java Agent Development Framework), Knime, NLTK, Rapid Miner, Open NLP, Orange (Python Tool), and text processing with REST API as tools over the product aspect database containing logs of a company selling drones and customer emails. The results revealed that the model is useful for businesses in terms of price discount recommendations for interested customers based on their sentiments (Rotovei, 2016).

Another study introduced a framework, combining a semantic oriented approach and a supervised machine learning model for better CEM by altering CRM and association analysis. The details of 143 students from LMS of King Abdul Aziz University were collected by using blackboard CRM and Twitter account of the Dean of E-learning. Moreover, distance education at King Abdul-Aziz University was used as social media data to conduct the study. The results revealed that Naïve Bayes better classified the sentiments with an accuracy of 95.05, as compared to SVM. The study attempted to identify the gap between criteria, moreover the needs of students were fulfilled and the role of social media was helpful with the CRM for the said purpose (AL-Rubaiee et al., 2018).

A study focused to generate decision-relevant knowledge from the User Generated Content (UGC) by utilizing the data mining techniques including machine learning and dictionary-based models. The 10-fold cross-validation was used for ML model evaluation. Mining was divided into three further tasks, namely properties, subjectivity, and sentiment recognition in terms of gained results which include SVM (with POS

Tagging) which was able to achieve the highest accuracy of 72.35%, while the dictionary-based method achieved 71.28% accuracy in terms of recognition of properties. In the case of recognition of subjectivity, the dictionary-based approach achieved 82.63% accuracy, while the ML SVM methods achieved 65.50% accuracy. Moreover, for the sentiment recognition, the SVM method (with word bigrams) achieved 76.80% accuracy, in contrast to it, the dictionary-based approach achieved 71.28% accuracy. The study concluded that the results of data mining and sentiment analysis were enhanced by combining machine learning and dictionary-based methods (Schmunk et al., 2013).

A research presented a hierarchical method to perform sentiment analysis of the feedback data written in Turkish language for an airline and further classified the output data by using Xgboost classifier. The customer data was collected by a private airline comprised of 14000 reviews, where 1070 were labeled, having 532000 words when preprocessed, were used in the study. A separated Xgboost classifier was trained over the output document vectors. Moreover, for the validation purpose, 10-fold cross-validation was used. The results showed that the model successfully obtained 92.5% of accuracy in sentiment analysis and 71.16% of accuracy in the categorization of the negative sentiments according to the areas for which reviews were recorded (Seyfioglu & Demirezen, 2017).

A study used Feature Ontology (FO) and Opinion Sentence (OS) and analyzed customer opinions against products by embedding a customer relational model with sentiment analysis. ETL (Extract, Transform, and Load) is used to extract sentimental words and nouns. For the pre-processing, part of speech (POS) tagging is used. The results showed that the model with an average 0.888 F1 score and 0.924 precision successfully performed the sentiment analysis task and by embedding customer details, such as feedback and product records and personal details developed a better recommendation system as well as increased sales (Yaakub et al., 2013).

Artificial Neural Networks (ANN), Random Forest, and SVM approaches were used by a research which proposed a sales prediction model in a (B2B) business-to-business environment. The sales data of an aircraft seller company, containing 28462 customers, was studied. The research used Multi Adaptive Regression Splines (MARS), parts-of-speech (POS) term, and Word2Vec for pre-processing and sent140, Vivekn, TextBlob, sentiment word sense disambiguation along with stanford CoreNLP classifiers to conduct the sentiment study. The results showed that neural networks had a (0.8148) F1 score and an accuracy of (0.8020), while SVM had an F1 score of (0.8035) and an accuracy of (0.7822). Moreover, the ANN had an accuracy of (0.8148) and (0.8020), and the Random Forest performed best with an accuracy of (0.8515) and an F1 score of (0.8514)(Rotovei & Negru, 2019).

A research was conducted in which computational intelligence methods were studied combined with a social CRM to improve the relationship management focused on social media channels of universities. The study utilized machine learning, lexicon-based, and hybrid approaches. About 41287 Facebook comments were used for the study. Topic modeling was performed by applying the latent dirichlet allocation technique and iFeel framework to perform sentiment analysis on English data. The results presented that the study was useful for the opinion analysis and helped in developing future policies (Cirqueira et al., 2017).

