Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 06, No. 03, September 2020.

SENTIMENT ANALYSIS AND CLASSIFICATION OF ARAB JORDANIAN FACEBOOK COMMENTS FOR JORDANIAN TELECOM COMPANIES USING LEXICON-BASED APPROACH AND MACHINE LEARNING

Khalid M.O. Nahar[1], Amerah Jaradat[1], Mohammed Salem Atoum[2] and Firas Ibrahim[3]

(Received: 7-Apr.-2020, Revised: 6-Jun.-2020, Accepted: 8-Jun.-2020)

ABSTRACT

Sentiment Analysis (SA) is a technique used for identifying the polarity (positive, negative) of a given text, using Natural Language Processing (NLP) techniques. Facebook is an example of a social media platform that is widely used among people living in Jordan to express their opinions regarding public and special focus areas. In this paper, we implemented the lexicon-based approach for identifying the polarity of the provided Facebook comments. The data samples are from local Jordanian people commenting on a public issue related to the services provided by the main telecommunication companies in Jordan (Zain, Orange and Umniah). The produced results regarding the evaluated Arabic sentiment lexicon were promising. By applying the user-defined lexicon based on the common Facebook posts and comments used by Jordanians, it scored (60%) positive and (40%) negative. The general lexicon accuracy was (98%). The lexicon was used to label a set of unlabeled Facebook comments to formulate a big dataset. Using supervised Machine Learning (ML) algorithms that are usually used in polarity classification, the researchers introduced them to our formulated dataset. The results of the classification were 97.8, 96.8 and 95.6% for Support Vector Machine (SVM), K-Nearest Neighbour (K-NN) and Naïve Bayes (NB) classifiers, respectively.

KEYWORDS

Jordan Telecom, Sentiment analysis, Lexicon-based, Polarity, Facebook comments, Machine learning, NLP.

1. INTRODUCTION

Language Processing (LP) is the field of computer science and artificial intelligence that mainly studies human-computer language interaction [1]. SA and opinion mining is a field of NLP that investigates and analyzes people's opinions, sentiments, evaluations, attitudes and emotions from written language. It is one of the most active research areas in NLP and is also widely studied in data mining, web mining and text mining [2][24].

The important part of information-gathering behaviour has always been to find out what other people think. With the growing availability and popularity of opinion-rich resources, such as online reviews and personal blogs, new opportunities and challenges arise, as people now can actively use information technologies to seek out and understand the opinions of others. Polarity classification can be applied in individual reviews to evaluate the goodness of a certain product [22][25]. The sudden eruption of activity in the area of opinion mining and SA, which deals with the computational treatment of opinion, sentiment and subjectivity in text, has thus occurred at least in part as a direct response to the surge of interest in new systems that deal directly with opinions as a first-class object [3].

To determine whether a sentence, text or any comment expresses a positive or negative sentiment, three main approaches are commonly used: the lexicon-based approach, machine learning approach and a hybrid approach. Figure 1 explains these approaches [4][29]. In this work, we implemented the lexicon-based approach. The reason behind choosing the lexicon-based approach is that both machine learning and hybrid approaches demand a labeled dataset for supervised learning. Also, Jordanians as

1. K. M. O. Nahar and A. Jaradat are with Computer Sciences Department, Yarmouk University, Irbid, Jordan. Emails: khalids@yu.edu.jo and ameera@yu.edu.jo
2. M. S. Atoum is with Computer Science Department, Irbid National University, Irbid, Jordan. Email: cs@inu.edu.jo
3. F. Ibrahim is with Information Systems and Networks Department, The World Islamic Sciences and Education University, Amman, Jordan. Email: firas.alzobi@wise.edu.jo

other Arabs use their dialects and modern Arabized words, letters, symbols, paronomasias and insinuations for expressing their opinions.

Companies (Zain, Orange and Umniah) interconnect through video, voice and data (mainly internet browsing and social media). The cost of communications provided by those companies is too low compared to neighbouring countries and the level of services provided is also very good, but Jordanians do express their opinions, feelings and sentiments about those companies regarding cost, coverage, offers, internet speed, …etc. These types of opinions may be an indicator of continuing or leaving one company to another or from offer to offer. Most of those opinions, feelings and sentiments are expressed using Jordanian different Arabic dialects in addition to lack of using Original Standard Arabic.

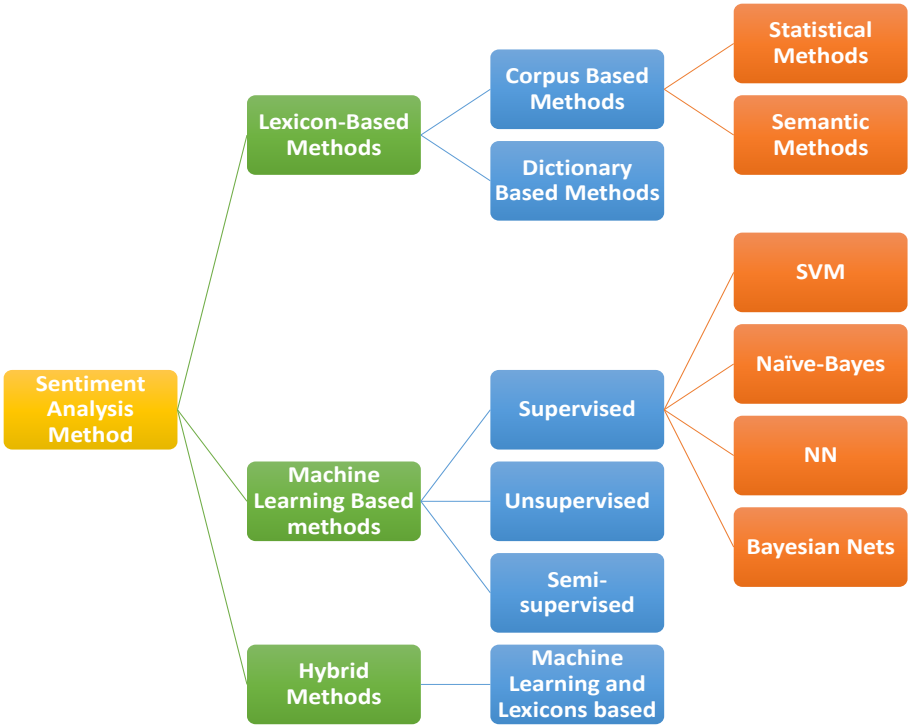


Figure 1. Main approaches of SA [4].

Customer churn analysis is a very common task in data analysis. It involves trying to predict whether customers will quit or continue the contract. It is crucial to the telecommunication companies to review and analyze their customers' feedback to enhance their provided services and avoid losing their contracts. NLP is a great method to automatically analyze sentiments and predict whether those sentiments are positive or negative as an early indicator for the quality of the provided services.

In this work, we are proposing an approach to predict customer satisfaction with the services provided by the telecommunication companies. The approach collects posts and comments from Facebook pages related to Jordanian telecommunication companies in order to find out the customer attitude toward these companies. After collecting and pre-processing the data, sentiment analysis is achieved using the Lexicons-Based Approach (LBA). Owing to the amount of data handled, the work involves automatic translation of English sentiment lexicon to create Arabic sentiment lexicons.

The paper is prearranged as follows. In Section 2, we review some of the previous research related to the field of SA. Then, in Section 3, we introduce the lexicon-based approach for creating the dataset. In Section 4, we apply supervised learning algorithms on the formulated dataset. In addition, we describe the supervised learning model on both KNIME and ORANGE software and show the experimental results and the evaluation of the anticipated method. Finally, we address the conclusions and discuss future works in Sections 5 and 6.

2. LITERATURE REVIEW

Many kinds of research have been devoted to the field of SA [26]-[28]. Unfortunately, few works consider discussing SA for the Arabic language. Moreover, research on Arabic sentiment analysis has not perceived noteworthy developments yet, typically due to the shortage of sentiment resources in Arabic [20]-[21].

Rehab M. Duwairi and Islam Qarqaz [5] carried out an experiment using Rapid miner, which is an open-source machine learning software, to perform SA in Arabic text. The dataset was collected from tweets and Facebook comments that address issues in education, sports and politics. In this study, the main issue was determining the polarity (positive, negative or neutral) of the given text. The authors applied two approaches: the machine learning approach and the lexicon-based approach. Three supervised classifiers (SVM, Naïve Bayes and K-NN) were applied on an in-house collected dataset of 2591 tweets/comments from social media to analyze the sentiment of Arabic reviews. Unfortunately, the dataset was not large enough to make strong conclusions.

