

Prototype Development and Pre-Commercialization Strategies for Mobile Based Property Analytics

<https://doi.org/10.3991/ijim.v13i10.11309>

Mohd Faizal Omar (✉), Nurul Husna Mahathir,
Mohd Nasrun Mohd Nawi, Faisal Zulhumadi
Universiti Utara Malaysia, Kedah, Malaysia
faizalomarr14@outlook.com

Abstract—Government agencies are constantly looking for viable strategies to strengthen their newly implemented policies for nation building. Grassroots or customer sentiments are very important to develop an inclusive policy with a mixed of bottom-up approach to incorporate customer opinion. However, due to the unique political and multiracial landscape in Malaysia, current commercial off-the-shelf Social Analytics are irrelevant to capture sentiments of the multi-lingual characteristics of Malaysian native speakers. Current Social Analytic tools lack the quality of analysis for foreign languages, such as Malay, which limits businesses to localize advertisement for a specific geographical area. Hence, this research was to develop a real-time social media analytic tool with sentiment analysis specifically in the Malaysian context, in order to engage and analyze customer reviews and opinions. The main purpose of this paper is to demonstrate an approach to utilize data from social media platforms, such as Facebook and Twitter, in gaining valuable insights to drive and improve marketing strategy in the property industry. This research developed a tool, namely Property Analytics, to assess public sentiments on specific property projects or services. The methodology and approach to enhance from lab scale to pre-commercialization activities are outlined in this paper. It is anticipated that this work is relevant to real world application, improve stakeholder’s decision making, and the people’s quality of life.

Keywords—Social Media Analytics, Sentiment Analysis, Pre-Commercialization

1 Introduction

The Malaysian digital landscape has shifted over the past couple of years, changing the way Malaysians interact with each other, how they form their opinions, and how they make decisions. According to “We are Social: Digital in 2018” report, findings indicated that 25million Malaysians are Internet users and 95.7% of them are active social media users. In fact, research showed that Malaysians spend 8.45 hours daily on various online media platforms[1]. Social media has removed the barrier between brands and consumers, letting them connect directly with each other. Social media facilitates real-time marketing and real-time marketing takes marketing one step further by allowing brands to share experiences with their audience. With billions of

existing social media posts and profiles, and more are being created daily, social media is one of the richest sources of information that organizations can use to aid in everything from loss prevention and asset security, to customer engagement and brand awareness. Moreover, social media gives an opportunity to be close to the user, allow organizations or companies to understand the current pulse of their customers, produce insights that facilitate decision making, and engage current and potential customers to drive business results. However, in Malaysia, this low-cost high-reach channel that has 25 million active users seems to be neglected.

Therefore, this paper presents a designed social media analytic tool specifically for the Malaysian context in order to engage and analyze public online reviews and opinions. Additionally, this research also generated pre-commercialization strategies in order to upscale this prototype tool. The main purpose of this prototype was to utilize data from social media platforms, such as Facebook and Twitter, in gaining valuable insights to drive and improve marketing strategies. Other than extracting data from Facebook and Twitter platforms, one special feature, which is the integration with Google Analytics data, will be added. With this tool, companies or organizations will be able to better understand online trends towards their products and services.

2 Social Media Analytics, Sentiment Analysis, and Property Analytics

Social media analytics refers to the approach of collecting data from social media sites and evaluating that data to make business decisions. This process goes beyond the usual monitoring or basic analysis of re-tweets or “likes” to develop an in-depth idea of the social consumer. For this prototype, three main features were included, which are sentiment analysis, real-time data monitoring, and reporting.

Sentiment analysis refers to the analysis of text in identifying feeling or emotion, an attitude or opinion. On social media, the sentiment of a post can be seen in the tone or emotion conveyed in brand mention. The opinionated text will then be grouped into a positive, neutral, and negative category. Reports on sentiment analysis can help the organizations to see how positively or negatively their brand is perceived on social media, based on the tone of mentions.

