A Sentiment Analysis of Egypt’s New Real Estate Registration Law on Facebook

Abdulfattah Omar
Department of English
College of Science and Humanities
Prince Sattam Bin Abdulaziz University, Saudi Arabia

Wafya Ibrahim Hamouda
Department of Foreign Languages
Faculty of Education
Tanta University
Tanta, Egypt

Abstract—In response to the increasing influence of social media networks on shaping the public opinion, sentiment analysis systems and applications have been developed to extract insights and gain an overview of the wider public opinion behind certain topics so as to support businesses, manufacturers, government agencies, and policymakers with their decisions and plans. Despite the importance of sentiment analysis in providing policymakers with effective mechanisms to understand the attitudes of customers and citizens which can be usefully used in decision-making processes and planning for the future, so far studies on sentiment analysis are very limited in Egypt. Much of the work is still done using survey tools such as questionnaires and polls to gather information about the citizens’ attitudes towards given issues and topics. Despite the effectiveness of such methods, citizens’ reflections on social media platforms and networks remain more powerful in providing comprehensive insights and overviews. Furthermore, social media-based sentiment analysis is usually more representative being based on larger numbers of participants, which has positive implications to reliability. Opinions expressed on social media are often the most powerful forms of feedback for businesses because they are given unsolicited. In light of this argument, this study seeks to provide a sentiment analysis of Egypt’s New Real Estate Registration Law on Facebook. To extract information about the users’ sentiment polarity (positive, neutral or negative), Facebook posts were used. The rationale is that Facebook is still the most popular social media platform in Egypt. Text classification was then used for classifying the selected data into three main classes/values: Positive, Negative, and Neutral. The findings indicate that sentiments expressed in the users’ posts and comments have a significant negative attitude towards the new law. Despite the effectiveness of the automatic evaluation and analysis of the sentiments and opinions of the users in social media concerning the new Real Estate Registration Law, linguistic approaches including Critical Discourse Analysis (CDA), functional linguistics, and semiotics need to be incorporated into sentiment analysis applications for gaining a better understanding of people’s attitudes towards specific issues.

Keywords—Egypt; Facebook; opinion; real estate registration law; sentiment analysis; social media

I. INTRODUCTION

The role of social media platforms and networks has grown drastically over the recent years. Today, social media platforms and networks are not only accessible and easy venues and channels for communication between individuals across the globe, they are among the most important tools for influencing the public opinion and decision making processes [1-3]. This is obviously reflected in the influential role social media played in the social and political movements in different countries in the world [4, 5].

In the Arab world, social media platforms and networks played a vital role during the so-called ‘Arab Spring events’ in political mobilization by calling for demonstrations and protests, publishing news and videos, expressing opinions, and political debate liberated from the authority’s control over the traditional media [6]. This role was definitely supported by the potentials of social media of reaching millions of users through very high interactive features within a very short period of time and forming discussion forums with large numbers of participants.

In recent years, there are many examples where social media influenced public opinion through forming a unified public opinion on specific issues as a result of the interaction between its users who belong to different cultures, but who believe in a common system of values. In 2014, for instance, the Canadians expressed their anger against Expedia Canada television advertisement on the social media networks especially Facebook and Twitter. The “Escape Winter: Fear” advertisement caused a stir on social media when the frequency of which it aired proved to be too much for some viewers [7]. Viewers from different social and cultural backgrounds placed tormented posts on Facebook and Twitter which was described as a public negative attitude towards the company.

Given the increasing importance of social media in public opinion, sentiment analysis has been developed in order to extract insights and gain an overview of the wider public opinion behind certain topics. In marketing, sentiment analysis is widely used to help brands, businesses and manufacturers address and handle customer anger, protect the brand image, and evaluate the effectiveness of their marketing campaigns [8-10]. In this sense, sentiment analysis has been one of the most common and fundamental measures of customers’ attitudes toward a brand by using different variables including language constructs, emotion, and so on. Manufacturers, businesses, and marketing agencies use sentiment analysis to learn more about how people feel about their business and goods, as well as to assess customer loyalty [11-14].
Recently, sentiment analysis has been widely used by government agencies and policy makers. Prior to the 2012 presidential election, for instance, the Obama administration used opinion analysis to gauge public response to policy announcements and campaign messages. The underlying principle is that understanding and interpreting citizens’ concerns in real-time can be usefully used in improving public services, building well-organized governments, and creating predictive models to anticipate bottlenecks in various public services [15, 16]. According to Arunachalam and Sarkar [17], sentiment analysis of the citizens’ posts on social media networks is best considered as the new eye of the government.

