Insights Discovery through Hidden Sentiment in Big Data: Evidence from Saudi Arabia’s Financial Sector

Young-Eun PARK¹, Yasir JAVED²

Received: April 01, 2020 Revised: April 11, 2020 Accepted: May 07, 2020

Abstract

This study aims to recognize customers’ real sentiment and then discover the data-driven insights for strategic decision-making in the financial sector of Saudi Arabia. The data was collected from the social media (Facebook and Twitter) from start till October 2018 in financial companies (NCB, Al Rajhi, and Bupa) selected in the Kingdom of Saudi Arabia according to criteria. Then, it was analyzed using a sentiment analysis, one of data mining techniques. All three companies have similar likes and followers as they serve customers as B2B and B2C companies. In addition, for Al Rajhi no negative sentiment was detected in English posts, while it can be seen that Internet penetration of both banks are higher than BUPA, rarely mentioned in few hours. This study helps to predict the overall popularity as well as the perception or real mood of people by identifying the positive and negative feelings or emotions behind customers’ social media posts or messages. This research presents meaningful insights in data-driven approaches using a specific data mining technique as a tool for corporate decision-making and forecasting. Understanding what the key issues are from customers’ perspective, it becomes possible to develop a better data-based global strategies to create a sustainable competitive advantage.

Keywords: Saudi Arabia, Financial Sector, Big Data, Social Media, Sentiment Analysis

JEL Classification Code: E37, F39, M10, M31, N25

1. Introduction

There was a time when Facebook was only used to share photos in dorm rooms, but the trend has changed with over 60 million active business pages in 2017, with a net income of around USD5.14 billion as of 2018 (Ellison & Boyd, 2013; Chen, De, Hu & Hwang, 2014; Park, Chaffar, Kim & Ko, 2017; Jiang, Muhammad & Mishal, 2018). Each individual as a consumer has now become a main entity of business and a contributor of social media data production by engaging in social media, and a firm has begun to use social media as a primary tool to build the company’s future strategy, particularly marketing strategy (Azam, 2015; Park, Allui & Alselaimi, 2017; Park & Alenezi, 2018). The participative characteristics of social media have changed the distribution of embellished resources (Baptista, Wilson, Galliers, & Bynghall, 2017) and reshaped the patterns of communication from pointed to multiple organizational environments (Huang, Baptista and Newell, 2015). In other words, social media has been essential to knowledge creation and management (Von Krogh, 2012), knowledge reuse (Majchrzak, Wagner & Yates, 2013), distribution leadership (Sutanto, Tan, Battistini, & Phang, 2011), and in facilitating interaction and internal collaboration (Razmerita Kirchner & Nabeth, 2014). Accordingly, it has emerged as a popular, powerful and convenient tool on a specific platform (Jiang, Muhammad, & Mishal, 2018).

The research motivation for this study is based on the implausible usage of social media such as a Facebook, Instagram, Twitter, YouTube, and others, which reach a significantly high level worldwide for business purposes. However, it does not mean the high level of social media usage can necessarily lead to increased productivity or revenue base (Roberto, Thierry & Offiong, 2016). This
study began with a question about potential link between social media and sustainable growth through increasing revenue base in business; furthermore, whether it is affecting the economic growth of a country or not. Even though there have been several papers looking at the relationship between social media and economic growth (Roberto et al., 2016; Park et al., 2017; Park & Alenezi, 2018), the positive and negative impacts of those relationships still remain controversial. In particular, apart from the macroscopic and broad relationships such as economic growth in a country, much research has not been carried out on whether social media influences the continuous growth, preference, and popularity of individual company.

As sustainability has gained attention in the last few decades, organizations are laying more stress on how to assess or measure sustainability or sustainable growth. Using sustainability as a metric typically means an augmentation of the traditional business reporting framework to be guided by social and environmental performance in addition to economic or financial performance as the Triple Bottom Line (Dess, Lumpkin, Eisner & McNamara, 2014). However, even though these three key performances as basic requirements should be measured, many organizations are developing in-house or specific measurement tools to find their own best practices in this regard.