Besides, the proposed prototype namely Property Analytics can provide a real-time method to engage with the community. With this feature, the companies in property and real estate in Malaysia can easily find out how many people write about their brand in a particular year, month, week, or day. In fact, real-time data monitoring may assist in identifying when Internet users are the most active. They can estimate the dynamics of changes in a buzz and distinguish events that influence these changes. What is more, organizations are able to view the share of voice for brands in a whole discussion. It is an easy way to recognize their biggest competition in digital media and compare their market position to others.

The third feature of the prototype is reporting where it lists down what people talk about, either positively or negatively, in relation to their brands or services. Further-

more, this report can be customized as a daily, weekly, or monthly report that is made exportable to Excel format.

In our previous research, a sentiment analysis model for both Malay and English languages was developed [2-5]. It emphasized on complex multilingual text data, including dialects in Malay and English, and the experiments conducted has proven that the model improved the accuracy level by 4.2% from the baseline model.

3 Methodology

As digital and social media users continue to grow exponentially, measuring a brand or public sentiment becomes an important element to arrive at the right strategy and decisions for many businesses and organizations. This is why there is a tremendous number of social analytics tool available in the market to cater for this need and many companies opt for this tool to gauge user sentiment towards their brands [6, 7]. However, most of the current analytics tool suffers from inaccuracy of sentiment results due to various reasons, such as the complexity of the language and non-dictionary based keywords [8]. According to various research studies and recent findings from FreshMinds, it was revealed that sentiment accuracy from various analytic tools, on average, was between 30-60%[1]. It is crucial to use the right approach in order to bring meaningful and more actionable insights from social media data and one cannot blindly rely on automated tools to process sentiment from social media data; which effectively can lead to wrong and detrimental strategies, decisions, and actions [10]. Therefore, there is a need to develop an analytic tool with a high degree of accuracy to ensure that this automated system works as best as possible.

The outline of this prototype development consisted of two phases:

- Sentiment analysis which involves data cleaning and sentiment classification and
- Analytics dashboard development for both mobile and desktop platforms. Figure 1 illustrates an overview of the prototype architecture.

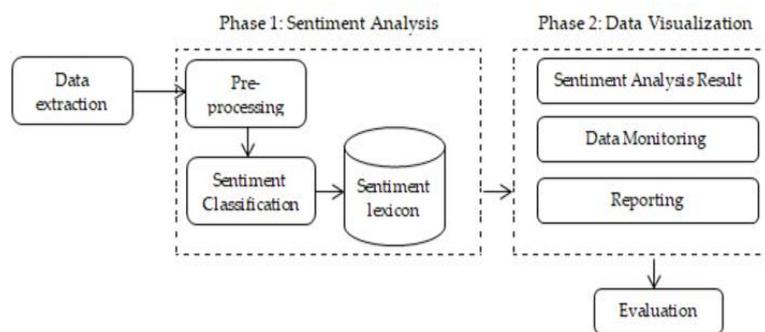


Fig. 1. Prototype Architecture

3.1 Data extraction

The data from Google Analytics, as well as Twitter and Facebook platforms shall be extracted. A few keywords related to a specific brand or services are identified and used to extract the data. Only reviews written in Malay and English languages will be collected.

Phase1: Sentiment analysis: During this phase, the data extracted shall be analyzed and classified into three categories, namely positive, neutral, and negative. Two levels of classifications will be performed, which are the overall sentiments and aspect-based sentiment analyses.

Phase2: Analytics dashboard development: In order to visualize the results of the analysis done in the first phase, a dashboard shall be developed using the Tableau application to display sentiment analysis results, data monitoring, as well as reporting. It consists of three components:

- The main dashboard shows overall sentiments of the analysis and aspect-based analysis—it includes the number of data being analyzed, the overall sentiments of positive, neutral, and negative classes, most trending words mentioned, and sentiments for each aspect/feature in the property domain;
- In the second component, real-time data monitoring consists of daily statistics, number of re-tweets, shares, and likes, as well as most frequent mention words on a particular subject is displayed;
- Finally, the third component shows the details of tweets or comments based on its feature/aspect and polarity that may help organizations to view details of what people talk about.