A methodology was presented in a study to preprocess the social media data in Brazilian Portuguese for performing sentiment analysis. The study focused on the most used steps to preprocess the text for sentiment analysis along with major terminologies in order to pre-process implementation in the sentiment analysis. Stop words removal was discovered as the most used step with the highest usage frequency of 46, while the phonetics, context, and verb tense pre-processing functions, with the least usage frequency of 3, were found to be the least used methods in the studied literature (Cirqueira et al., 2019).

A comparative study was conducted on SVM with Term Frequency-Inverse Document Frequency (TFIDF) as a feature extractor and the SentWordNet was performed against the sentiment analysis. The dataset for the study consisted of 259 samples, derived from Google PlayStore and Apple AppStore. The study showed that the SVM model scored highest in both cases. For an imbalanced dataset, the accuracy was 75.75% and for balanced dataset it was 89.06%. The SentWordNet achieved 55.81% accuracy in case of imbalanced dataset and 51.59% accuracy for balanced dataset. The study was performed on an Indonesian oriented dataset and after the analysis it was determined that for non-English languages the SentiWordNet is not a good option for translating any language into English when the words of that language do not have synset (Fikri & Sarno, 2019).

Tusar and Islam (2021) performed a comparative study on machine learning techniques and natural language processing technique to measure customer satisfaction and loyalty. The study determined that SVM and logistic classifier together with the BOW performed with the highest score of 77%. The other models performed with the accuracies of 74%, Random Forest 74% for the Multinomial Naïve Bayes, and Random Forest accordingly (Tusar & Islam, 2021).

In a study, the churn behavior prediction in telecom sector was performed. Churn prediction is used for customer retention; therefore it is an important aspect in terms of market competition. A chi-squared test was performed to validate the defined factors. High tariffs were the 1st ranked factors relative to its importance in predicting the churn behavior (Arthur et al., 2012).

The relationship between customer experience and customer relationship were studied to examine customer loyalty in telecom sector. The decomposition of customer experience aspect was based on three factors. These factors include brand, charging, and the core service. The results showed that positive customer experience impacts loyalty. The study concluded that convenient sampling technique and expansion of variables inevitably limits generalization of the findings (Imbug et al., 2018)

Wang et al. (2022), in their research, studied the Twitter Sentiment Analysis (TSA) techniques and based on the significance, the publications in the said domain were grouped into different categories. The characteristics of Twitter as a social media platform and why it is a primary data source for sentiment analysis were studied as well. The SVM and NN are two wide implementations of the linear classifiers for class prediction. A survey of more than 60 publications was conducted and it was cleared that machine-learning-based method is the most used and popular method (Wang et al. 2022).

Amin et al. (2019) introduced the data transformation approaches for cross-company churn prediction in telecom sector, where a company uses the data of another company in case of unavailability of enough data of the target company for customer churn prediction. The Naïve Bayes achieved the highest accuracy score of 51%, 51%, and 51.3% based on the data transformation techniques of raw, log, and box-cox, respectively. The main contribution of the study was to use the data of one company to generate insights for another company for which there exists no historical data (Amin et al., 2019).

A research investigated sentiment analysis at document level as well as sentence level for the Arabic text. The study discussed the challenges in sentiment mining for the (non-English) Arabic. About 2238 sentences from 44 documents were given to the SVM classifier. It was concluded that the chunking approach with 4 chunks gave the highest accuracy of 87% and 8 chunks gave the lowest score (Farra et al., 2010).

An NN based approach was studied to perform customer churn prediction in cellular network services. The study determined that customer service calls are the most important attributes that affect the customer satisfaction with the highest importance score of 0.25. The other three most important attributes that are highly effective are international plan, day minutes, and the day charges with the scores of 0.13, 0.12, and 0.11, respectively. It was concluded that the customers who make customer service calls, international plans (calling), and make frequent calls in daytime are supposed to be the potential churners (Sharma & Panigrahi, 2011).

A competitive analysis was performed in a pizza industry by using social media text mining. The study aimed to develop social media competitive intelligence for three major pizza chains. A three interlinked step analysis model was proposed including text pre-processing, text processing/analysis, and actionable intelligence. A total number of 17,951 reviews were collected from Facebook and Twitter. The SPSS Clementine`s linguistic and NVivo 9 software distributions were used to find the data patterns. The study concluded that a better online experience could help build brands in the online communities because of the vast number of customers interacting on the same platform, thus can attract other customers (He et al., 2013).