Rehab M. Duwairi [6] used classification for SA. After extracting Arabic tweets, the author applied Naïve Bayes (NB) and Support Vector Machine (SVM) classifiers. SVM and NB classifiers were used on a big dataset that consists of almost 22500 Arabic tweets. The experiments involved comparing the lexicon values without the dialect lexicon to the values with converting dialectical words into MSA words. The results show the great impact of the dialect lexicon on the F-measure of the positive and negative classes as well as the Macro-Precision, Macro-Recall and F-Measures. The results were limited by the storage deficiency of the Rapid miner software used.

Ahmad A. Al Sallab et al. [7] concentrated on a deep learning framework to analyze the sentiment of Arabic text with features based on the developed Arabic sentiment lexicon with standard lexicon features. One supervised classifier (SVM) and four unsupervised classifiers (DNN, DBN, DAE and RAE) were applied on a dataset of 3795 entries. Results show that RAE produces the best accuracy.

Haifa K. Aldayel and Aqil M. Azmi [8] proposed a hybrid approach combining semantic orientation and SVM classifiers. The used data passed through pre-processing operations to be ready to a lexical-based classifier, then the output data became a training data for the machine learning classifiers. The proposed approach used 1103 tweets. The experimental results show better F-measure and accuracy of the hybrid approach.

Hala Mulki et al. [9] proposed two classification models to analyze the Arabic sentiment of 3355 tweets written with MSA and Arabic dialects. Authors considered the sentiment classification of Arabic tweets through two classification models: supervised learning-based model and unsupervised learning-based (lexicon-based) model. The conducted experiments showed better F-score and Recall values using the supervised learning-based model. On the other hand, the unsupervised learning-based (lexicon-based) model achieved better results if the stemming did not assign the lookup process.

Nora Al-Twairesh et al. [10] collected a corpus of Arabic tweets by collecting over 2.2 million tweets. Authors presented the sequence of operations used in collecting and constructing a dataset of Arabic tweets: cleaning and pre-processing the collected dataset included filtering, normalization and tokenization. Later, with the help of annotators, the dataset was labeled with (positive, negative, mixed, neutral or indeterminate). Then, the data was classified using the SVM classifier and provided as a benchmark for future work on SA of Arabic tweets.

Hassan Najadat et al. [11] applied four supervised classifiers on a dataset of 4227 posts' texts from the Facebook pages to determine the efficiency of the main three telecommunication companies in Jordan: Orange, Zain and Umniah, based on the SA of customers who use social media, especially Facebook. The results were promising. However, the accuracy without sampling was better than that with sampling.

Leena Lulu and Ashraf Elnagar [12] proposed neural network models from different deep learning classifiers for the automatic classification of Arabic dialectical text. The proposed approach used the manually annotated Arabic online commentary (AOC) dataset that consists of 110 K labeled sentences. This approach yielded an accuracy of 90.3%.

Assia Soumeur et al. [13] and [19] focused on opinions, sentiments and emotions based on various

Facebook pages' posts written in Algerian dialect. The authors applied two types of neural network models: MLP and CNN, in addition to Naïve Bayes to classify comments as negative, positive or neutral. After considering the pre-processing steps, both models achieved good accuracy results with a slightly better accuracy using the CNN model. This indicates obtaining higher accuracies using deep learning models in general.

Jalal Omer Atoum and Mais Nouman [14] focused on SA of social media users' tweets written in Jordanian dialect. After a sequence of pre-processing steps, the dataset was labeled with positive, negative or neutral. The study applied two supervised classifiers, Naïve Bayes and SVM, on the tweets. The conducted experiments involved experimenting with different factors. The results show higher accuracy values using the SVM classifier. Results also show that using stems and root trigrams on balance data enhances the accuracy. In summary, Table 1 provides a comprehensive and comparative overview of the studied literature for the research from [5]-[14].

Table 1. A comprehensive and comparative overview of the studied literature.

Paper	Dataset	Approach	Features	Classifiers	Results (Accuracy)
[5]	Tweets Facebook comments	Supervised machine learning Lexicon-based	N-grams in words	SVM NB K-NN	Best precision with SVM = 75.25% Best recall with K-NN = 69.04%
[6]	Tweets	Hybrid (supervised and unsupervised)	Words	SVM Naïve Bayes	F1-score of SVM = 87% F1-score of NB = 88%
[7]	LDC - ATB	Hybrid (supervised and unsupervised)	ArSenL-sentence ArSenL-Lemma Raw words	DNN DBN DAE RAE	SVM = 45.2% DNN = 39.5% DBN = 41.3 % Deep auto DBN = 43.5% RAE = 74.3%
[8]	Tweets	Hybrid (orientation and machine learning techniques)	Unigrams Bigrams Trigrams	SVM Lexical-based	F1-score = 84% Accuracy = 84.01%
[9]	Tweets	Supervised learning Lexicon-based	Higher-order N-grams (compared to unigrams up to trigrams)	Naïve Bayes SVM (LIBSVM) Baseline	F1-score of SVM = 0.384 F1-score of NB = 0.284 F1-score of baselines = 0.249 Avg. Accuracy = 0.454
[10]	Tweets	Supervised learning	TF-IDF (term frequency – inverse document frequency) Term frequency Term presence	SVM (linear kernel) Two-way, three-way, four-way classifications	Two-way: (Term presence) = 62.27% Three-way: (Term presence) =58.17% Four-way: (Term presence) = 54.69%
[11]	Facebook posts	Supervised learning	Words	Naïve Bayes SVM DT K-NN	Best accuracy with SVM without sampling: (Avg.) 93.7% Best accuracy with SVM with sampling: (Avg.) 73.54%

[12]	Arabic online commentary (AOC)	Unsupervised learning	Lexical-based words	LSTM CNN BLSTM CLSTM	Best Accuracy = 90.3%
[13]	Facebook comments	Hybrid (supervised and unsupervised)	Word (text from comments)	MLP CNN Naïve Bayes	MLP = 81.6% CNN = 89.5% NB: Before pre-processing = 60.11% After pre-processing = 71.73%
[14]	Tweets	Supervised Semantic orientation	N-grams in word (unigrams up to N-gram)	SVM Naïve Bayes	Accuracy of NB with trigram = 55% SVM with trigram = 76%

In addition to previously summarized literature in Table 1, Saif M. Mohammad et al. [16] applied two different approaches to automatically generate several large sentiment lexicons. The first generating method was using distant supervision techniques on Arabic tweets and the second method was translating English sentiment lexicons into Arabic using a freely available statistical machine translation system. The authors provide a comparative analysis of the new and old sentiment lexicons in the downstream application of sentence-level SA.

Rehab Duwairi and Mahmoud El-Orfali [18] approached SA in Arabic text using three perspectives. First, investigating several alternatives for text representation; in particular, the effects of stemming feature correlation and n-gram models for Arabic text on SA. Second, investigating the behaviour of three classifiers; namely, SVM, Naïve Bayes and K-nearest neighbor, with SA. Third, analyzing the effects of the characteristics of the dataset on SA.

3. METHODOLOGY

Recently, many researchers have devoted efforts to studying the platforms of social media [30]-[31]. The interest in studying social media is due to the rapid growth of its contents as well as its impact on people's behaviour [15]. A major part of their studies focused on SA and opinion mining.

3.1 Collecting Sentiment Lexicons

In the lexicon-based approach, big efforts focused on the English sentiment lexicons [16] while little focus was placed on Arabic sentiment lexicons. On the other hand, most of these efforts focused on solving special problem statements. Arabic language sentence flow is a challenging issue due to many reasons; for example, Arabic sentences are full of using negations, modals, intensifiers and diminishers. Moreover, the Arabic language is very rich in prepositions, conjunctions, connected pronouns, object pronouns, demonstrative pronouns, relative pronouns, pronouns, paronomasia, insinuation and many other issues that require a lot of work when computerizing the Arabic language understanding. Table 2 shows Arabic language controls.

Negators change lexicons from negative to positive and the other way around. For example, the word (سعيد, happy) is a 100% positive sentiment word, but if it is preceded with any suitable negation as (غيرسعيد، ليس سعيداً، لا يعتبر سعيداً), then it is negative (not necessarily, maybe neutral).

The same is valid for the negative sentiment words if preceded with any suitable negation; for example, (ضارّ) is a 100% negative sentiment word, but when negated like (غير ضارّ، ليس ضارً، لا يعتبر ضارّاً), then it is positive (not necessarily, maybe neutral). Arabic negations applied are such as (لا، لم، ما، ليس، غير، لن، دون، مش مو). In diminishers, the word might be negative by itself, but in the sentence, it gives a positive connotation (كريم في الخيرات، بخيل في الإساءة الى الناس). Other examples are (حلو مُرّ لين شرس عصبي طيّع). The Arabic language is rich in paronomasia words and sentences that give many meanings, like (رقيق) which has the meaning of (slave), (thin) or (gentle). The insinuations are close to paronomasias, where the speaker may say positive words, but he/she means negative ones, like (اتقوا الله).

