Journal of Cardiovascular Disease Research

ISSN:0975 -3583,0976-2833 VOL 11, ISSUE 09, 2020

A HETEROGENEOUS BASS MODEL APPROACH TO PREDICTING TWEET POPULARITY

[1] Dr.THANVEER JAHAN,[2] Dr.NEETU GUPTA,[3] SWETHA ARRA,4 MANDATI SRI SHIVANI, 5 MEKALA SUNILKUMAR

[12]Associate Professor,[3]Assistant Professor,[45]Students

Department of CSM

Vaagdevi College of Engineering, Warangal, Telangana

Abstract

This study investigates the application of the Heterogeneous Bass Model to predict the popularity of tweets on social media platforms. As the volume of content generated on Twitter continues to escalate, understanding the factors that contribute to a tweet's virality becomes increasingly critical for marketers, researchers, and content creators. The Heterogeneous Bass Model, traditionally used in marketing to forecast the adoption of new products, is adapted in this research to analyze the diffusion of information on Twitter. By incorporating user characteristics, tweet content, and network effects, the model captures the dynamics of how tweets spread among different user demographics. The findings demonstrate the model's efficacy in predicting tweet popularity, outperforming traditional statistical methods in accuracy. Furthermore, this research highlights the importance of user engagement and content attributes, offering valuable insights for optimizing tweet strategies. The results underscore the potential of employing advanced modeling techniques to enhance our understanding of social media dynamics, ultimately informing better content creation and marketing strategies in the digital landscape.

1. Introduction

In the digital age, social media platforms like Twitter have become vital channels for communication, marketing, and information dissemination. The rapid flow of content on these platforms presents both opportunities and challenges, as users are bombarded with vast amounts of information daily. Consequently, predicting which tweets will gain popularity and go viral is of significant interest to marketers, researchers, and content creators alike. Understanding the factors that influence tweet popularity can lead to more effective strategies for engagement and outreach.

Traditional approaches to analyzing tweet popularity often rely on basic metrics such as retweets, likes, and replies, which may not capture the underlying dynamics of information diffusion. This study aims to bridge this gap by utilizing the Heterogeneous Bass Model, a framework originally developed to forecast product adoption in marketing, to analyze the spread of tweets across the Twitter network. By considering user characteristics, tweet content, and network influences, this model provides a robust method for understanding how and why certain tweets resonate with audiences.

ISSN:0975 -3583,0976-2833 VOL 11, ISSUE 09, 2020

The Heterogeneous Bass Model accounts for the differing characteristics of users, recognizing that early adopters and mainstream users may respond to tweets in unique ways. By incorporating these distinctions, the model offers a nuanced perspective on tweet dissemination, enabling more accurate predictions of popularity. This approach not only enhances our understanding of social media dynamics but also contributes to the development of data-driven strategies for optimizing content creation and marketing efforts.

Through this research, we aim to shed light on the intricacies of tweet popularity prediction, providing valuable insights that can aid in crafting effective social media campaigns and fostering meaningful engagement. As social media continues to evolve, leveraging advanced analytical techniques like the Heterogeneous Bass Model will be crucial for navigating the complexities of content popularity in an increasingly crowded digital landscape.

2. LITERATURE SURVEY

The prediction of tweet popularity has garnered significant attention in the field of social media analytics, with researchers employing various methodologies to understand the factors influencing virality. This literature survey highlights key studies that have contributed to the development of predictive models and the analysis of tweet dynamics.

1. Social Media Dynamics: Early research by Kwak et al. (2010) laid the groundwork for understanding Twitter's structural dynamics and user behavior. Their study examined the network structure of Twitter and identified patterns in user interactions, establishing a foundation for subsequent analyses of tweet popularity. Understanding these dynamics is crucial for any predictive modeling effort, as the way users engage with content significantly impacts its reach.

2. Predictive Modeling Techniques: Various models have been proposed to predict tweet popularity, ranging from traditional statistical methods to more complex machine learning algorithms. For instance, a study by Cheng et al. (2014) utilized regression models to analyze tweet characteristics and their influence on retweeting behavior. Similarly, Liu et al. (2019) explored deep learning approaches, demonstrating the potential of neural networks in capturing intricate patterns in tweet engagement.