A brand reputation analysis was performed in a study based on the customer satisfaction of mobile phone service providers by using the Twitter sentiment analysis. A total number of 10,000 tweets, against the three major service providers in Indonesia, were extracted. Customer satisfaction was measured against five services: 3G, 4G, voice, Internet, and short messaging. The SVM performed better in terms of processing time and accuracy and achieved the highest accuracy score of 82.40% after applying 10-fold cross validation. NB and DT performance were 78.90% and 72.90%, respectively (Vidya et al., 2015).

Jianqiang and Xiaolin (2017) executed a comparative research on the pre-processing techniques used for Twitter sentiment analysis. The study used five different datasets having a total number of 24,561 tweets for the study including Stanford Twitter Sentiment Test (STS-Test), Stanford Twitter Sentiment Gold (STS-Gold), SemEval2014, Sentiment Strength Twitter Dataset (SS-Twitter), and the Sentiment Evaluation Dataset (SE-Twitter). The results showed that the choice of classifier affects the performance of the model. The SVM performance was improved in the cases of different pre-processing methods.

Pakistan Telecommunication Authority (PTA) is the regulatory authority of telecommunication in Pakistan. According to the statistics provided by official website of PTA, the number of active mobile cellular subscribers in Pakistan was stated to be approximately 184 million as of August 12, 2022. This data highlights the substantial presence and usage of mobile technology and telecom sector in Pakistan, implying that the telecommunication sector holds significant potential for exploration and development.

Table 1 represents a detailed critical analysis of the above-mentioned studies in contrast to the problem addressed, proposed solution, future work, methodology used, data sources, size of the data, classification types, the algorithms being used, and the accuracy scores of the model.

Table 1

Critical Analysis of the Existing Literature

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Sajid et al., 2020)	How to improve the quality of service in the healthcare industry using patient's feedback?	Sentiment analysis of free-text data collected via PFP (web application) against patient experience	Further Classification Categories, emotion understanding from text, profiling user data for better understanding of health i.e. psyche profile	Analytical Research	Patients Feedback Platform (PFP)	8000	Binary-positive & negative	MLP	88%
(Ranjan et al., 2018)	How to predict the growth of telecom companies in India?	Real-time Sentiment analysis of Twitter data for customer feedback	multi-source data gathering	Analytical Research	Twitter	153,651	Binary-positive & negative	SVM, NB, CNN	93%
(Ullah et al., 2019)	How to measure reputation for telecom companies?	Sentiment analysis of Twitter data (Retweet and Favorites) using a polarity equation.	Developing an Arabic language dataset to perform multilingual SA.	Analytical Research	Twitter	5000	Multiclass Classification	SVM, Naïve Bayes, Decision Tree	95% , 87%, 68% Respectively
(Abdullah et al., 2021)	How to perform customer churn prediction?	Sentiment analysis using Call Detail Record (CDR) of telecom customers.	To investigate behavior, change of churn customers.	Qualitative Research	Call Detail Records of Companies	1000+	Multiclass Classification	Random Forest	88.63%

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Khatoon, 2017)	How to develop competitive intelligence for telecom companies using Twitter data?	Twitter data analysis to assess historical and spatial changes.	Expanding the domain specific data size for more accurate analysis.	Qualitative Research	Twitter	3000	Multiclass Classification	SentiWord Net (SWN)	-
(Alamsyah & Bernatapi, 2019)	How to improve telecom Customer Experience Management?	Identifying service dimensions using topic modeling on Twitter data.	Increasing dimensions to find which affects the customer experience more.	Qualitative Research	Twitter	7249	Multiclass Classification	Naïve Bayes	82%
(Hermansyah & Sarno, 2022)	What service aspects affect the customers? And how they can be used to improve telecom services?	Topic modeling using TF-ICF, and Sentiment Analysis on Twitter Data.	Expanding TF-ICF for adding more dimensions to represent every aspect for better analysis.	Qualitative Research	Twitter	-	Multiclass Classification	SVM, Random Forest	96.3% and 96.2%
(Al-Ansari, 2019)	How emotion choreography enhances CRM for the telecom sector?	Sentiment analysis on Twitter data containing telecom relevant Tweets.	Developing a SA tool that can support Arabic Language, extend study for other sectors as well.	Qualitative Research	Twitter	83,981	Multiclass Classification	NVivo data analytics Software	-
(Srivastava et al., 2019)	Advantages and disadvantages of CRM and its types.	A detailed description of the components.	-	Descriptive Study	-	-	-	-	-
(Song & Liang, 2021)	How to improve the CRM aspect of businesses?	Applying data mining on CRM and Twitter data	Including other social media platforms for study	Quantitative Research	Twitter, CRM	-	Multiclass, Binary	K-Mean, MLP, SVM	86%