Evaluation: Face Validation and User Interface (UI) design validation shall be conducted to validate the proposed tool, while Usability and Test cases will be conducted for UI improvement using heuristics evaluation. Stakeholders such as managers/technical experts in government initiative projects, such as PRIMA and PPAM will be invited for the workshop.

4 Planning for Pre-Commercialization Activities

The researchers for this study conducted prototype market analysis with similar commercial products in the market. The unique features of the prototype and strategies to position this prototype from lab scale and extend it to pre-commercialization are described.

Benchmarking with commercial tools

Most of the Sentiment Analysis (SA) systems in the literature only concentrate on mathematical models, but not for real-world usage. Thus, this research moves a step forward by implementing an SA system in a form of an IT prototype. Moreover, even though there are various social media analytic tools available in the market, their analysis is mainly focused on the English language only. This study is proposed to handle text written in the Malay language as well, in order to cater for the Malaysian

context. Table 1 below shows a comparison of the proposed prototype with two leading social media analytics tools in Malaysia: -

Table 1. Property Analytics in comparison with other Sentiment Analysis Products

	Product A	Product B	Our proposed analytics tool
Platform covered	Facebook, Twitter	Facebook Twitter	Facebook, Twitter
Language covered	English	Malay, English	Malay, English
Real-time	√	-	√
Features offered	Sentiment Analysis Data Monitoring Reporting	Sentiment Analysis Reporting	Sentiment Analysis Data Monitoring Reporting
Exportable reports	√	-	√

Property analytics unique features

The anticipated unique features of this developed prototype shall include:

- sentiment analysis task that covers up to aspect/feature level analysis,
- extra information included, such as real-time data monitoring and detail reporting made exportable to Excel format, using other social analytic elements such as top tweets, word cloud, tweet trends, etc., and
- mobile based dashboard for data visualization.

Property analytics Pre-commercialization strategy

The following activities have been planned to be conducted as strategies to move the application from prototype form to pre-commercialized stage:

- Activity 1: Grant acquisition, prototype development, and documentation/closing.
- Activity 2: IP protection, document preparation, and registration.
- Activity 3: Product/Market Fit. Market validation will be conducted.
- Activity 4: Product revision and documentation. A prototype will be revised based on market input from Activity 3, and it will be packaged and documented.
- Activity 5: Searching for potential partners and negotiation. Due diligence and business negotiation with potential partner will be performed.
- Activity 6: Closing deals and contract agreement. All parties will negotiate for commercialization details such as IP rights, licensing/outright terms, etc.
- Activity 7: Up-scaling the prototype (if required). Bridging grants are needed.
- Activity 8: Set up a start-up company.
- Activity 9: Operation expenditure of start-up. Require seed funding from university or other venture capitalists/angels.
- Activity 10: Sustainability and expansion.

5 Discussion and Conclusion

The authors of this paper described the prototype development approach and its related pre-commercialization strategies. Social media data can be leveraged to evaluate project/services in property or real estate industries. The main feature to be provided is text mining and sentiment analysis. Data will be extracted from Facebook and Twitter platforms. Instead of using the traditional Natural Language Processing techniques in performing sentiment analysis, the prototype will utilize our own constructed sentiment lexicon that had been developed from previous research which covers bilingual sentiment words (Malay and English) for sentiment classification.

The prototype development will consist of two major activities which are sentiment analysis and analytics dashboard development (mobile and desktop). This proposed tool will yield a distinct feature from other competitors where it will be able to analyze customer opinions in two languages, i.e., Malay and English, with better accuracy. Besides that, this tool will include special features such as the integration with Google Analytics data, real-time data monitoring, and reporting. This prototype will enable stakeholders to understand online trends towards their brands, services, or policies. It is vital for stakeholders to embrace and leverage the use of analytics to effectively track, assess, and optimize their marketing and decision-making efforts. In addition, local authority can also monitor activities or threats from analytic data for a better cybersecurity system.

