

Project Synopsis on Twitter Opinion Mining

BACHELOR OF TECHNOLOGY

COMPUTER SCIENCE AND ENGINEERING



DELHI GLOBAL INSTITUTE OF TECHNOLOGY

Bhagdurgarh, Jhajjar, Haryana 124201

Affiliated to

Maharshi Dayanand University

Submitted by:

Bazgha Razi

MDU Registration No:

191380214

Guided by:

Mr. Kapil Singh

TABLE OF CONTENT

S.No.	Topic	Page No.
1.	Introduction	3
2.	Proposed Method	3
3.	Technologies Used	5
4.	Application area of the Project	6
5.	Work implementation in Phases	7
6.	Future Scope	8
7.	References	9

INTRODUCTION

Twitter Opinion Mining:

Twitter opinion mining analyses the sentiment or emotion of tweets. It uses natural language processing and machine learning algorithms to classify tweets automatically as positive, negative, or neutral based on their content. It can be done for individual tweets, or a larger dataset related to a particular topic or event.[6]

PROPOSED METHOD:

As our goal is to achieve sentiment analysis for data provided from Twitter. We are going to build a classifier which consists of different machine learning classifiers. Once our classifier is ready and trained, we are going to follow the steps shown below

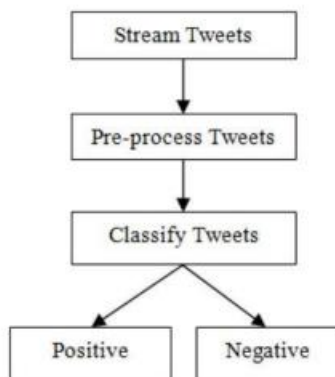


Figure 4.1 Process to classify tweets using build classifier

Figure: Process to classify tweets using build classifier [12]

Step-1 First we are going to stream tweets in our build classifier with the help of Tweepy library in python.

Step-2 Then we pre-process these tweets, so that they can be fit for mining and feature extraction.

Step-3 After pre-processing we pass this data in our trained classifier, which then classify them into positive or negative class based on trained results.

Since, Twitter is our source of data for analysis. We are going to stream the tweets from twitter in our database. For this we are going to use Twitter Application.[9]

TECHNOLOGIES USED

SOFTWARE REQUIREMENTS:

- Windows 7,10,11
- VS Code (Python IDE)
- Git & Github

HARDWARE REQUIREMENTS:

- Monitor: 1024 X 720 display
- RAM: 8GB or more
- Speed: 2.7GHZ and more
- Cache: 512KB
- Processor: Intel i5 or more

LIBRARIES AND MODULE USED

- Python
- matplotlib
- numpy
- pandas
- nltk
- scikit-learn

APPLICATION AREA OF TWITTER OPINION MINING

Applications that use Reviews from Websites:

Today Internet has a large collection of reviews and feedbacks on almost everything. This includes product reviews, feedbacks on political issues, comments about services, etc. Thus, there is a need for a sentiment analysis system that can extract sentiments about a particular product or services. It will help us to automate in provision of feedback or rating for the given product, item, etc. This would serve the needs of both the users and the vendors.

Applications in Business Intelligence:

It has been observed that people nowadays tend to look upon reviews of products which are available online before they buy them. And for many businesses, the online opinion decides the success or failure of their product. Thus, Sentiment Analysis plays an important role in businesses. Businesses also wish to extract sentiment from the online reviews in order to improve their products and in turn their reputation and help in customer satisfaction.

Applications in Smart Homes:

Smart homes are supposed to be the technology of the future. In future entire homes would be networked and people would be able to control any part of the home using a tablet device.

Recently there has been lot of research going on Internet of Things (IoT). Sentiment Analysis would also find its way in IoT. Like for example, based on the current sentiment or emotion of the user, the home could alter its ambiance to create a soothing and peaceful environment.[9]

Sentiment Analysis can also be used in trend prediction. By tracking public views, important data regarding sales trends and customer satisfaction can be extracted.

