Project Proposal
Twitter Clustering and Visualization

Maksim Podkorytov, Kobra Amiri Zirtol, Rituparna Khan

I. Introduction

II. Literature Survey

Social media is becoming increasingly popular and provides a great platform for communicating individual thoughts. Twitter works as a great platform and there has been numerous analysis based on tweets. Andrei Sechelea et al.[1] has worked on clustering Twitter data based on opinion, trends and events in a specific geographic location. They used Harvester API for pre-processing large amounts of tweets. For clustering, they have developed an algorithm combining k-means, Density based spatial clustering of algorithms with noise algorithm and a consensus matrix. Finally, for visualization of clusters they used a graphing software Gephi. They provided 2D and 3D diagrams for visualization. Alan Ritter et al.[2] developed a model to cluster unsupervised twitter dialogues. Two models were considered, namely Conversation model and Conversation+Topic model. Muqtar Unnisa et al.[3] did opinion mining on twitter data based on unsupervised technique. Spectral clustering algorithm was used compared to k-means as it uses several mathematical concepts. Shreya Ahuja et al.[4] used K-means for clustering as the data is quite large and Euclidean distance needs to computed in less time for making the analysis efficient.

III. Details of Algorithm

IV. Expected Experiments and Analysis

V. Timeline

References

[1] Z. D. Sechelea, Huu, "Twitter data clustering and visualization," in *2016 23rd International Conference on Telecommunications (ICT)*, 2016, pp. 1–5.

[2] B. D. Alan Ritter, Colin Cherry, "Unsupervised modeling of twitter conversations," in *HLT '10 Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics)*, 2010, pp. 1–9.

[3] A. A. Muqtar Unnisa and S. Raziuddin, "Opinion mining on twitter data using unsupervised learning technique," *International Journal of Computer Applications*, vol. 148, no. 12, pp. 12–189, 2016.

[4] S. A. . G. Dubey, "Clustering and sentiment analysis on twitter data," in *2017 2nd International Conference on Telecommunication and Networks (TEL-NET)*. IEEE, 2017, pp. 1–5.