Managing the Social Influence of Public Figures on Social Media

Hsiao-Ting Tseng
National Central University
httseng@mgt.ncu.edu.tw

Chia-Lun Lo
Fooyin University
allenlo.tw@gmail.com

Ting-Ann Shih
National Central University
stephenlove52@gmail.com

Abstract

Social media is critical to personal branding, and social advertising is commonly used to enhance brand communication and develop stronger personal brands. Public figures invest their resources in social advertising to improve their social influence within social networks. However, an effective social media advertising strategy has yet to be clearly defined. This study investigated effective social media advertising strategies and provided precise guidelines for social advertising investment for individuals managing personal brands. The study focused on political figures. The study applied machine learning to sort social advertisements and identify effective advertising patterns. Finally, fan pages were clustered to identify the most influential cluster. This study offers insights into personal branding on social media, which may help social investors refine their social influence marketing strategies.

Keywords: Personal branding, social advertising, social influence management, clustering

1. Introduction

The widespread use of social media has led to the emergence of a new online retail model called social commerce, which includes group buying and sales driven by social networks. Facebook is a crucial channel for promoting products and services. In addition to large companies and brands, individuals in the social ecosystem who focus on personal image or impression management have also been leveraging social media's ability to influence others and expand their social reach (Scheitle & Ecklund, 2017).

As social competition intensifies, acquiring individual differentiation on social media to enhance visibility has become vital. Consequently, strategies to preserve personal brand value within social media have grown increasingly crucial for individuals, particularly public figures.

Social media has emerged as the prime platform for showcasing personal image, values, and professional expertise. Various influential figures, including celebrities, experts, and politicians, must manage their social media presence to present and cultivate their required personal brand. Effective social media advertising strategies directly impact the establishing and development of individual brands.

Facebook is a significant social platform for companies or individuals to promote products and services. Many people manage personal dedicated fan pages on Facebook. With access to exclusive audience insights and data analytics, these pages can assist in precise message delivery and further advertising and promotion efforts. Complete fan pages can serve as an extension of your image, allowing you to showcase more professional expertise, values, and brand characteristics. We can blend personal and professional elements within the fan page, establishing a more comprehensive image. Despite the emergence of social networking sites, which has led to the prominence of influencers wielding a level of influence over the public comparable to that of celebrities in traditional media, there remains a dearth of research focused on socially influencing one of the most influential segments – politicians.

Good impression management is critical to personal branding to ensure election success (Jacobson, 2020). As is widely recognized, traditionally conservative politicians have historically crafted their image through conventional media. However, to engage a broader spectrum of constituents, there is a growing necessity to embrace social media for interacting with voters and shaping one's image. While some studies have delved into the brand value of politicians on social media(Lalancette, Raynauld, 2019; Fakas & Bene, 2021; Zhuravskaya, Petrova, Enikolopov, 2020; Ernst, Esser, Blassnig, Engesser, 2019; Krishen, Dwivedi, Bindu, Kumar, 2021; Reisach, 2021), they often require a substantial amount of time for statistical analysis. Swiftly and efficiently gauging media effectiveness on online

URI: https://hdl.handle.net/10125/106725
978-0-9981331-7-1

social media platforms to inform subsequent action plans remains an area that has yet to be thoroughly explored.

Therefore, this study utilizes machine learning techniques to analyze how politicians manage their branding, identifying favored patterns. This is achieved through the following three steps:

1. After segmenting a dataset of political advertisements on Facebook, the study sorts and analyzes the data and then conducts clustering to aid in recognizing various types of ads.
2. By utilizing machine learning techniques to segment advertisements, this study aims to identify the most effective advertising patterns on the Facebook platform. Different performance differences among various groups can be recognized through grouping advertising patterns, leading to the formulation of impactful advertising strategies.
3. To explore the patterns and audience engagement of the "boosted posts" subtype of advertisements on the Facebook platform. The level of audience engagement for this type of advertisement will be assessed, providing insights into its effectiveness in assisting fan page owners in achieving enhanced audience interaction.

We delve into the advertising patterns employed by politicians on Facebook and analyze the correlation between these patterns and the social influence of their fan pages. This analysis aims to confirm the impact of advertising patterns on social influence while ensuring the implementation of precise social advertising investment strategies.

2. Literature review

2.1. Personal branding on social media

The “personal branding” concept was introduced in 1997 (Peters, 1997) and has garnered considerable research attention. The scope of celebrities has expanded beyond actors and singers to include various fields, such as celebrity chefs, athletes, and businesspeople (Moulard et al., 2015). Researchers have examined personal branding in the contexts of professionals from different domains, including marketing practitioners (Rangarajan et al., 2017), academics, artists, and politicians (Scheidt et al., 2020).