Despite the importance of sentiment analysis in providing policymakers with effective mechanisms to understand the attitudes of customers and citizens which can be usefully used in decision-making processes and planning for the future, so far studies on sentiment analysis are very limited in Egypt. Much of the work is still done using survey tools such as questionnaires and polls to gather information about the citizens’ attitudes towards given issues and topics. Despite the effectiveness of such methods, citizens’ reflections on social media platforms and networks remain more powerful in providing comprehensive insights and overviews. Furthermore, social media-based sentiment analysis is usually more representative being based on larger numbers of participants, which has positive implications to reliability. Opinions expressed on social media are often the most powerful forms of feedback for businesses because they are given unsolicited. In light of this argument, this study seeks to provide a sentiment analysis of Egypt’s New Real Estate Registration Law on Facebook.

In August 2020, the House of Representatives in Egypt (the Egyptian Parliament) approved a draft law submitted by the Government regarding the amendment of some provisions of Law No. 114 of 1946 regulating the real estate registry in all of its articles, by adding an updated Article No. 35 (bis) in order to protect the real estate wealth and its rights, according to the government officials [18]. However, the amendments meant stopping the introduction of utilities for all unregistered buildings as of March 6, 2021. By the law, electricity, water and gas companies and other companies, agencies, Ministries and government departments are not allowed to transfer facilities and services to properties unless they are officially registered in one of the registry offices. This amendment was published in the Official Gazette on September 5, 2020, to come into effect in six months, i.e. on March 6, 2021.

On September 9, 2020 the Prime Minister indicated that the government is currently working on setting up a digital system to create a certificate and a national number for each apartment and real estate, and apartments will only be dealt with. On February 22, 2021, the Minister of Justice confirmed that all government agencies will not deal with unregistered real estate, meaning that facilities will not be connected to such unregistered units starting from March 6, 2021 [19].

As a result, there was a lot of discussion, controversy, and uncertainty on social media networks in Egypt. Some considered the issuance of the law in its new form an achievement that makes each apartment and real estate a national number, and accurately determines the size of the country’s real estate wealth, as well as gives real estate an actual and substantial value. Others called on the government to review the controversial law and simplify the procedures for registering housing units [20].

This paper discusses the controversy that took place on social media networks with the enforcement of the new amendment on March 6, 2021. It seeks to address the following research questions:

1) What were the sentiments, opinions, and attitudes of the Egyptians towards the new Real Estate Registration Law as expressed on their posts, comments, and replies on Facebook?

2) What is the effectiveness of automated sentiment analysis in evaluating the opinions and responses of citizens on social media towards regulations and government policies as represented in the analysis of the sentiments, opinions, and attitudes of the Egyptians towards the new Real Estate Registration Law as expressed on their posts, comments, and replies on Facebook?

The rest of the paper is organized as follows. Section 2 is an introduction to the concept of sentiment analysis. Section 3 is a brief survey of the development of sentiment analysis systems and applications in Arabic. Section 4 describes the methods and procedures of data collection, preparation, and classification. Section 5 reports the results of the study. Section 6 is conclusion.

II. SENTIMENT ANALYSIS

Sentiment analysis is a relatively new discipline of knowledge. It emerged around the year 2000 with the development of social networks and platforms. Nevertheless, it has been very active since its inception. Sentiment analysis has been associated with different disciplines including behavior analysis, business administration, data mining, information retrieval, marketing, Natural Language Processing (NLP), text classification, and text mining, and web mining [21-24].

Apparently, the recent years have witnessed an accelerating development of sentiment analysis due to the explosive growth of the web and the advent of social media platforms and networks over the past twenty years. Sentiment analysis, also called opinion mining, Pozzi, et al. [21] argue, has come to be extensively used to evaluate people’s opinions, sentiments, appraisals, judgments, attitudes, and emotions toward products, services, organizations, individuals, events, issues, or topics as expressed in written text, emotions, and so on through their posts and comments on social media platforms and networks.