Table 2. Arabic language controls.

Controls	Meaning in Arabic	Example in Arabic	Action
Negators	النافيات	لا، لم، ما	Manipulate
Modals	الأفعال الناقصة	أصبح، أمسى	Remove
Intensifiers	التوكيد	تكرار الكلمة	Remove
Diminishers	المتناقضات	محب مبغض	Phrases
Prepositions	حروف الجر	من، الى	Remove
Conjunctions	حروف العطف	ثم، أو	Remove
Connected Pronouns	ضمائر متصلة	الهاء في نفسه	Remove
Demonstrative Pronouns	ضمائر الإشارة	هذا، هذه	Remove
Relative Pronouns	الأسماءُ الموصُولة	الذي، التي	Remove
Pronouns	الضمائر	هو، هي	Remove
Paronomasia	التورية	أعينيَّ جودا	Phrases
Insinuation	التلميح	لا والله كريم!	Phrases

The approaches used to create Arabic sentiment lexicons can be broadly divided into three categories [32]. The first and most used approach is strongly based on the automatic translation of English sentiment lexicons and resources, either for all the entries in the lexicon or only for some parts. The second approach relies on choosing seed sentiment words and then identifying the words that occur alongside the seed words, using either statistical measures or simply using conjugates. The third approach involves human effort in manually extracting sentiment words, either from Arabic dictionaries or from datasets collected for Arabic SA and subsequently labeling these words with their polarities (positive, negative, neutral) [17].

In this work, we created Arabic sentiment lexicons through the automatic translation of English sentiment lexicons and the manual extraction of sentiment words. Next, we describe each of these resources.

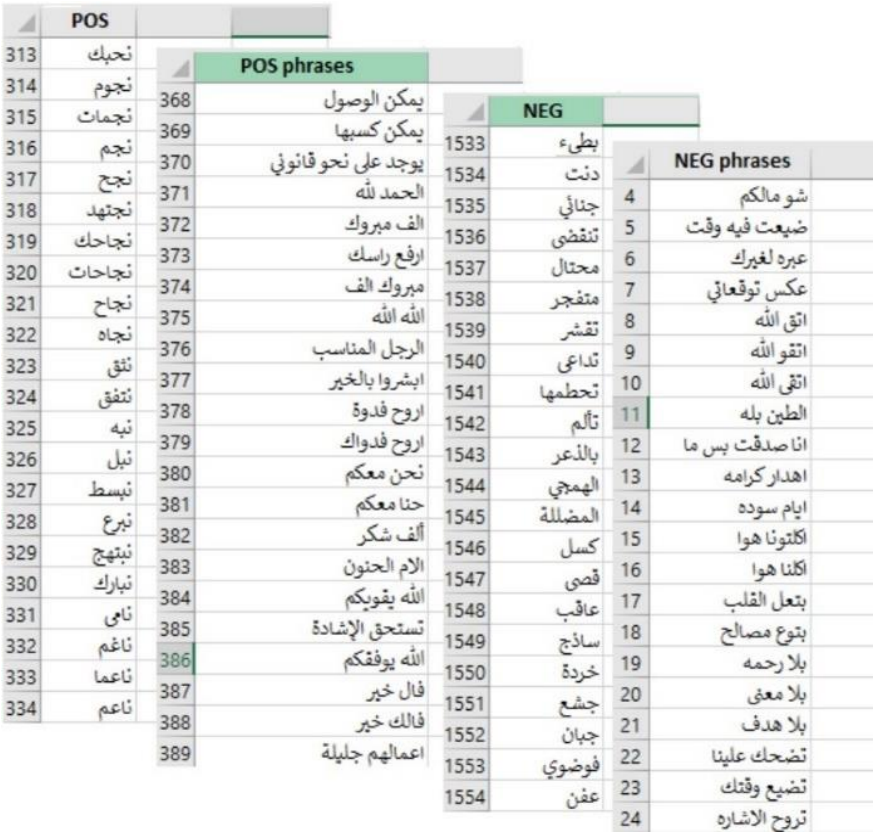
3.1.1 Manually Prepared Lexicon

Our work concentrates on the comments of Facebook users related to social, public issues. The researchers had to add more words and phrases to the available lexicons because of the dialect phrases and words used by commenters as well as using negators, intensifiers, paronomasias and insinuations.

The Arabic Sentiment Lexicon comprises 333 negative phrases, 369 positive phrases, 4956 negative words and 2145 positive words, in addition to a largely manipulated negation applied on negative and positive words (39648 negative negations and 17160 positive negations). These negations were made through the concatenation of applicable negations with the sentiment's words and phrases. Figure 2 shows samples from the generated lexicons.

3.1.2 English Translated Sentiment and Emotion Lexicon

There are many English lexicons translated into Arabic, but they are hardly free of mistranslation or have different synonyms [33]-[34]. In this work, we used some lexicons from Saif M. Mohammad's collection [16]. In this collection, the author used automatic translations of English sentiment lexicons into Arabic. The study in [16] reveals that about 88% of the automatically translated entries are valid for Arabic and around 10% invalid entries were the result of gross mistranslation. 40% of the invalid entries occur due to translation into a related word and about 50% occur due to differences in how the word is used in Arabic. The translations were often the word representing the predominant sense of the word in the English source [16].



	POS
313	نحبك
314	نجوم
315	نجمات
316	نجم
317	نجح
318	نجتهد
319	نجاحك
320	نجاحات
321	نجاح
322	نجاه
323	نثق
324	نتفق
325	نبه
326	نبل
327	نبسط
328	نبرع
329	نبتهج
330	نبارك
331	نامى
332	ناغم
333	ناعما
334	ناعم

	POS phrases
368	يمكن الوصول
369	يمكن كسبها
370	يوجد على نحو قانوني
371	الحمد لله
372	الف مبروك
373	ارفع راسك
374	مبروك الف
375	الله الله
376	الرجل المناسب
377	ابشروا بالخير
378	اروح فدوة
379	اروح فدواك
380	نحن معكم
381	حنا معكم
382	ألف شكر
383	الام الحنون
384	الله يقويكم
385	تستحق الإشادة
386	الله يوفقكم
387	فال خير
388	فالك خير
389	اعمالهم جليلة

	NEG
1533	بطيء
1534	دنت
1535	جنائي
1536	تنقضى
1537	محتال
1538	متفجر
1539	تقشر
1540	تداعى
1541	تحطمها
1542	تألم
1543	بالذعر
1544	الهمجي
1545	المضللة
1546	كسل
1547	قصى
1548	عاقب
1549	ساذج
1550	خردة
1551	جشع
1552	جبان
1553	فوضوي
1554	عفن

	NEG phrases
4	شو مالكم
5	ضيعت فيه وقت
6	عبره لغيرك
7	عكس توقعاتي
8	اتق الله
9	اتقو الله
10	اتقى الله
11	الطين بله
12	انا صدقت بس ما
13	اهدار كرامه
14	ايام سوده
15	اكلتونا هوا
16	اكلنا هوا
17	بتعل القلب
18	بتوع مصالح
19	بلا رحمه
20	بلا معنى
21	بلا هدف
22	تضحك علينا
23	تضيع وقتك
24	تروح الاشاره

Figure 2. Sample from the lexicons.

All those lexicons are very powerful in terms of Arabic sentiments words and would be helpful if researchers were to mainly analyze texts written correctly according to the Arabic linguistic structure. Unfortunately, most comments were written quickly, without correct wording (misspelling) or informed prior thinking. The Facebook post was to get the opinions of users regarding services provided by Jordanian Telecommunication Companies (Zain, Orange and Umniah) and subjects (persons) of this study who put their comments on the post seemed to be in a hurry and with no concentration when writing their comments. Most of those subjects used their own words to describe the service or to talk about their own experience with the company using sentences not always free of improvised created words, slang vocabularies and inclusion. Saif M. Mohammad and his team [16] provided huge size files that need a lot of work and approval before applying them. However, we used what seems to fit and applies to the conducted experiments.

3.2 Formulating Labeled Dataset

Related Facebook comments were collected from the Facebook[123] pages of the Jordan telecom companies with some reviews from different related comments. Figure 3 shows some sample comments from Facebook regarding the services provided by the major telecommunication companies in Jordan.