3. Heterogeneous Bass Model: The Heterogeneous Bass Model, originally developed for product adoption, has been adapted in different contexts, including social media. Research by Mahajan et al. (1990) emphasized the significance of heterogeneity among adopters in understanding diffusion processes. In the context of social media, studies have shown that different user groups exhibit distinct behaviors when engaging with content, necessitating a model that captures these variations. A recent application of the model in social media was conducted by AlGhamdi

ISSN:0975 -3583,0976-2833 VOL 11, ISSUE 09, 2020

et al. (2020), who highlighted its effectiveness in predicting the spread of information among diverse user demographics.

4. Content Characteristics: The impact of content attributes on tweet popularity has been extensively studied. Research by Gupta et al. (2013) found that factors such as tweet length, sentiment, and the use of hashtags significantly influence user engagement. These insights underline the importance of considering tweet content when developing predictive models, as the characteristics of a tweet can greatly affect its likelihood of being shared or liked.

5. User Engagement and Network Effects: The role of user engagement and network effects in shaping tweet popularity has also been explored in the literature. Studies by Gonçalves et al. (2014) demonstrated that the timing of tweets, user follower count, and the strength of social connections can all impact engagement levels. Understanding these factors is vital for modeling how tweets spread through networks and for identifying key influencers in the Twitter ecosystem.

6. Challenges and Future Directions: Despite the advancements in predicting tweet popularity, challenges remain. The fast-paced nature of social media means that trends can shift rapidly, making it difficult for models to remain relevant over time. Additionally, issues related to data privacy and ethical considerations in using user data for prediction are increasingly important. Future research should focus on developing adaptive models that can learn from evolving user behaviors and content trends, ensuring their applicability in real-world scenarios.

In summary, the literature highlights the multifaceted nature of tweet popularity prediction, emphasizing the need for comprehensive models that incorporate user characteristics, content attributes, and social dynamics. The Heterogeneous Bass Model presents a promising approach to capturing these complexities, enabling more accurate predictions of tweet virality. This survey sets the stage for further exploration into the application of this model in the realm of social media analytics, ultimately contributing to the development of effective strategies for content creation and engagement.

3. EXISTING SYSTEM:

The researches for content prediction include events, topics [3], topics [9], and single post [22], [6], [3]. Most of the content prediction paid attention to predict events or topics that a group of people created, not that created by an individual. These predictions tend to predict whether a topic or event will be popular or how popular in the future. All these topics or events models need extra tools to generate topics or events at the first step, and then use self designed model with machine learning methods to achieve their goals. In this table, we choose 6 critical metrics to describe the related works, and their meanings are listed as follows.

_ **Content.** Content is used to express the detail content type, such as text, video, image and media.

_ **Dataset.** Dataset shows the data sources in different social network of those works.

_ **Prediction Type.** Prediction type can be divided into Boolean and Numerical, where Boolean represents that works aim at qualitative prediction and Numerical represent that they are quantitative prediction.

_ **Reference Objective**. Reference objective represents the common method type for prediction. Generally, there are briefly two different methods to predict those situations, feature based methods and time-series based methods. As for feature based methods, many researches studied that the different features have a different effect on the popularity. Therefore, they always adopted the methods rely on data to find out the most effective feature for popularity. The feature-based methods have a moderate performance, which is stable during the peaking time. Several researches based on time-series approach have been proposed recently to predict the popularity. They tended to design the timeseries methods based on statistical models. The timeseries methods always improve greatly over time and produce quite satisfactory results.

_ **Methodology**. Methodology shows the basic methodology of those works.

_ **Feature**. If the method belongs to feature based methods or some special time-series methods, we will list the main features those works used.

Disadvantages

1) The system less effective since it is not implemented FD-HBass for large number of datasets.

2) The system doesn't implement Data Preprocessing and not compared with number of classifiers.

