JKAU: Comp. IT. Sci., **Vol. 12** No. 1, pp: 37 – 50 (1444 A.H. / 2023 A.D.)
DOI: 10.4197/Comp.12-1.4

Using Sentiment Analysis of Arabic Tweets to Fine-Tune CRM Structure

Hamed AL-Rubaiee[1] and Khalid Hamad Alomar[2]
[1]*University of Bedfordshire*, Bedfordshire, United Kingdom
[2]Information *Systems Department, Faculty of Computing and Information Technology*, *King Abdulaziz University*, Jeddah, Saudi Arabia
[1] Rubaiee@study.beds.ac.uk, [2] khalomar@kau.edu.sa

***Abstract*— Understanding how customers perceive the services they receive has always been crucial to a business's success. It is widely accepted that Customer Relationship Management (CRM) and Customer Experience Management (CEM) have both been shown to aid businesses in making better decisions by providing them with better information. Unfortunately, in real-world business applications, there are distinctions between customer opinions collected through customer relationship management (CRM) and the real customer opinions gathered via social media for customer experience management (CEM). It is critical to close this gap between the two to match customer expectations. Both CRM and CEM have been recognised as tools that can help firms make better decisions. However, until recently, the problem of integrating unstructured data from CEM directly into CRM was largely ignored, and, particularly in the Arabic language, it has remained unsolved. Thus, the goal of this study is to offer a framework for CRM modification via CEM that integrates a semantic orientation method and supervised machine learning. Sentiment analysis was utilised to enhance the CRM structure to add important aspects that were missing. The results of this study show that businesses could undoubtedly make more informed decisions by expanding their CRM structure to incorporate more of the issues discussed by customers on social media. Surprisingly, the negative class had the most label matching between CRM and CEM.**

***Keywords*— customer relationship management, customer experience management, sentiment analysis, social media, Arabic language**

I. Introduction

In the new global economy, understanding customers' perceptions of service provided has become a central issue for businesses. CRM and CEM can both benefit from each other to help companies make better decisions [1]. CRM is primarily motivated by how a business views its customers, while CEM is primarily concerned with what customers think of the business [2]. Consequently, organisations currently manage relationships through a variety of processes and modern technology, with the primary goal of trying to retain their customers and continuing to develop positive relationships with them. For example, CEM allows businesses to better understand customer experience through reviews and feedback, whereas CRM uses data mining to extract hidden information, allowing businesses to better understand customer behaviour and take proactive measures to provide excellent service [3]. This is significant, as the rapid growth of social media and microblogging has heightened concerns with customer opinion analysis. Clearly, both CRM and CEM nowadays require the support of social media to maximise their effectiveness. Analysing social opinions through sentiment analysis allows organisations and companies to evaluate and monitor public opinions toward the services they provide [4-6].

Currently, in real-world applications, disagreements still abound between the real opinions gathered in CEM via social media and the opinions predicted by CRM. A company may confront difficult decisions because of mismatches between CRM-predicted opinions and actual opinions gathered via social media platforms in CEM. Until now, integrating the unstructured data (textual data) from CEM into CRM has required further investigation [7]. Notably, in the case of Arabic literature, Arabic dialects lack a lexicon [8]. Additionally, the language utilised in social media communication is a synthesis of Modern Standard Arabic [9], Arabic dialects (AD), and colloquial Arabic. [10]. Furthermore, there is a huge variety of Arabic dialects among the Arab countries. Since accurate labelling is fundamental for unstructured feedback, these factors create difficulties when attempting to analyse Arabic text. To address these issues, a framework for updating CRM via CEM via Twitter is presented and demonstrated using a real-world business case. Thus, the goal of this study is to offer a framework to utilise sentiment analysis of Arabic tweets to fine-tune CRM structure.

The classification of sentiment naturally includes cases such as “known knowns,” “unknown knowns,” “known unknowns,” and “unknown unknowns.” In the case of known knowns, for example, a single tweet contains only positive or negative keywords, so the annotator doesn’t get confused and the labelling is reliable. In the case of “known unknowns”, for example, a single tweet could have both positive and negative keywords, which would confuse the annotator and make labelling less accurate. Due to unknown unknowns and unknown knowns, human-based labelling is subject to error. Therefore, a data annotator must grasp a subjective text better than an objective one to accurately identify it [11]. This increased understanding would extend the known unknowns and subsequently improve the quality of labelling. Unreliable labelling comes from tweets that contain both positive and negative terms, which may confuse someone conducting analysis. Thus, including a neutral class would help a person to express uncertainty more freely. In addition, human-based labelling produces many errors due to unknown unknowns and unknown knowns. Thus, operators unintentionally establish incorrect classifications. This kind of mistake could show up as sounds in subsequent stages of sentiment analysis. However, labelled messages generally contain noise, especially extended comments with several sentences, like blogs [12]. Although the suggested relabelling strategy will not remove classification noise, it will assist operators in identifying potential labelling conflicts and reveal instances of unknown unknowns and unknown knowns.

The paper is organised as follows. Section 2 establishes the context for the research. The framework for CRM and CEM integration is discussed in Section 3. Section 4 uses Mubasher as an example to show how CRM should be adjusted based on CEM at the phrase level. Then, Section 5 creates a sentiment analysis for CEM based on CRM criteria. The sentiment analysis experiment for CEM is described in Section 6. Section 7 examines classification using CRM. Section 8 explains how to update CRM using CEM. Finally, Section 9 brings this study to a close.

II. Background

Nowadays, a wealth of customer feedback may be captured on social media sites like Twitter [13-15]. Sentiment analysis is utilised to examine public opinion, and it can be carried out on a variety of different levels [16]. A large and growing body of literature has investigated sentiment analysis on various levels, such as the document, sentence, and phrase level. The goal of document-level sentiment analysis is to categorise the text and determine its sentiment [17, 18]. In contrast, sentence-level sentiment analysis is more in-depth [11]. The sentence expresses a single point of view, whether positive, negative, or neutral [17]. A different approach is to use targeted positive or negative thoughts to analyse sentiment at the phrase level. Additionally, several scientists have expanded sentiment analysis by adding two additional levels [19]: Comparative sentiment analysis seeks to identify comparison decision phrases and extract the chosen entity from each viewpoint. The dictionary-based strategy is more effective for lexicon acquisition since it can start with a small collection of acceptable feeling words and then expand it using synonyms and antonyms.

