

# Sentiment Analysis

Ruchi Bhatt\* and Dr. Prinima Gupta

Department of Computer Science and Engineering, Manav Rachna University,  
Faridabad, India; bhattruchi16@gmail.com

## Abstract

**Background/objectives:** Opinions and its related concepts such as sentiments, evaluations, attitudes, and emotions are the subjects of study of sentiment analysis and opinion mining. **Methods/findings:** This research introduced the field of sentiment analysis and opinion mining and surveyed the current state-of-the-art. After that, we discussed the widely studied topic sentiment analysis based on Twitter and need of opinion spam detection. We discussed "Sentiment Analysis on section-377", in which we would discuss about the LGBT community, about the pride parades, about the action taken by the Supreme Court, and about the sentiments or opinion of people towards the section-377 through a poll.

**Keywords:** Sentimental Analysis, Pride Parade, Twitter, Opinion Spamming, LGBT Community

## 1. Introduction

Sentiment analysis, which is also known as opinion mining,<sup>1</sup> is the analysis or field of study of people's opinion, reviews on any topic, appraisals, sentiments, attitudes evaluations and emotions towards entities such as products, organisations, services, issues, and events.<sup>2,3</sup> It is basically the representation of good and bad responses of the people regarding any topic and their attributes.

Sentiment analysis is the mining of the text which identifies information. The social media and its outlet play a very important role in this area. Social media outlets include blogs, forum, and twitter.<sup>4</sup> Nowadays, the networking sites have gained so much attention that we count the reviews and scale ratings based on these reviews that we receive from these sites. Sentiment analysis always helps people to understand a product, brand, or service while monitoring online conversations or reviews.

Sentiment analysis contains key aspects of the brand's product and service that customer uses and underlying intentions and reactions concerning those aspects. The focus is on methods that seek to address the new challenges raised by sentiment-aware applications.<sup>5</sup> This study gives a broader view to make wise or smart decisions and also the products are measured quickly based on this approach.

## 2. Different Level of Analysis

- 1. Document Level:** In this level, the job (or task) goes to the whole opinion document and determines whether a whole document conveyed a good opinion or a bad opinion. At this level, the task is to examine or describe whether a whole opinion document expresses a positive or a negative sentiment. There is an assumption at this level of analysis which is each and every document is supposed to express opinions on a single entity, and therefore, document level is not applicable for those documents which evaluate or compare multiple entities.<sup>6</sup>
- 2. Sentence Level:** In this level, the task goes to the sentences. This task is to classify whether each sentence conveyed a positive, negative, or neutral sentiment. Neutral generally means no opinion.
- 3. Entity and Aspect Level:** Both the level mentioned above, sentence level and the document level, examines do not discover what people exactly like or do not like. Earlier, aspect level was known as feature level. Aspect level directly looks at the opinion itself instead of looking at how language constructs. Basically, it is based on the plan that an opinion consists of a sentiment and a target. There is a limited use of opinion or

\*Author for correspondence

we can say the opinion being identified is of limited use when that opinion does not consist of any target. To understand the sentiment analysis problem better, we need to realise that how much the opinion targets matters or the importance of opinion targets.

### 3. Case Study of Sentiment Analysis Based on Twitter

Twitter is a popular social networking site for microblogging and social networking, counting hundreds of millions of active users and daily published messages.<sup>7</sup> Twitter connects millions of people throughout the world. In twitter, people share their reviews or feelings regarding movies or product or any entity.

Twitter users include friends, followers, @mentions and RTtweets, users connected based on following, friend, RT, or mentoring. Some special features common on twitter enabling connectivity are direct/indirect connections, global visibility, instant sharing. Tweet property and availability includes publicity, visibility, and limited space.