While balancing the principles of triple bottom line to achieve sustainability and creating sustainable competitive advantages overall is a desired goal in an entire organization, it is not sure whether it is achievable in every project. This is the reason why we need to investigate an indicator to predict and measure the growth potential of a firm from the perspective of customers. Thus, this research provides us with meaningful and invaluable implications to suggest or develop a new barometer in order to recognize the future sustainability of the enterprises with big data from customer-side besides triple bottom line to fill the gap between them.

2. Literature Review

2.1. Financial sector and Sustainable Growth in Saudi Arabia

Saudi Arabia has the largest economy, accounting for 38% of GDP in the Middle East and North Africa (hereafter, MENA), and has the greatest economic and political impact in this region (Kavoossi, 2000; Bassam, 2015; Part et al., 2017; Park & Alenezi, 2018). Although economic growth subsided in both advanced and emerging G20 economies in 2018, according to Saudi Arabian Monetary Authority (hereafter, SAMA), overall, the Saudi economy remains quite resilient against disorder from growing uncertainty in the global economy and made a recognizable recovery in the same period, driven by the oil sector. A low level of public debt and high levels of foreign reserves have shown that there is abundant fiscal capacity to counter an economic downturn.

Domestic output grew by 2.2 percent point in 2018 compared to a decline of 0.7 percent point in the previous year. Especially, growth in the private sector remains well below the most recent period of economic expansion in Saudi Arabia (SAMA, 2019). In particular, among private sectors, finance and insurance, and manufacturing sectors grew at active rates while the construction sector continued to shrink. This is driven by both demand and supply factors as doubt prevails about the sector’s promising growth potential (see Figure 1).

Accordingly, the role of leading companies such as Aramco, SABIC and other financial institutions in the Saudi economy is becoming incredibly important in alignment with the current development goals of the Saudi Vision 2030 (Part et al., 2017; Park & Alenezi, 2018). According to Aljazira Capital in 2018, the Saudi banking sector’s balance sheet grew 2.1% point. The Saudi banking sector has 12 listed and other non-listed banks. In terms of the balance sheet size, National Commercial Bank (NCB) is the largest bank in the Kingdom with assets of more than SAR 459.1bn. It accounts for 20.4% of the total market, followed by Al Rajhi bank with asset base of SAR 356.4bn and 15.9% market share, Samba with total assets worth of SAR 228.3bn, Riyad Bank with asset base of SAR 225.9bn. It accounts for 10.2% and 10.1% of total banking assets, respectively.

Meanwhile, BUPA is the biggest player in the insurance industry with 32.7% market share in Gross Written Premiums paid the highest claims of SAR 1.83bn, followed by Tawuniya at 13.9%, MEDGULF at 9.7%, Al Rajhi Takaful at 8.1%, AXA-Cooperative at 4.7%, and others at 30.9% as reported by AlJazira Capital 2019 shown in Figure 2.
The year 2018 witnessed favorable developments in the asset composition of the banking sector in Saudi Arabia. Based on that, the role of Saudi Government became more important like that the government incentives of the international players can play in achieving balanced socio-economic growth, diversity and sustainability. Accordingly, many studies have looked at the Saudi economy or leading companies to figure out the main, critical success factors, but most of them remain qualitative research using the methods such as case studies in the specific contexts, not empirical or data-driven studies through numeric data. Moreover, there is little business research on the Saudi financial sector. With those limitations, this study began to cover the gap.

2.2. Predictive Data Mining with Social Media

Forecasting refers to the process of making predictions for the future based on past and present historical data and most commonly by analysis of trends. As diverse predictive techniques improve the accuracy of the prediction rate, the use of predictive techniques through data mining has been on the rise in various fields with different perspectives. Data mining is the process of extracting data from a huge volume of data set from various different sources and transforming it into a recognizable patterns and useful information (Choi, Choi, Kim, & Lee, 2019; Park et al., 2017; Park & Alenezi, 2018).