The major challenge in analysing Arabic sentiments derives from the lack of an Arabic lexicon for sentiment terms. Arabic text also has a different structure than other languages. Arabic is written right to left, with varying letter forms. Furthermore, gender and number in Arabic words are precise, and their suffixes can vary accordingly. Arabic text does not contain capital letters, meaning additional grammatical rules are necessary to recognise nouns, abbreviations, and acronyms. [20, 21]. These factors make it difficult to interpret Arabic text. Due to the lack of research on Arabic sentiment analysis, this discipline is in its infancy. Similarly, precise labelling of unstructured feedback is important to understanding how CEM works. Doing this takes considerable time and involves a lot of unclear input from humans. Humans may not be able to figure out which words are positive and which are negative, which makes it more difficult for them to label things. By establishing strong interactions with key customers, CRM attempts to increase shareholder value [22]. CRM increases opportunities by utilising data to generate a more complete picture of customers and marketing initiatives. Additionally, organisations want technology and applications to enhance client relationships and comprehend the information created by consumer data. CRM’s goal is to build strong customer relationships to improve customer retention and profitability while also potentially improving the company's position in the marketplace [20]. Due to today's market competition, it is critical for companies to account for customer support and retention [23]. CRM has grown into a significant industry with significant technological advancement [23]. Customer data and information technology are now the core of a successful CRM strategy. The internet has considerably enhanced marketing opportunities and revolutionised how firms communicate with their customers [24]. Data mining is becoming an essential area of CRM, and customers' behaviour has been studied extensively. To understand customer behaviour, Rygielski et al. [25] demonstrated the critical nature of the customer lifecycle framework. To save time, according to Achabal et al. [26], efficient data-mining methods must be used. Traditional CRM has been replaced by social CRM in the last year, which may help manage client connections online [27-29]. In another study conducted by Sari et al. [30], the researchers employed sentiment analysis to assess the level of service quality based on online customer reviews. The researchers were able to reach a classification accuracy of 90%. Further, customer tweets were used to analyse the experience of internet service providers [31]. In addition, some studies investigated how customers feel about chatbots in a variety of retail settings, as well as the impact chatbots have on customers' perceptions and expectations of other types of service interactions with online human agents [32].

III. A Framework for CRM Adjustment via CEM

This study presents a paradigm for CRM adjustment at the sentence level via CEM. The framework combines an approach based on semantic orientation with supervised machine learning. The framework includes sentiment analysis for CEM, CRM classification, and CEM data tuning. The three processes have similar data modelling components. The first process uses unstructured social media data, and the second uses a structured customer database. The proposed process begins with data collection using social media platforms. It then combines CRM and CEM by adjusting and optimising CRM to match the expectations of social media users. Specifically, the goal is to close the gap between users' social media opinions and their CRM-predicted beliefs. Therefore, association rules were created to reveal the connections between seemingly unrelated data. Then, an unsupervised data-mining method will be used to connect data from social media and CRM (association).

A. *Analysis of Sentiment in the Context of CEM*

Sentiment analysis for the CEM process utilises social media to identify crucial features represented by a semantic schema for CEM. The procedure begins with the collection of tweets via Twitter's streaming API. The tweets gathered are based on terms associated with this work's case study. Then, to decrease noise in the gathered tweets, text-processing techniques are applied in accordance with the chosen ontology model. Finally, according to their sentiment polarity, tweets are categorised as positive, negative, or neutral.

The CRM classification method, on the other hand, extracts crucial properties represented by a semantic schema for CRM from an existing customer database. The procedure begins with the collection of structure data via database API. Techniques for text processing based on an existing CRM model are then used to minimise the noise in the collected tweets. Finally, it uses supervised learning to extract essential features, just like the CEM method.

B. *Adjustment of CRM via CEM*

Fine adjustment means optimising and updating the CRM structure to better accommodate social media customer feedback. Adjustment starts with asking users on social media about general aspects or specific information on CRM. As a result, CEM is constructed at the sentence level. Resultingly, CEM will have incomplete CRM detail coverage and matching. Some CRM elements can be restructured to better match social media users' expectations and bridge the CRM-CEM gap. Fig. 1 shows the proposed framework for CRM-CEM integration. The framework process starts by analysing the unstructured social media input for CEM, then classifies the structured customer database from CRM. Then, CEM structured data are automatically labelled by CEM or vice versa. Once these two processes have been conducted, fine adjustment of CRM through CEM takes place.

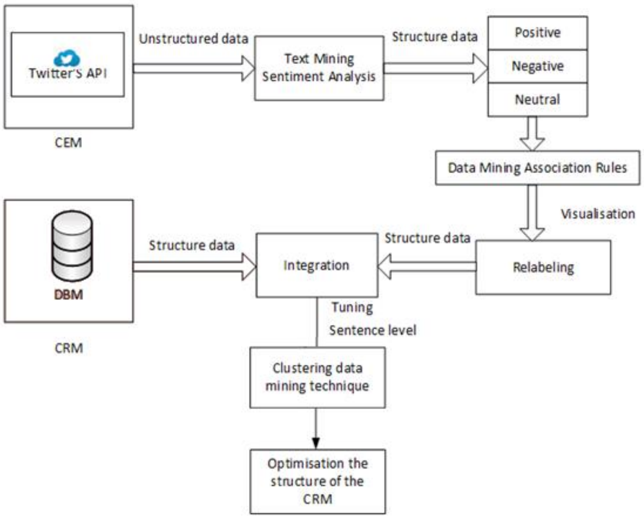

Fig. 1. The proposed framework.

IV. Case Study: Using CEM at the Sentence Level to Optimise Mubasher's CRM

The capital markets in Saudi Arabia are overseen by the Capital Market Authority (CMA), the Saudi government's financial regulatory body. The CMA was founded to safeguard Saudi Arabian investors and the public against unfair and fraudulent commercial activities. Currently, the CMA supervises 192 publicly traded firms on the Saudi Stock Exchange, or Tadawul. Mubasher is a large financial services organisation based in the Middle East that provides a variety of services. Mubasher can receive real-time prices for a wide range of products with the CMA's approval. Mubasher has released several other products in addition to Mubasher Pro 10. Mubasher uses a classic legacy CRM that provides extremely little information about users, transactions, and activities. Customers were asked broad questions regarding various aspects of the services they received to modernise this CRM. These questions were suggested by Mubasher (such as questions regarding the customer's subscription and the Mubasher Pro 10's usability). The goal of this project was to incorporate critical features missing from the existing CRM based on user feedback. Because of the incompatibilities between historical CRM and CEM, CRM can only be adjusted at the sentence level.