In twitter, each user can choose to follow a number of other users, known as followees, and users can be followed by other users, known as followers. In twitter, user can follow anyone whether it is a Prime Minister or a President or any political leader or any Bollywood or Hollywood star without any limitations and without any discrimination. As Twitter provides microblogging service, hence anyone can use twitter to publish short messages counting a maximum of 140 tweets (or characters).<sup>8</sup> These short messages may consist of facts, thoughts, opinions, and other media. Tweets are classified as either positive or negative concerning a query term.<sup>2</sup>

It is possible to use references to discussion topics through the “#” symbol and through the “@” symbol, it is possible to tag other users (i.e., introduce references to other users).

There are three types of data – User type, Tweet type, and Friend type.

1. User type: It represents users’ profiles. User\_id, name, location, num\_followers, and num\_tweets can be obtained from Twitter.
2. Tweet type: It represents posted messages. Tweet\_id, User\_id, message, and date can be obtained from Twitter.
3. Friend type: It represents the “follow” relationships among users.

Follow means number of users want to connect with other user.

As a start, a tweet may contain many elements which are not significant for classification and can thus be dropped through a filtering process. Various filters are defined to polish the message.

<<RT @greatman omg today is the #best day of my life. #excited #joyful #lifeisbeautiful: D http://w.co/4AEP918bV7>>

A first filter eliminates the tokens which are not in use. Removed tokens include the following: “RB” sequence which indicates a retweet; @ character and the whole following user name: # symbol, but not the following topic name, which is kept in message. A second filter applies the language specific rules. In this, unknown words can be removed, which may not appear in any other tweet (e.g., “lifeisbeautiful”). Finally, a third filter separates all the punctuation symbols from the text and organises them as single-character words.

There are very few people who give positive outputs towards the brand or product they are referring to, which generally happens at twitter. Countable number of people are included in negative as well as positive reviews but many in the neutral favor of the entity and its attributes. Here are some observations that are calculated on the percentage basis of the twitter negative, positive, and neutral sentiments as shown in Figure 1.

### 4. Opinion Spam Detection

The principle highlight of online networking is that with no limitation it permits anybody from any place on the planet, without any discrimination, to express their point of view or to share their thoughts and beliefs and conclusion without showing their actual character and

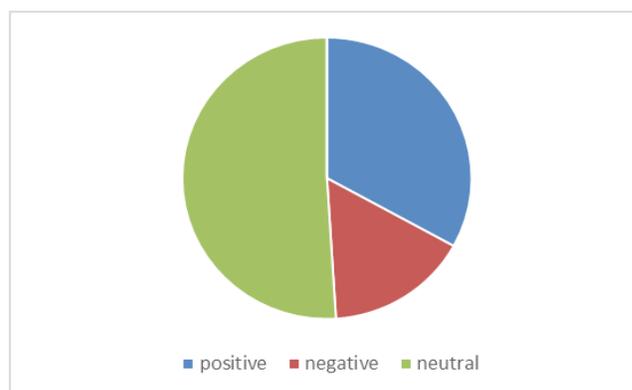


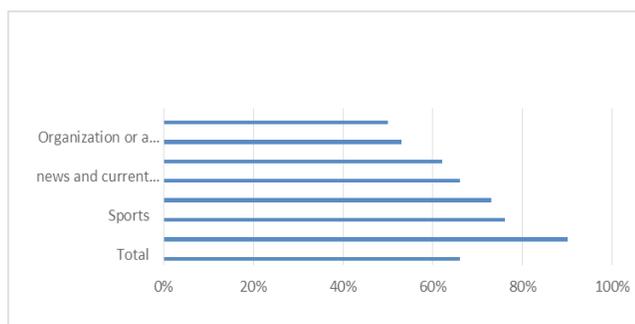
Figure 1. Twitter sentiment.

without the dread of undesirable impacts. Subsequently, these feelings are exceptionally significant. In this case, this hideousness comes with a cost. It permits individuals with concealed plan or malignant aims to effortlessly divert the framework to give individuals the feeling that they are autonomous individuals for people in general and post counterfeit sentiments or wrong data to elevate or to dishonor the target items, administrations, associations, or people without uncovering their actual expectations, or the individual or the association that they are covertly working for. Such people are called supposition spammers and their exercises are called assessment spamming.