After loading and labelling the collected comments, the researchers applied some text pre-processing on them. Pre-processing is a vital step for getting sentimental text. The main tools that were used in this step are Microsoft Excel and KNIME. Microsoft Excel was used in preparing and proofing the lexicon words, concatenation of negations with lexicon words, removing modals, prepositions, pronouns and intensifiers, as well as other text operations, like refinement and trimming of demonstrative pronouns.

[1] https://m.facebook.com/OrangeJordan

[2] https://m.facebook.com/zainjordan

[3] https://m.facebook.com/Umniah

	A
1	امنية كويسة
2	امنيه ماشي حالها
3	اورانج كويسه بشكل عام
4	زين خط تلفوني كويسه بس مش ممتازه و اورنج عنا خط النت صاير سئ جدا وبطئ وعم افكر اغير بس اشوف بديل مناسب
5	خطي معمول عليه حظر بث ومشاركة ما بسمح بالبث والمشكلة فيه 10 جيجا والله احيانا بروحو باخر الشهر وعالفاضي بسبب انه الشبكة بطيئة
6	خطوط الاردن كلها تكلفتها رخيصة مقارنة بالبلدان المجاورة هاك الامارات بتكلف كثير
7	اتصالات زين على البلدان المجاورة مثل السعودية تكلفتها منطقية ومناسبة
8	اسخف عروض واسخف خدمات شفتها بحياتي
9	الحقيقة انه احسن من هيك شركة ما في
10	الخدمة مريحة والتكلفة مناسبة سرعة الانترنت مقبولة
11	الافضل والاحسن والاقوى شركة اورانج ولا منازع
12	هيك الشركات ولا بلاش انا جربت كل الشركات والله تعامل ما شاء الله قمة الاحترام
13	زين بحسه أفضل شي بعد عدة تجارب
14	زين والسبب تغطية اوسع على مستوى المملكة
15	على الرغم من انه عروض الشبكات الاخرى افضل ولكن شو الفايدة من انه يكون معك بيانات اكثر ونسبة التحميل ضعيفة
16	ملاحظة احيانا تجد في بعض المناطق اورانج او امنية اقوى ولكنها مناطق محدودة
17	زين مكالمات ونت ممتاز بس بسحبو نت كتير
18	منيه بس طبعا حسب المنطقة قوة او ضعف النت وكل الشبكات نفس المشكلة
19	اورانج الشبكة ممتازة و الزين ممتازة كمان
20	امنية وكثير بيجنن والحزم بقعدو معي شهر كامل والشبككه ممتازه

Figure 3. Sample from the collected comments.

Meanwhile, KNIME was used in the analysis and labeling. The main pre-processing steps that were performed are the following steps:

- Removing punctuations including apostrophe, colon, comma, dash, ellipsis, exclamation point, hyphen, parentheses, period, question mark, quotation mark and semicolon.
- Replacing some letters and words with other letters or words that are known to the system. Here are some examples: (أ/آ/إ/ٱ by ا), (ئ/ي by ى), (ليست/ليستا/لستما/ليسا/لسنا//لستم لستن /ليسوا/لست/لسن by ليس), (ة by ه), (ؤ by و), removing (الـ) by replacing it with space, replacing (لا يعتبر/لا يعد by ليس).
- Removing of propositions, such as: (من، الى، عن، على، في).
- Tagging positive phrases, like (الف مبروك، والله نشامى) and negative words, such as: (اتقوا الله، حسبي الله فيكم، عديم الفائدة).
- Tagging of negations.
- The priority of tagging was as follows: negative phrases, positive phrases, negated negative, negated positive, negative and positive. Note that after applying Dictionary Tagger in KNIME on a phrase or word, it will not be changed. Mostly, the researchers focused on negative phrases and words, since this solution follows the lexicon-based approach to perform sentiment analysis. A recent and important work can be referred to in [2].
- Filtering the tagged words, so that only sentiment words are included in the counting of sentiments.
- Creating a bag of words that separate each group of sentiment words individually.
- Counting the frequency of each group.
- Calculating the result using Equation 1 and if the value gained from Equation 1 is greater than or equal the mean (result), then the comment is POSITIVE; otherwise, the comment is NEGATIVE.

$$Results = \frac{POS_{words} - NEG_{word}}{TOTAL_{words}} \quad (1)$$

Figure 4 represents the previously mentioned steps for pre-processing, analyzing and labelling of collected comments, then formulating a labeled dataset.

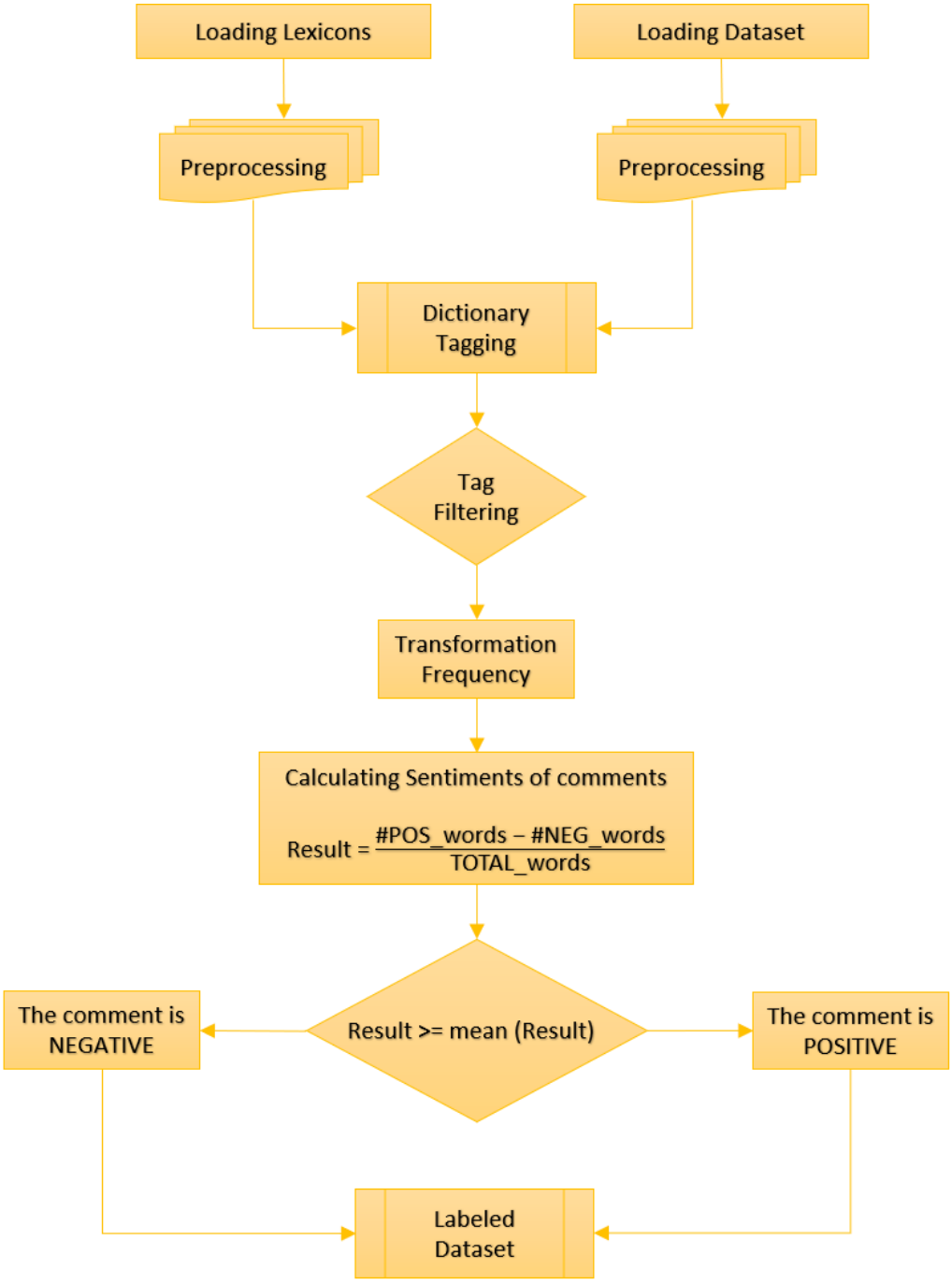


Figure 4. Labeling dataset and pre-processing flow.

In this research, we have applied the KNIME Analytics Platform for analysis and labeling based on the dictionary tagger. KNIME is a software built for fast, easy and intuitive access to advanced data science, helping organizations drive innovation. KNIME[4] Analytics Platform became one of the leading open solutions for data-driven innovation, designed for discovering the potential hidden in

[4] Knime official website: https://www.knime.com/about

data, mining for fresh insights or predicting new features. Figure 5 shows the KNIME analyzing and labeling model.