Personal branding typically involves image building, positioning, and assessment (Khedher, 2014). Shafiee et al. (2020) explored brand positioning factors and indicators in offline spaces, such as physician clinics, and online areas, such as social media platforms. In addition, Geva et al. (2019) explored the influence of the online environment on personal brand performance and behaviors related to personal branding, such as sharing content on social networks. Although numerous studies have investigated personal branding in a variety of contexts (Labrecque et al., 2011; Moulard et al., 2015; Rangarajan et al., 2017; Wroblewski and Grzesiak, 2020), few have explored the intersection of personal branding and advertising, particularly within the realm of online social media.

2.2. Social influence management

As the popularity of social networking sites continues to grow, online social diffusion and social influence have garnered increasing research attention. Zheng et al. (2012) highlighted the role of user relationships in shaping behavior on platforms such as Twitter and Facebook. Users can influence their friends, either explicitly or implicitly, by sharing content. Studies have conducted empirical analyses to elucidate information-spreading patterns and determine the extent of social influence (Bakshy et al., 2011; Eagle et al., 2009). Several studies have focused on maximizing influence in online social networks through social network analyses (Kempe, 2003; Wang et al., 2018; Wang et al., 2021; Zhang et al., 2017). However, these studies have primarily focused on identifying the most influential groups or individuals in the network without exploring how to become significant. Some other studies have aimed to identify influencers within online social networks. Rios et al. (2019) argued that influencers play a vital role in the functioning specific social networks known as communities of practice. They propose a semantic-based filter to identify influencers who contribute to the network. Similarly, Quan et al. (2020) proposed an approach to identify categorical influencers on microtext-based social media.

Social networking sites offer advertisers a quicker and more direct means to reach their target audience through social network advertising. Social network advertising enables advertisers, including individuals and brands, to penetrate different subnetworks within the more extensive social network. Consequently, existing solutions for maximizing social influence may not effectively provide a competitive advantage. Furthermore, existing influence maximization models primarily rely on graph theory and may overlook the properties of messages exchanged between nodes in the social network.

2.3. Machine learning in social media

Machine learning has gained considerable popularity in various industries. Machine learning can be divided into three categories: reinforcement learning, supervised learning, and unsupervised learning (Alpaydin, 2020). Each type has distinct characteristics and can be applied to different scenarios and tasks.

Reinforcement learning is suitable for sequential scenarios, aiming to formulate effective action policies to achieve goals. Its success spans video game development (Mnih et al., 2015), medical diagnosis (Pineau et al., 2009; Zhao et al., 2011), autonomous vehicles (Kiran et al., 2022), and influence maximization in social networks (He et al., 2021; Li et al., 2022). Supervised learning encompasses classification and regression. Classification algorithms are used for prediction and pattern recognition, applied in cancer classification (Ressom et al., 2008), predictive maintenance (Garg et al., 2015; Li et al., 2014; Susto et al., 2015), asset value prediction (Pai and Wang, 2020), stock price prediction (Tsai and Wang, 2009), and quality classification in social commerce ads (Trehan and Sharma, 2020). Unsupervised learning discovers patterns in input data, with clustering applied in customer segmentation (Tsai et al., 2015; Sun et al., 2021) and offering potential in social commerce. Despite extensive use in various fields, machine learning's utilization in personal social marketing remains underexplored (Sun et al., 2021).

In personal social commerce, individuals can exhibit and market their products or services through social media platforms while actively engaging in interactions and dialogue with fellow users. Consequently, social media platforms constitute a vast repository of information, encompassing even political subjects. Therefore, in past studies, many investigations within politics and elections have centered on harnessing social media to analyze and prognosticate electoral polls. Many studies in politics and elections have used machine learning approaches to focus on social media to analyze and predict election outcomes in previous research.

In a study by Brendan, Rammath, Bryan, and Noah (2010), an analysis was undertaken on an expansive dataset of 1 billion Twitter messages spanning 2008 to 2009. This endeavor entailed establishing correlations between political perspectives and the frequencies of sentiment-associated words within Twitter messages. Meanwhile, another study by Andraik, Timm, Philipp, and Welpe (2010) explored Twitter's role in political deliberation, delving into the fidelity of online messages on this platform as accurate reflections of offline political sentiments. The researchers conducted meticulous content analysis on over 100,000 statements referencing political parties or politicians. Diego and Karin (2014) ventured into predicting shifts in voting inclinations by analyzing sentiment time series extracted from comments on news articles encompassing three distinct Brazilian elections. The landscape was further enriched by an accessible system (Shawn and Larry, 2013) capable of prognosticating the voting choices of Twitter users during the 2012 U.S.A. presidential election. Marco and Popescu (2011) also introduced a robust machine- learning framework designed for the large-scale classification of social media users, categorized according to dimensions of interest. This innovative framework leveraged automatically generated prototypical words with Latent Dirichlet Allocation (LDA) (David, Andrew, Jordan, 2003) for enhanced classification accuracy.

In our study, we focused on Facebook fan pages. Specifically, we gathered and analyzed messages from the fan pages of various Taiwanese politicians. This choice was based on all posts on these fan pages being publicly accessible, including interactions from fans and users.