Perspective spamming has turned into a noteworthy issue. Aside from people who give counterfeit conclusions in audits and gathering talks, there are additional business organisations and associations that are in the matter of composing counterfeit surveys and fake web journals for their customers for making benefits. A few prominent instances of phony surveys have been accounted for in the news. These days, it is imperative to recognise such spamming exercises and the spammers to guarantee that the sentiments on the web are a confided in a wellspring of the profitable and right data. In contrast to the extraction of constructive and contrary sentiments, assessment spams identification is not only an NLP issue as it includes the investigation of individuals' posting practices. In this manner, it is likewise an information mining issue as shown in Figure 2

## 5. Case Study of Sentiment Analysis: Section-377

India has made significant progress in recent years to further equality people. From the partial decriminalisation of homosexuality 50 years ago to the remarkable progress



**Figure 2.** Automated accounts post the majority of tweeted links pf popular websites across the range of domain.

in advancing equality for LGBT+ people. Discrimination of homosexuality or LGBT+ community is one of the most highlighted issues and it has gained much attention in recent year. LGBT+ people can now enjoy the right to marry, to start a family, and to change their legal gender to match their identity. LGBT+ people face higher inequalities in health satisfaction and outcomes, discrimination, and hate crime, bullying and harassment in education and at work.

LGBT community is a community which describes the culture of lesbian, gay, bisexual and transgender and along with heterosexual and also describes gender identity and or people's sexual orientation. LGBT community celebrates the peace, pride, diversity, sexuality, and individuality.

### 5.1. Pride Parades

Pride parades, which earlier known to be as "gay liberation protests", focuses on celebrating oneself, without being shamed of their identity, imbibing the feelings of strength and unity among people. Every human being has the right to equality, to express their feelings and thoughts with freedom. When was the last time someone was hurt by equality? To support the rights of the LGBT+ community several pride marches take place around the globe. The pride parade comes into action and supports LGBT+ rights, their community and culture. These parades aim at raising issues, spreading awareness and fighting for the LGBT+ rights. It encourages people to come forth and be proud of who they are without fearing anything.

These parades majorly happen over the month of June. In these parades, people come together to rally shouting slogans, and displaying placards, massive floats, dancers, music and so on to celebrate and send out a message of love and equality. Many celebrities, politicians, and sportsperson join in support of these parades. These parades and protests have different objectives of their own, to celebrate proudly about their identity, to fight against the issues and discrimination faced by them, or to stand up for equal rights.

The LGBT+ pride parades have been going on since 1970. The stonewall riots were the beginning of pride marches. The stonewall inn is a popular gay bar where these riots first took place. Over the last 45 years, the parades have been named differently such as – Christopher Street West in 1972, Gay Freedom Day from 1973–1980,<sup>10</sup> and the San Francisco Lesbian, Gay, Transgender, Bisexual

Pride Parade and Celebration from 1995 till now, have played a pivotal role in raising issues and legalising the fundamental human right to love and be loved, without regard to cast, creed, gender, color, politics, or sexuality.<sup>11</sup>

Over the years, the LGBT+ community has been recognised and many rights have been passed, to support them such as same sex marriage is legalisation, domestic partnerships, and anti-discrimination laws have been passed in most of America, Central Africa, East Asia, Europe, and Australia. Still there are 73 countries that do not support the LGBT community and have not legalised same sex marriage, instead they have strict laws and punishment against them. Various countries are recognising the need to open up and come forward to support these parades, the LGBT+ community and their rights. These parades encourage people to love, celebrate, and identify themselves freely.

## 5.2. LGBT Community No Longer a Crime in India

In 2013, Delhi High Court delivered the verdict on Section-377 that the consensual acts by adults of the same sex as offence and treat them as a criminal and provides for life in prison.<sup>12</sup>

In 2016, Supreme Court is ready to hear the petition again and this case continues for approximately 1.5–2 years.

