IJITCE

ISSN 2347- 3657

# International Journal of Information Technology & Computer Engineering


Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 3, 2024

# Cross-Platform Reputation Generation System Based on Aspect-Based Sentiment Analysis

[1]P MOUNIKA, [2]Y.CHAITHANYA SAI

[1](Assistant Professor), MCA, DNR college(A) PG courses Bhimavaram
[2]MCA, scholar, DNR college(A) PG courses Bhimavaram

## ABSTRACT

The active growth of Internet-based applications such as social networks and e-commercewebsites leads people to generate a tremendous amount of opinions and reviews about products and services.Thus, it becomes very crucial to automatically process them. Over the last ten years, many systems have beenproposed to generate and visualize reputation by mining textual and numerical reviews. However, they haveneglected the fact that online reviews could be posted by malicious users that intend to affect the reputation ofthe target product. Besides, these systems provide an overall reputation value toward the entity and disregardgenerating reputation scores toward each aspect of the product. Therefore, we developed a system thatincorporates spam filtering, review popularity, review posting time, and aspect-based sentiment analysis togenerate accurate and reliable reputation values. The proposed model computes numerical reputation valuesfor an entity and its aspects based on opinions collected from various platforms. Our proposed system alsooffers an advanced visualization tool that displays detailed information about its output. Experiment resultsconducted on multiple datasets collected from various platforms (Twitter, Facebook, Amazon . . . ) show theefficacy of the proposed system compared with state-of-the-art reputation generation systems.

Machine learning is an important component of the growing field of data science. Through the use of statistical methods, different type of algorithms is trained to make classifications or predictions, and to uncover key insights in this project. These insights subsequently drive decision making

ISSN 2347–3657
Volume 12, Issue 3, 2024

within applications and businesses, ideally impacting key growth metrics.

Machine learning algorithms build a model based on this project data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of datasets, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

## I. INTRODUCTION

Having easy access to the web has radically changed the way people interact with brands and products. From physical products to online services, people tend to instantly share their opinions and reviews on various platforms on the Internet. A recent research experiment1 shows that consumers are more willing to share a review when the experience they have had evokes emotions, whether positive or negative. This large volume of consumers' reviews holds insightful information about the quality of the product/service, therefore analyzing them will help consumers make a better judgment toward the targeted item. In the past few years, a new subfield of natural language processing (NLP) called reputation generation has been well-established as an area of interest. The main focus of reputation generation systems is to produce a numerical value in which an entity is held based on mining customer reviews and their numerical ratings.

Over the last decade, many reputation generation systems have been proposed [1]–[8] to generate and visualize reputation of online products and services based on fusing and mining textual and numerical reviews. However, these systems have not taken into consideration (1) extracting and processing reviews from various platforms, (2) filtering reviews written by potential spammers, (3) generating a numerical reputation value toward each aspect of the target product, and, (4) providing an advanced reputation visualization tool for a better decision-making process. Thereby, we designed and built an upgraded reputation generation model that overcomes the shortcomings of the previous systems in order to compute and visualize the reputation of an entity (product, movie, hotel, restaurant, service) with consistent reliability. The proposed system collects and processes data from both e-commerce and social media platforms. Then, a spam filtering system is applied to eliminate spam reviews and prepare the cleaned output for aspect-based sentiment analysis (ABSA), where aspects of the target entity are extracted from the reviews with their sentiment polarities. Later, the time and popularity features of the reviews are exploited along with the ASBA results to finally generate a reputation value of each aspect of the target entity as well as the overall reputation value using mathematical formulas. The system also proposes an analytical dashboard that displays in-depth information about the reputation of the target entity. In this manner, this study addresses the following research question: with the consideration of review popularity, review time, spam filtering, and ABSA, can the proposed reputation model offer better results in terms of generating and visualizing reputation than state-of-the-art (SOTA) systems? This paper is organized as follows. Section 2 presents the related work concerning the previous reputation generation systems as well as the ABSA models. Section 3 presents the preliminaries. Section 4 describes our proposal. Section 5 details the experiments. Section 6 presents the discussion. And finally, Section 7 concludes this paper.