Iglesias and Moreno [22] assert although sentiment analysis can be applied to different data sources including surveys, opinion polls, and focus group discussions, the advent of social media systems has given sentiment analysis more value and added to its popularity. In other words, social media networks, platforms, and systems have provided sentiment analysts with rare opportunities to construct

www.ijacsat.org
organized and actionable knowledge through defining automated tools that can extract subjective information [25]. In other words, the prolific data on the social media networks, platforms, and systems have made it possible for organizations to automate the analysis of opinions expressed in digital form [24].

In recent years, different sentiment analysis systems have been developed to score and quantify the sentiments towards issues, goods, services, and so on. These have been developed using different software languages including R and Python [26]. In these model, the objective is to identify emotions, such as: happiness, fear, anger, etc. Glossaries are usually generated using machine learning algorithms to define these feelings. To put it into effect, these are lists of words with corresponding emotions associated with them. Here it should be noted that when using these glossaries, the problem of different emotions conveyed by words appears, especially that humans can express their emotions in different ways.

III. RELATED WORK

Arabic sentiment analysis was first introduced in 2008. Progress on Arabic sentiment analysis research was very slow. During this early period which extended until approximately 2015, very few studies were done. This can be attributed to the idea that the main bulk of research in sentiment analysis was traditionally focused on English. Very few studies were done on other languages due to the lack of resources that support the analysis of sentiments in other domains [27, 28].

During this early or first stage, Arabic sentiment analysis studies were concerned with developing sentiment analysis systems taking into consideration the peculiar linguistic features of Arabic. In order to support sentiment analysis applications in Arabic, Rushdi-Saleh, et al. [27] built a corpus of 500 movie reviews collected from different web pages and blogs in Arabic. Similarly, Abdul-Mageed, et al. [29] attempted to address the linguistic challenges that are always associated with sentiment analysis through developing a sentiment analysis system based on sentence-level that takes into consideration the morphological system of Arabic. The proposed model was able to achieve higher levels of performance.

To bridge the gap between sentiment analysis theory and applications, Elhawary and Elfeky [30] developed an Arabic sentiment analysis system that enables businesses and corporations in the Middle East to better understand the experiences of their Arab customers in the region. The proposed system was largely based on lexicon-based methods and it showed good performance compared to the English sentiment analysis. Despite the relative success of these studies during this early stage to introduce sentiment analysis in Arabic for the first time, research and applications in Arabic sentiment analysis were very limited.

The rise of the social media in the recent years in the Arab world, however, has fueled interest in sentiment analysis [31]. In this regard, there has been a relative increase in sentiment analysis studies over the recent years. Boudad, et al. [32] assert that sentiment Analysis in Arabic has grown drastically in the last five or six years. Despite these recent developments and, the plentiful data resources on social media platforms and networks, and the growth of Arabic text content on the web, Arabic sentiment analysis is still facing serious challenges in terms of available tools and reliable systems.

In their evaluation of sentiment analysis in Arabic, Elghazaly, et al. [33] argue that sentiment analysis in Arabic is still very challenging for different reasons, including the fact that Arabic is not a case-sensitive language and does not use capital letters. Furthermore, Arabic is a highly inflected language and has a very flexible morphological system that make it very difficult for machine learning systems to perform appropriately [34-36]. Likewise, Nassif, et al. [37] stress that “despite the plentiful online Arabic content, the research on ANLP in general and sentiment analysis in particular is still suffering from the lack of tools and resources available as compared to English”. They attribute the problem mainly to the morphologically complex system of Arabic which has negative impacts on the functionality of different NLP systems including sentiment analysis.

These challenges have been clearly reflected on the sentiment analysis applications in marketing, product analysis, social media monitoring, brand reputation management, decision-making processes, and measuring public opinion in the Arab countries. This study seeks to bridge this gap in the literature through evaluating the sentiments of the Egyptian citizens on social media towards regulations and government policies as represented in the analysis of the sentiments, opinions, and attitudes of the Egyptians towards the new Real Estate Registration Law as expressed on their posts, comments, and replies on Facebook.

IV. METHODS AND PROCEDURES

To extract information about the users’ sentiment polarity (positive, neutral or negative), Facebook posts were used. The rationale is that Facebook is the most popular social media platform in Egypt according to recent statistics, as shown in Figure 1.