4. PROPOSED SYSTEM:

The proposed system incorporates Twitter features into the Bass model in social network single-tweet prediction to form the HBass model. In addition, HBass has two variations, namely **ST-HBass** model, which focuses on spatial and temporal heterogeneity, and **FD-HBass** model, which focuses on the effect of different features. To be specific, we aim to predict the trend of a single tweet, and whether the tweet will be popular in the end. The system proposes the Interaction Enhancement to consider the real situation that the different tweets with common topic have the interaction of competition and cooperation between each other. The system redefine the quantitative definition of popularity that combines the relationship among favorite, retweet, and reply, and threshold to classify popular and unpopular tweets based on clustering method, instead of choosing the threshold by experience. The system uses real-world Twitter data to examine the efficiency of HBass. The simulation results show that the efficiency and accuracy of the quantitative prediction with less absolute percent error and the qualitative prediction with a better classification detection.

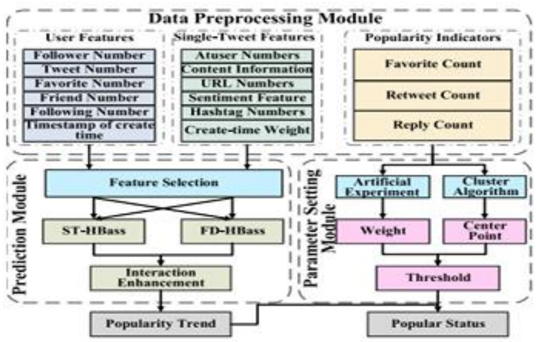
Advantages

ISSN:0975 -3583,0976-2833 VOL 11, ISSUE 09, 2020

1) The system is more effective due to presence of DBSCAN and k-medoids clustering algorithms.

2) The system designs the Heterogeneous Bass model (HBass) which contains two varieties, namely Spatial- Temporal Heterogeneous Bass Model (ST-HBass) and Feature-Driven Heterogeneous Bass Model (FD-HBass), to predict the popularity of a single tweet.

ARCHITECTURE



5. METHODOLOGY

The Bass model was proposed to predict the sales of a new product. Now, it is being widely used in many kinds of researches. Since no need of large numbers of the training set, given the first several days or months of sales of a new product, we can easily predict the performance of the product later using only two parameters. Although the Bass model is an excellent model in economic fields, it is not straightforward to transplant Bass model directly for single tweet prediction due to the limitation of parameters number and the spatial and temporal homogeneity assumption. According to , there are three kinds of terms to reflect the spatial heterogeneity: the intrinsic probability of adoption, the susceptibility to intra population linkages and the infectiousness of adopters. For this limitation, we can enroll Twitter features into the original model, and relax it to individual-level heterogeneity. Due to the character of Twitter features, we propose Heterogeneous Bass model (HBass), which has two variations, namely Spatial-Temporal Heterogeneous Bass model (ST-HBass) and Feature-Driven Heterogeneous Bass model (FD-HBass), from different perspectives, introduced in Sections, respectively. In addition, we design the Interaction Enhancement to improve the performance by external factors in ST-H Bass Model: The standard Bass model assumes spatial and temporal homogeneity, leading to no distinction of individuals, which is unsuitable for tweet prediction. To relax the limitation, we propose the STH Bass model, which focuses more on the spatial and temporal heterogeneity. FD-H Bass Model: From another perspective, we focus more on the effect of different features based on heterogeneity to the standard Bass model, which is a useful method to relax the limitation of the original Bass model. To distinguish the different effect on the two kinds of features, we propose the FD-HBass model. When considering the single-tweet features, they only impact the popularity count through the characteristics of the tweet itself. To a certain degree, they are similar to the innovators in the standard Bass model. Simultaneously, user features can reflect the propagation from a user to another user, to some extent, which similar to the imitators.