3. Methodology

3.1. Research design

Enhancing personal image activities and strategies are bolstered by advancements in community technology and the rapid expansion of social media users. Social media provides several advantages, including facilitating interpersonal connections, enabling open conversations, self-expression, and idea-sharing (Bryman, 2012). This study collected multiple data sources from the popular social media platform Facebook. We applied cluster analysis of machine learning techniques to gain insights into advertising on social networks.

The research process is shown in Figure 1. First, we constructed a data warehouse and performed preprocessing procedures to sort and organize the collected data. Then, we applied a clustering algorithm to generate advertising data and boosted post segments.

An effective advertising pattern was identified based on the advertising results and boosted post segmentation on Facebook. Subsequently, a clustering analysis of fan pages was conducted to determine the different ways of influence within social advertising (Figure 1).

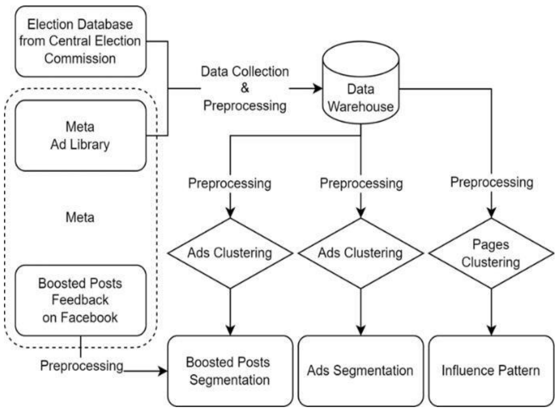

Figure 1. Research Process.

3.2. Data preprocessing

3.2.1. Data.

The data initial dataset was obtained from the Meta Ad Library Report, which includes advertisements related to public issues, such as elections. The dataset covers promotions from November 11, 2019, to May 5, 2022, during which many crucial polls were held.

The columns in the dataset from the Meta Ads Library Report include page ID, page name, disclaimer, amount spent, and number of ads in the library. The original dataset from the Meta Ad Library Report includes advertisements related to politics, public issues, and elections. However, the present study focused on politicians who aimed to enhance their social influence to win elections. To narrow down the dataset, we screened the politicians' fan pages from the original dataset by comparing them against the database of political candidates from the Central Election Commission, which includes all the politicians who have ever participated in elections. This comparison helped identify the politicians who participated in elections and excluded those appointed rather than elected.

Following the screening procedure, we collected fan pages that specifically featured the full names of politicians who had participated in elections. We also excluded fan pages that were not managed directly by the politicians themselves, even if they contained the full names of the politicians. Subsequently, we queried the Meta Ad Library API for each politician's fan page and combined all the advertisements found into a new dataset focused on politician advertising. This dataset encompassed 20 distinct data attributes. To gather feedback on the "boosted posts," we developed a web crawler to retrieve information from Facebook. This data encompassed various metrics such as the count of comments and shares and the number of reactions (*like*, *love*, *care*, *haha*, *wow*, *sad*, and *angry*).

3.2.2. Data preprocessing.

The data preprocessing comprised the data cleaning, integration, and transformation procedures. In the data cleaning step, data related to deleted fan pages was removed from the dataset. Additionally, four advertisements that were still active but had yet to meet specific deadlines were addressed. However, because other attributes of these advertisements, such as spend and impressions, can only reflect the situation, such as the data download time, the missing value in the four advertisements was manually filled with the date when the data was downloaded. We extracted all variables needed for each experiment in the data integration and transformation phase. Furthermore, normalization was performed in the clustering process to ensure a balanced weight for each variable.

All relevant attributes from the Meta Ad Library API were included in the first experiment except for the estimated_audience_size column, a new metric still under development. The demographic distribution variable from the Meta Ad Library API had 1357 instances of missing values. However, these missing values consistently displayed a minimum impression level, ranging from 0 to an upper limit of 999. These missing values might represent characteristics; they were retained in their original state. In other words, nine variables for advertising segmentation sourced from demographic distribution remained unpopulated.

In the second experiment focused on boosted post segmentation, data from the web crawler and Meta Ad Library API were combined. Boosted posts are a subtype of advertising, and some variables remained the same as in the first experiment. Posts can be promoted multiple times. However, each boosted post is recorded as a separate advertisement. To address this, the boosted posts for the same post were combined into a single data entry, and a "times" variable was added to indicate the number of times a post was boosted. Furthermore, the cost and impressions of each boosted post were aggregated, and the total number of regions reached and days delivered were computed.

Data from the Meta Ad Library Report and web crawler were combined with preprocessing fan page clustering. The exact page with different disclaimers was recorded as multiple entries in the report. To address this, data from the same page were combined, and the disclaimer column was transferred to account for different releases. Moreover, the total spending and total number of advertisements on the same page were calculated through aggregation.