V. Using CRM Criteria to Conduct a Sentiment Analysis for CEM

Twitter is a popular form of mobile-friendly, real-time mass communication [33]. The initial step in the opinion-mining model is to collect data. Then, the text of tweets was cleaned up (by removing retweets, mentions, links or URLs, and hashtags) before being delivered to the classifier. Hashtags are used before, within, or after the body of a tweet on Twitter [34]. Subject hashtags are used as search terms on Twitter to locate additional tweets on a certain subject [35]. For instance, certain hashtags were made to celebrate important events such as Mubasher's products, "#مباشر_برو_١٥عام" (#MubasherPro15Years). A hashtag was created in this experiment to solicit user feedback on and reactions to Mubasher's products. In addition, sentiment analysis of customers' opinions was conducted on Twitter at the sentence level.

VI. Conducting a Sentiment Analysis Experiment for CEM

Sentiment analysis module of Arabic tweets is summarised in Fig. 2. Mubasher's products were first assessed by customers using a hashtag. The supervised machine-learning method was then used to label data represented by feature vectors as positive or negative. The classifier then utilised the vectors as training data to figure out which combination of specific attributes corresponded to which class.

The "Retrieve" operator was used to load the dataset. Then, using "Process documents from data," a word vector list was created from the text. Next, using the "Set role" operator, the text was sorted into all classes. Finally, a nested operator doing X-validation verified the correctness. RapidMiner, which contains a list of operators for text processing, was used to process the text. These operations analysed Arabic text, tokenising and deleting stop words, in order to attain our vector list [36].

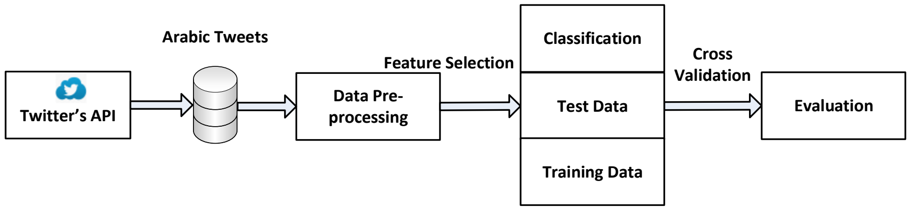

Fig. 2. The analysis of Arabic tweets.

A. Description and Pre-Processing of Data

Twitter's official developers' API was used to collect opinions from users on Mubasher's page. Negative, positive, and neutral tweets were manually categorised from the tweets that had been downloaded. In addition to tweet-text and tweet-label, the company provided two properties (data-type=text). The RapidMiner wizard imported the data, the features of which were typed. That data type remained unchanged. Nominal scales were employed to label variables that did not have a numerical value [37]. Then, a label column for RapidMiner was created using the operator "Set role". There were 2,058 tweets in the dataset. A list of roughly 20 frequently occurring Saudi dialect phrases representing positive, negative, and neutral groups was then produced (see Table I).

TABLE I. FREQUENTLY USED ARABIC STANDARD SAUDI DIALECT TERMS FOR THE GRADES.

Positive Sentiment	English Translation	Negative Sentiment	English Translation
عطاء	liberality	انقطاع	abruption
تطور	growth	مشكلة	trouble
جميل	magnificent	بطيئ	slow
محترف	professional	غير دقيق	not accurate
ممتاز	excellent	السعر غالي	the price is high
سريع جدا	very fast	غير جدير	ineligible
منصة احترافية	professional platform	غير ناجح	unsuccessful

The data were manually labelled by two Mubasher workers who were familiar with Mubasher's goods. Positive tweets were assigned a value of "1," negative tweets were assigned a value of "-1," neutral tweets were assigned a value of "0," and irrelevant tweets were eliminated. A total of 1,331 tweets were gathered, as 727 irrelevant tweets were eliminated during the labelling process. In addition, repeat cross-validation was used 10 times to divide datasets into nine parts for training data and one for testing data. Normalised versions of the tweets were created by removing hashtags, URLs, retweets, and other special characters during the labelling process. After that, RapidMiner was used to perform preliminary processing on the data. Some words, letters, and signs were later replaced. For instance, some letters that are a challenge to spell correctly such as Hamza "ء" (أ ، إ ، آ) to "ا", and (لآ ، لأ ، لإ) to "لا" were replaced. In addition, special characters were replaced, such as % with percentage. Further, some words in which people elongate one of the letters to emphasise the meaning were replaced, such as very "جداااا" with "جدا".

Following that, the five pre-processing processes occurred. Using the RapidMiner software's "Process documents from data" operator:

- Each tweet was tokenised, converting it into a collection of tokens;
- Stop words in Arabic, such as "إلى " and "من," which translate as "to" and "from," were eliminated.
- words were restored to their basic forms by deleting suffixes and prefixes.
- Filter token by length was used to remove extra words and make two word.
- N-grams were used to show a group of N items from a text.

The dataset was then transformed into a word vector using the operator "Process documents from data." Data mining with NB and SVM was used to create tweet identification models after word vectors were generated using the weighting techniques TF-IDF (term frequency–inverse document frequency) and BTO (binary term occurrence). The evaluation was completed by utilising precision and recall methods. This study used C-SVM for classification, and the linear kernel type was selected. C was set to its default value of zero,

TABLE II: ACCURACY OF SVM AND NB CLASSIFIERS WITH AND WITHOUT N-GRAMS FOR POSITIVE AND NEGATIVE CLASSES.

Classifier	Weighting Schemes	Accuracy	Recall	Precision	Classification Error
NB	BTO	٨٣	٨٢.2	٨٠.8	16.9 %
	TF-IDF	82.5	٨٠.8	٨٠.1	17.5 %
SVM	BTO	٨٣.8	٧٧.3	٨٥.8	16.2%
	TF-IDF	٨٤.9	٧٩.10	٨٦.3	15.1%
NB with N-gram	BTO	٨٤.1	٨١.9	٨٢.3	15.8 %
	TF-IDF	٨٣.9	٨١.5	٨٢.0	16.1 %
SVM with N-gram	BTO	٨٣.9	٧٦.4	٨٨.2	16.1%
	TF-IDF	84.5	77.2	89.0	15.5%

TABLE III: ACCURACY OF SVM AND NB CLASSIFIERS WITH AND WITHOUT N-GRAMS FOR ALL CLASSES.