Finally in 2018, five-judge constitution bench delivered the verdict on section-377, which is headed by the Chief Justice Dipak Misra, reverses the court's 2013 judgement and owes an apology, on the behalf of Delhi High Court, for the decisions taken in the past which made them to live a life full of fear and discrimination.

Chief Justice Dipak Misra said that "History owes an apology to these people and their families. Homosexuality is a part of human sexuality. They have the right of dignity and free of discrimination. Consensual sexual acts of adults are allowed for the LGBT+ community."

## 5.3. Final Verdict of Court on Section-377

Supreme Court said that "any discrimination on the basis of sexual orientation violates fundamental rights. History owes apology to the LGBT+ people and their families. They were made to live a life full of fear. Section-377 had become a weapon for the harassment for LGBTs and subject them to discrimination."

The final verdict was delivered on 6 September 2018 which says that "people have no control on sexual orientation and it is natural." Supreme Court reads down 158 years old colonial law and allow LGBT+ community among consenting adults in private.

## 5.4. Public Opinion Regarding LGBT Rights

Public opinion is very complex in India regarding LGBT rights. It took many years to accept and support LGBT+ people (or community).

According to a 2016 poll conduct by the International LGBT and Intersex Association, there are only 35% Indians who were in the favor of legalising same-sex marriage and further 55% were opposed. A survey was conduct by the Varkey Foundation found that among adults, approx. 18–21 year olds, support for the same-sex marriage was higher at 53%. So basically, more than half of the population, all over the world, were against LGBT+ community and only the youth were in favor of or supports LGBT+ community.

In 2017, the percentage of people, who were in the favor of legalising the same-sex marriage, was increased and the increment of this percentage is good. According to the 2017 survey conduct by ILGA, 58% of people were in favor of LGBT+ rights and they agreed that the LGBT+ people should have or enjoy the same rights as the straight people have. Approximately, 39% of Indians said that the LGBT people or the people who are in same-sex relationship should be charged as criminals, while a plurality of 44% Indians disagreed. Additionally, 59% of Indians are in the favor that the LGBT+ people should be protected from the workplace discrimination as shown in Figures 3 and 4.

In 2018, approximately 81% of people were in the support of legalising the same-sex marriage. After the final verdict declared by the Supreme Court, everyone

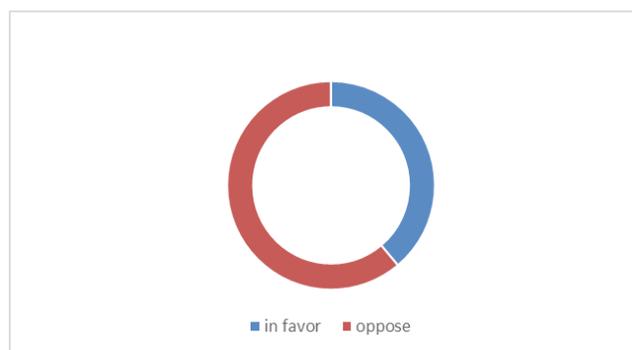
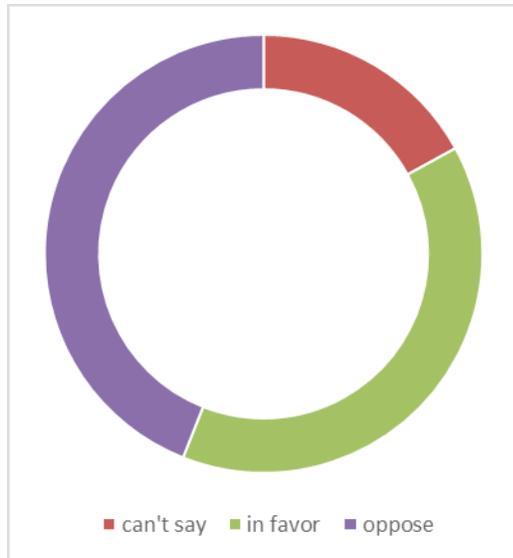