3.3. Data analysis tools and techniques.

Clustering is a suitable method for analyzing unlabeled data. Because the datasets in this study were unlabeled, we only considered clustering techniques for analysis. Tan et al. (2019) revealed different categories of clustering, namely density-based, graph-based, and prototype-based clustering. The K-means approach was chosen due to its applicability to high-dimensional, sparse data and its effectiveness in customer segmentation research.

The K-means algorithm proceeds in several steps. K initial centroids are initially selected, where K is a user-defined parameter. Next, data points are assigned to their nearest centroid, resulting in the formation of K clusters. The centroids for each cluster are then recomputed, and finally, the assignment and centroid recomputation steps are repeated until the centroids no longer change. K-means clustering has been adopted in several customer segmentation studies (Shin & Sohn, 2004; Tsai et al., 2015). Both advertising and customer segmentations subdivide large amounts of high-dimensional data and make decisions based on the segmentation results. Therefore, the two have a certain degree of conceptual similarity. We believed that K-means clustering suited this study and adopted it for all clustering experiments.

Dolnicar et al. (2014) suggested an adequate sample size of at least seventy times the number of variables. In this study, each clustering experiment met the required minimum sample size.

This study employed a personal computer for data collection and analysis using Anaconda as the development environment. Anaconda was chosen for its package support and suitability for Windows. A Python environment was established using Python 3.8.12, and Jupyter Notebook was used for programming. Additional packages like CSV, Pandas, Json, Requests, Selenium, Bs4, and Emoji were used for data processing and collection.

For analysis, Weka, a Java-based machine learning software, was utilized. Clustering, specifically K-means, was chosen as it aligned with the research goals. The K value for K-means was determined through the Elbow method by comparing SSE.

The study ensured that the data size met the required minimum for clustering. Overall, the research's technical approach involved precise programming, data extraction, and appropriate clustering methods per its objectives.

4. Results and discussions

4.1. Segmentation results of advertisements.

The dataset comprised 34,533 advertisements, and almost all advertisements had creative bodies. In the demographic reach section, the average audience ratio aged 35 to 64 was higher than that of other age groups. The average male audience ratio was more than twice the female audience ratio.

The result of the first experiment conducted on the advertising dataset yielded three advertising segments. The number of clusters was determined by investigating the clustering results for two to eight clusters, resulting in three segments. The advertisements in these three segments were 25,842, 3682, and 5009, respectively.

Segment 1 comprised advertisements without link titles and descriptions. These advertisements had the lowest average spend and shortest average delivery period and reached the least number of regions among the three segments. Moreover, they had the lowest average number of impressions.

Segment 2 included advertisements with link titles but needed link captions and descriptions. This segment had the highest average impressions among the three segments, indicating it is the most effective advertising segment. However, it also had the highest average advertising spend due to increased impressions.

Segment 3 comprised advertisements with link captions, with most also having link descriptions and titles. Moreover, they had the longest average delivery period and the least average number of delivery platforms among the three segments.

The average impressions per dollar were 7.96, 10.22, and 6.93 for Segment 1, Segment 2, and Segment 3, respectively, indicating that advertisements in Segment 2 are the most effective and the most cost-efficient among the three segments.

The age distribution of each audience segment was like that in the original data. The average ratios of the age groups 35–44, 45–54, and 55–64 years were higher than those of other groups. Segment 1 had a higher reach ratio in the 65+ age group but was the least effective segment overall. Therefore, focusing only on older audiences may not be an efficient strategy. The most effective and efficient segment, segment 2, had a higher reach ratio in the middle and older age groups than other segments. However, the reach ratios were the lowest among the three segments in all three younger age groups. Therefore, focusing on these two age groups may enhance advertising performance. Segment 3, however, had the highest average reach ratios in the age groups 18–24, 25–34, and 35–44 years.

These results reveal that an advertising strategy is more effective if more regions are reached, and more money is spent. However, advertising on more platforms and increasing the duration of advertisements could have improved advertising performance. Similar to Zhang et al. (2021), we can conduct a cluster analysis of the advertisements on Facebook to identify their characteristics. In addition, we can further apply these characteristics to the business strategies of organizations and even individuals, as suggested by Gull et al. (2014).

4.2. Boosted posts and feedback

The analysis revealed a total of 19,663 boosted posts. These boosted posts have creative bodies, but only a few include link captions, descriptions, and titles. The number of clusters was determined by examining the clustering results across two to eight clusters. Ultimately, the analysis identified three segments. The boosted posts in these three segments were 3504, 4819, and 11,340, respectively.

The results aid in identifying the segment that generates a greater response from the audience and garners more impressions. In Segment 1, almost all boosted posts included link titles, and the ratio of having link captions and descriptions was also the highest among the three segments. Furthermore, boosted posts in segment 1 had the highest average number of times they were boosted, impressions, delivery period, and spending among the three segments. Moreover, they had the highest average number of comments and shares. Thus, segment one can be considered the most effective for boosted posts.

Segment Two and Segment Three had boosted posts with link titles; only a few included link captions and descriptions. Boosted posts in Segment 2 had the highest average delivery regions and the second-highest average number of comments and shares among the three segments.