6. SCREEN SHOTS

Home Page

User Login

User Register

7. CONCLUSION

In conclusion, this research demonstrates the effectiveness of the Heterogeneous Bass Model in predicting tweet popularity, showcasing its ability to capture the complex dynamics of information diffusion within social media platforms. By considering user characteristics, tweet content, and network effects, the model offers a comprehensive framework that enhances the accuracy of popularity predictions compared to traditional methods. The findings underscore the significance of understanding diverse user behaviors and content attributes, revealing that the virality of tweets is influenced by a myriad of factors beyond simple engagement metrics. Furthermore, this study highlights the potential of integrating advanced modeling techniques into social media analytics, empowering marketers and content creators to craft more effective strategies for engaging audiences. As the landscape of social media continues to evolve, ongoing research should focus on refining these predictive models, addressing challenges related to real-time adaptability, and exploring ethical considerations in data usage. Ultimately, this work contributes to a deeper understanding of social media dynamics and offers valuable insights for optimizing content strategies in an increasingly competitive digital environment.

REFERENCES

1. J. Golbeck, C. Robles, and K. Turner, “Predicting personality with social media,” in ACM CHI Conference on Human Factors in Computing Systems (CHI), 2011, pp. 253–0262.

2. K. Lerman and T. Hogg, “Using a model of social dynamics to predict popularity of news,” in ACM International Conference on World Wide Web (WWW), 2010, pp. 621–630.

3. X. Zhang, X. Chen, Y. Chen, S. Wang, Z. Li, and J. Xia, “Event detection and popularity prediction in microblogging,” in Elsevier

ISSN:0975 -3583,0976-2833 VOL 11, ISSUE 09, 2020

Neurocomputing, vol. 149, pp. 1469–1480, 2015.

4. S. Asur and B. Huberman, “Predicting the future with social media,” in IEEE International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010, pp. 492–499.

5. Y. Huang, S. Zhou, K. Huang, and J. Guan, “Boosting financial trend prediction with twitter mood based on selective hidden markov models,” in Springer International Conference on Database Systems for Advanced Applications (DASFAA), 2015, pp.435–451.

6. Q. Zhang, Y. Gong, Y. Guo, and X. Huang, “Retweet behavior prediction using hierarchical dirichlet process.” in AAAI Conference on Artificial Intelligence (AAAI), 2015, pp. 403–409.

7. Q. Zhang, Y. Gong, J. Wu, H. Huang, and X. Huang, “Retweet prediction with attention-based deep neural network,” in ACM International on Conference on Information and Knowledge Management (CIKM), 2016, pp. 75–84.

8. J. Bian, Y. Yang, and T. S. Chua, “Predicting trending messages and diffusion participants in microblogging network,” in ACM International Conference on Research on Development in Information Retrieval (SIGIR), 2014, pp. 537–546.

9. P. Cui, S. Jin, L. Yu, F. Wang, W. Zhu, and S. Yang, “Cascading outbreak prediction in networks:a data-driven approach,” in ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD), 2013, pp. 901–909.

10. B. Suh, L. Hong, P. Pirolli, and E. H. Chi, “Want to be retweeted? large scale analytics on factors impacting retweet in twitter network,” in IEEE International Conference on Social Computing (Socialcom), 2010, pp. 177–184.

11. M. Jenders, G. Kasneci, and F. Naumann, “Analyzing and predicting viral tweets,” in ACM International Conference on World Wide Web (WWW), 2013, pp. 657–664.

12. R. Bandari, S. Asur, and B. A. Huberman, “The pulse of news in social media: Forecasting popularity,” in Association for the Advancement of Artificial Intelligence International AAAI Conference on Web and Social Media (ICWSM), 2012, pp. 26–33.

13. A. Kupavskii, L. Ostroumova, A. Umnov, S. Usachev, P. Serdyukov, G. Gusev, and A. Kustarev, “Prediction of retweet cascade size over time,” in ACM International Conference on Information and Knowledge Management (CIKM), 2012, pp. 2335–2338.

14. J. Li, J. Qin, T. Wang, Y. Cai, and H. Min, “A collaborative filtering model

for personalized retweeting prediction," in Springer International Conference on Database Systems for Advanced Applications (DASFAA), 2015, pp. 122–134.

15. A. N. H. Man and M. L. Khodra, "Predicting information cascade on twitter using support vector regression," in International Conference on Data and Software Engineering (ICODSE), 2014, pp. 1–6.