Classifier	Weighting Schemes	Accuracy	Recall	Precision	Classification Error
NB	BTO	٦7	٦٥.1	٦٣.8	30.0%
	TF-IDF	٧٠.1	٦٣.6	٦3	29.9%
SVM	BTO	٧٨.1	٦٢.5	٨٠.2	21.9%
	TF-IDF	٧٨.5	٦٣.5	٧٨.0	21.5%
NB with N-gram	BTO	٧٣.4	٦٧.2	٦٦.4	26.6%
	TF-IDF	73.6	٦٦.2	٦٥.9	26.4%
SVM with N-gram	BTO	٧6	٥٨.7	٧٦.0	24.0%
	TF-IDF	٧٥.7	٥٨.0	٧٥.5	24.3%

TABLE IV: FOR ALL CLASSES WITH AND WITHOUT AN N-GRAM IN THE FIRST SUBSET.

Classifier	Weighting Schemes	Accuracy	Recall	Precision	Classification Error
NB	BTO	٦٧.3	٦٣.1	٦١.2	32.7%
	TF-IDF	٦٧.5	٦٢.5	٦1	32.5%
SVM	BTO	٧٤.1	٥٩.5	٧٢.9	25.9 %
	TF-IDF	٧٦.7	٦٤.4	٧٧.3	23.3 %
NB with N-gram	BTO	٧١.9	٦٤.9	٦٤.8	28.1%
	TF-IDF	٧٢.1	٦٤.5	٦٤.9	28%
SVM with N-gram	BTO	٧٤.6	٥٨.9	٨٠.1	25.4 %
	TF-IDF	٧٥.7	٥٩.9	٨١.6	24.2 %

Epsilon was set to its default value of 0.001, and the cache size was set to 80 megabytes.

B. Experimental Findings

To elicit thoughts about Mubasher's goods, the described methodology for assessing the sentiment of tweets in Saudi Arabic was applied. To identify tweets based on their sentiment polarity as positive or negative, machine-learning algorithms for CEM creation and natural language processing was utilised. To begin, we examined the dataset's document pre-processing. Second, we employed both NB and SVM with weighting techniques that varied (TF-IDF and BTO). Third, results for token N- grams were generated using the proposed sentiment analysis model [38]. Finally, the data were labelled by humans, and the classification accuracy of the results was improved using a neutral class. Table II describes the changes required to obtain the highest level of accuracy in sentiment analysis.

Following that, the data were labelled by experts in the field, and a neutral class was established to allow classification between the positive and negative groupings. The SVM method concentrated on the newly established neutral class. Table III shows the NB and SVM classifiers' accuracy, precision, and recall. Finally, the experiment shows that using the N-gram feature with both schemas (TF-IDF and BTO) increased NB performance.

C. Reproducing the Experiment

This experiment's main goal was to show the suggested method's success in Arabic-language sentiment analysis, especially with the neutral class. This is vital to the first proposed component's success. The data were also split into two sections to test SVM and Bayesian techniques.

TABLE V. FOR ALL CLASSES WITH AND WITHOUT AN N-GRAM IN THE SECOND SUBSET.

Classifier	Weighting Schemes	Accuracy	Recall	Precision	Classification Error
NB	BTO	٥٩.9	٥٦.6	٥٧	40.1%
	TF-IDF	60.9	٥٧.1	٥٧.3	39.1%
SVM	BTO	٦٨.4	٥٩.5	٦٧.2	31.6%
	TF-IDF	٦٩.8	٦١.0	٦٨.1	30.2%
NB with N-gram	BTO	٦٤.9	٦٠.4	٦٠.6	35.1%
	TF-IDF	64.7	٥٩.8	٥٩.9	35.3%
SVM with N-gram	BTO	٦٦.6	٥٦.3	٦٠.5	33.4%
	TF-IDF	٦٩.9	٥٩.5	٧2	30.1%

TABLE VI. ACCURACY OF ALL CLASSES IN THE ORIGINAL AND NEW CLASSIFICATION FOR SVM.

	true neutral		true positive		true negative		class precision	
	original	new	original	new	original	new	original	new
pred. neutral	48	48	4	2	6	8	83%	81%
pred. positive	129	129	744	749	154	156	73%	72%
pred. negative	21	21	7	10	211	206	88%	87%
class recall	24.2%	24.3%	98.5%	98.4%	56.9%	55.7%		

TABLE VII. ACCURAY OF ALL CLASSES IN THE ORIGINAL AND NEW CLASSIFICATION FOR NB.

	true neutral		true positive		true negative		class precision	
	original	new	original	new	original	new	original	new
pred. neutral	91	70	70	36	47	23	43.8%	54.3%
pred. positive	62	72	621	657	82	91	81.1%	80.1%
pred. negative	45	51	64	68	242	256	69%	68.3%
class recall	46%	36.3%	82.3%	86.3%	65.2%	69.2%		

TABLE VIII. COMPARISON OF SVM AND NB IN THE ORIGINAL AND NEW LABELLING.

Labelling Dataset	Classifier	Accuracy	Recall	Precision	Classification Error
Original	SVM	٧٦.74	٦٤.40	٧٧.26	23.26 %
	NB	٧٢.05	٦٤.48	٦٤.85	27.95%
New	SVM	75.31	58.63	81.85	24.69%
	NB	74.25	63.92	67.40	25.75%

1) *First Subset:* The first subset contained 665 tweets (189 positive, 377 negative, and 99 neutral tweets). Table IV compares the performance of the NB and SVM classifiers with and without the N-gram feature set to 2: cross-validation = 10, and stratified sampling was used.

2) Second Subset: The second subset contained 666 tweets (189 positive, 378 negative, and 99 neutral tweets).

Table V compares the accuracy, precision, and recall of the NB and SVM classifiers using the N-gram feature set. Additionally, this table compares the NB and SVM classifiers' accuracy, precision, and recall across all classes to the results obtained using the N-gram feature.

As a result, the dataset was split into two distinct time periods, with and without the N-gram feature and neutral class. Separating the data improved classification accuracy, whereas adding a neutral class reduced classification accuracy. SVM achieved the highest accuracy in both cases, with and without the N-gram feature.

Table VI and VII show the TF-IDF schema and N-grams respectively for the original and modified SVM and NB. Table VIII displays SVM performance accuracy with the TF-IDF schema before and after relabelling. Using SVM with the

TABLE IX: PHRASE “IT-WORKS” [مباشر - شغال] IN THE NEUTRAL CLASS DOCUMENTS.