Boosted posts in segment 3 exhibited the poorest performance regarding the number of comments, shares, and emotional reactions. However, their average spending was the highest. Therefore, this segment was defined as insufficient boosted posts.

The average impressions per dollar were 9.31, 9.59, and 7.68 for segment 1, segment 2, and segment 3, respectively. Although the boosted posts in Segment 1 were the most effective among the three segments, they could have been more efficient regarding the cost of impressions. By contrast, segment 2 was the most cost-efficient.

The average number of *like* reactions was considerably higher than the other reactions in each segment. Giuntini et al. (2019) reported that a *like* reaction can indicate positivity, negativity, or neutrality, making it all-encompassing. This may explain why the number of *like* reactions was the highest among all reaction types observed. In the most effective segment, Segment 1, the number of reactions was the highest for all emotions except *like* and *wow*. However, although Segment 2 had the lowest spending, the number of reactions to each emotion was higher than in Segment 3.

Kim and Yang (2017) proposed that reactions, comments, and shares represent different levels of engagement on social media. Emotions can be expressed with a simple click of an emoji, whereas leaving a comment requires typing. Therefore, comments are considered to have higher engagement than emotions. Sharing a post involves displaying it on one's profile page, contributing to self-presentation. The boosted posts in Segment 1 had the largest average number of comments and shares. Regarding the average number of shares, considered the highest level of audience engagement, Segment 1 outperformed Segment two and Segment three more than twice and four times, respectively. In addition to being the most effective in generating impressions, Segment 1 was also the most effective in audience engagement based on the highest number of comments, shares, and emotions. Segment 2, which had the lowest spending, ranked second out of the three segments regarding the number of comments and shares and had nearly twice as many shares as Segment 3. Therefore, segment two can be considered the most efficient in terms of engagement and cost- effectiveness. Clustering analysis can be applied to multimedia content and advertisements on social networking sites like Facebook. The results of this research align with similar research by Bhat and Malaganve (2022), which explored Facebook advertising characteristics through cluster analysis. These characteristics can be further used in the business strategies of organizations or even individuals, as demonstrated by Gull et al. (2014).

4.3. Correlation between the advertising pattern and social influence

The dataset analyzed covered 531 fan pages. More than half of these fan pages belong to the category of politicians, and almost half of them have been verified by Facebook.

After page clustering, three clusters of pages were identified. The pages in each cluster were 231, 235, and 65, respectively. Among the three clusters, cluster 1 exhibited the lowest average number of disclaimers, advertisements, and followers. The average total spending on advertisements was also the lowest in this cluster. Moreover, none of the fan pages in cluster 1 had a verified badge by Facebook. By contrast, all the fan pages in cluster 2 have been verified by Facebook. Furthermore, cluster 2 exhibited the highest average number of disclaimers, advertisements, and followers and the highest average total spending. The values of the attributes of cluster 3 were between cluster 1 and cluster 2, which is the medium across the analyzed variables.

Research has shown that central nodes in online social networks tend to possess considerable social influence (Zhang et al., 2011). The number of followers a fan page has can measure its centrality in the social network. That is, the page with more followers has more social influence. Thus, the fan pages in cluster 2 have the highest social influence on online social networks among the three clusters.

In cluster 1, 70% of the fan pages were categorized as those of politicians, and nearly 20% of pages belonged to other page types. In cluster 2, with the highest social influence, more than 90% of the fan pages belonged to the politician category, with less than 5% belonging to other category types. In cluster 3, all the fan page types were categorized as public figures. These findings indicate that the fan pages in cluster 2, which had the greatest social influence, were predominantly concentrated in the politician category.

Figure 2 illustrates the distribution of advertising segments within each cluster of fan pages. The fan pages in cluster 2 had the highest number of advertisements. Across all clusters, the percentages of segment 2, which is the most effective and efficient advertising segment, were the lowest among the three segments. Specifically, the rates of advertisements in Segment 1 were 9.52%, 11.23%, and 10.22 %for each page cluster. The cluster with the highest social influence had the highest percentage of advertisements in segment 2. These findings suggest that ads' distribution considerably affects fan pages' social impact and that there is substantial room for improvement in the social advertising strategies employed by fan pages.

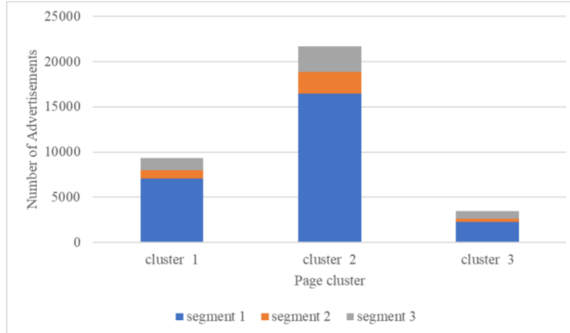

Figure 2. Distribution of advertising segments in each cluster of fan pages.

