Sentiment Analysis Using Machine Learning Approaches on Social Media Data

Dharmendra Dangi*, Amit Bhagat

Department of Mathematics and Computer Applications, MANIT Bhopal

*Corresponding author

Abstract

Background: Twitter is a microblogging service that allows users to send updates (tweets) to their friends (followers) [1]. It has grown into a substantial sentiment database. Sentiment analysis is the study of emotions and ideas expressed in text [2]. Opinion mining is another name for sentiment analysis. Sentiment analysis of this massively generated data is extremely beneficial for expressing the public's opinion or ideas on a certain issue or trend. Customers can use sentiment analysis to search for products, companies can keep track of and analyse public sentiment about their brands, and it can be used for a variety of other purposes [3].

Objective: To find and classify a person's sentiment towards a certain source of document or text.

Methodology: In this study, we are creating dataset using Twitter API and saving it for experiments. These experiments can be done on variety of public tweet sentiment datasets using ensembles classifier formed various classifiers using majority voting such as Multinomial Naïve Bayesian, Support Vector Machine, Bernoulli Naive Bayes and Logistic Regression can improve classification accuracy and comparison of different algorithms with each other when used alone for sentiment analysis [4].

Result and Discussion: Five machine learning classifiers Logistic Regression, Linear SVC, Multinomial Naive Bayes, Bernoulli Naive Bayes and Ensemble Classifier have been used to obtain the sentiment of twitter data, out of which the results of the Logistic Regression Classifier has obtained 82.92% in all parameters.

Future Work: In the future, we will collect real-time data and use deep learning like CNN, LSTM and their ensemble methods to improve outcomes.

References

[1] Da Silva, Nadia FF, Eduardo R. Hruschka, and Estevam R. Hruschka Jr. "Tweet sentiment analysis with classifier ensembles." Decision Support Systems 66 (2014): 170-179.

[2] Agarwal, Apoorv, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca J. Passonneau. "Sentiment analysis of twitter data." In Proceedings of the workshop on language in social media (LSM 2011), pp. 30-38. 2011.

[3] Dangi Dharmendra, Amit Bhagat, and Dheeraj K. Dixit. "Sentiment analysis of social media data based on chaotic coyote optimization algorithm based time weight-AdaBoost support vector machine approach." Concurrency and Computation: Practice and Experience: e6581, 2021.

[4] Zucco, Chiara, Barbara Calabrese, Giuseppe Agapito, Pietro H. Guzzi, and Mario Cannataro. "Sentiment analysis for mining texts and social networks data: Methods and tools." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 10, no. 1 (2020): e1333.