Accordingly, the number of studies involved is increasing sharply, in particular, in the field of Business, Medical science, Politics, and Computer science as well (Abramowitz, 1988; Amigobulls, 2015; Fishbein, Azjen, & Hinkle, 1980; Franch, 2013; Lewis-Beck & Rice, 1984; Nadeau, Lewis-Beck, & Belanger, 2009; Park et al., 2017; Park & Alenezi, 2018; Park, 2019; Whiteley, 2005; Whiteley, Sanders, Steward, & Clarke, 2011). This big development in data mining techniques sheds light on the synergy effect of building big data using social media. As social media users or customers of general goods and services, we produce a large amount of social media data every single day. Social media is now more than just networking and exchange of opinions; it has become the main subject of producing quality data (Azam, 2015; Becker & Lee, 2019; Nguyen, 2018; Park et al., 2017; Park & Alenezi, 2018; Park, 2019).

In the early days, predictive techniques in the medical field were useful in determining the survival rate of cancer patients. Moreover, health forecasting generally became a valuable tool for predicting future health events or situations, for example, demands for healthcare needs and health services overall (Majer, 2011; Soyiri & Reidpath, 2013; Javed et al. 2015). One of the major concerns is to predict population health consequence in terms of the number of events occurring within a space of time; for instance, the forecasting of life expectancy and health expectancies (Majer, 2011). In such an effort, it facilitates preventive medicine and health care intervention strategies, by pre-indicating health service providers to take appropriate reassuring actions to minimize risks and manage demand.

Meanwhile, politics has long used several prediction tools such as poll data to predict the outcome of elections. However, lately, many countries, in addition to the USA and the UK, are predicting election results through social media reactions ahead of elections. This is favored with traditional poll data techniques because it is very positive for improving the accuracy and increasing the prediction rate of election results in advance (Franch, 2013). Accordingly, election forecasting has been widely studied with importance in political science field (Abramowitz, 1988; Campbell & Garand, 2000; Fishbein et al. 1980; Franch, 2013; Lewis-Beck & Rice, 1984; Nadeau et al. 2009; Whiteley, 2005; Whiteley et al. 2011).

Especially, forecasting in business became very crucial matter to respond to threats and take advantage of opportunities, particularly, prepare for the future strategy proactively by predicting the future or future outcomes (Dess et al. 2014). As a result, being able to anticipate outcomes allows companies to generate sustainable competitive edge in this competitive world. With this incredible importance of forecasting, the explosion in number of social media users and social media data make it possible to use a variety of predictive techniques using data mining. Predictive analytics
through social media data has been powerful and useful in predicting various situations or trends, for instance, Arthur (2015) has found the positive relationship between Facebook popularity measured in Facebook like and consumer brand stock prices and also predicting the recurrence of trading and the stock value for the next day derived from social media data such as twitter (Amigobulls, 2015).

Park et al. (2017) have found how to predict Arab consumers’ preferences on the Korean entertainment contents using Facebook data in Middle Eastern and North African areas, and recently, Park and Alenezi (2018) have studied social media data-driven insights by predicting the popularity of Saudi multinational enterprises with Facebook and Twitter data. Besides, Park (2019) also investigated a big data of news channels and social media for determining a corporate strategy using a text mining (see Figure 3).

Based on those existing researches, this study aims to discover the new phenomena or trends by analyzing customers’ sentiment and preference. Sentiment analysis through smart language algorithms is worked by identifying and quantifying the positive and negative feelings within our words. Using the sentiment analysis, it is possible to associate the business world with the evidence of social media to test the relationship between social media and its real impact on preference of the company. For this, there is a need for a predictive system to systematically assess consumers’ preference and feedback for financial companies and find the key factors to demand. Due to this data-driven approach, a theoretical framework with hypothesis testing is not applicable for this study. Rather, the analysis will be based on extracting dataset from social media in order to predict the future preference. Thus, among several types of predictive data mining techniques, the classification through sentiment analysis was used for this study.