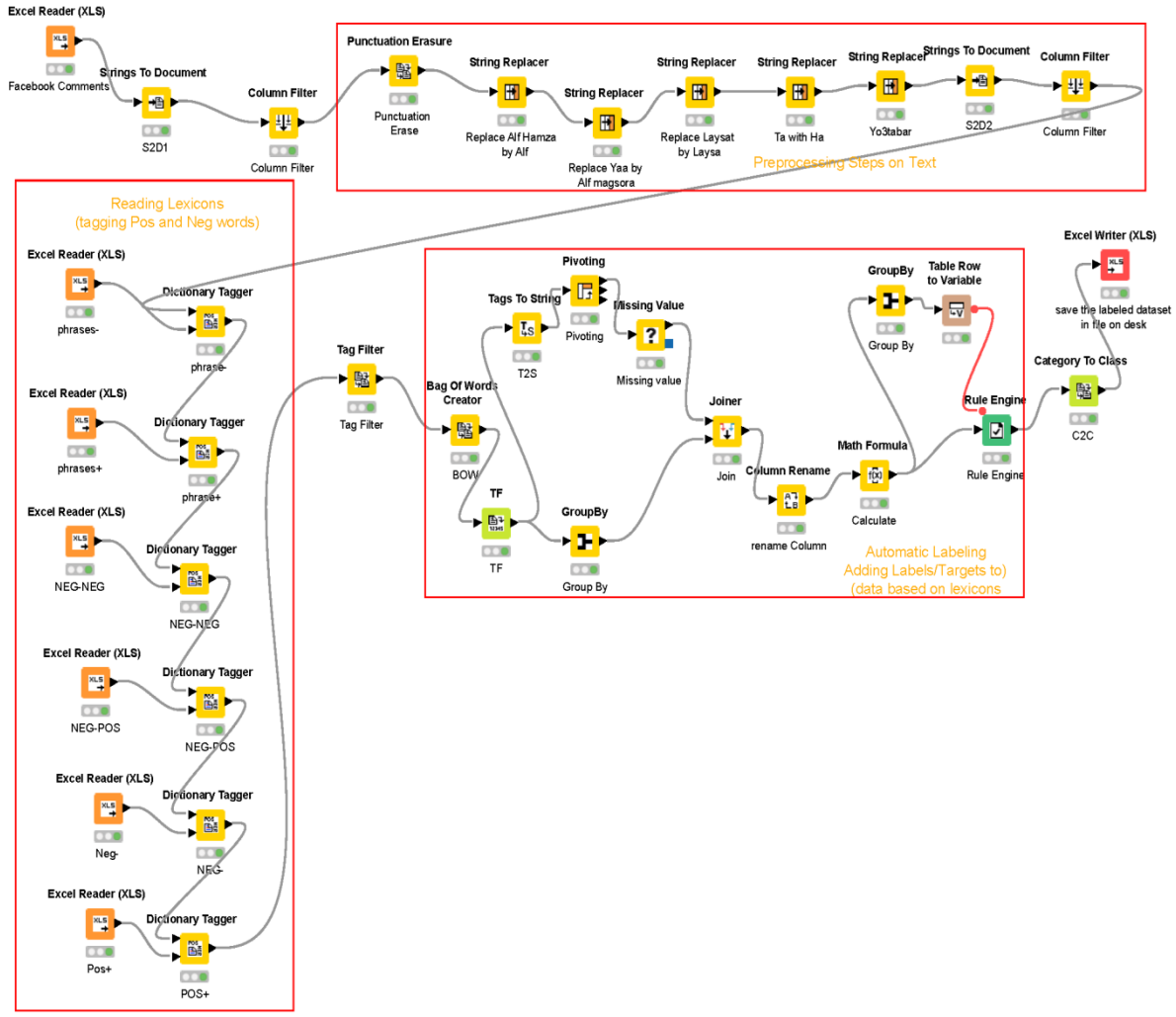


Figure 5. Analyzing and Labeling in KNIME.

From the model in Figure 5, KNIME parts, nodes, extensions and components of the model along with its purpose are as follows:

- "Excel Reader (XLS)" node is designed for reading Text Data from Excel files. Here, we manually collected the dataset from three Facebook pages. Then, we inserted it into the Excel sheet in addition to other lexicon files.
- "Strings to Document" node converts the specified strings into documents. The input is a data table containing string cells and the output is a table containing the strings of the data of the input table as well as the created documents in an additional column.
- "Column Filter" node filters certain columns from the input table and allows the remaining columns to pass through the output table.
- "Punctuation Erasure" node removes all punctuation characters of terms contained in the input documents. Input is the table that contains the documents before pre-processing and output is the table that contains the documents after pre-processing.
- "String Replacer" node is for replacing and removing pre-positions and other punctuations previously listed. It replaces values in string cells if they match a certain wildcard pattern. Input is arbitrary data and the output column contains the SA (NEG, POS) of each data column (comment). The SA results are calculated using Equation 1.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 06, No. 03, September 2020.

The labeling process in this manner was unsupervised. It was depending on the collected data, either from Arabic lexicons or English translated lexicons after refinement and filtering.

4. SUPERVISED MACHINE LEARNING

The previously described workflow in section 3 resulted in labeled data (unsupervised). To test the accuracy of the proposed approach, we used the labeled dataset to formulate a Machine Learning model using ORANGE software to implement the classifier. ORANGE is an open-source data visualization, machine learning and data mining toolkit. It features a visual programming front-end for explorative data analysis and interactive data visualization. ORANGE is a component-based visual programming software package for data analysis. ORANGE's components are called widgets and they range from simple data visualization, subset selection and pre-processing, to empirical evaluation of learning algorithms and predictive modeling. In ORANGE, visual programming is implemented through an interface in which workflows are created by linking predefined or user-designed widgets, while advanced users can use ORANGE[5] as a Python library for data manipulation and widget alteration.

Experiments were conducted on the 1300 comments to produce a Predictive Machine Learning Model (PMML) (supervised) using the PMML predictor. A text pre-processing step is necessary to remove all unnecessary and misleading words. Then, we experiment with the pre-processed text on Machine Learning. Figure 6 shows the workflow of the ML model implemented using ORANGE software. After that, we tested the collected dataset using SVM, NB and K-NN algorithms, since they are the most used algorithms in this context.

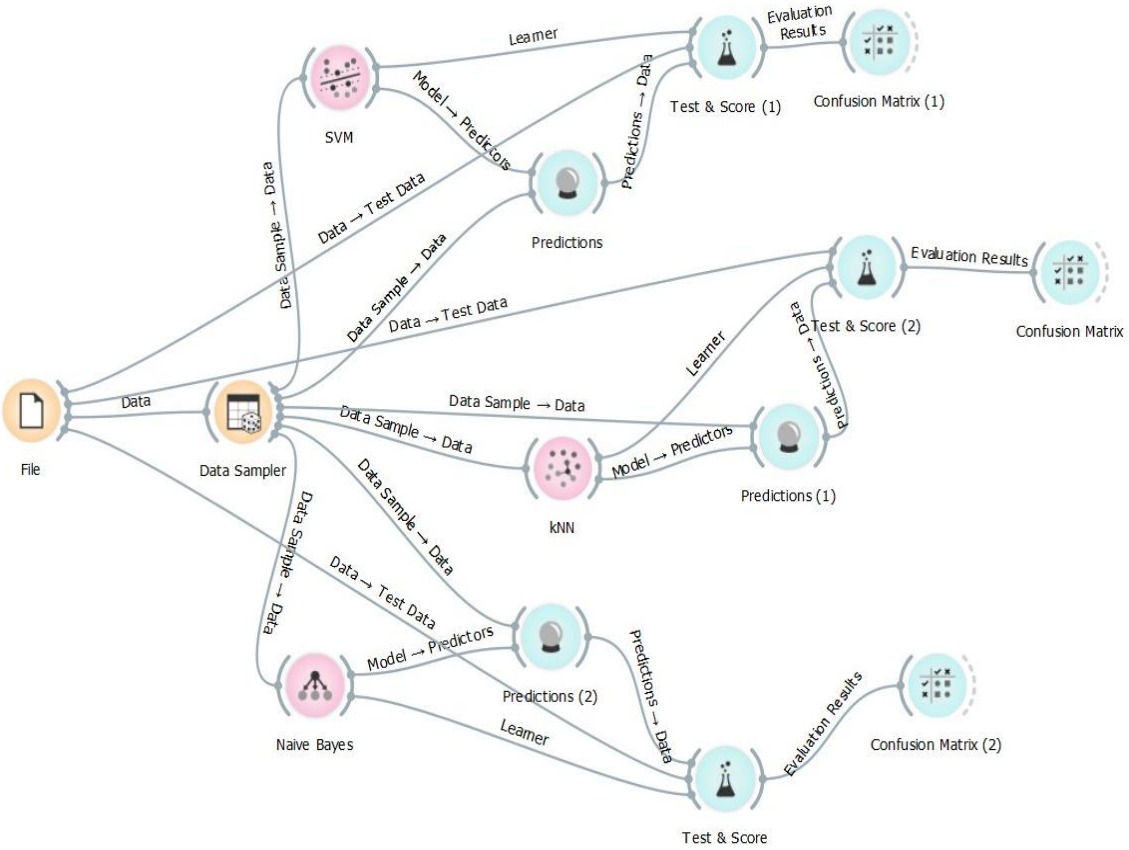


Figure 6. ORANGE-workflow (ML model of SVM, NB and K-NN).

