

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/338711391>

NLP BASED SERVICE RECOMMENDATION SYSTEM

Conference Paper · December 2019

CITATIONS

0

READS

438

5 authors, including:



Uroosa Shaikh

Mehran University of Engineering and Technology

2 PUBLICATIONS 0 CITATIONS

[SEE PROFILE](#)



Isma Farah Siddiqui

Mehran University of Engineering and Technology

80 PUBLICATIONS 178 CITATIONS

[SEE PROFILE](#)



Sorath Burio

Mehran University of Engineering and Technology

2 PUBLICATIONS 0 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Intuitive Implementation of Alpha Beta pruning using Tic-Toc-Toe [View project](#)



Crowdsourced Machine Learning Based Recommender for Software Design Patterns International Journal of Computer (IJC) [View project](#)

NLP BASED SERVICE RECOMMENDATION SYSTEM

Engr. Uroosa Shaikh, Isma Farah Siddiqui*, Qasim Ali Arain, Sorath Buriro, Nasreen Babar

Department of Software Engineering,
Mehran University of Engineering and Technology,
Jamshoro, Sindh, Pakistan

[Emails: suroosa@rocketmail.com, isma.farah@faculty.muet.edu.pk]

*Corresponding Author : Isma Farah Siddiqui

Abstract

The growing popularity of social networks provides the availability of user sentiments which has a significant impact on buying decisions, public opinion. Millions of users share reviews on twitter about products every day. Based on these reviews and experiences they share any comment or complaint on any product and express their thoughts in terms of positive or negative sentiment. This paper focuses on the opinion of the users about products and services what people think and feel. This paper aims to provide a service recommendation approach that extracts the data from Twitter API Tweepy. The extracted data are then processed by the NLTK tool. NLTK will remove the words with POS (Part_Of_Speech) tags. AlchemyAPI is used to extract information from it and API will perform sentiment analysis of words with positive and negative words in the document created by AlchemyAPI. If the PNP ratio of positive words is greater than the ratio of negative words, then it will recommend users about the product that they should use it. The main purpose of this research is to propose a system that classifies opinions about products and services from the consumers' perspective and experience on Twitter for the recommendation.

Keywords: NLTK, AlchemyAPI, recommendation system, Tweepy.

1. Introduction

The growing popularity of social networks provides availability of user sentiments which has significant impact on buying decisions and public opinion. Now a day's most of products/services are shared on social media to help people about products and Social media has each type of reviews about products and services. Millions of users share opinions on social media about different aspects of life every day. On twitter, people discuss about a particular product, service in the form of tweet. Twitter is social networking site in which people share their ideas based on their experiences. They share a comment or complaint on any product and express their thoughts in terms of positive or negative

sentiment[3]. There is large number of reviews available, so it is difficult for a customer to find out usefulness of reviews in huge crowd of positive and negative comments.

2. Recommendation System

The Recommendation system is an information filtering system which recommends the products or services to the user. It considers users' requirements, interests and explores targeted information. The recommender system is a kind of information filtering systems, it is used to predict a users' rating or preference for an item.

The three categories of Recommendation systems are:

(i) **Collaborative Filtering:** Numeric rating to find the nearest neighbors to the target users and generates recommendation.

(ii) **Content based Filtering:** Recommends text based items which are similar to previous users.

(iii) **Hybrid Filtering:** this approach uses both collaborative and content based filtering.

In this paper we propose a system that takes the reviews about products and services from twitter to recommend best product or services based on classification. At first, data is extracted from twitter API Tweepy and extracted data are processed by NLTK tool, it removes the words with POS tags after that AlchemyAPI performs sentiment analysis of words with positive and negative words in the document created by AlchemyAPI.

Then according to sentiment analysis of positive negative words PN ratio is counted. PN ratio is a calculation of total positive and negative reviews. According to PN ratio, if ratio of positive words is greater than the ratio of negative words, then it will recommend users about product that they should use it.

3. Related Work

VibhuJawa, et al [1] proposed a model working on the basics of Interest Graph in conjunction with sentimental analysis to compute the correlation between different entities and provide recommendation, which ranges from whom to follow on Twitter and what to buy online.

Salama Hassan Shaikh, et al [2] worked on business Intelligence and used Dynamic LM Classifier Algorithm to classify reviews and predict the sales performance after that vendor may take necessary steps afterwards.

M. Trupthi, et al [3] provided an interactive automated system which predicts the sentiment of the reviews/tweets of the people posted in social media using Hadoop, which can process the huge amount of data and they used real time sentimental analysis on the tweets and provide time based analytics to the user.

Shayan Raju, et al [4] proposed video recommendation system on YouTube and

provide framework to develop a solution in a general sense, without specifying it to a certain area of content or interest, which can act as a form of personalized recommendation system for the users, and can deliver content to them without the need of having to search for the item.

4. Methodology

This paper presents a methodology in which system classifies opinions about products and services from reviews that are extracted from Twitter API Tweepy.

The extracted data are processed by NLTK tool as follows:

i. **NLTK Tool:** It performs the text classification and removes the words with POS tags.

ii. **AlchemyAPI** extracts information from NLTK tool and API will perform sentiment analysis of words with positive and negative words in the document. AlchemyAPI creates the list of positive and negative keywords in the document.

iii. **PN ratio** is the calculation of total positive and negative words. If the PN ratio of positive words is greater than the ratio of negative words, then it will recommend users about product that they should use it.