The results indicate that fan pages with higher spending and total advertisement and disclaimers have more significant social influence on Facebook. Furthermore, the verified badge Facebook provides is a crucial determinant of social impact. Finally, setting fan pages within the politician category enhances the social influence of personal brands in the political field. Corporate advertising is advantageous for building one's power and brand image (Dehghani & Turner, 2015), and personal-based advertising management can yield similar benefits.

5. Conclusion and future study

5.1. Conclusion and contributions

The importance of political figures establishing and maintaining their image on social media is becoming increasingly evident. In today's digital age, social media has become one of the primary platforms for interacting with voters, conveying political messages, and shaping public impressions. By effectively utilizing social media, political figures can directly communicate with voters, share their policies and stances, and showcase their personality traits and leadership abilities. This not only helps in building closeness and trust but also attracts more supporters and followers.

This study provides a general guideline for influencing investment on social media platforms. It identified effective advertising patterns based on spending and regions reached, emphasizing their influence on advertising performance. By following this guideline, individuals and brands can make more precise investments. Moreover, the study identified the most effective boosted posts regarding impressions and audience engagement. Boosted posts with higher frequencies, spending, and delivery durations are more effective. Additionally, the most efficient segment of boosted posts corresponds to the regions with the highest reach. Based on these findings, we suggest that individuals adjust their approach to boosting posts according to their budget to strike the best balance between cost and advertising performance. Regarding page clustering, we found that the characteristics of fan pages with the highest social influence were higher total advertising numbers, advertising amounts, and disclaimer numbers. Therefore, authenticity verification through these factors was identified as a key factor in advertising performance in this study.

This study's findings contribute to research and the public and private sectors. The research provides a valuable social network advertising guideline for industry personnel, particularly those focused on personal branding. Government officials, including those responsible for managing personal brands and government organizations, can also benefit from this research. It enabled them to adapt their existing strategies, utilize limited resources more effectively, and invest more accurately in online social media platforms to maximize social influence. Following our research findings, efficient social media enables political figures to disseminate information and influence a large audience rapidly. By posting various forms of content, political figures can present a more diverse and enriched image to cater to the needs and preferences of different audiences, thereby enhancing their presence and influence in the public eye.

For scholars, this research provides insights into a relatively underexplored field—influence investment in online social networking platforms. Furthermore, this research establishes a link between two fields rarely explored by scholars, namely personal branding and social network advertising, thereby contributing to efforts at fostering interdisciplinarity.

5.2. Limitations and future research direction

This study has three limitations.

The first limitation is data accessibility. The privacy policies of social networking platforms such as Meta restrict access to certain information, including responses from other users, under public posts on fan pages. Consequently, this research's existing data mining techniques may need to capture the complete picture. Moreover, meaningless or negative comments are also counted as users' active responses, which can lead to overestimating the effectiveness of advertising on social media. To address this limitation, future research can employ advanced data mining techniques and sentiment analysis to screen for positive responses, enabling a more accurate assessment.

The second limitation is that the algorithm was as- is in the absence of any further optimization. Analyzing the differences between algorithms was beyond this study's scope. Therefore, future research can explore the applicability of various algorithms in the social media advertising segmentation scenario and even design a model specifically for this scenario.

Finally, this research primarily focused on the domain of politics. The results of using social advertising for personal branding may vary across different fields. Therefore, to gain a more comprehensive understanding, future research can explore other fields of personal branding and compare the outcomes between different fields.

References

Alpaydin, E. (2020). *Introduction to machine learning* (4 ed.). The MIT Press.

Andranik Tumasjan, Timm O. Sprenger, Philipp G. Sandner, and Isabell M. Welpe. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. In Proceedings of the 4th International Conference on Weblogs and Social Media (ICWSM '10). AAAI, 178-185

Bakshy, E., Hofman, J. M., Mason, W. A. & Watts, D. J. (2011). Everyone's an influencer: quantifying influence on Twitter. In WSDM '11: Proc. Fourth ACM International Conference on Web Search and Data Mining 65–74.

Brendan O'Connor, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. (2010). From tweets to polls: Linking text sentiment to public opinion time series. In Proceedings of the 4th International Conference on Weblogs and Social Media (ICWSM '10). AAAI, 122-129

Bryman, A. (2012). Social research methods: Oxford University Press.

Bhat, P., & Malaganve, P. (2022). A Study on Clustering Facebook Multimedia Based on Metadata—A Comparative Analysis. In R. Agrawal, J. He, E. Shubhakar Pilli, & S. Kumar (Eds.), *Cyber Security in Intelligent Computing and Communications* (pp. 73- 83). Springer Singapore.

David M. Blei, Andrew Y. Ng, Michael I. Jordan. (2003). Latent Dirichlet Allocation. J. Mach. Learn. Res. 3 (January 2003), 993-1022.

Dehghani, M., & Turner, M. (2015). A research on effectiveness of Facebook advertising on enhancing purchase intention of consumers. *Computers in Human Behavior*, *49*, 597-600.

