IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 12, May 2023

A Machine Learning Approach for Opinion Mining Online Customer Reviews

Akanksha Kohakade[1], Govinda Jogdand[2], Pratiksha Kohokade[3], Kavita Kadam[4], Prof. A. Kadam[5]
Students, Department of Computer Engineering[1,2,3,4]
Professor, Department of Information Technology[5]
Sinhgad Institute of Technology, Lonavala, Maharashtra, India

Abstract: *This study was conducted to apply supervised machine learning methods in opinion mining online customer reviews. First the machine collects reviews as a input then train machine learning models to find out which model is most compatible with the training dataset and then apply this model to forecast opinions for the collected dataset. The results showed that Logistic Regression (LR), Support Vector Machines (SVM) ,Neural Network, NaviaBayes, Random forest and decision tree methods have the best performance in opinion mining . This study is valuable for applications of opinion mining in the field of business and organizations. it helps customers to find best products.*

Keywords: Navie Bayes, Support Vector Machine, Logical Regression, Neural Network, Decision Tree, Random Forest

I. INTRODUCTION

Now a days people are buying a product by using online technologies. Social networks and online review websites allow customers to give their opinions on products or services through reviews. By using online customer reviews on the product people can choose a best product and also business enterprises can easily understand customer purchase behaviour, as well as their interests and satisfactionlevel on product or service quality. Opinion mining has become the subject of studies in different areas. Currently, the community of scientists have lots of studies on opinion mining methods as well as the application of opinion mining at different levels. The studies used a hybrid method of machine learning and lexical based and supervised learning. Research methodology on opinion mining is old and each method has its advantages and disadvantages and none of them are accurate. the application of lexical based method is a big challenge for researchers. The objective of this study is to review studies onopinion mining and propose the application of machine learning method in opinion mining customer reviews. The method of knowledge discovery indatabases is applied to check whether the product is good or bad. Then, the study conducts data pre-processing and training using machine learning methods to find the most suitable model with the training data sets and apply this model to forecast opinions.

II. PROPOSED WORK

Opinion classification is a text mining technique natural language processing (NLP).Machine learning method plays an important role in opinion mining. Opinion mining at the document and sentence level is used to determine whether a statement is positive or negative. The objective of this study is to review studies on opinion mining and propose the application of machine learning method in opinion mining customer reviews in Vietnamese. The method of knowledge discovery in databases is applied to this study in which 39,976 tourists' reviews on hotels in Vietnam are collected through Agoda.com. Then, the study conducts data pre-processing and training using machine learning methods to find the most suitable model with the training data sets and apply this model to forecast opinons for the entire dataset.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 12, May 2023

III. UML DIAGRAM

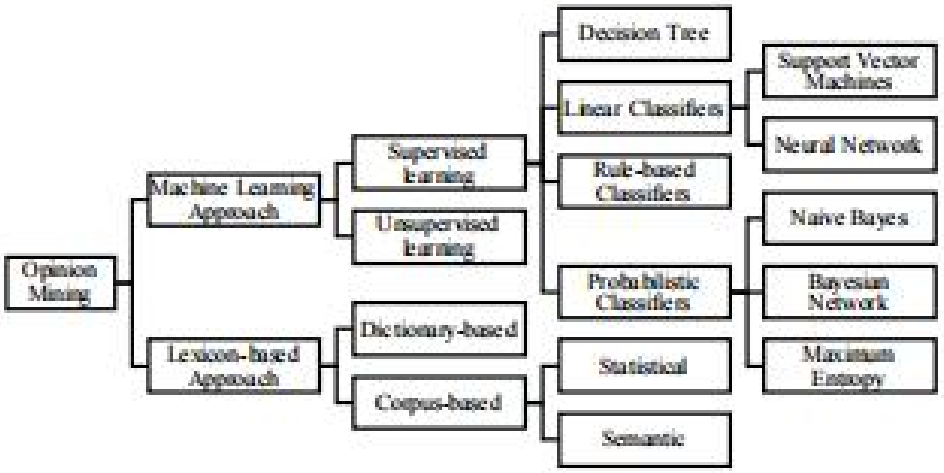

Figure 1 : System Architecture

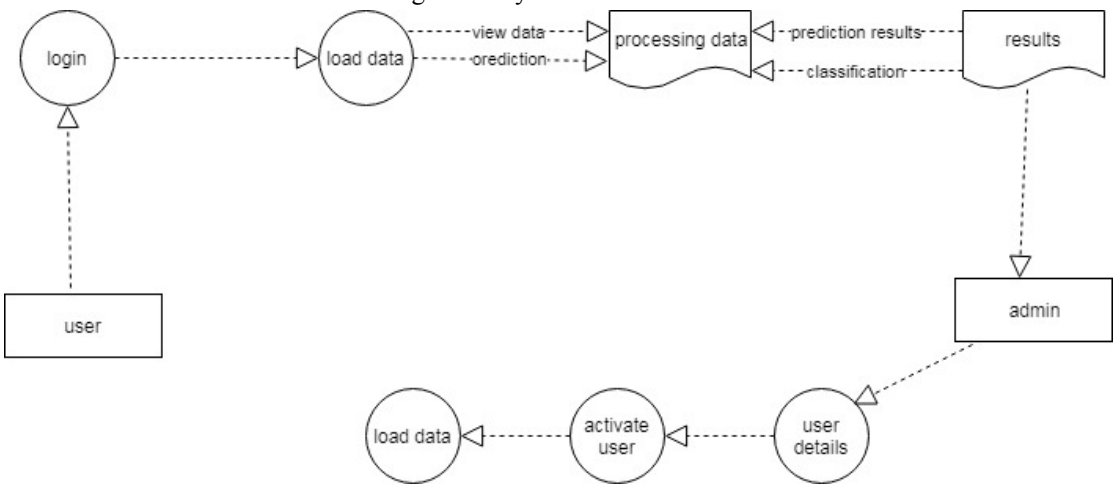

Figure 2: Data Flow Diagram

IV. HARDWARE AND SOFTWARE REQUIREMENTS

4.1 Software Requirements

- System: Intel i5 6 core.
- Coding Language:Python
- Tool: PyCharm, Visual Studio Code
- Database: MYSQL