The following are the summary of the expected services that can be provided by Property Analytics, which are to:

- Outsource of social media monitoring services
- Measure sentiments for specific brands/events/campaign/policies
- Tailor-made report and analysis
- Monitor sentiment at specific locations
- Brand/event/campaign strategy and execution
- Organizational reputation management
- Control and evaluate government policies
- Assist government agencies in gauging people's expectations.

6 Acknowledgement

The authors of this paper would like to express deepest gratitude to Universiti Utara Malaysia for providing research funding (SO Code: 13959).

7 References

- [1] S. R. Department, (2019). Average time spent using online media in Malaysia in Q2 and Q3 2018, by activity (in hours per day). Available: <https://www.statista.com/statistics/803614/daily-time-spent-using-online-media-by-activity-malaysia/>

- [2] S. Abdul-Rahman, N. S. Sobri, M. F. Omar, A. M. Benjamin, and R. Ramli, "Graph coloring heuristics for solving examination timetabling problem at Universiti Utara Malaysia," 2014, pp. 491-496. <https://doi.org/10.1063/1.4903627>
- [3] N. H. Mahadzir, M. F. Omar, and M. N. M. Nawawi, "Towards sentiment analysis application in housing projects," *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 8, pp. 145-148, 2016. <https://doi.org/10.1063/1.4960900>
- [4] N. H. Mahadzir, M. F. Omar, and M. N. M. Nawawi, "A sentiment analysis visualization system for the property industry," *International Journal of Technology*, vol. 9, pp. 1609-1617, 2018. <https://doi.org/10.14716/ijtech.v9i8.2753>
- [5] N. H. Mahadzir, M. F. Omar, and M. N. M. Nawawi, "Semantic similarity measures for Malay-English ambiguous words," *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 10, pp. 109-112, 2018.
- [6] W. Fan and M. Gordon, *The Power of Social Media Analytics* vol. 57, 2014.
- [7] M. Mahrt and M. Scharnow, "The Value of Big Data in Digital Media Research," *Journal of Broadcasting & Electronic Media*, vol. 57, pp. 20-33, 2013/01/01 2013. <https://doi.org/10.1080/08838151.2012.761700>
- [8] S. Yoo, J. Song, and O. Jeong, "Social media contents-based sentiment analysis and prediction system," *Expert Systems with Applications*, vol. 105, pp. 102-111, 2018/09/01/ 2018. <https://doi.org/10.1016/j.eswa.2018.03.055>
- [9] T. Padmapriya, S. V. Manikanthan, "LTE-A Intensified Voice Service Coder using TCP for Efficient Coding Speech", *International Journal of Innovative Technology and Exploring Engineering*, Vol. 8, issue 7s, 2019.
- [10] S. Shayaa, N. I. Jaafar, S. Bahri, A. Sulaiman, P. S. Wai, Y. W. Chung, et al., "Sentiment Analysis of Big Data: Methods, Applications, and Open Challenges," *IEEE Access*, vol. 6, pp. 37807-37827, 2018. <https://doi.org/10.1109/access.2018.2851311>

8 Authors

Mohd Faizal Omar is a senior lecturer at the Department of Decision Science, School of Quantitative Sciences (SQS), Universiti Utara Malaysia. His research mainly focuses on development of decision support system (DSS) in construction discipline. faizalomarr14@outlook.com

Nurul Husna Mahadzir is a PhD student at School of Quantitative Sciences (SQS), Universiti Utara Malaysia. Her expertise is in the fields of data mining, social network analysis, theory of computation, and public opinion. nhusna84@gmail.com

Mohd Nasrun Mohd Nawawi is a senior lecturer at School of Technology Management and Logistics (STML), Universiti Utara Malaysia. He received his PhD in Construction Project Management at University of Salford, United Kingdom in 2012. nasrun@uum.edu.my

Faisal Zulhumadi is a PhD student at School of Technology Management and Logistics (STML), Universiti Utara Malaysia. His expertise is in the fields of artificial intelligence, case-based reasoning, and technology management in nanotechnology. faisal@uum.edu.my

Article submitted 2019-07-16. Resubmitted 2019-08-06. Final acceptance 2019-08-07. Final version published as submitted by the authors.