Original Labelling	Original Arabic tweets with English translation	New Labelling
Neutral	مباشر ماهو شغال عندي زين كم مطلوبه التصنيع يا اهل مباشر “Mubasher is not working with me I need to know how much Zain share price.”	Negative
Neutral	هل يوجد برنامج مباشر برو يشتغل علي نظام الماك “Is there a mubasher program running on the Mac?”	Neutral
Neutral	ياخي علمني وشلون يشتغل مباشر على الايباد “He taught me how to work mubasher on the iPad”	Negative
Neutral	انا مثلك ودقيت على مباشر قالي اتصال موبايلي تحميله ضعيف جدا وقال غير الخط الى الاتصالات او اي خط اخر يشتغل “I'm just like you and I called on mubasher, they told me Mobily's connection is very weak and they said I have to change the line to STC line that can works.”	Negative

TABLE X: PHRASE “IT-WORKS” [مباشر - شغال] IN THE POSITIVE CLASS DOCUMENTS.\

Original Labelling	Original Arabic tweets with English translation	New Labelling
Positive	انا عندي مباشر برو ١٠ ونزلت نسخت الاندرويد وشغال على نوت ٤ “I have a mubsher Pro 10 and I downloaded the version of the android and its work on Note 4”	Positive
Positive	أنا شغال على مباشر برو ٩ أنصحك قفله وقفل النت وشغلهم الاثنين من جديد “Mubasher Pro 9 works I advise you to have to lock it and lock the net and lunch them again”	Positive
Positive	سلام انا عندي مباشر من موقع تداول الراجحي شغال ١٠٠% “Salam, I have mubasher from the site of trading Al Rajhi 100% working”	Positive
Positive	شغال عندي مع مباشر من قبل الافتتاح “Mumasher is working with me from the start”	Positive
Positive	مباشر برو شغال عندي من بنك الرياض من سنين زي اللوز “Mubasher Pro has been working for me from Riyad Bank since the years of almonds”	Positive
Positive	مباشر برو **شغال** ممتاز الحمدلله إذا واجهتك أي مشكلة أو عندك أي أستفسار أتصل على الدعم الفني ويتجاوب معك سريع “Mubasher Pro is working perfectly if you have any problem or you have any questions contact technical support and respond quickly”	Positive
Positive	مباشر شغال عندي ممتاز “Mubasher is performing well”	Positive

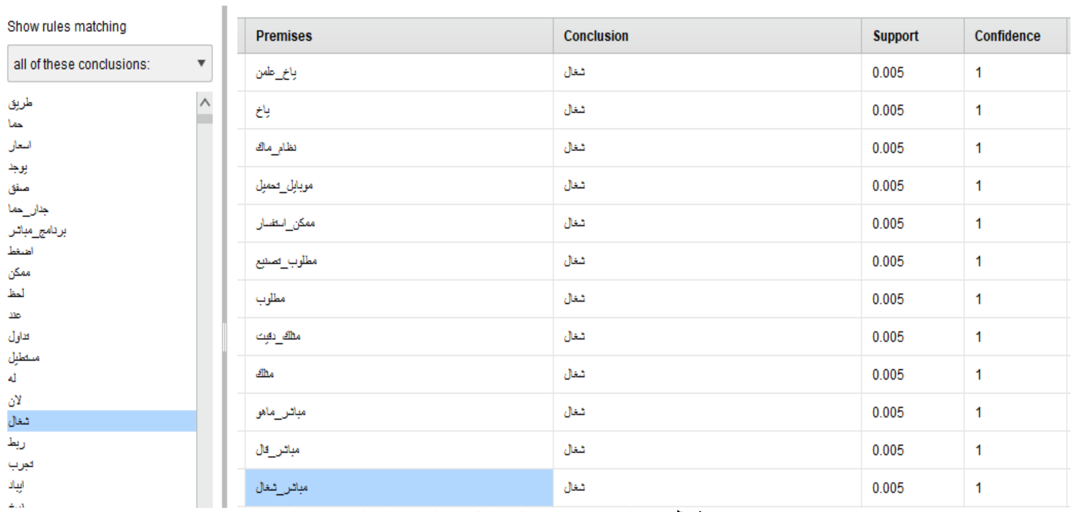

Fig. 3: The correlation rules of the feature “Works” (شغال) in neutral class.

TF-IDF schema and the N-gram feature did not increase classification accuracy. The original labelling dataset had 76.74, 64.40, and 77.26 per cent SVM accuracy, recall, and precision, with a classification error of 23.26 per cent. The new labelling dataset's SVM accuracy, recall, and precision were 75.31, 58.63, and 81.85 per cent respectively, with a classification error of 24.69 per cent.

D. Relabelling Experiment

The process of relabelling begins with a matrix representation of the text data to illustrate the frequency with which each phrase occurs within the three classifications. For all data classes, the relabelling technique captures the link between the feature selection and new words. For example, when users described Mubasher interruption. The feature “it-works” (شغال) is a negative sentiment and it can be happened in positive and neutral class (see Table IX).

TABLE XI: OCCURRENCE OF THE FEATURE “IT-WORKS” (شغال)

Feature	Occurrence	Neutral	Positive	Negative
شغال	80	8	13	59

Fig. 3 displays the association rules for the feature “it-works”(شغال) in all classes with minimum support and confidence thresholds. In addition, Fig. 3 demonstrates the most significant “it-works”(شغال) rules, such as [مباشر] <-- [شغال] (support: 0.071 confidence: 1. The term (Mubasher) correlated with the feature “it-works” [شغال] to compose positive phrases.

The next step is to determine the frequency of the most common phrases associated with the feature “it-works” (شغال) using the wordlist representation. Table X demonstrates that phrases having the feature "it-works" appeared in every class.

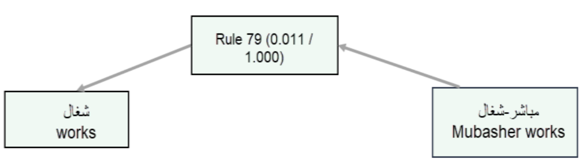

Fig. 4: The association rules for the feature “it-works” “شغال ”.

TABLE XII :PHRASE FOR THE FEATURE “IT-WORKS” (شغال) IN THE ALL DAT

Feature	Neutral	Positive	Negative
مباشر - شغال	24	7	13

From Table X, "Mubasher is working" appeared in positive, negative, and neutral classes, and "it-works" (شغال) is considered negative sentiment towards the Mubasher product. Consequently, both the neutral and positive classes become questionable classes. Therefore, "it-works" (شغال) needs further inquiry to determine the association rules in both classes. Two scenarios result: First scenario: extracting "Works" association rules from the neutral class. Visualization-generated association rules for minimum support and confidence thresholds.

Fig. 4 indicates that “it-works” (شغال) may occur with multiple rules that have little support and confidence. Based on the first scenario, the rule here is [مباشر- شغال] <-- [شغال] (support: 0.011 confidence: 1), which represents the phrase “Mubasher is working” [مباشر- شغال] that is illustrated in the wordlist matrix in the neutral class. According to Fig. 4, the rule] <-- [افضل] [مباشر - شغال (support: 0.079 confidence: 1) which represents the phrase “mubasher is working” occurred in the neutral class. Thus, the next search is for the phrase “mubasher is working” [مباشر- شغال] in the neutral class documents. Table XI shows the phrase “mubasher is working” [مباشر- شغال] occurred in four documents.