4.2 Hardware Requirements

- System :Intel i5 6 core.
- Hard Disk :500 GB.
- Monitor :15'' LED.
- Input Device : Keyboard, Mouse
- Ram :16GB

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 12, May 2023

V. APPLICATIONS

- Extract Nouns, Adjectives, Verbs and Adverbs on the remaining reviews by using dictionary approach.
- Identify frequent words by using Apriori frequent item set mining Algorithm.
- Perform Sentiment Analysis on the frequent words using SentiWordNet.
- Provide visualization.

VI. CONCLUSION

This study has conducted a theoretical background on opinion mining methods, opinion classification techniques and proposed the application of supervised machine learning method for automatic opinion mining. Experimental results show that LR, SVM and NN are the best among the training methods. This study is valuable as a reference for applications of opinion mining in socioeconomic fields. However, this study still has some limits that can be adjusted in future studies. Firstly, in terms of data collection, this study only collects customer reviews about hotels on Agoda.com. The study may expand to collect reviews about any products or services on ecommerce websites or social networks. Secondly, in terms of the scale, this study only classifies customer reviews on a 2-level scale: positive and negative. More level scales may be applied in the next study (for example, on a 5-level Likert scale). Thirdly, in terms of opinion classification technique, this study only uses supervised machine learning method. It will give better results with a hybrid method of supervised machine learning and lexicon based. However, currently, there are not many tools that support processing Vietnamese as well as English. Finally, this research is just limited to the classification of opinions. The extended research's directions will focus on the application of opinion mining in behavior, sentiment, and shopping preference analysis as well as products and services quality assessment, which has more practical implications for entrepreneurs and customers.

ACKNOWLEDGMENT

The authors would like to thank professor of the Department of Information Technology in Sinhgad institute of technology, Lonavala Prof. A. Kadam for their time and efforts he provided throughout the project and his advice and suggestions were really helpful to us.

REFERENCES

[1] A. Dhokrat, S. Khillare, and C. N. Mahender, "Review on techniques and tools used for opinion mining,"International Journal of Computer Applications Technology and Research, vol. 4, no. 6, pp. 419-424,2015.
[2] B. Liu, "Sentiment analysis and opinion mining," Synthesis Lectures on Human Language Technologies, vol. 5, no. 1, pp. 1-167, 2012.
[3] B. Pang and L. Lee, "Opinion mining and sentiment analysis," Foundations and trends in information retrieval, vol 2., no. 12, pp. 1-135, 2008.
[4] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," Journal of machine Learning research, pp. 993-1022, 2003.
[5] F. Wanner, C. Rohrdantz, F. Mansmann, D. Oelke, and D. A. Keim, "Visual sentiment analysis of RSS news feeds featuring the us presidential election in 2008," In VISSW, 2009.
[6] G. Stylios, D. Christodoulakis, J. Besharat, M.Vonitsanou, I. Kotrotsos, A. Koumpouri, and S.Stamou, "Public opinion mining for governmental decisions," Electronic Journal of e-Government, vol. 8, no. 2, pp. 203-214, 2010.
[7] H. Binali, V. Potdar, and C. Wu, "A state of the artopinion mining and its application domains,"International Conference on Industrial Technology,pp. 1-6, 2009.
[8] J. Lee, D. H. Park, and I. Han, "The different effects of online consumer reviews on consumers' purchaseintentions depending on trust in online shopping malls: An advertising perspective," Internet Research, vol. 21, no. 2, pp. 187–206, 2011.
[9] J. Shawe-Taylor and S. Sun, "A review of optimization methodologies in support vector machines,"Neurocomputing, vol. 74, no. 17, pp. 3609-3618, 2011.
[10] S. K. Yadav, "Sentiment analysis and classification: A survey," International Journal of Advance Research in Computer Science and Management Studies, vol. 3,no. 3, pp. 113-121, 2015.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 12, May 2023

[11] S. M. Mudambi and D. Schuff, "What makes a helpfulreview? A study of customer reviews on Amazon.com," MIS quarterly, vol. 34, no. 1, pp. 185-200, 2010.
[12] S. Sun, C. Luo, and J. Chen, "A review of naturallanguage processing techniques for opinion mining systems," Information Fusion, vol. 36, pp. 10-25,2017.
[13] W. Medhat, A. Hassan, and H. Korashy, "Sentimentanalysis algorithms and applications: A survey," Ain Shams Engineering Journal, vol. 5, no. 4, pp. 1093-1113, 2014.246

BIOGRAPHY

- Akanksha D. Kohakade -An Undergraduate Scholar pursuing Bachelors of Engineering in Computer from Sinhgad Institute of Technology. She is working under the guidance of Prof. A. Kadam
- Govinda R. Jogdand - An Undergraduate Scholar pursuing Bachelors of Engineering in Computerfrom Sinhgad Institute of Technology. He is working under the guidance of Prof. A. Kadam
- Pratiksha J. Kohokade - An Undergraduate Scholar pursuing Bachelors of Engineering in Computerfrom Sinhgad Institute of Technology. She is working under the guidance of Prof. A. Kadam
- Kavita A. Kadam - An Undergraduate Scholar pursuing Bachelors of Engineering in Computer from Sinhgad Institute of Technology. She is working under the guidance of Prof. A. KadamBabar