Reviewing Flipkart Product Comments using Methods Based on Sentiment Analysis

Ishan Raj[1], Manveer Singh Jhala[2], Pravin Vilas Kadam[3], Prof. Amruta Patil[4]

Students, Department of Computer Engineering[1,2,3]

Guide, Department of Computer Engineering[4]

Sinhgad Inst. Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

Abstract: *With the explosion of social networking sites, blogs and review sites a lot of information is available on the internet. This data contains sentiments and opinions about a wide variety of product features of such products. This form of opinion and feedback is crucial to the companies developing these products as well as the companies that want to develop these products better than the competitive products. Sentiment Analysis is the process of analyzing all this information, retrieving opinions about such products and services, and labeling them as good or bad. In other words, finding out whether the product is good or bad or positive or negative. The most crucial parts of any review of any particular product is the numeric rating and the description provided along with that product. In this project, we will take into consideration both these factors for product reviews to make a final decision on a classifier that is best suited for product reviews analysis. We have collected reviews based on the features that best describe the emotions and opinion for each review, we have built a feature set of 1000 features, and with this limited feature set, we will find out which classifier obtains the best result on review type data. To decide the best classifier, we perform evaluations on it, by running a wide variety of data set generators and calculating the resubstituting and generalization mistakes and errors for each classifier. We then use the mean of these findings to calculate the paired Student's t-test to relatively compare the output of the classifiers. Based on the findings of this evaluation, we can state which is the best suited classifier for our model.*

Keywords: Filpcart.

REFERENCES

[1]. B. Liu, "Sentiment analysis and opinion mining," Synthesis Lectures on Human Language Technologies, vol. 5, no. 1, pp. 1–167, May 2012.

[2]. B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10, ser. EMNLP'02. Stroudsburg, PA, USA: Association for Computational Linguistics, 2002, pp. 79–86.

[3]. V. Ng, S. Das gupta, and S. M. N. Arifin, "Examining the role of linguistic knowledge sources in the automatic identification and classification of reviews," in Proceedings of the COLING/ACL on Main Conference Poster Sessions, ser. COLING-ACL '06. Stroudsburg,PA, USA: Association for Computational Linguistics, 2006,pp. 611–618

[4]. J. Zhao, K. Liu, and G. Wang, "Adding redundant features for crfs-based sentence sentiment classification," in Proceedings of the Conference on Empirical Methods in Natural Language Processing, ser.EMNLP '08. Stroudsburg, PA, USA: Association for Computational Linguistics, 2008, pp. 117–126.

[5]. P. Melville, W. Gryc, and R. D. Lawrence, "Sentiment analysis ofblogs by combining lexical knowledge with text classification," in Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD'09. New York,NY, USA: ACM, 2009, pp. 1275–1284.

[6]. A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts,"Learning word vectors for sentiment analysis," in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human

Language Technologies - Volume 1, ser. HLT'11.Stroudsburg, PA, USA: Association for Computational Linguistics,2011, pp. 142–150.

[7]. Alexander Pak and Patrick Paroubek, "Twitter as a corpus for Sentiment Analysis and Opinion Mining", Proceedings of the Seventh conference on International Language Resources and Evaluation, pp. 1320-1326, 2020

[8]. Das, Sourav, and Anup Kumar Kolya (2019)"Sense GST: Text mining & sentiment analysis of GST tweets by Naive Bayes algorithm." Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), IEEE, 239-244.

[9]. Garg, Neha, and Rinkle Rani (2020) "Analysis and visualization of Twitter data using k-means clustering." International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, 670-675.

[10]. Mishra, Prerna, Ranjana Rajnish, and Pankaj Kumar. (2019) "Sentiment analysis of Twitter data: Case study on digital India." International Conference on Information Technology (InCITe)-The Next Generation IT Summit on the Theme-Internet of Things: Connect your Worlds. IEEE, 148 – 153.

[11]. Devika, R., and S. Revathy. (2020) "Survey on clustering techniques in Twitter data." Second International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 1073-1077.

[12]. McCallum, Andrew, and Kamal Nigam. "A comparison of event models for Naive Bayes text classification." AAAI-98 Workshop on learning for text categorisation. Vol. 752. 1998.

[13]. Kotsiantis, Sotiris B., I. Zaharakis, and P. Pintelas. "Supervised machine learning: A review of classification techniques." (2007): 324.

[14]. Zhu, Maoran, et al. "Identification of Opinion Leaders in Social Networks Based on Sentiment Analysis: Evidence from an Automotive Forum." (2016).

[15]. Kim, Soo-Min, and Eduard Hovy. "Determining the sentiment of opinions." Proceedings of the 20th international conference on Computational Linguistics. Association for Computational Linguistics, 2004.

[16]. Pang, B., L. Lee, and S. Vaithyanathan, Thumbs up?: "sentiment classification using machine learning techniques", Acl-02 Conference on Empirical Methods in Natural Language Processing, 2002, p. 79--86..