The term “Mubasher is working” appears in four documents, as seen in Table XI. Therefore, it has been resent to the expert who initially labelled the paper. Moreover, based on the format of the aforementioned documents, it is clear that they pertain to user discussions regarding mubasher products and platforms, hence the appropriate class for these documents is the neutral class.

Nevertheless, based on the second scenario, the rule of interest is [شغال] -- > [مباشر - شغال] (support: 0.011, confidence: 1), which represents the phrase “mubasher is working” [مباشر- افضل] in the negative class of the wordlist matrix.Next is a search for the positive class documents for the phrase “mubasher is working” [مباشر- شغال]. Table XII displays the occurrences of the phrase “mubasher is working” [مباشر- شغال] in seven documents.

As a result, the document was sent back to the expert who originally labelled it. In total, 18 of 42 analysed tweets were relabelled.

The new labelling dataset performed better using NB and the TF-IDF schema. For example, the precision of the neutral class increased from 43.75 to 54.26 per cent after relabelling. Furthermore, recall was 45.96 per cent for the initial neutral class categorisation but 36.27 per cent for the new classification after relabelling. Furthermore, recall was 45.96 per cent for the initial neutral class categorisation but 36.27 per cent for the new classification after relabelling. To summarise the comparison, the results show that combining NB with the TF-IDF schema and the N-gram feature for the new classification can increase performance. The original labelling dataset's NB accuracy, recall, and precision were 72.05, 64.48, and 64.85 per cent respectively, with a classification error of 27.35 per cent. The new labelling dataset has NB accuracy, recall, and precision of 74.25, 63.92, and 81.85 per cent respectively, with a classification error of 25.75 per cent.

VII. CRM CLASSIFICATION

CRM data created from structured consumer records might be analysed using a similar approach. CRM investment aims to develop solutions that automate a business's sales process. Additionally, maintaining consumer data in databases and utilising data-mining techniques might assist in identifying client patterns. As a result, businesses may identify their most loyal clients and develop new marketing strategies. The experiment collected data from customer records that had been segmented into several CRM groups depending on their perspective on the business. Mubasher's data, for example, were tagged according to the following criteria. Individuals who had never traded on Mubasher's platform were classified as "neutral." A user was classified as "negative" if he or she began trading and then ceased for whatever reason. If the user kept trading on Mubasher's platforms, the label was "positive." The label was created by comparing the start date of each user to his or her most recent order. If the customer's final order was placed prior to the final week of the data collection period (i.e., he or she did not enter a single transaction during that week), he or she must have discontinued trading earlier. Mubasher's technique for CRM data classification is summarised in Fig. 5.

A. Experimental Findings and Evaluation

The proposed approach for classifying Mubasher's CRM data required an understanding of the data and then classification to determine the data's value in the intended CRM. It was decided to identify each customer record as negative, positive, or neutral based on company criteria to determine whether the CRM categorisation findings were compatible with the previous outcome from the first experiment's opinion mining. Mubasher's CRM classification model utilises its database's API to acquire structured CRM data. It applies normalisation techniques to rescale the values of the selected attributes to fit inside a specified range based on the current CRM data. Consequently, 488 records were extracted from Mubasher's CRM, and each record was labelled according to Mubasher's criteria (156 positive, 119 negative, and 213 neutral). The normalised, duplicate-free, null-free dataset was loaded into RapidMiner using RapidMiner. The accuracy, precision, and recall of the NB and SVM classifiers are shown in Fig. 6. To summarise Mubasher's CRM studies, NB attained the highest level of accuracy due to NB's inherent advantages.

VIII. ADJUSTING CRM WITH CEM

In general, schema integration is a generic term that refers to the act of conceptually merging structures from various databases into a single coherent schema [39]. Integration is used to retrieve data from many sources that do not share the same schema. For instance, this research began by mining the Twitter network for information on tweets (text documents in which users can express their viewpoints) (CEM data). Additionally, a second database stores client data gathered through Mubasher's CRM. The critical component lacking in traditional CRM is a mechanism for incorporating social media inputs. This Mubasher case study intended to link clients and their social media remarks. It is necessary to build Arabic sentiment analysis and then connect it to CRM data to collect consumer opinions from Twitter. This study advances CRM by optimising

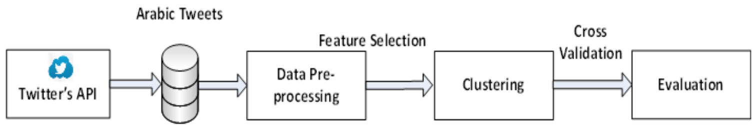

Fig. 5. The process of Arabic tweet-clustering analysis.

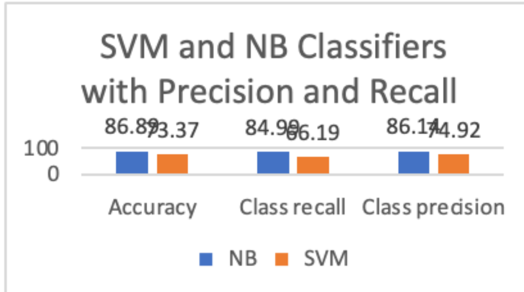

Fig. 6. SVM and NB classifiers with precision and recall.

and updating the CRM framework to accommodate social media customer feedback. The main subjects that should be included in the CRM framework were clustered.

A. *Experiment in Adjustment of CRM by CEM*

After the integration processes were completed, text-mining tools were offered in the form of a framework to gather users' text details to identify those components of each CRM customer record that were lacking, which could result in the CRM structure being updated. The procedure of Arabic tweet-clustering analysis is depicted in. Tweets that were included in the main dataset were kept. RapidMiner was then used to perform five pre-processing steps:

1- Substitution of English terms with corresponding Arabic ones.
2- Splitting tweets into token-based whitespace characters.
3- Removing Arabic stop words.
4- Removing suffixes and prefixes using an Arabic stemmer, returning words to their roots.
5- Removal of useless words (set to four).

In the absence of a predicted class, clustering was used. The purpose was to compute semantic similarity between words using tweets. The analysis started with TF-IDF and then found the shortest distances between centroids inside the corpus.