![Figure 1: Social Media Stats in Egypt - February 2021](image-url)
According to a recent report released by Facebook, Egypt had 50,220,000 Facebook users in February 2021, accounting for 47.4 percent of the country's total population. The majority of the users (62.4 percent) were male users. The largest user group was between the ages of 25 and 34. (15,800,000). Men lead by 4,200,000 users aged 25 to 34, which is the largest gap between men and women, as shown in Figure 2.

This large user base combined with a significant event such as the enforcement of a new law that has direct impacts on the life of millions of citizens makes Facebook data ideal to understand the attitudes of the Egyptian citizens towards he issue. The Arabic key terms ‘التمهيد العقاري’ translated as ‘registry office’ and ‘تسجيل الشهر العقاري’ translated as ‘real estate registration’ were searched to gather relevant data over the period February 23-March 04, 2021.

The decision to use keywords rather than hashtags was based on the premise that hashtags are usually used by more experienced users who are at least somewhat familiar with the idea of hashtags and therefore have more experience than other users. We attempted to make our data sample as inclusive and representative as possible for generalizability and reliability purposes.

A corpus of 16478 documents was built. These were collected from the posts, comments on the posts, and replies to the posts on the real estate registration law. Picture posts were also transcribed and included. All texts were then tokenized and normalized for addressing variation in text length. Stemming, which is a normal procedure in standard classification systems, was not executed to make use of the linguistic richness of the Arabic derivational and inflectional morphemes [38]. A matrix was then generated including all the 16478 documents. The matrix included 60,968 words/variables.

One major problem with the corpus, however, was the high dimensionality of the data. This is referred to in text classification applications as the curse of dimensionality [39-42]. To address the problem, two steps were taken. First, a stop-word list with all the function words including articles, prepositions, and pronouns, was executed. Second, term weighting was carried out using Term Frequency Inverse Document Frequency (TF-IDF) to keep only the most distinctive variables within the data matrix. This is shown as follows.

In document classification applications, function words are classified irrelevant and noisy terms since they do not carry lexical significance. Accordingly, content/lexical word indexing is always recommended as an effective approach for a reliable text classification performance [43]. The premise is that the inclusion of such irrelevant variables is useless and has adverse impacts on the accuracy and reliability of classification applications [44]. It is a standard practice in text classification applications to remove function words using stop-word lists so that only bearing content words are only retained [45]. The default method for removing function words is a linguistic one. This is usually executed through the implementation of a stop-word list where all function words are identified and removed. In our case, the removal of function words through the execution of stop-word lists had the effect of reducing the matrix into 38,793 variables.

A TF-IDF analysis was then carried out to identify the most distinctive variables within the data collection, as shown in Figure 3. Based on the TF-IDF analysis, only the highest 200 variables with TF-IDF values were retained.

Text classification was finally used for classifying the selected data into three main classes/values: Positive, Negative, and Neutral. Text classification can be simply defined as the task of automatically sorting a set of documents into a number of classes or categories where each is given a label [46, 47]. Classification relies on priori reference structures that divide the space of all possible data points into a set of classes that are usually, but not necessarily, nonoverlapping [48]. A text classification task starts by discovering and finding groups that have similar content then organizing our perceptions of these groups into categories. In other words, clustering places documents into natural classes while classification places them into predefined known ones [49].
The classification of the data was carried out using the lexical-based model developed by Tabodada et al. in 2011. In this model, sentiment-bearing words (including adjectives, verbs, nouns, and adverbs), are extracted and used to calculate the semantic orientation of each document/post [50]. In our case, the classification of the data into the three classes: positive, negative, or neutral sentiment is achieved through a three-level analysis, as shown in Figure 4.

At the document level, also known as message level, the classifier or classification system categorizes the polarity of the post as a collective body. For example, given a product review, the system decides whether the text message is overall positive, negative, or neutral. It is assumed that the message/document conveys only one point of view on a single topic. At the sentence level, the classifier or classification system determines the polarity of each sentence contained in a post. The assumption is that each sentence, in a given post, denotes a single opinion on a single entity. Finally, entity and aspect level analysis is more comprehensive than message and sentence level analysis. It is based on the concept that an opinion is made up of two parts: a sentiment and a target of opinion [51].