3. Research Methods and Materials

There is a number of sources from which data has been collected for the companies based on criteria such as the company must be listed in the TOP10 list issued by Ministry of commerce–Saudi Arabia. Secondly, the companies must be actively present on social media, and the company is contributing to the National Transformation Plan 2020 (NTP-2020) for the Kingdom of Saudi Arabia. Lastly, the companies must have a presence in the Middle East. Based on those criteria, we selected three companies, National Commercial Bank (NCB), Al Rajhi Bank, and BUPA Arabia that fulfilled the above-mentioned criteria. Later, each companies’ social presence has such as most common social media platforms Facebook and Twitter. The analysis of each post is compared to followers and how followers are contributing in terms of promoting the brand and how it is contributing in reaching multiple audiences. Similarly, while looking at Figure 2, the top companies in terms of market share are National Commercial Bank (NCB), Al Rajhi Bank, and BUPA Arabia that also contributed to the selection of companies.

The aim of this study is to examine customers’ sentiment and feedback in order to predict future popularity or preference for financial companies in Saudi Arabia. For that, we used a sentiment analysis, one of data mining techniques, which are popular and an overwhelming tools in the process of knowledge discovery in a big data to extract patterns and find future trends or directions. The data was collected from the financial companies that are located and mainly operating in the Kingdom of Saudi Arabia. Among them, two companies are selected from the banking sector and other company is selected from the insurance sector. The selection of companies was made using the following criterial fulfillment (see Table 1).

The data was collected from start till October 2018 using a tool developed in PHP and JavaScript that used the API provided by Facebook and Twitter to extract the data from each company separately (Park & Alenezi, 2018). Using this tool, the sentiment analysis was carried out for this research. The sentiment analysis refers to the prediction of the sentiment of people from their posts. It helps in predicting the overall popularity as well as tells the perception of people about the company by understanding customers’ emotion and feedback. This data is open to the public as it is acquired from the pages the companies post for the audience to access. However, it can be seen from our results that penetration of the market is not more than the gulf area. The collected data focused on the post reach, the post made, and the total likes.

4. Results and Discussion

This study shows the result of the quantitative analysis on Facebook and Twitter content. Table 2 shows the presence
of selected three companies (NCB, Al Rajhi, and BUPA) on Facebook and Twitter till October 2018. It can be seen that all three companies have similar likes and followers as they serve customers as B2B and B2C companies as well. A number of followers indicate the people who like to get particular company news and announcements. These followers may be mostly public customers, including investors, stockholders, competitors or job seekers.

As shown in Figure 4, the number of followers on Twitter in Al Rajhi and BUPA is above 1000% with heavy regional impact, while the number of Facebook posts in NCB is more than Al Rajhi and BUPA. Also, Twitter has the highest penetration in the market compared to Facebook and it is contributing to a higher reach in the market. But compared to the banking sector, the insurance company has a bigger penetration due to the service sector they cater to. As seen from Table 2, that number of followers in terms of Facebook is less, but in terms of Twitter, the insurance company surpasses all. In terms of followers, Al Rajhi bank has above 1 million followers that means a single post will reach at least 1 million people.

Table 1: The list of companies with selection criteria

<table>
<thead>
<tr>
<th>Company Name</th>
<th>Model Type</th>
<th>Regional Presence</th>
<th>Social Media Presence</th>
<th>Top 10 Ranked</th>
<th>Contributing to NTP 2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Commercial Bank (NCB)</td>
<td>Banking sector</td>
<td>KSA, International</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Al Rajhi bank</td>
<td>Banking sector</td>
<td>KSA, International</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>BUPA Arabia</td>
<td>Insurance sector</td>
<td>KSA, International</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 2: The selected companies’ presence over Facebook and Twitter

<table>
<thead>
<tr>
<th></th>
<th>NCB</th>
<th>Al Rajhi</th>
<th>BUPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likes</td>
<td>369,155</td>
<td>593,653</td>
<td>593,489</td>
</tr>
<tr>
<td>Follow</td>
<td>370,757</td>
<td>590,078</td>
<td>592,623</td>
</tr>
<tr>
<td>Twitter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tweets</td>
<td>202K</td>
<td>43.8K</td>
<td>3.91K</td>
</tr>
<tr>
<td>Following</td>
<td>3</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>Followers</td>
<td>847K</td>
<td>1.14M</td>
<td>309K</td>
</tr>
<tr>
<td>Moments</td>
<td>13</td>
<td>12</td>
<td>1106</td>
</tr>
<tr>
<td>Likes</td>
<td>131</td>
<td>90</td>
<td>1106</td>
</tr>
<tr>
<td>Videos and Photos</td>
<td>4916</td>
<td>5185</td>
<td>2815</td>
</tr>
</tbody>
</table>