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Díaz-Martín et al., 2015)	How do social networks generate customers and customer engagement on Social Media platforms?	Using data mining on Twitter social media	Finding other types of electronic word-of-mouth(eWOM) and their frequency of sharing among the customers	Analytical Research	Twitter	4000	Multiclass	KNN, Decision Tree, SVM, NB	91%
(Mohbey, 2020)	How to study User behavior in general elections in India using social media data?	Multiclass classification deep learning framework for Twitter data Sentiment analysis	improving the accuracy of traditional supervised models with large-sized data	Analytical Research	Twitter	50,000	Multiclass	Logistic Classifier, DT, SVM, Random Forest, DL	99%
(Seki, 2016)	How public sentiments are studied against local government policies?	Twitter data sentiment analysis approach for polarity study.	A real-time approach for efficient utilization of time constraints can be developed	Quantitative Research	Twitter, Survey	-	Binary-positive & negative	PMI-IR 2	87%
(Di et al., 2018)	What are the impacts of CRM on a company and how sentiment analysis can contribute to better CRM?	Identifying different factors of CRM relating to the businesses. Utilized AI with NetBase platform for sentiment analysis.	Different attributes relating to customer relationship enhancement to be identified.	Quantitative Research	Blogs, Twitter, Forums, and News	1927623	Binary-positive & negative	AI, Netbase	78.20%

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Griesser & Gupta, 2019)	How to generate emotional intensity using sentiment analysis? And how do different Sentiment Analysis models perform?	Multimodal Sentiment analysis method on Twitter data for finding emotional intelligence	Improving correlation of Psycholinguistics method	Analytical Research	Twitter	10,000	Binary-positive & negative	Lexicon-Based Model, ML, Psycholinguistics	-
(Rotovei & Negru, 2017)	How to perform lost/won classification of complex deals using CRM feedback data?	identified attributes of complex sales by Sentiment analysis of CRM Data.	Identifying what are the best pre-processing methods for better result generation on large-sized CRM notes?	Analytical Research	CRM	28462	Binary positive & negative	SVM	93%
(Capuano et al., 2021)	How does multi-lingual sentiment analysis contribute to better CRM?	Using multiple pre-trained models for Sentiment Analysis with a language detection step.	Improving model accuracy excluding CRM operator intervention for making it fully autonomous.	Quantitative Research	Multiple sources	30,000	Multiclass	NLP	89%
(Al Asaad & Rotovei, 2020)	How Sentiment analysis be useful for developing a Decision Support System?	Finding overall and product-oriented Semantic analysis of product reviews for generating insights to help in DSS.	Improving model accuracy	Analytical Research	Amazon Review Dataset	1128438	Binary-positive & negative	Multinomial NB, NLP aspect-based Classifier	(64% MNB), (55% NLP)

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Pavaloaia et al., 2019)	How does multi-source data acquisition help in finding user preferences?	Applying different statistical methods for finding correlation on datasets.	-	Quantitative Research	6 different Social Media Platforms	-	Binary-positive & negative	Chi-Square and *t*-test	75%
(Rotovei, 2017)	How sentiment analysis on product level help in supply chain management and price discount recommendation?	Sentiment analysis along with the multi-agent system over product aspect database	-	Analytical Research	Product Aspect Database	-	Recommendation System	Knime, NLTK, RapidMiner	-
(AL-Rubaiee et al., 2018)	How does altering CRM and association analysis enhance Customer Experience Management (CEM)?	Comparing CRM and Social Media data by applying sentiment analysis for finding the relation between them.	Will increasing the number of attributes impact model accuracy?	Quantitative Research	Twitter, CRM	-	Association Analysis	NB, SVM	(95% NB),(81 %SVM)
(Schmunk et al., 2013)	Can decision-making knowledge be generated from User-generated data?	Combined Machine learning and Dictionary-based method for sentiment analysis are applied for generating relevant information	Improving model accuracy	Qualitative Research	user-Generated content	208	Multiclass	SVM, SVM (with Bigrams)	76%