B. *Results and Evaluation of Experiments*

After implementing the clustering procedure in four clusters, as desired by Mubasher. The primary topic for each cluster was Customer Service (training), Price, Session Interruption (Tec), and Comparison with Another Product. The average distance between cluster centroid centres was calculated using the cluster distance performance task. The results indicate that the best grouping was accomplished by reducing the distances between centroids using IDF's TF-N-gram feature. Fig. 7 depicts the differences between CRM and CEM labelling. In Cluster 0, 36 tweets had identical labels in CRM and CEM, whereas 77 had distinct labels. In Cluster 1, 16 tweets shared the CRM and CEM labels, while 45 had unique labels. Fifty-nine tweets in Cluster 2 shared the same CRM and CEM labels, whereas 174 had unique labels. In Cluster 3, 17 tweets had identical labels in CRM and CEM, whereas 64 had distinct labels.

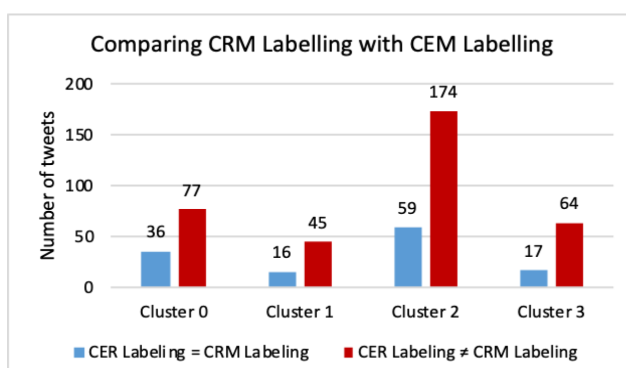

Fig. 7. Comparing labelling CEM with CRM.

The similarities between CRM and CEM labelling are described in Fig. 8. Cluster 0 has the highest similarity score of 32 per cent (a difference of 47 per cent between CRM and CEM labelling), followed by Cluster 1, which has a similarity score of 26 per cent; Cluster 2, which has a similarity score of 25 per cent; and Cluster 3, with 21 per cent. By and large, around 70 per cent of the labelling is different between CRM and CEM. In conclusion, according to the prior data, the greatest degree of label matching between CRM and CEM occurred in the negative class. A possible explanation for this might be that negative sentiments are typically more obvious to people compared to positive and neutral sentiments. Additionally, when discussing uncertain subjects like customer service, the neutral class is essential. On the other hand, certain topics, such as Mubasher's price and comparison to other products, are unimportant. As a result, professionals should refrain from classifying certain mindsets as neutral.

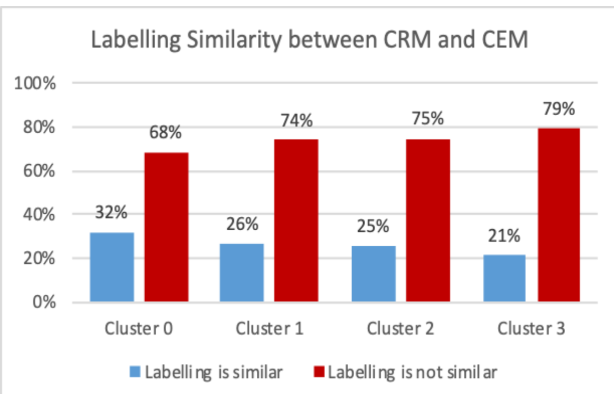

Fig. 8. Similarity of CRM and CEM.

IX. Conclusion

The purpose of this study was to determine whether social media may aid experts in comprehending customers' perspectives and, potentially, alter experts' impressions of consumers. As a result, a framework for CRM and CEM integration was suggested. Among the study's more noteworthy conclusions is that social media can significantly improve CRM by expanding its structure to include more elements highlighted by customers in their social media posts. The second significant conclusion was that, despite the study's modest sample size, knowledge gained via social media may help organisations make better decisions. The findings of this study have several important implications for future practice, such as building an upgraded application that automates the process of Arabic dialect relabelling. Additional points would be necessary to validate the suspicious class and minor rules stated in the text. This could aid in the future development of a robot dictionary.

References

[1] AlHarbi, A., C. Heavin, and F. Carton, Improving customer oriented decision making through the customer interaction approach. Journal of Decision systems, 2016. 25(sup1): p. 50-63.

[2] Chen, I.J. and K. Popovich, Understanding customer relationship management (CRM): People, process and technology. Business process management journal, 2003.

[3] Ranjan¹, J. and V. Bhatnagar, Critical success factors for implementing CRM using data mining. Journal of Knowledge Management Practice, 2008. 9(3).

[4] Saif, H., Y. He, and H. Alani. Semantic sentiment analysis of twitter. in International semantic web conference. 2012. Springer.

[5] Saif, H., et al., Contextual semantics for sentiment analysis of Twitter. Information Processing & Management, 2016. 52(1): p. 5-19.

[6] Younis, E.M., Sentiment analysis and text mining for social media microblogs using open source tools: an empirical study. International Journal of Computer Applications, 2015. 112(5).

[7] Medjani, F. and S.J. Barnes, Understanding the Implementation of Social Customer Relationship Management in the North African Context: An Integrated Theory Perspective. Journal of Global Information Technology Management, 2021. 24(4): p. 299-318.

[8] Badaro, G., et al. A Large Scale Arabic Sentiment Lexicon for Arabic Opinion Mining. Association for Computational Linguistics.

[9] Lemsara, A., S. Ouadfel, and H. Fröhlich, PathME: pathway based multi-modal sparse autoencoders for clustering of patient-level multi-omics data. BMC bioinformatics, 2020. 21(1): p. 1-20.

[10] Sadat, F., F. Kazemi, and A. Farzindar. Automatic identification of arabic dialects in social media. ACM Press.

[11] 11. Haddi, E., Sentiment analysis: text, pre-processing, reader views and cross domains. 2015, Brunel University London.

[12] Yoshida, S., et al. Sentiment analysis for various SNS media using Naïve Bayes classifier and its application to flaming detection. in 2014 IEEE Symposium on Computational Intelligence in Big Data (CIBD). 2014. IEEE.

[13] Bollen, J., H. Mao, and X. Zeng, Twitter mood predicts the stock market. Journal of computational science, 2011. 2(1): p. 1-8.

[14] Martin, V. Predicting the french stock market using social media analysis. in 2013 8th International Workshop on Semantic and Social Media Adaptation and Personalization. 2013. IEEE.

[15] Zhang, W. and S. Skiena. Trading strategies to exploit blog and news sentiment. in Fourth international aAAI conference on weblogs and social media. 2010.