WORK IMPLEMENTATION IN PHASES

PHASE 1:

- Defining scope for the project, idea brainstorming.

PHASE 2:

- Add all the necessary components using some libraries and modules.

PHASE 3:

- Collect dataset and code for twitter opinion mining.

PHASE 4:

- Evaluation and testing of the twitter opinion mining.

FUTURE SCOPE

- **Multi-lingual support:** Due to the lack of multi-lingual lexical dictionary, it is currently not feasible to develop a multi-language based sentiment analyser.
Further research can be carried out in making the classifiers language independent. The authors have proposed a sentiment analysis system with support vector machines, similar approach can be applied for our system to make it language independent.
- **Interpreting Sarcasm:** The proposed approach is currently incapable of interpreting sarcasm. In general sarcasm is the use of irony to mock or convey contempt, in the context of current work sarcasm transforms the polarity of an apparently positive or negative utterance into its opposite.[3]
The main goal of this approach is to empirically identify lexical and pragmatic factors that distinguish sarcastic, positive and negative usage of words.
- Analysing sentiments on emoji/smiley.

REFERENCES

- [1] 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.
- [2] R. Parikh and M. Movassate, "Sentiment Analysis of User-Generated Twitter Updates using Various Classification Techniques", CS224N Final Report, 2009.
- [3] Go, R. Bhayani, L.Huang. "Twitter Sentiment Classification Using Distant Supervision". Stanford University, Technical Paper, 2009.
- [4] L. Barbosa, J. Feng. "Robust Sentiment Detection on Twitter from Biased and Noisy Data". COLING 2010: Poster Volume, pp. 36-44.
- [5] Bifet and E. Frank, "Sentiment Knowledge Discovery in Twitter Streaming Data", In Proceedings of the 13th International Conference on Discovery Science, Berlin, Germany: Springer, 2010, pp. 1-15.
- [6] Agarwal, B. Xie, I. Vovsha, O. Rambow, R. Passonneau, "Sentiment Analysis of Twitter Data", In Proceedings of the ACL 2011 Workshop on Languages in Social Media, 2011 , pp. 30-38.
- [7] Dmitry Davidov, Ari Rappoport. "Enhanced Sentiment Learning Using Twitter Hashtags and Smileys". Coling 2010: Poster Volume pages 241-249, Beijing, August 2010.
- [8] Po-Wei Liang, Bi-Ru Dai, "Opinion Mining on Social Media Data", IEEE 14th International Conference on Mobile Data Management, Milan, Italy, June 3 - 6, 2013, pp 91-96, ISBN: 978-1-494673-6068-5.

[9] Pablo Gamallo, Marcos Garcia, "Citius: A Naive-Bayes Strategy for Sentiment Analysis on English Tweets", 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, Aug 23-24 2014, pp 171-175.

[10] Neethu M,S and Rajashree R," Sentiment Analysis in Twitter using Machine Learning Techniques" 4th ICCCNT 2013,at Tiruchengode, India. IEEE – 31661.

[11] 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. 417–424, Association for Computational Linguistics, 2002.

[12] J. Kamps, M. Marx, R. J. Mokken, and M. De Rijke, "Using wordnet to measure semantic orientations of adjectives," 2004.

[13] R. Xia, C. Zong, and S. Li, "Ensemble of feature sets and classification algorithms for sentiment classification," Information Sciences: an International Journal, vol. 181, no. 6, pp. 1138–1152, 2011.

[14] ZhunchenLuo, Miles Osborne, TingWang, An effective approach to tweets opinion retrieval", Springer Journal onWorldWideWeb,Dec 2013, DOI: 10.1007/s11280-013- 0268-7.

[15] Liu, S., Li, F., Li, F., Cheng, X., &Shen, H.. Adaptive cotraining SVM for sentiment classification on tweets. In Proceedings of the 22nd ACMinternational conference on Conference on information & knowledgemanagement (pp. 2079-2088). ACM,2013.