## 2.EXISTINGSYSTEM

Poria *et al.* presented the first deep learning approachfor the AE task in opinion mining. The authors employeda 7-layer deep convolutional neural network to tag eachword in the textual opinions as either aspect or non-aspectword. The authors also proposed a set of heuristic linguisticpatterns and integrated them with the deep learning classifierwhich significantly improves the accuracy compared withthe previous SOTA methods. In [19], the authors proposedan attention-based long short-term memory (LSTM) [20]for aspect-level sentiment classification. The idea is to learn aspect embeddings and let aspects participate in computingattention weights.

The proposed model can focus on differentparts of a sentence when different aspects are given so thatthey are more competitive for aspect-level classification. Theproposed model achieved better results compared with thestandard LSTM on the SemEval 2014 Task 4 dataset [21].In [22], Wei and Toi improved the deficiencies of the previousLSTM approaches by proposing convolutional neuralnetworks [23] and gating mechanisms (GCAE) based model,which has been proved to be more accurate and efficient.The novel Gated Tanh-ReLU Units can selectively output thesentiment features according to the

ISSN 2347–3657
Volume 12, Issue 3, 2024

provided aspect or entity.The architecture of the proposed model is much simpler thanthe attention layer used in the previously existing models.

The experiments on SemEval datasets show a performanceimprovement compared with the LSTM based models. Theauthors in [24] proposed an interactive multi-task learningnetwork (IMN) capable of jointly learning multiple relatedtasks simultaneously at both the token-level as well as thedocument-level. The IMN introduces a message passingmechanism that allows informative interactions betweentasks, enabling the correlation to be better exploited. Experimentson three benchmark datasets, taken from SemEval2014and SemEval 2015 [25] show that IMN outperforms otherbaselines by large margins. Since most existing methodsignore the position information of the aspect when encodingthe sentence, authors in [26] proposed a hierarchicalattention-based position-aware network (HAPN), whichincludes position embeddings to learn the position-aware representationsof sentences to generate the target-specific representationsof contextual words. HAPN achieved the SOTAperformance on SemEval 2014 dataset compared with theprevious methods.

## 3.PROPOSED SYSTEM

This system aims at generating a reputation value towardonline entities (movies, hotels, restaurants, services, etc.)and computing a satisfaction score toward each aspect ofthe target entity by processing textual and numerical datacollected from multiple platforms. Proposed system presents itsarchitecture. First, we start by collecting users' reviews fromdifferent platforms such as Twitter, Amazon, YouTube, etc.Next, an automatic spammers filtering system is employed to detect and eliminate unwanted spam reviews. Then, we applya SOTA ABSA model to users' textual reviews in order tocompute a score based on the sentiment orientation of theextracted aspects from those reviews. Further, we calculate apopularity score and a time score based on statistical featuresextracted with the textual reviews. Finally, we compute areputation value based on the previously calculated scores,and we propose a new user-friendly visualization interfacethat displays in-depth details about the reputation of the targetentity.

International Journal of Information Technology & Computer Engineering

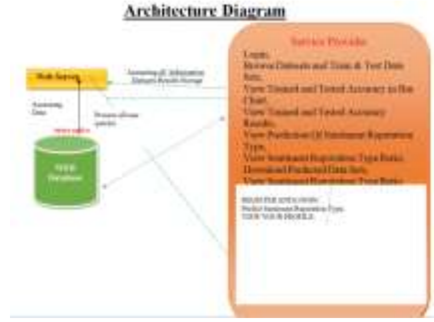
ISSN 2347–3657

Volume 12, Issue 3, 2024

One of the important features of the proposed system isthe ability to collect and process data from various platforms. Previous reputation generation systems gather necessarydata from either e-commerce websites such as Amazon, TripAdvisor, or social media platforms such as Twitter and Facebook. In this work, we decided to normalize the featuresof all platforms in order to create a single merged dataset byclassifying the platforms on the Internet into two types: thefirst type provides the accessibility of extracting the textualreview with the number of likes received for that reviewsuch as Amazon, YouTube, etc. The second type provides theaccessibility of extracting the textual review with the numberof likes received for that review along with the number oftimes the review was shared among the network such asTwitter, Facebook, etc.