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Seyfioglu & Demirezen, 2017)	How polarity categorization against the area of review be performed using SA?	Sentiment analysis of CMR data, xgboost classifier for categorization of the polarity result	-	Analytical Research	CRM	14000	Binary-positive & negative	Xbgoost	92.50%
(Yaakub et al., 2013)	How sentiment analysis with a customer rational model embedded together will be useful in developing a recommendation system?	Feature Ontology and Opinion Sentences used for model development and customer opinion analysis	-	Analytical Research	-	-	Binary-positive & negative	ETL, POS, FO, OS	92.40%
(Rotovei & Negru, 2019)	How sentiment analysis be useful in complex sales prediction in a B2B environment?	Sentiment analysis along with different classifiers are used for complex sales prediction	Classification of complex deals with more parameters	Analytical Research	CRM	28,462	Binary-positive & negative	NN, SVM, ANN, Random Forest, TextBlob, MARS, Vivekn	80%,78%,80%,85% respectively
(Cirqueira et al., 2017)	How does sentiment analysis of Social CRM data help in developing policies?	Analyzing social media and CRM data with ML, Lexicon-based, and Hybrid approaches for SA.	Combining multiple data sources.	Quantitative Research	Facebook	41287	Topic Modeling	ML, Lexicon-based, Hybrid approaches	-

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Cirqueira et al., 2019)	What is the most used pre-processing step in Sentiment Analysis?	Frequency analysis of literature on sentiment analysis for finding the most used pre-processing step	-	Quantitative Research	Research Literature	4,049	Frequency Analysis	-	-
(Fikri & Sarno, 2019)	How do feature extractor/vectorization method performs with balanced and imbalanced datasets? Is SentiWordNet database efficient for languages other than English?	A comparative sentiment analysis of multiple languages using SVM and SentiWordNet was developed.	Using other translation mediums.	Qualitative Research	Google PlayStore and Apple AppStore Reviews	259	Classification Problem (On Balanced and Imbalanced Dataset)	SVM and SentlWord Net	(SVM: 75.75% & 89.06%) and (Sent WordNet : 55.81% & 51.59%)
(Tusar & Islam, 2021)	NLP and ML what is best for Sentiment Analysis? How using different vectorization techniques: BOW and TF-IDF performance changes?	A comparative analysis of NLP and ML algorithms is performed against multiple feature extraction methods on Twitter data.	Building a generalized model for similar dataset to improve model performance.	Qualitative Research	Twitter data	14,640	Classification Problem	SVM, Logistic Classifier, Random Forest, and Multinomial Naïve Bayes	77%, 77%, 74% and 74% respectively

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Arthur et al., 2012)	What are the factors behind Telecom customer churn?	Using sentiment analysis to determine customer switching behavior and satisfaction.	-	Qualitative Research	Questionnaire	261 Samples	Binary Classification	Chi-Squared	-
(Imbug et al., 2018)	How customer experience impacts the telecom customer loyalty?	Partial Least Square (PLS) approach for Structural Equation Modelling (SEM) was performed.	Other variables should be considered which could impact the customer experience.	Qualitative Research	Questionnaire	248 Samples	Classification Problem	Partial Least Square	-
(Wang et al., 2022)	Which social media platform can be considered as primary data source for SA, and why what classification models are mostly used?	A survey of more than 60 publications was conducted to answer the questions addressed.	Expending Twitter SA usage for multiple domains.	Qualitative Analysis	Research Publication	60 Publications	Frequency Analysis	-	-
(Amin et al., 2019)	How is data of another company used for churn prediction when there is not enough data available for target company?	Utilizing two different datasets for churn prediction using Machine learning models	Application of ensembles and using the same approach for other domains than telecom.	Qualitative Research	Two Publicly available datasets	3333 and 18,000 samples	Binary Classification	Naïve Bayes	51.3%

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Farra et al., 2010)	Can chunking approach improves classifiers and what are the challenges in SA for Non-English languages?	Two level classification model was introduced for SA, introduction of a dynamic document partition pre-processing step for improving model accuracy.	Using the output of the classifiers on a hierarchical classifier instead of providing known class labelled dataset.	Qualitative Research	Customized Dataset	2238 sentences & 44 documents	Binary Classification	SVM	87%
(Sharma & Panigrahi, 2011)	What is the customer satisfaction affecting attributes for the cellular network services?	Neural Network approach for customer churn prediction and finding the attributes that are highly affective.	Involving a better feature selection model. Using conventional ML models such as SVM for creating hybrid models. Applying the developed method for other sectors as well.	Qualitative Research	Publicly available (Irvine Database)	-	Multiclass Classification and Frequency Analysis	MLP	92%
(He et al., 2013)	Can social text mining help develop competitive intelligence?	A comparative study of different competitors in business industry was studied.	Finding innovative ways for turning positive feedback into 'Buy' by expanding the scope from consumers to their friends.	Quantitative Research	Social Media Data	17,951	Multiclass Classification and Frequency Analysis	NVivo and SPSS software distributions	-