Diego Tumitan and Karin Becker. 2014. Sentiment-Based Features for Predicting Election Polls: A Case Study on the Brazilian Scenario. In Proceedings of the IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies - Volume 02 (WI-IAT '14), Vol. 2. IEEE Computer Society, Washington DC, USA, 126- 133.

Dolnicar, S., Grün, B., Leisch, F., & Schmidt, K. (2014). Required Sample Sizes for Data-Driven Market Segmentation Analyses in Tourism. *Journal of Travel Research*, *53*(3), 296-306.

Eagle, N., Pentland, A., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(36), 15274-15278.

Ernst, N., Esser, F., Blassnig, S., & Engesser, S. (2019). Favorable opportunity structures for populist communication: Comparing different types of politicians and issues in social media, television and the press. The International Journal of Press/Politics, 24(2), 165-188.

Farkas, X., & Bene, M. (2021). Images, politicians, and social media: Patterns and effects of politicians' image-based political communication strategies on social media. The international journal of press/politics, 26(1), 119-142.

Garg, A., Vijayaraghavan, V., Tai, K., Singru, P. M., Jain, V., & Krishnakumar, N. (2015). Model development based on evolutionary framework for condition monitoring of a lathe machine. *Measurement*, *73*, 95-110.

Geva, H., Oestreicher-Singer, G., & Saar-Tsechansky, M. (2019). Using retweets when shaping our online persona: Topic modeling approach. *MIS Quarterly*, *43*(2), 501-524.

Giuntini, F. T., Ruiz, L. P., Kirchner, L. D., Passarelli, D. A., Dos Reis, M. D. D., Campbell, A. T., & Ueyama, J. (2019). How Do I Feel? Identifying Emotional Expressions on Facebook Reactions Using Clustering Mechanism. *IEEE Access*, *7*, 53909-53921.

Gull, K. C., Angadi, A. B., Seema, C. G., & Kanakaraddi, S. G. (2014, 21-22 Feb. 2014). A clustering technique to increase marketing tactics by looking out the key users taking Facebook as a case study. 2014 IEEE International Advance Computing Conference (IACC),

He, Q., Wang, X. W., Zhao, Y., Yi, B., Lu, X. J., Yang, M. Z., & Huang, M. (2021). Reinforcement-Learning-Based Competitive Opinion Maximization Approach in Signed Social Networks. *IEEE Transactions on Computational Social Systems*.

Jacobson, J. (2020). You are a brand: social media managers' branding and "the future audience". Journal of Product & Brand Management, 29(6), 715-727.

Kempe, D. J. K., Ev´ a Tardos. (2003). Maximizing the Spread of Influence through a Social Network.

Khedher, M. (2014). Personal Branding Phenomenon. *International Journal of Information, Business and Management*, *6*(2), 29-40.

Kim, C., & Yang, S. U. (2017). Like, comment, and share on Facebook: How each behavior differs. *Public Relations Review*, *43*(2), 441-449.

Kiran, B. R., Sobh, I., Talpaert, V., Mannion, P., Sallab, A. A. A., Yogamani, S., & Perez, P. (2022). Deep Reinforcement Learning for Autonomous Driving: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, *23*(6), 4909-4926.

Krishen, A. S., Dwivedi, Y. K., Bindu, N., & Kumar, K. S. (2021). A broad overview of interactive digital marketing: A bibliometric network analysis. Journal of Business Research, 131, 183-195.

Labrecque, L. I., Markos, E., & Milne, G. R. (2011). Online Personal Branding: Processes, Challenges, and Implications. *Journal of Interactive Marketing*, *25*(1), 37-50.

Lalancette, M., & Raynauld, V. (2019). The power of the political image: Justin Trudeau, Instagram, and celebrity politics. American behavioral scientist, 63(7), 888-924.

Li, H., Xu, M. T., Bhowmick, S. S., Rayhan, J. S., Sun, C. S., & Cui, J. T. (2022). PIANO: Influence Maximization Meets Deep Reinforcement Learning. *IEEE Transactions on Computational Social Systems*.

Li, H. F., Parikh, D., He, Q., Qian, B. Y., Li, Z. G., Fang, D. P., & Hampapur, A. (2014). Improving rail network velocity: A machine learning approach to predictive maintenance. *Transportation Research Part C- Emerging Technologies*, *45*, 17-26.

Marco Pennacchiotti and Ana-Maria Popescu. (2011). Democrats, republicans and starbucks afficionados: user classification in twitter. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '11). ACM, New York, NY, USA, 430-438.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529-533.

Moulard, J. G., Garrity, C. P., & Rice, D. H. (2015). What Makes a Human Brand Authentic? Identifying the Antecedents of Celebrity Authenticity. *Psychology & Marketing*, *32*(2), 173-186.