Unit: K – thousands, M- Million

Sentiment analysis reveals people’s perception or preference about the company or it measures the company’s marketing campaign and its posts (See Table 3). No negative sentiment was detected for Al Rajhi on English posts while it can be seen that Internet penetration of both banks is higher.
than BUPA, which is mentioned rarely in few hours. This result shows sentiment analysis provides insight into the real mood behind customers’ posts or messages by identifying and quantifying the positive and negative feelings or emotions within their words. In terms of content posted on the Internet by companies reveal customer engagement, and only Al Rajhi seems to be better, which shows it has more followers. Thus, it can be seen that content does matter about the posts and it is predicted that Al Rajhi is likely to penetrate local customers compared to other banks.

5. Conclusions and Limitations

This study shows the sentiment analysis detects the emotional tone behind text, offering key insights into customer perception and satisfaction. Accordingly, it can lead to investigate sustainable growth potential of Saudi companies pertaining to financial sectors with social media data to predict the future preference of customers in the long term. In particular, this intensive research in the financial sector, which has a difficulty to access company information due to Saudi’s conservative business culture and environment, makes it more valuable and unique. Many studies have been conducted about Saudi companies, but most of them were only qualitative studies such as case studies, or simply testing causality or causal relationship between variables. However, this study reveals the hidden sentiment of customers embedded in big data, which gives profound insights into predicting the direction of Saudi financial firms and looking ahead to sustainable growth by reducing customer churn and handling multiple customers with tracking overall customer satisfaction. Accordingly, this study suggests the following valuable implications.

First, this study assumes that Saudi banks and insurance companies play an important role in predicting the future economic development of Saudi Arabia as leading and key indicators of its economy. In this regard, this study focuses on the response of customers surrounding these major companies, and thus throws a very important intuition in estimating the far future of Saudi Arabia. Second, it is shown that social media data, such as Facebook and Twitter, discussed in this study, play a critical role as a company’s decision-making tool, rather than just simple communication tool which is establishing and supporting relationships between people. We are now living in an era of data-driven decision making. Therefore, it is necessary to apply a big data from various global markets to the most traditional financial firms such as banks and insurance companies to survive in the fierce global markets.

In other words, by looking at the changing trends of consumers’ preferences or tastes, traditional banks and insurance companies can try to change their image in the future which was fixed as a traditional entity. For example, by identifying key emotional triggers, new consumers can lead to new trends or preferences through social media data, and new understanding of consumers’ sentiment about the company. So, financial firms can develop innovative financial products or services, which can be reflected in their lifestyles, sniper service spaces and complex cultures in banks and insurance companies, and finally it will lead to upselling opportunities. It shows that it is possible to create tremendous possibilities for various strategic tools for expansions of companies according to specific cultural concepts and trends.

Finally, this research presents meaningful insights in data-driven approaches using a data mining technique as a tool for corporate strategic decision-making and forecasting. Data mining techniques have been used in various fields, such as the medical field and politics, based on computer science. But now, data mining has become a very popular and important tool for driving more professional management in business. This study shows that diverse data mining techniques are not just methods for forecasting, but also are the most critical tools for planning current marketing activities and implementing future strategies.

Despite those academic achievements and valuable implications, this study has the following limitations. First, there are several banks and insurance companies in Saudi Arabia, but this study concentrates on only three top companies. Therefore, it would be very insightful if we do cross-industry comparative research that expands the scope of the data in the future or expands the data area itself. Second, the sentiment analysis, is one of the popular data mining tools used in this study. However, there are many other data mining techniques from which to draw insights using big data. Accordingly, diverse research results can be obtained with more holistic perspectives if we approach them to compare the results using more data mining techniques. Third, if the scope of this study is extended to other regions beside Saudi Arabia since the selected companies, while operating in the Kingdom, their main target, have also a presence in the entire Middle East, North Africa and the Asia region, more meaningful results would be produced.