Research paper reference	Problem	Solution	Future Work	Methodology	Data source	Dataset size	Class	Algorithm Used	Accuracy
(Vidya et al., 2015)	Can twitter sentiment analysis help in finding brand reputation? What are the 5 major services contributing into customer satisfaction?	Sentiment analysis on Twitter based data was performed using different machine learning techniques to measure customer satisfaction.	To use an unsupervised feature representation for the learning.	Qualitative Research	Twitter	10,000	Multiclass Classification	SVM, DT and NB	82.40%, 72.90% and 78.90% respectively
(Jianqiang & Xiaolin, 2017)	How text-preprocessing impacts the SA performance?	Classifiers were used against the SA to measure their performance against accuracy parameters.	-	Qualitative Research	Five Different Databases	24,561 Tweets	Multiclass Classification	Logistic Classifier, SVM, NB, & Random Forest	SVM Highest against parameters
(Pakistan Telecom Authority, n.d.)	-	Statistical data of telecom users in Pakistan.	-	-	-	-	-	-	-

Conclusion

The current study attempted to conduct a survey in order to access the current state-of-the art to measure customer satisfaction against a company and find out the criteria of satisfaction for the customers to improve CRM aspect. Moreover, it also compared the level of satisfaction with other companies to develop a market competitive strategy employing sentiment analysis. The findings would be beneficial for research community and would also help the researchers to identify techniques, methods, datasets, and tools being used for improving the CRM aspect. The researchers can leverage this information to enhance their understanding of customer satisfaction and develop effective strategies for managing customer relationships. Ultimately, the current study contributed to the field of CRM and assisted in advancing the overall understanding and implementation of effective CRM practices.

References

Abdullah, B., Alosaimi, N., & Almotiri, S. (2021). Reputation measurement based on a hybrid sentiment analysis approach for Saudi telecom companies. *International Journal of Advanced Computer Science and Applications*, *12*(6), 929–937. https://doi.org/10.14569/IJACSA.2021.01206107

Al Asaad, B., & Rotovei, D. (2020, September 1–4). *Business decisions support using sentiment analysis in CRM systems* (Paper presentation). Proceedings - 2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, Timisoara, Romania. https://doi.org/10.1109/SYNASC51798.2020.00057

Alamsyah, A., & Bernatapi, E. A. (2019). *Evolving customer experience management in internet service provider company using text analytics* (Paper presentation). 2019 International Conference on ICT for Smart Society (ICISS), Bandung, Indonesia. https://doi.org/10.1109/ICISS48059.2019.8969828

Al-Ansari, J. A. (2019). *From sentiment analysis to choreography of emotions: Social media analysis for improved customer relationship management (CRM) in the Omani telecom sector* [Doctoral dissertation, University of Strathclyde]. https://stax.strath.ac.uk/concern/theses/7s75dc45x

AL-Rubaiee, H., Alomar, K., Qiu, R., & Li, D. (2018). Tuning of customer relationship management (CRM) via customer experience management (CEM) using sentiment analysis on aspects level. *International Journal of Advanced Computer Science and Applications*, *9*(5), 300–312. https://doi.org/10.14569/IJACSA.2018.090540

Amin, A., Shah, B., Khattak, A. M., Lopes Moreira, F. J., Ali, G., Rocha, A., & Anwar, S. (2019). Cross-company customer churn prediction in telecommunication: A comparison of data transformation methods. *International Journal of Information Management*, *46*, 304–319. https://doi.org/10.1016/j.ijinfomgt.2018.08.015

Arthur, Y. D., Ahenkrah, K., & Asamoah, D. (2012). Determinants analysis of customer switching behavior in Ghanaian telecommunication industry, an exploratory and inferential study. *International Journal of Business and Social Research*, 105–112.