## 4.SYSTEM DESIGN

Architecture Diagram



## 5.CONCLUSION

In this paper, we proposed a reputation system capable ofgenerating numerical reputation values for a specific item(product, movie, service, hotel, etc.) and its aspects basedon opinions and reviews expressed online. The contributionof this work revolves around four components that were notexploited in previous systems. The first one is cross-platformcompatibility, where the proposed system can collect andprocess opinions from different platforms (Facebook,Amazon, Twitter, TripAdvisor, etc.) as well as managingand standardizing those platforms' features. The second one is opinion spam filtering, where the spam opinionsare detected and eliminated based on spammers' behavior features, keeping only authentic opinions. The third one isemploying a SOTA aspect-based sentiment-analysis modelnamed LCF-ATEPC in order to extract and analyze theaspects within the textual opinions. Finally, we incorporatedthe previous results with a calculated review time scoreand review popularity score using mathematical formulasto obtain a reputation value for the targeted entity as wellas the reputation values of the entities' aspects. In additiona holistic reputation visualization is provided within thesystem that displays the detailed output results of the reputationgeneration process. To assess the effectiveness of ourreputation system, we invited 32 participants and 3 expertsto choose the best performing system out of four SOTAreputation systems by giving numerical satisfaction scoresto each system. Our reputation system achieved the highestaverage satisfaction scores from both users and experts. In thefuture, we propose to investigate the effectiveness of ourproposed system by attempting to generate more than thenumerical reputation values, such as extending the system toautomatically generate a textual summary of the benefits anddrawbacks of the targeted entity. Also, we intend to extendthis system to be used in multilingual content.

## 6.REFERENCES

[1] Abdel-Hafez, Y. Xu, and D. Tjondronegoro, ``Product reputation model: An opinion mining based approach,'' in *Proc. 1st Int. Workshop Sentiment Discovery Affect. Data*, vol. 917, London, U.K., Jun. 2013, pp. 16_27.
[Online]. Available: https://eprints.qut.edu.au/58118/

[2] U. Farooq, A. Nongaillard, Y. Ouzrout, and M. A. Qadir, ``A featurebased reputation model for product evaluation,'' *Int. J. Inf. Tech-nol. Decis. Making*, vol. 15, no. 6, pp. 1521_1553, Nov. 2016, doi: 10.1142/S0219622016500358.

[3] Z. Yan, X. Jing, and W. Pedrycz, ``Fusing and mining opinions for reputation generation,'' *Inf. Fusion*, vol. 36, pp. 172_184, Jul. 2017, doi: 10.1016/j.inffus.2016.11.011.

[4] A. Benlahbib and E. H. Nfaoui, ``A hybrid approach for generating reputation based on opinions fusion and sentiment analysis,'' *J. Organiza-tionalComput. Electron. Commerce*, vol. 30, no. 1, pp. 9_27, 2020, doi: 10.1080/10919392.2019.1654350.

[5] E. I. Elmurngi and A. Gherbi, ``Building sentiment analysis model and compute reputation scores in E-commerce environment using machine learning techniques,'' *Int. J. Organizational Collective Intell.*, vol. 10, no. 1, pp. 32_62, Jan. 2020.

[6] A. Benlahbib and E. H. Nfaoui, ``Aggregating customer review attributes for online reputation generation," *IEEE Access*, vol. 8, pp. 96550_96564, 2020, doi: 10.1109/ACCESS.2020.2996805.

[7] A. Gupta, S. Priyani, and R. Balakrishnan, ``Customized reputation generation of entities using sentiment analysis," *World J. Eng.*, vol. 18, no. 4, pp. 596_605, Jul. 2021, doi: 10.1108/WJE-09-2020-0470.

[8] A. Boumhidi and E. Nfaoui, ``Leveraging lexicon-based and sentiment analysis techniques for online reputation generation," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 6, pp. 274_289, 2021, doi: 10.22266/ijies2021.1231.25.

[9] V. M. Pradhan, J. Vala, and P. Balani, ``A survey on sentiment analysis algorithms for opinion mining," *Int. J. Comput. Appl.*, vol. 133, no. 9, pp. 7_11, Jan. 2016.

[10] A. Tripathy, A. Anand, and S. K. Rath, ``Document-level sentiment classi _cation using hybrid machine learning approach," *Knowl. Inf. Syst.*, vol. 53, no. 3, pp. 805_831, 2017.