In summary, Figure 7 illustrates our proposed work from unsupervised labeling using a lexicon-based approach till supervised learning verification of the labeled comments using the ML model.

[5] ORANGE official website: https://orange.biolab.si/workflows/

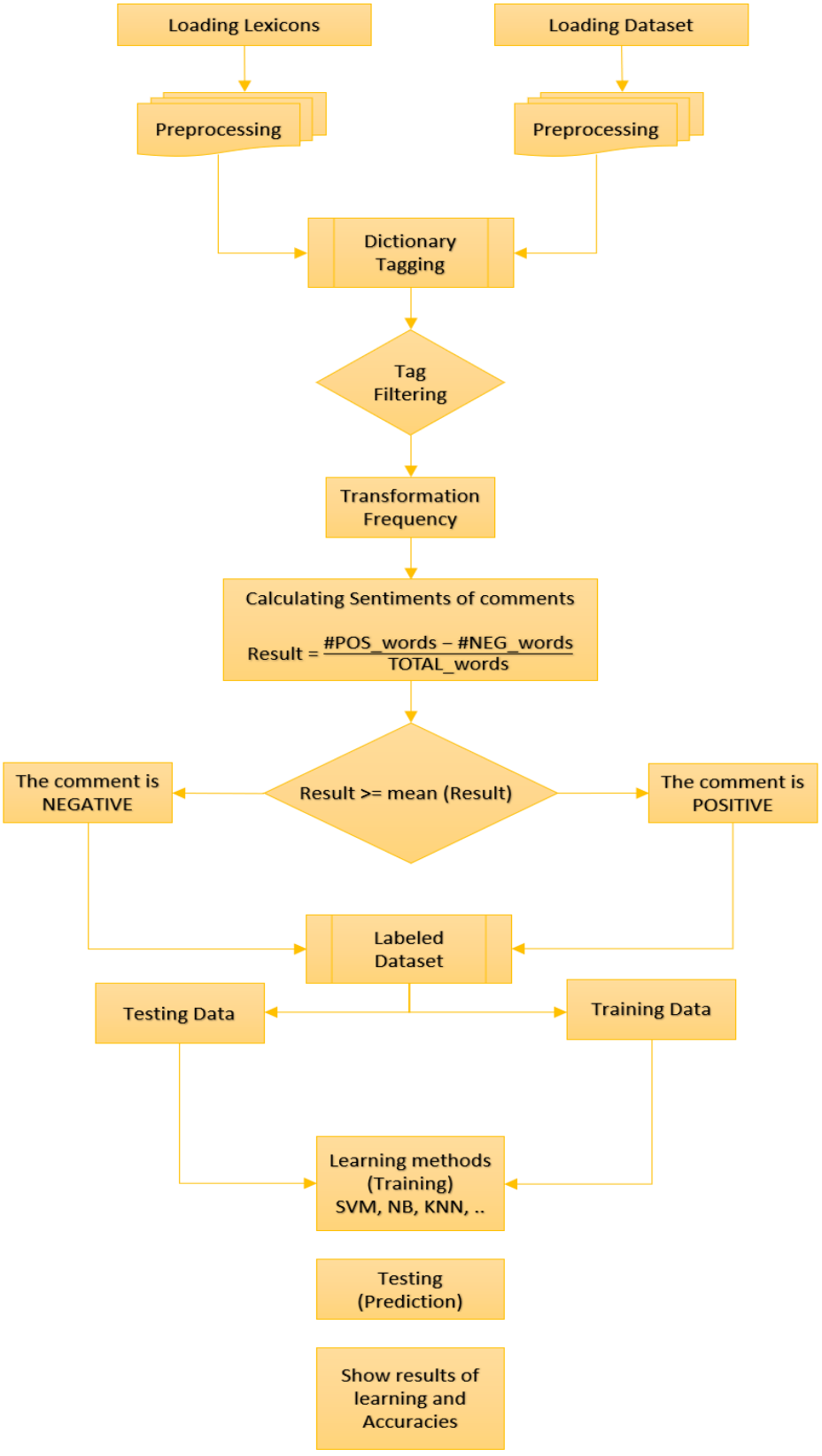


Figure 7. Methodology flow diagram.

5. RESULTS

This paper has addressed SA in Arabic comments. First, we implemented the lexicon-based approach for identifying the polarity of the provided text. Lexicons were two types of pure Arabic lexicons and refined and filtered English translated lexicons. The data samples are from local Jordanian people

commenting on a public issue related to the services provided by the main telecommunication companies in Jordan. Second, we used the resulting labeled dataset frequently used ML algorithms for classification of comments in the absence of lexicons. The workflow in Figure 7 shows the whole process starting with importing the data through labelling it and ending with classification and results in SA. The procedure involves applying a user-defined lexicon based on the common Facebook posts and comments used by Jordanians, which resulted in a (60%) positive comments and (40%) negative ones. The total accuracy of lexicon-based labeling was calculated through a comparison between the achieved results and the ones achieved through manually labeled comments by experts. The general accuracy of lexicon-based labeling was (98%). Higher accuracy values can be accomplished by adding more words and phrases to our lexicons.

Figure 8 shows a sample of the results. The SA column (F) is automatically produced using the KNIME Software nodes (Tag Filter node in Figure 5). Researchers may see the partial results comment by comment with their sentiments using (Document view node).

	A	B	C	D	E	F
1	comment	pos words	neg words	total words	result	sentiment
2	""امنيه وخط متميز نت واتصالات عل كل الشبكات مرتاح كثير""	4	0	4	1	POS
3	""امنيتى ان يحيا اطفال العالم والمسلمين وخاصه ابناء غزه بسلام""	4	0	4	1	POS
4	""امنيه - جيده""	2	0	2	1	POS
5	""امنيه اتوقع اسوا خط بالعالم""	0	2	2	-1	NEG
6	""امنيه احسن عروض""	2	0	2	1	POS
7	""امنيه النت ضعيف كثير""	0	2	2	-1	NEG
8	""امنيه النت فيه بمنطقتى شبه عدم مابستخدمه نهائى بستخدم النت الارضي""	0	4	4	-1	NEG
9	""امنيه النت كتير سيئ بس يخلص الاشتراك ما برجعلهم""	0	4	4	-1	NEG
10	""انه لشيء جميل الاشتراك بخدمات زين المميزه""	2	0	2	1	POS
11	""اورانج ،،، اسوا نت فايبر""	0	2	2	-1	NEG
12	""اورانج الخدمه سيئه جدا وكتير وقدمنا شكوى اكتر من مره وما تغير عليها شي""	4	4	8	0	NEG
13	""اورانج الشبكه ممتازه و الزين ممتازه كمان""	4	0	4	1	POS
14	""اورانج النت بطيء""	0	2	2	-1	NEG
15	""اورانج بس النت قليل""	0	2	2	-1	NEG
16	""اورانج بس سيئه جدا من ناحيه حزم البيانات سرقه عينك عينك مشكووووك فيها هالشركه""	2	4	6	-0.33	NEG
17	""اورانج كتير منيح خدمات وعروض واشتراكات فواتير شهريه ونطاق النت سريع""	4	0	4	1	POS
18	""اورانج كثير كويس""	2	0	2	1	POS
19	""اورانج كويسه بشكل عام""	4	0	4	1	POS
20	""بس النت ضعيف جدا جدا جدا""	0	2	2	-1	NEG
21	""بس لو النت ما يخلص بسرعه""	2	2	4	0	NEG
22	""زين انت مثل الكحل للعين""	2	0	2	1	POS
23	""زين بعض الاحيان تكون ضعيفه وخاصه شبكه النت""	0	2	2	-1	NEG
24	""زين خدماتهم حلوه بس بالنت الويرلس فاشل""	2	2	4	0	NEG
25	""زين شايفته كويس من ناحيه التغطيه وسرعه النت لكن بحسهم نصابين بحزم البيا	4	0	4	1	POS

Figure 8. Sample results of labeling.