As a final step, polarized words were counted. It was obvious that the number of negative word appearances is much greater than the number of positive word appearances, as shown in Figure 5.

V. RESULTS AND DISCUSSIONS

The results indicate that the vast majority of the users (around 80%) had negative sentiments towards the new law. This is clearly reflected in the negative words including No and reject they used in reference to the law. Many of the posts referred to the law as illogical. They explained that it is illogical during the world financial crisis caused by COVID-19 to impose additional fees on the citizens. Many of the negative sentiments focused on the concept that the new law imposes heavy financial burdens and requires several bureaucratic steps. Furthermore, many posts also rejected any link between the new law and the access to the very basic facilities including gas, water, and electricity.

Positive sentiments towards the amended law, in turn, stressed that the new law regulates and protects individual ownership. It also addresses many of the inherent limitations within the old law. These implied different arguments such as “About 95 per cent of Egypt’s real estate wealth is unregistered, and that the government hopes that all citizens register their real estate units, with the objective of securing their properties”, “Registering real estate properties in registry offices helps in securing the citizens’ ownership rights, and also leads to abolishing the bad phenomenon of illegal construction and slum areas”, and “Registration contributes to increasing the value of properties”. These are shown in Fig. 6.
Referring to the research questions, it can be claimed that automatic sentiment analysis was usefully used in evaluating the sentiments of the Egyptians towards the new Real Estate Registration Law as expressed on their posts, comments, and replies on Facebook. Through the classification of the sentiments into Positive, Negative, and Neutral, it was obvious that the vast majority of the citizens were against the law. These findings represent good opportunities for policymakers and legislators to know more about people’s sentiments, attitudes, and perceptions.

Despite the effectiveness of automatic sentiment analysis in providing a general idea about the perceptions and sentiments of citizens towards the issue, this automatic approach cannot deal effectively with all the data sorts collected for the purpose of the study. It was obvious that many of the posts implied judgments, arguments, agreements, disagreements, and even political stances as reflected on the declarations by the Minister of Justice on the issue who accused opposing movements of using the issue for their own agendas, as shown in Figure 7.

It was also obvious that many of the negative posts on the issue were characterized by humor. The comics below are based on Egyptian movies where users made some funny scripts on the issue, as shown in Figure 8.

In this sense, it is challenging for automatic approaches of sentiment analysis to define accurately the sentiments of the users or individuals towards given issues.

Elghazaly, et al. [33] agree that it is still a challenging task to perform sentiment analysis appropriately in Arabic due to the fact that automatic sentiment analysis is largely context-free which is not appropriate for Arabic where the use of words and expressions is very subjective. They explain that the word "(a spare tire)" was used as negative reference to former Egyptian President Mohammed Morsi during the 2012 Presidential elections. The absence of context in sentiment analysis thus poses serious challenges for the performance of sentiment analysis systems. Nassif, et al. [37] agree that there is a need for more efforts to address the subjective data in Arabic sentiment analysis applications.
Given that automatic approaches to sentiment analysis do not consider context, it becomes important to incorporate more sophisticated language approaches including functional linguistics, Critical Discourse Analysis (CDA), and semiotics into sentiment analysis applications.

VI. CONCLUSION

This study adopted Sentiment analysis methods to evaluate the sentiments, opinions, and attitudes of the Egyptians towards the new Real Estate Registration Law as expressed on their posts, comments, and replies on Facebook. It was obvious that the majority of the citizens had negative sentiments and attitudes towards the new law. Despite the effectiveness of the automatic sentiment analysis tools of the sentiments of social media users towards the issue, these automatic tools cannot effectively evaluate many posts including judgments, arguments, and irony. In this regard, it is important to further understand the contextual aspects of the data. Although this study was limited to the evaluation of the users’ opinions towards the new Real Estate Registration Law, the results have positive implications to the sentiment analysis applications for obtaining a general idea about the perceptions of citizens concerning specific issues.

ACKNOWLEDGMENT

We take this opportunity to thank Prince Sattam Bin Abdulaziz University in Saudi Arabia alongside its Scientific Deanship, for all technical support it has unstintingly provided to the fulfillment of the current research project.

REFERENCES


