Data Mining Using K-Mean Clustering and K-Nearest Neighbor Classification

Prof. Shivangi D. Matieda, Prof. Mani Butwall, Prof. Diksha Durgapal
Department of Computer Science and Engineering
ITM Universe, Vadodara, Gujrat, India
s.borasia@gmail.com, s.borasia@gmail.com, s.borasia@gmail.com

Abstract: *Understudies casual discussions via web-based networking media (e.g., Twitter, Facebook) shed regard into their instructive encounters—sentiments, emotions, and worries concerning the training procedure. data from such unenlightened things will offer profitable data to teach understudy learning. Examining such data, in any case, will be testing. The complexity of understudies' encounters mirrored from net based mostly life content needs human elucidation. In any case, the developing size of data knowledge requests programmed information investigation systems. During this paper, we have a tendency to design a piece method to include each subjective investigation and large-scale data mining ways. We have a tendency to focus on building understudies Twitter presents to get problems and problems in their instructive encounters. We have a tendency to originally direct a subjective examination on tests taken from around twenty-five thousand tweets known as building understudy's college life. We have a tendency to discover building understudy's expertise problems, as an example, substantial investigation load, absence of social commitment, and lack of sleep. In light of those outcomes, we have a tendency to actually do a multi-name arrangement calculation to order tweets mirroring understudies' problems.*

Keywords: Data Mining, Index Terms—K-Nearest Neighbor, K-Mean Algorithm, Facebook, Twitter

REFERENCES

[1]. P. D. Turney, Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews, in Proceedings of the 40th annual meeting on association for computational linguistics, pp. 417424, Association for Computational Linguistics, 2002.

[2]. A.Pak and P. Paroubek. Twitter as a Corpus for Sentiment Analysis and Opinion Mining". In Proceedings of the Seventh Conference on International Language Resources and Evaluation, 2010, pp.1320-1326.

[3]. Po-Wei Liang, Bi-Ru Dai, Opinion Mining on Social MediaData", IEEE 14th International Conference on Mobile Data Management, Milan, Italy, June 3- 6, 2013, pp 91-96, ISBN: 978-1-494673-6068- http://doi.ieeecomputersociety.org/10.1109/MDM.2013.

[4]. T. T. Thet, l.-C. Na, C. S. Khoo, and S. Shakthikumar, "Sentiment analysis of movie reviews on discussion boards using a linguistic approach," in Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion. ACM, 2009, pp. 81-84.

[5]. Hussein, D.-M.E.D.M. A survey on sentiment analysis challenges. Journal of King Saud University Engineering Sciences (2016), http://dx.doi.org/10.1016/j.jksues.2016.04.002

[6]. A. Kowcika and Aditi Guptha Sentiment Analysis for Social Media, International Journal of Advanced Research in Computer Science and Software Engineering, 216-221,Volume 3,Issue 7, July 2013.

[7]. G. Vinodini and RM.Chandrashekaran, Sentiment Analysis and Opinion Mining: A Survey, International Journal of Advanced Research in Computer Science and Software Engineering, 283-294, Volume 2, Issue 6, June 2012.

[8]. Apoorv Agarwal Boyi Xie Ilia Vovsha Owen Rambow Rebecca Passonneau, Sentiment Analysis of Twitter Data,Columbia University,New York.

[9]. Pablo Gamallo and Marcos Garcia A Naive-Bayes Strategy for Sentiment Analysis on English Posts

Proceedings of the 8th InternationalWorkshop on Semantic Evaluation (SemEval 2014), pages 171175, Dublin, Ireland, Aug 23-24 2014.

[10]. Harry Zhang "The Optimality of Naive Bayes", FLAIRS2004 conference. (available online: PDF (http:/ / www. cs. unb. ca/ profs/ hzhang/ publications/FLAIRS04ZhangH. pdf)).