1	Comments	Pos-Words	Neg-Words	Total Words	Results	Sentiment
2	اقسم بالله لولا اشتراكي السنوي ولا الان بفصل اشتراكي بس المشكلة اني دفعت لسنة كاملة اقرف من هيك خدمة ما في والنت بطيء وسيء والحقيقة اني خلص اكتفيت منهم	2	10	12	-0.66666667	?
3	الخدمات ممتازة عروض حلوة ومفيدة تكلفة ما شاء الله مريحة مع اورانج	10	0	10	1	?
4	النت قوي وسريع لكن بساعات المساء تضعف قليلا	8	2	10	0.6	?
5	انا بشكر شركة زين لانها تحفظ حق المستخدم فعلا	6	0	6	1	?
6	افضل شركة هي امنية	2	0	2	1	?
7	امنية والله العكس تماما بطيء جدا تغطية معدومة	0	8	8	-1	?
8	امنية عديمة الجدوى على الفاضي بس شغل سحب مصاري الله لا يربحهم بشيء	6	2	8	0.5	?
9	8 اشهر وانا انتظر عرض ينفع والان لقيت العرض 1000 جيجا ب12 دينار والله مفخرة	4	0	4	1	?
10	كويس جدا خط اورانج يعطي نت 5 جيجا استخدام شخصي مريح جدا والسرعة ممتازة	6	0	6	1	?
11	اتمنى انكم تعملو تحديث للعروض يعني هيك اشي ب3 دنانير يكون 15 جيجا مثلا ويكون يعمل مشاركة اهم اشي واشكركم مقدما اقتراح بس	4	0	4	1	?

Figure 9. Sample of unlabeled comments.

In Figure 8 and by referring to Figure 5, column A is the comment, columns B and C are sentiments that have resulted from the (TF node). This data is grouped (Group By) and columns are pivoted (Pivoting) to produce column D (total sentiments). The (Math formula node) is calculating the result (E) and finally with (Rule Engine node) a new column is added (F), which is our targeted LABELING

(Sentiment of the comment: POS or NEG). Figure 9 shows a sample of unlabeled comments provided to the lexicon-based model, whereas the labeling results (SA) are shown in Figure 10.

	Comments	Pos-Words	Neg-Words	Total Words	Results	Sentiment
2	اقسم بالله لولا اشتراكي السنوي ولا الان بفصل اشتراكي بس المشكلة اني دفعت لسنة كاملة اقرف من هيك خدمة ما في والنت بطيء وسيء والحقيقة اني خلص اكتفيت منهم	2	10	12	-0.66666667	NEG
3	الخدمات ممتازة عروض حلوة ومفيدة تكلفة ما شاء الله مريحة مع اورانج	10	0	10	1	POS
4	النت قوي وسريع لكن بساعات المساء تضعف قليلا	8	2	10	0.6	NEG
5	انا بشكر شركة زين لانها تحفظ حق المستخدم فعلا	6	0	6	1	POS
6	افضل شركة هي امنية	2	0	2	1	POS
7	امنية والله العكس تماما بطيء جدا تغطية معدومة	0	8	8	-1	NEG
8	امنية عديمة الجدوى على الفاضي بس شغل سحب مصاري الله لا يربحهم بشيء	6	2	8	0.5	NEG
9	8 اشهر وانا انتظر عرض ينفع والان لقيت العرض 1000 جيجا ب12 دينار والله مفخرة	4	0	4	1	POS
10	كويس جدا خط اورانج يعطي نت 5 جيجا استخدام شخصي مريح جدا والسرعة ممتازة	6	0	6	1	POS
11	اتمنى انكم تعملو تحديث للعروض يعني هيك اشي ب3 دنانير يكون 15 جيجا مثلا ويكون يعمل مشاركة اهم اشي واشكركم مقدما اقتراح بس	4	0	4	1	POS

Figure 10. Unlabeled comments after labeling.

About 1300 comments were collected from Facebook pages and provided to the lexicon model for labeling. The accuracy achieved was 98% based on some experts and by expressing some comparisons with other labeled comments. Unfortunately, this model is still restricted to the availability of the words or phrases in the lexicons. It is considered as unsupervised learning that depends on a mathematical counting formula.

On the other hand, researchers can use the resulting labeled dataset to build an ML model that will efficiently classify any newly outlet comments, which is considered a supervised learning model. To perform the classification, we applied three classifiers; namely, SVM, NB and K-NN using ORANGE software tool. The model for the three famous classifiers and its details are shown in Figure 6. The accuracy results were very promising. Table 3 shows the accuracy results of those classifiers. Table 3 makes clear that all classifiers provided good results, but superiority was for the SVM classifier since it is powerful when dealing with binary classification problems.

Table 3. Classification results.

Classifier	**Accuracy**
Support Vector Machine (SVM)	**97.9 %**
k-nearest neighbor (K-NN)	**96.8 %**
Naïve Bayes (NB)	**95.6 %**

6. CONCLUSIONS

This paper has focused on SA of Facebook Arabic comments for Jordanian telecom companies. The output of our work was an Arabic Sentiment Lexicon, which comprises 333 negative phrases and 369 positive phrases. Besides, the researchers have collected 4956 negative words and 2145 positive words in addition to a largely manipulated negation applied to negative and positive words. Most of the phrases and words came from the Jordanian dialect and MSA in addition to the applicable sentiment words from the English sentiments translated into Arabic.

The researchers implemented the lexicon-based approach for identifying the polarity of each of the provided Facebook comments. Data samples are from local Jordanian people commenting on a public issue related to the services provided by the main telecommunication companies in Jordan (Zain, Orange and Umniah). The produced results regarding the evaluation of Arabic sentiment lexicon were promising. When applying the user-defined lexicon based on the common Facebook posts and comments used by Jordanians, it scored (60%) positive and (40%) negative. The general accuracy of the lexicon was (98%). The lexicon was used to label a set of Facebook comments to formulate a big dataset of unlabeled comments. Using supervised Machine Learning (ML) algorithms that are usually used in polarity classification, the researchers introduced them to our formulated dataset. The results of the classification were 97.8, 96.8 and 95.6% for Support Vector Machine (SVM), K-Nearest

Neighbour (K-NN) and Naïve Bayes (NB) classifiers, respectively. It is worthy to note that without applying Arabic language grammar rules and Arabic sentence structure, any lexicon would fail in such a task because of issues related to the Arabic language.

7. FUTURE WORK

The formulated lexicons can be improved by adding new phrases and words related to sentiments that will improve the accuracy and quantity of labeling. The paper highlights the need to have a dedicated website for uploading lexicons and datasets collected by researchers in the field of NLP which may be helpful in this context. Moreover, there are other fields in NLP that rely on the lexicon approach, which makes this work exploited in other tasks. To overcome some of the challenges of Arabic sentiment analysis, we are considering the use of recourses as SenticNet [23].

REFERENCES

[1] M. D. Devika, C. Sunitha and A. Ganesh, "Sentiment Analysis: A Comparative Study on Different Approaches," Procedia Comput. Sci., vol. 87, no. June, pp. 44–49, 2016.

[2] B. Liu, Sentiment Analysis (Introduction and Survey) and Opinion Mining, [Online], Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.244.9480&rep=rep1&type=pdf, 2012.

[3] R. K. Bakshi, N. Kaur, R. Kaur and G. Kaur, "Opinion Mining and Sentiment Analysis," Proc. of the IEEE 3rd Int. Conf. Comput. Sustain. Glob. Dev. (INDIACom), vol. 2, no. 1, pp. 452–455, 2016.

[4] A. Chandra Pandey, D. Singh Rajpoot and M. Saraswat, "Twitter Sentiment Analysis Using Hybrid Cuckoo Search Method," Inf. Process. Manag., vol. 53, no. 4, pp. 764–779, 2017.

[5] R. M. Duwairi and I. Qarqaz, "Arabic Sentiment Analysis Using Supervised Classification," Proc. of the 1st International Workshop on Social Networks Analysis, Management and Security (SNAMS - 2014), pp. 579–583, Barcelona, Spain, August 2014.

[6] R. M. Duwairi, "Sentiment Analysis for Dialectical Arabic," Proc. of the 6th Int. Conf. Inf. Commun. Syst. (ICICS), DOI: 10.1109/IACS.2015.7103221, pp. 166–170, Amman, Jordan, 2015.

[7] A. Al Sallab, H. Hajj, G. Badaro, R. Baly, W. El Hajj and K. Bashir Shaban, "Deep Learning Models for Sentiment Analysis in Arabic," Proceedings of the Second Workshop on Arabic Natural Language Processing, DOI: 10.18653/v1/W15-3202, pp. 9–17, Beijing, China, 2015.