Capuano, N., Greco, L., Ritrovato, P., & Vento, M. (2021). Sentiment analysis for customer relationship management: An incremental learning approach. *Applied Intelligence*, *51*(6), 3339–3352. https://doi.org/10.1007/s10489-020-01984-x

Cirqueira, D., Fontes Pinheiro, M., Jacob, A., Lobato, F., & Santana, A. (2019). *A literature review in preprocessing for sentiment analysis for Brazilian Portuguese social media* (Paper presentation). Proceedings - 2018 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2018, Santiago, Chile. https://doi.org/10.1109/WI.2018.00008

Cirqueira, D., Pinheiro, M., Braga, T., Jacob, A., Reinhold, O., Alt, R., & Santana, Á. (2017, August 23–26). *Improving relationship management in universities with sentiment analysis and topic modeling of social media channels: Learnings from UFPA* (Paper presentation). Proceedings - 2017 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2017, Leipzig, Germany. https://doi.org/10.1145/3106426.3117761

Di, D., Economiche, S., Aziendali, E. D., & Administration, B. (2018). *"Marco Fanno" Corso Di Laurea Magistrale in "CRM and sentiment analysis, the supporting role of leading-edge technologies"* [Master's Thesis, Universita' Degli Studi Di Padova]. https://thesis.unipd.it/bitstream/20.500.12608/25456/1/Michela_Chiarato.pdf

Díaz-Martín, A., Suplet, M., & Menéndez, H. (2015). Using Twitter to engage with customers: A data mining approach. *Internet Research: Electronic Networking Applications and Policy*, *25*. https://doi.org/10.1108/IntR-11-2013-0249

Farra, N., Challita, E., Assi, R. A., & Hajj, H. (2010, December 13). *Sentence-Level and document-level sentiment mining for arabic texts* (Paper presentation). 2010 IEEE International Conference on Data Mining Workshops Sydney, NSW, Australia. https://doi.org/10.1109/ICDMW.2010.95

Fikri, M., & Sarno, R. (2019). A comparative study of sentiment analysis using SVM and SentiWordNet. *Indonesian Journal of Electrical Engineering and Computer Science*, *13*(3), 902–909. https://doi.org/10.11591/ijeecs.v13.i3.pp902-909

Griesser, S. E., & Gupta, N. (2019, June 14). *Triangulated sentiment analysis of tweets for social CRM* (Paper presentation). Proceedings - 6th Swiss Conference on Data Science, SDS 2019, Bern, Switzerland. https://doi.org/10.1109/SDS.2019.000-4

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, *33*(3), 464–472. https://doi.org/10.1016/j.ijinfomgt.2013.01.001

Hermansyah, R., & Sarno, R. (2022). *Sentiment analysis based on quality aspects in effort to improve quality of Indhome product and services PT Telkom Indonesia Tbk* (Paper presentation). 3rd International Conference on Business and Management of Technology (ICONBMT 2021), Surabaya, Indonesia. https://doi.org/10.2991/aebmr.k.211226.005

Imbug, N., Ambad, S. N. A., & Bujang, I. (2018). The influence of customer experience on customer loyalty in telecommunication industry. *International Journal of Academic Research in Business and Social Sciences*, *8*(3), 103–116. https://doi.org/10.6007/IJARBSS/v8-i3/3909

Jianqiang, Z., & Xiaolin, G. (2017). Comparison research on text pre-processing methods on twitter sentiment analysis. *IEEE Access*, *5*, 2870–2879. https://doi.org/10.1109/ACCESS.2017.2672677

Khatoon, S. (2017). Real-Time twitter data analysis of Saudi telecom companies for enhanced customer relationship management. *International Journal of Computer Science and Network Security*, *17*, 141–147

Mohbey, K. K. (2020). Multi-class approach for user behavior prediction using deep learning framework on twitter election dataset. *Journal of Data, Information and Management*, *2*(1), 1–14. https://doi.org/10.1007/s42488-019-00013-y

Pakistan Telecom Authority. (n.d.). *Telecom indicators*. Retrieved August 12, 2022, from https://www.pta.gov.pk/en/telecom-indicators/1

Pavaloaia, V. D., Teodor, E. M., Fotache, D., & Danileţ, M. (2019). Opinion mining on social media data: Sentiment analysis of user preferences. *Sustainability (Switzerland)*, *11*(16), Article e4459. https://doi.org/10.3390/su11164459