[16] Mouthami, K., K.N. Devi, and V.M. Bhaskaran. Sentiment analysis and classification based on textual reviews. in 2013 international conference on Information communication and embedded systems (ICICES). 2013. IEEE.

[17] Liu, B., Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 2012. 5(1): p. 1-167.

[18] Pang, B. and L. Lee, Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2008. 2(1–2): p. 1-135.

[19] Ganapathibhotla, M. and B. Liu. Mining opinions in comparative sentences. in Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008). 2008.

[20] Ahmad, K. and Y. Almas. Visualising sentiments in financial texts? in Ninth International Conference on Information Visualisation (IV'05). 2005. IEEE.

[21] Boudad, N., et al., Sentiment analysis in Arabic: A review of the literature. Ain Shams Engineering Journal, 2018. 9(4): p. 2479-2490.

[22] Bolton, R.N. and C.O. Tarasi, Managing customer relationships, in Review of marketing research. 2007, Emerald Group Publishing Limited. p. 3-38.

[23] Galitsky, B. and J.L. de la Rosa, Concept-based learning of human behavior for customer relationship management. Information Sciences, 2011. 181(10): p. 2016-2035.

[24] Ngai, E.W., Customer relationship management research (1992-2002) An academic literature review and classification. Marketing intelligence & planning, 2005. 23(6): p. 582-605.

[25] Rygielski, C., J.-C. Wang, and D.C. Yen, Data mining techniques for customer relationship management. Technology in society, 2002. 24(4): p. 483-502.

[26] Achabal, D.D., et al., A decision support system for vendor managed inventory. Journal of retailing, 2000. 76(4): p. 430-454.

[27] Choudhury, M.M. and P. Harrigan, CRM to social CRM: the integration of new technologies into customer relationship management. Journal of Strategic Marketing, 2014. 22(2): p. 149-176.

[28] Budiardjo, E.K., Building Social CRM Framework on Enterprise Architecture Framework Using Value Chain Process Approach. 2014.

[29] Woodcock, N., A. Green, and M. Starkey, Social CRM as a business strategy. Journal of Database Marketing & Customer Strategy Management, 2011. 18(1): p. 50-64.

[30] Sari, P.K., A. Alamsyah, and S. Wibowo. Measuring e-Commerce service quality from online customer review using sentiment analysis. in Journal of Physics: Conference Series. 2018. IOP Publishing.

[31] Alamsyah, A. and E.A. Bernatapi. Evolving customer experience management in internet service provider company using text analytics. in 2019 International Conference on ICT for Smart Society (ICISS). 2019. IEEE.

[32] Tran, A.D., J.I. Pallant, and L.W. Johnson, Exploring the impact of chatbots on consumer sentiment and expectations in retail. Journal of Retailing and Consumer Services, 2021. 63(C).

[33] Chang, H.-C., A new perspective on Twitter hashtag use: Diffusion of innovation theory. Proceedings of the American Society for Information Science and Technology, 2010. 47(1): p. 1-4.

[34] Alsaedi, N., P. Burnap, and O.F. Rana, A combined classification-clustering framework for identifying disruptive events, in In Proceedings of the 6th ASE International Conference on Social Computing (SocialCom ’14). 2014. p. 1-10.

[35] Sankaranarayanan, J., et al., TwitterStand: news in tweets, in Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. 2009, ACM: Seattle, Washington. p. 42-51.

[36] Rapidminer. Rapidminer. 2016 [cited 2017 1 December]; Available from: https://rapidminer.com.

[37] 37. Beniwal, S. and J. Arora, Classification and feature selection techniques in data mining. International journal of engineering research & technology (ijert), 2012. 1(6): p. 6.

[38] Tripathi, G. and S. Naganna, Feature selection and classification approach for sentiment analysis. Machine Learning and Applications: An International Journal, 2015. 2(2): p. 1-16.

[39] Gal, A., et al. A model for schema integration in heterogeneous databases. in Database Engineering and Applications Symposium, 2003. Proceedings. Seventh International. 2003. IEEE.

استخدام تحليل المشاعر للتغريدات العربية لضبط هيكلة إدارة علاقات العملاء

[١]حامد الربيعي و [٢]خالد حمد العمر

[١]جامعة بدفوردشير، بدفورشير، المملكة المتحدة

[٢]كلية الحاسبات وتقنية المعلومات، جامعة الملك عبدالعزيز، جدة، المملكة العربية السعودية

[١]Rubaiee@study.beds.ac.uk, [٢]khalomar@kau.edu.sa

المستخلص. لطالما كان فهم كيفية إدراك العملاء للخدمات التي يتلقونها أمرًا حاسمًا لنجاح الأعمال. ومن المعروف على نطاق واسع أن إدارة علاقات العملاء (CRM) وإدارة تجربة العملاء (CEM) قد ثبت أنهما يساعدان الشركات في اتخاذ قرارات أفضل من خلال تزويدهم بمعلومات أفضل. لسوء الحظ، في تطبيقات الأعمال الفعلية، هناك اختلافات بين آراء العملاء التي يتم جمعها من خلال إدارة علاقات العملاء (CRM) وآراء العملاء الحقيقية التي تم جمعها عبر وسائل التواصل الاجتماعي لإدارة تجربة العملاء (CEM). من الأهمية بمكان سد هذه الفجوة بين الاثنين لمطابقة توقعات العملاء. تم التعرف على كل من CRM و CEM كأدوات يمكن أن تساعد الشركات على اتخاذ قرارات أفضل. ومع ذلك، حتى وقت قريب، تم تجاهل مشكلة دمج البيانات غير المهيكلة من إدارة تجربة العملاء مباشرة في CRM إلى حد كبير، ولا سيما في اللغة العربية، ظلت دون حل. وبالتالي، فإن الهدف من هذه الدراسة هو تقديم إطار عمل لتعديل CRM عبر CEM الذي يدمج طريقة التوجيه الدلالي والتعلم الآلي الخاضع للإشراف. تم استخدام تحليل المشاعر لتعزيز هيكل CRM لإضافة الجوانب المهمة التي كانت مفقودة. تظهر نتائج هذه الدراسة أن الشركات يمكنها بلا شك اتخاذ قرارات أكثر استنارة من خلال توسيع هيكل CRM الخاص بهم لدمج المزيد من القضايا التي يناقشها العملاء على وسائل التواصل الاجتماعي. والمثير للدهشة أن الفئة السلبية لديها أكثر تصنيف مطابق بين CRM و CEM.

الكلمات المفتاحية- إدارة علاقات العملاء، إدارة تجربة العملاء، تحليل المشاعر، وسائل التواصل الاجتماعي، اللغة العربية.