[8] H. K. Aldayel and A. M. Azmi, "Arabic Tweets Sentiment Analysis - A Hybrid Scheme," J. Inf. Sci., vol. 42, no. 6, pp. 782–797, 2015.

[9] H. Mulki, H. Haddad, M. Gridach and I. Babaoğlu, "Tw-StAR at SemEval-2017 Task 4: Sentiment Classification of Arabic Tweets," Proc. of the 11th International Workshop on Semantic Evaluations (SemEval-2017), DOI: 10.18653/v1/S17-2110, pp. 664–669, Vancouver, Canada, 2018.

[10] N. Al-Twairesh, H. Al-Khalifa, A. Al-Salman and Y. Al-Ohali, "AraSenTi-Tweet: A Corpus for Arabic Sentiment Analysis of Saudi Tweets," Procedia Comput. Sci., vol. 117, pp. 63–72, 2017.

[11] H. Najadat, A. Al-Abdi and Y. Sayaheen, "Model-based Sentiment Analysis of Customer Satisfaction for the Jordanian Telecommunication Companies," Proc. of the 9th International Conference on Information and Communication Systems (ICICS 2018), DOI: 10.1109/IACS.2018.8355429, 2018.

[12] L. Lulu and A. Elnagar, "Automatic Arabic Dialect Classification Using Deep Learning Models," Procedia Comput. Sci., vol. 142, pp. 262–269, 2018.

[13] A. Soumeur, M. Mokdadi, A. Guessoum and A. Daoud, "Sentiment Analysis of Users on Social Networks: Overcoming the Challenge of the Loose Usages of the Algerian Dialect," Procedia Computer Science, vol. 142, pp. 26-37, 2018.

[14] J. O. Atoum and M. Nouman, "Sentiment Analysis of Arabic Jordanian Dialect Tweets," Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 2, pp. 256–262, 2019.

[15] M. Bouaziz and T. Ohtsuki, "Multi-class Sentiment Analysis in Twitter: What If Classification Is Not the Answer," IEEE Access, vol. 6, pp. 64486-64502, 2018.

[16] S. Kiritchenko, S. M. Mohammad and M. Salameh, "SemEval-2016 Task 7: Determining Sentiment Intensity of English and Arabic Phrases," Proc. of the 10th International Workshop on Semantic Evaluation (SemEval-2016), DOI: 10.18653/v1/S16-1004, San Diego, California, 2016.

[17] A. Al-Thubaity, Q. Alqahtani and A. Aljandal, "Sentiment Lexicon for Sentiment Analysis of Saudi Dialect Tweets," Procedia Computer Science, vol. 142, pp. 301-307, 2018.

[18] R. M. Duwairi and M. El-Orfali, "A Study of the Effects of Pre-processing Strategies on Sentiment Analysis for Arabic Text," Journal of Information Science, vol. 40, no. 4, pp. 501–513, 2014.

[19] M. Mataoui, O. Zelmati and M. Boumechache, "A Proposed Lexicon-based Sentiment Analysis Approach for the Vernacular Algerian Arabic," Res. Comput. Sci., vol. 110, pp. 55-70, 2016.

[20] S. L. Lo, E. Cambria, R. Chiong et al., "Multilingual Sentiment Analysis: From Formal to Informal and Scarce Resource Languages," Artif. Intell. Rev., vol. 48, pp. 499–527, 2017.

[21] O. Oueslati, A. I. S. Khalil and H. Ounelli, "Sentiment Analysis for Helpful Reviews Prediction," International Journal of Advanced Trends in Computer Science and Engineering, vol. 7, no. 3, pp. 34-40, DOI: 10.30534/ijatcse/2018/02732018, 2018.

[22] E. Cambria, "Affective Computing and Sentiment Analysis," IEEE Intelligent Systems, vol. 31, no. 2, pp. 102-107, DOI: 10.1109/MIS.2016.31, 2016.

[23] E. Cambria, S. Poria, D. Hazarika and K. Kwok, "SenticNet 5: Discovering Conceptual Primitives for Sentiment Analysis by Means of Context Embeddings," Proc. of the 32nd AAAI Conference on Artificial Intelligence (AAAI-18), pp. 1795–1802, New Orleans, Louisiana, USA, 2018.

[24] I. Al-Agha and O. Abu-Dahrooj, "Multi-level Analysis of Political Sentiments Using Twitter Data: A Case Study of the Palestinian-Israeli Conflict," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 5, no. 3, December 2019.

[25] S. Kumar, V. Koolwal and K. K. Mohbey, "Sentiment Analysis of Electronic Product Tweets Using Big Data Framework," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 5, no. 1, April 2019.

[26] A. Al-Shaikh, R. Al-Sayyed and A. Sleit, "A Case Study for Evaluating Facebook Pages with Respect to Arab Mainstream News Media," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 3, no. 3, December 2017.

[27] A. Bouziane, D. Bouchiha, N. Doumi and M. Malki, "Toward an Arabic Question Answering System over Linked Data," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 4, no. 2, August 2018.

[28] E. Al-Shawakfa, "A Rule-based Approach to Understand Questions in Arabic Question Answering," Jordanian Journal of Computers and Information Technology (JJCIT), vol. 2, no. 3, Dec. 2016.

[29] Oueslati et al., "A Review of Sentiment Analysis Research in Arabic Language," Future Generation Computing Systems, vol. 112, pp. 408-430, 2020.

[30] K. Nahar, N. Alhindawi, O. Al-Hazaimeh, R. M. Al-Khatib and A. Al-Akhras, "NLP and IR-based Solution for Confirming Classification of Research Papers," Journal of Theoretical and Applied Information Technology, vol. 96, pp. 5269-5279, 2018.

[31] M. Alshanaq, K. Nahar and K. Halawani, "AQAS: Arabic Question Answering System Based on SVM, SVD and LSI," Journal of Theoretical and Applied Information Technology, vol. 97, pp. 681-691, 2019.

[32] K. Nahar, M. Alshanaq, A. Manasrah, R. Alshorman and I. Alazzam, "A Holy Quran Reader/Reciter Identification System Using Support Vector Machine," International Journal of Machine Learning and Computing, vol. 9, pp. 458-464, 2019.

[33] K. Nahar, A. Al Eroud, M. Barahoush and A. Al-Akhras, "SAP: Standard Arabic Profiling Toolset for Textual Analysis," International Journal of Machine Learning and Computing, vol. 9, pp. 222-229, 2019.

[34] K. Nahar, R. Al-Khatib, M. Alshanaq, M. Daradkeh and R. Malkawi, "Direct Text Classifier for Thematic Arabic Discourse Documents," International Arab Journal of Information Technology, vol. 17, (Published Online), 2020.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 06, No. 03, September 2020.

ملخص البحث:

تحليل الآراء إحدى التقنيات التي تستخدم لتحديد القطبية (موجب، سالب) لنصّ معين باستخدام تقنيات معالجة اللغات الطبيعية. ويُعدّ فيسبوك مثالاً على منصّات التواصل الاجتماعي الأوسع استخداماً بين المواطنين والمقيمين في الأردن للتعبير عن آرائهم بخصوص قضية عامة محوريّة.

في هذه الورقة، تم استخدام منهجية قائمة على المعاجم من أجل تحديد قطبية عدد من التعليقات المنشورة على فيسبوك. البيانات الخاضعة للتحليل هي لأشخاص أردنيين يعلّقون على قضية عامّة تتعلق بالخدمات التي تُقدمها شركات الاتصالات الرئيسية العاملة في الأردن (زين، وأورانج، وأمنية). وكانت النتائج المتعلقة بتقييم معجم الآراء باللغة العربية واعدة. وبتطبيق المعجم القائم على البوستات والتعليقات المنشورة على فيسبوك التي يعبر بها المستخدمون الأردنيون عن آرائهم، مثلث التعليقات الإيجابية ما نسبته 60% والتعليقات السلبية ما نسبته 40%، وكانت النسبة العامة لدقة المعجم 98%.

وقد استخدم المعجم لوسْم مجموعة من التعليقات غير الموسْومة على فيسبوك لتشكيل مجموعة بيانات كبيرة. وباستخدام عدد من خوارزميات تعلُّم الآلة التي يشيع استخدامها لتحديد القطبية، تم تطبيق هذه الخوارزميات على مجموعة البيانات التي جرى تشكيلها في هذه الدراسة. وكانت نتائج التصنيف 97.8%، و96.8%، و95.6% لمصنِّفات (SVM، وK-NN، وNB) على الترتيب.