References

Abdulaziz B.A. (2013). A survey of social media users in Saudi Arabia to explore the roles, motivations and expectations...
toward using social media for social and political purposes. 
Master Thesis, Arkansas State University.

presidential election outcomes. PS: Political Science and 


Azam, O. (2015). Social media impact on arab spring, a comparison 
study between four middle eastern countries. Master Thesis, 
Hawaii Pacific University.

Social Media and the Emergence of Reflexiveness as a New 
322-336.DOI.org/10.1016/j.lrp.2016.07.005

Bassam, A. (2015). Does Saudi Arabia’s economy benefit from 
foreign investments? Benchmarking: Bradford, 22(7), 1214- 
1228.

Media for Corporate Reputation Management. Journal of Asian 
org/10.101310/j.jfabe2019.vol6.no1.231

Campbell, J. E. & Garand, J. C. (Eds.). (2000). Before the vote: 
Forecasting American national elections. Thousand Oaks CA: 
Sage.

crowds: The value of stock opinions transmitted through social 

Smart City Evolution in South Korea: Findings from Big Data 
Analytics. Journal of Asian Finance, Economics and Business, 7(1), 
no4.179

(2014). Strategic Management: text and cases. New York, NY: 
McGraw-Hill Education.

network sites. In The Oxford Handbook of Internet Studies, ed. 

Fishbein, M., Azjen, I., & Hinkle, R. (1980). Predicting and 
understanding voting in American elections: Effects of external 
variables. In I. Azjen and M. Fishbein (Eds.), Understanding 
and predicting behavior, 176-195. Englewood Cliffs, NJ: 
Prentice Hall.

ambidexterity as a new capability to manage social media 
communication within organizations. The Journal of Strategic 
jsis.2015.03.002

Diabetic cases in KSA through search trends and Creating 
Cyber Diabetic Community, International Conference on 
Recent Advances in Computer Systems, Atlantis Press

Social Influence Processes and Psychological Factors to 
Measure Pervasive Adoption of Social Networking Sites: 
Evidence from Pakistan, Emerging Markets Finance and Trade, 
54(15), 3485-3499, DOI: 10.1080/1540496X.2017.1417834

Kavoossi, M. (2000). The globalization of business and the Middle 
East: Opportunities and constraints. Westport, CT: Quorum 
Books.

elections. Legislative Studies Quarterly, 9(30), 475-486. DOI: 

shaping on knowledge for organizational improvement with 
wikis. MIS Quarterly, 37(2), 455-A412. DOI: 10.25300/ 
MISQ/2013/37.2.07

Majer I. (2011). Modeling and forecasting health expectancy: 
theoretical framework and application. In: Netspar Discussion 

forecasting in the United Kingdom: A two-step model. Journal 
doi.org/10.1080/17457280903074726

Nguyen, T. (2018). The impact of hallyu 4.0 and social media 
on korean products purchase decision of generation C in Vietnam. 
Journal of Asian Finance, Economics and Business, 5(3), 81- 

Park, Y. E., Allui, A., & Alselaimi, R. (2017). Determinants of 
entry modes choice for MNEs: Exploring major challenges 
and implications for Saudi Arabia. 1st AUE International 
org/10.1007/978-3-030-01662-3

Park, Y. E., Chaffar, S., Kim, M.S., & Ko, H.Y., (2017). Predicting 
Arab consumers preferences on the Korean contents 
DOI:10.15722/jds.15.4.201704.33

Park, Y. E. (2018). The endless challenges of KIA motors for 
globalization: A case study on Kia in Saudi Arabia. International 
Journal of Industrial Distribution & Business, 9(9), 45-52. DOI: 

Park, Y. E. & Alenezi, M. (2018). Predicting the Popularity of Saudi 
Multinational Enterprises Using a Data Mining Technique. 
Journal of Management Information and Decision Sciences, 
21(1), 1-14. https://www.abacademies.org/articles/predicting- 
the-popularity-of-saudi-multinational-enterprises-using-a- 
data-mining-technique-7791.html

Park, Y.E. (2019). Data Empowered Insights for Sustainability 
of Korean MNEs. Journal of Asian Finance, Economics and