Ranjan, S., Sood, S., & Verma, V. (2018, August 30–31). *Twitter sentiment analysis of real-time customer experience feedback for predicting growth of Indian telecom companies* (Paper presentation). 2018 4th International Conference on Computing Sciences (ICCS), Jalandhar, India. https://doi.org/10.1109/ICCS.2018.00035

Rotovei, D. (2016, September 24–27). *Multi-agent aspect level sentiment analysis in CRM systems* (Paper presentation). Proceedings - 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, Timisoara, Romania. https://doi.org/10.1109/SYNASC.2016.068

Rotovei, D., & Negru, V. (2017, September 21–24). *Improving lost/won classification in CRM systems using sentiment analysis* (Paper presentation). Proceedings - 2017 19th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, Timisoara, Romania. https://doi.org/10.1109/SYNASC.2017.00038

Rotovei, D., & Negru, V. (2019). *Data driven sales prediction using communication sentiment analysis in B2B CRM systems* (Paper presentation). Proceedings - 21st International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, Timisoara, Romania. https://doi.org/10.1109/SYNASC49474.2019.00032

Sajid, A., Awais, M., Mehmood, M. A., Batool, S., Shahzad, A., & Zafar, A. (2020). Patient's feedback platform for quality of services via "Free Text Analysis" in healthcare industry. *EMITTER International Journal of Engineering Technology*, *8*(2), 316–325. https://doi.org/10.24003/emitter.v8i2.502

Schmunk, S., Höpken, W., Fuchs, M., & Lexhagen, M. (2013). Sentiment analysis: Extracting decision-relevant knowledge from UGC. In. Xiang & I. Tussyadiah (Eds.), *Information and Communication technologies in Tourism 2014* (pp. 253–265. Springer International Publishing. https://doi.org/10.1007/978-3-319-03973-2_19

Seki, Y. (2016, May 18–20). Use *of twitter for analysis of public sentiment for improvement of local government service* (Paper presentation). 2016 IEEE International Conference on Smart Computing (SMARTCOMP), St. Louis, MO, USA. https://doi.org/10.1109/SMARTCOMP.2016.7501726

Seyfioglu, M. S., & Demirezen, M. U. (2017, September 3–6). *A hierarchical approach for sentiment analysis and categorization of Turkish written customer relationship management data* (Paper presentation). Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, Prague, Czech Republic. https://doi.org/10.15439/2017F204

Sharma, A., & Panigrahi, P. K. (2011). A neural network based approach for predicting customer churn in cellular network services. *International Journal of Computer Applications*, *27*(11), 26–31. https://doi.org/10.5120/3344-4605

Song, D., & Liang, C. (2021, March 26–28). *Application research of data mining technology in customer relationship management* (Paper presentation). 2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), Changsha, China. https://doi.org/10.1109/AEMCSE51986.2021.00244

Srivastava, S. K., Chandra, B., & Srivastava, P. (2019). The impact of knowledge management and data mining on CRM in the service industry. In V. Nath & J. K. Mandal (Eds.), *Nanoelectronics, circuits and communication systems* (Vol. 511, pp. 37–52). Springer. https://doi.org/10.1007/978-981-13-0776-8_4

Tusar, T. H. K., & Islam, T. (2021). A comparative study of sentiment analysis using NLP and different machine learning techniques on US airline twitter data, *arXiv*, Article e2110.00859. https://doi.org/10.48550/arXiv.2110.00859

Ullah, I., Raza, B., Malik, A. K., Imran, M., Islam, S. U., & Kim, S. W. (2019). A churn prediction model using random forest: Analysis of machine learning techniques for churn prediction and factor identification in telecom sector. *IEEE Access*, *7*, 60134–60149. https://doi.org/10.1109/ACCESS.2019.2914999

Vidya, N. A., Fanany, M. I., & Budi, I. (2015). Twitter sentiment to analyze net brand reputation of mobile phone providers. *Procedia Computer Science*, *72*, 519–526. https://doi.org/10.1016/j.procs.2015.12.159

Wang, Y., Guo, J., Yuan, C., & Li, B. (2022). Sentiment analysis of twitter data. *Applied Sciences*, *12*(22), Article e11775. https://doi.org/10.3390/app122211775

Yaakub, M. R., Li, Y., & Zhang, J. (2013). Integration of sentiment analysis into customer relational model: the importance of feature ontology and synonym. *Procedia Technology*, *11*(Iceei), 495–501. https://doi.org/10.1016/j.protcy.2013.12.220