Peters, T. (1997). *The Brand Called You*. Retrieved 05/26 from https://www.fastcompany.com/28905/brand-called-you

Pineau, J., Guez, A., Vincent, R., Panuccio, G., & Avoli, M. (2009). TREATING EPILEPSY VIA ADAPTIVE NEUROSTIMULATION: A REINFORCEMENT LEARNING APPROACH. *International Journal of Neural Systems*, *19*(4), 227-240.

Quan, T. T., Mai, D. T., & Tran, T. D. (2020). CID: Categorical Influencer Detection on microtext-based social media. *Online Information Review*.

Rangarajan, D., Gelb, B. D., & Vandaveer, A. (2017). Strategic personal branding—And how it pays

off. *Business Horizons*, *60*(5), 657-666.
Reisach, U. (2021). The responsibility of social media in times of societal and political manipulation. European journal of operational research, 291(3), 906-917.
Ressom, H. W., Varghese, R. S., Zhang, Z., Xuan, J. H., & Clarke, R. (2008). Classification algorithms for phenotype prediction in genomics and proteomics. *Frontiers in Bioscience-Landmark, 13*, 691-708.
Rios, S. A., Aguilera, F., Nunez-Gonzalez, J. D., & Grana, M. (2019). Semantically enhanced network analysis for influencer identification in online social networks. *Neurocomputing*, *326*, 71-81.
Scheidt, S., Gelhard, C., & Henseler, J. (2020). Old Practice, but Young Research Field: A Systematic Bibliographic Review of Personal Branding. *Front Psychol*, *11*, 1809.
Scheitle, C. P., & Ecklund, E. H. (2017). The influence of science popularizers on the public's view of religion and science: An experimental assessment. *Public Understanding of Science*, *26*(1), 25-39.
Shafiee, M., Gheidi, S., & Khorrami, M. S. (2020). We are proposing a new framework for personal brand positioning. *European research on management and business economics*, *26*(1), 45-54.
Shawn O'Banion and Larry Birnbaum. (2013). Using explicit linguistic expressions of preference in social media to predict voting behavior. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM '13). ACM, New York, NY, USA, 207-214.
Shin, H. W., & Sohn, S. Y. (2004). Segmentation of stock trading customers according to potential value. *Expert Systems with Applications*, *27*(1), 27-33.
Sun, Z. H., Zuo, T. Y., Liang, D., Ming, X. G., Chen, Z. H., & Qiu, S. Q. (2021). GPHC: A heuristic clustering method to customer segmentation. *Applied Soft Computing, 111*, Article 107677.
Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2015). Machine Learning for Predictive Maintenance: A Multiple Classifier Approach. *Ieee Transactions on Industrial Informatics*, *11*(3), 812-820.
Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). *Introduction to Data Mining* (2nd ed.).
Trehan, D., & Sharma, R. (2020). Assessing advertisement quality on C2C social commerce platforms: an information quality approach using text mining. *Online Information Review*.
Tsai, C. F., Hu, Y. H., & Lu, Y. H. (2015). Customer segmentation issues and strategies for an automobile dealership with two clustering techniques. *Expert Systems*, *32*(1), 65-76.
Tsai, C. F., & Wang, S. P. (2009). Stock price forecasting by hybrid machine learning techniques. Proceedings of the international multiconference of engineers and computer scientists,
Wang, F., Jiang, W., Li, X., & Wang, G. (2018). Maximizing positive influence spread in online social networks via fluid dynamics. *Future Generation Computer Systems*, *86*, 1491-1502.
Wang, F., She, J., Ohyama, Y., Jiang, W., Min, G., Wang, G., & Wu, M. (2021). Maximizing positive influence in competitive social networks: A trust-based solution. *Information Sciences*, *546*, 559-572.
Wroblewski, L., & Grzesiak, M. (2020). The Impact of Social Media on the Brand Capital of Famous People. *Sustainability*, *12*(16), Article 6414.
Zhang, K., Du, H., & Feldman, M. W. (2017). Maximizing influence in a social network: Improved results using a genetic algorithm. *Physica A: Statistical Mechanics and its Applications*, *478*, 20-30.
Zhao, Y. F., Zeng, D. L., Socinski, M. A., & Kosorok, M. R. (2011). Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. *Biometrics*, *67*(4), 1422-1433.
Zhang, X., Habibi Lashkari, A., & Ghorbani, A. A. (2021). Classifying and clustering malicious advertisement uniform resource locators using deep learning. *Computational Intelligence*, *37*(1), 511-537.
Zhang, Y. C., Liu, Y., Zhang, H. F., Cheng, H., & Xiong, F. (2011). The research of information dissemination model on online social network. *Acta Physica Sinica*, *60*(5), Article 050501.
Zheng, X., Zhong, Y., Zeng, D., & Wang, F. Y. (2012). Social influence and spread dynamics in social networks. *Front Comput Sci*, *6*(5), 611-620.
Zhuravskaya, E., Petrova, M., & Enikolopov, R. (2020). Political effects of the Internet and social media. Annual review of economics, 12, 415-438.