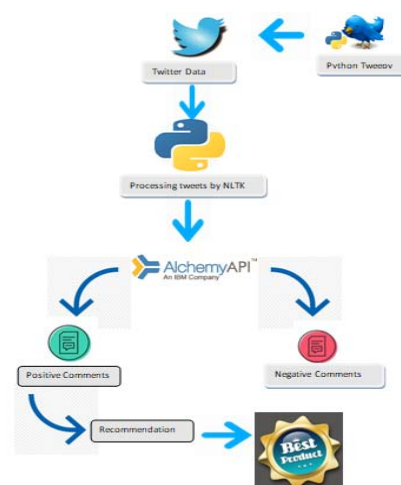


Fig. 1. Methodology Flow chart

As shown in Fig. 1, twitter reviews are collected through twitter API Tweepy by following process:

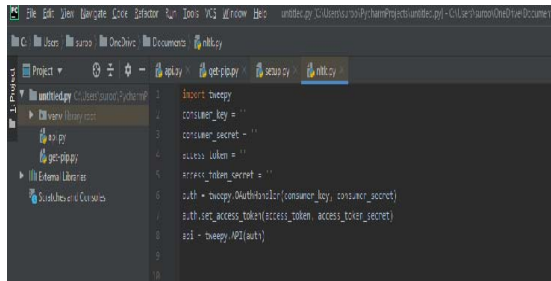


Fig. 2. Extraction of Data through TweepyAPI

After that, reviews are processed through NLTK toolkit.

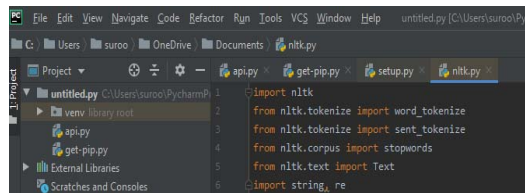


Fig. 4. Processing of reviews through NLTK

After importing Tweepy data, NLTK tool remove punctuations, tokenized the string, removed stop words. Filtered data which is created by NLTK will be extracted by AlchemyAPI and create a document for positive negative words. Fig. 3 & 4 show the implementation of our model. If the PN ratio of positive words is greater than the ratio of negative words, then it will recommend users about product that they should use it, as shown in Fig. 5 below.

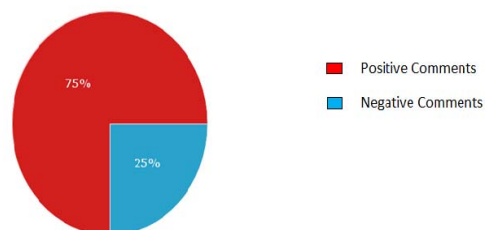


Fig. 5. PN Ratio

5. Conclusion and Future Work

This paper presents a system that classify opinions about products and services from consumers' perspective and experience on Twitter for recommendation. The system is useful to all customers for all kinds of products and services. Our model is also beneficial to know about best product and service.

6. References

- [1] Jawa, V., &Hasija, V. (2015). A Sentiment and Interest Based Approach for Product Recommendation. 2015 17th UKSim-AMSS International Conference on Modelling and Simulation (UKSim). doi: 10.1109/uksim.2015.26
- [2] Shaikh, S. H., & Lobo, L. M. R. J. (2016). Revealing insights for sales based on analysis of Twitter product reviews. 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC). doi: 10.1109/icgtspicc.2016.7955303
- [3] Trupthi, M., Pabboju, S., & Narasimha, G. (2017). Sentiment Analysis on Twitter Using Streaming API. 2017 IEEE 7th International Advance Computing Conference (IACC). doi: 10.1109/iacc.2017.0186
- [4] S. Raju and G. Poravi, "Recommender System for Generic User Preferences for Online Content," 2018 3rd International Conference for Convergence in Technology (I2CT), 2018.
- [5] P. R. Mala and S. S. Devi, "Product response analytics in Facebook," 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), 2017.
- [6] F. B. Kharrat, A. Elkhleifi, and R. Faiz, "Recommendation system based contextual analysis of Facebook comment," 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), 2016.
- [7] A. Goel, J. Gautam, and S. Kumar, "Real time sentiment analysis of tweets using Naive Bayes," 2016 2nd International Conference on Next Generation Computing

- Technologies (NGCT)*, 2016.
- [8] A. K. Soni, "Multi-lingual sentiment analysis of Twitter data by using classification algorithms," *2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 2017.
 - [9] OpenNLP Python with NLTK <https://opennlp.apache.org/> (Last Accessed on 31/10/2019)
 - [10] IBM Watson AlchemyAPI <https://www.ibm.com/watson/alchemy-api.htm> (Last Accessed on 31/10/2019)
 - [11] Tweepy API <https://www.tweepy.org/> (Last Accessed on 31/10/2019)
 - [12] Domoto, E., Shibata, J., Fukushima, M., & Okuhara, K. (2018, December). *Multivariate Analysis of Text Data from Image Information Acquired by Wearable Device*. In 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS) (pp. 405-408). IEEE.