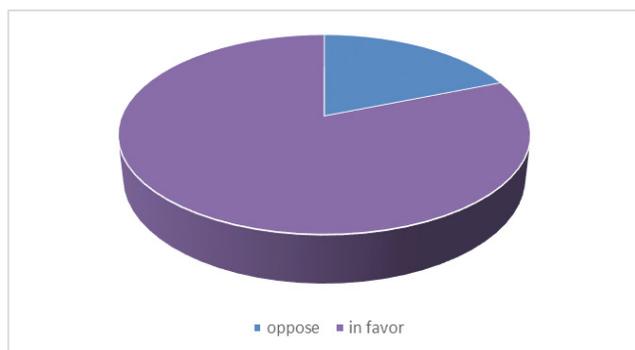
Figure 3. LGBT+ community in India.



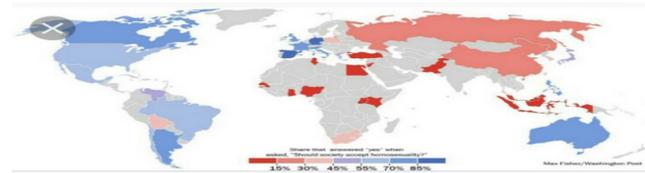
**Figure 4.** LGBT+ people charged as criminals.

shows their support through social media like Instagram, Facebook, and twitter. Many celebrities shows their support by sharing the rainbow colored picture which is the sign of humanity, equality, peace, dignity, individuality, and love. Every Indian praised the Supreme Court verdict. Every Indian were happy and proud to be an Indian and ready to say bye to Section-377.

There are still 73 countries, mostly in the Middle East, Africa, and Asia,<sup>13</sup> that do not support the LGBT+ community, and have not legalised same sex marriage. Instead of supporting the LGBT+ community or legalising same sex marriage, they have strict laws and punishments against them. It was (or is) very hard for LGBT+ community to live in this cruel world who treated them as a criminal just because of the stupid reason that these people are not straight. Everybody needs to understand that to be a gay or lesbian or transgender or bisexual is as natural as to be straight as shown in Figures 5 and 6.



**Figure 5.** LGBT+ community in India.



**Figure 6.** Should society accept homosexuality.

## 6. Summary and Conclusion

Due to many challenging research problems and a wide variety of practical applications, the research in the field has been very active in recent years. It has spread from computer science to management science, as opinions about products are closely related to profits.

This research paper first defines the sentiment analysis. It then discussed the three levels of sentiment analysis. First level is document level which objective is to determine whether an opinion document (e.g., a review) expresses a good or bad sentiment. Second level is sentence level which describes whether a sentence expresses a positive or negative or a neutral opinion. Last level is aspect level which conveys the feature of the sentence or document.

After that, we discussed the widely studied topic sentiment analysis based on twitter. In today's world, people express their views, opinions or feelings regarding a product or service provided by any organisation through social media or World Wide Web. So, basically World Wide Web or social networking plays a central role in everyone's life as it is easy to use and with no limitations, it permits everybody, without any discrimination and without showing their actual identity, can share their point of views regarding any social cause, products, politics or any entity and its attribute and Twitter is the one of the best platform for this.

Last but not least, we discussed about the need of opinion spam detection in which we had assessed the quality of reviews. More and more people are relying on the opinions on the Web for decision making. To ensure the trustworthiness of such opinions, combating opinion spamming is an urgent and critical task.

After that, we discussed the widely studied topic "Sentiment Analysis on section-377", in which we would

discuss about the LGBT community, about the pride parades, about the section-377 passed by the Supreme Court, and about the sentiments or opinion of people towards the section-377.

## Acknowledgement

We would like to express our thanks of gratitude to Accendere Knowledge Management Services for providing us a platform and opportunity to pursue the research.

## References

1. Fang X, Zhan J. Sentiment analysis using product review data. Licensee Sprinkler; 2015.
2. Liu B. A survey of opinion mining and sentiment analysis. University of Illinois at Chicago; 2012. P. 415–63.
3. Banati H, Bhattacharyya S, Mani A, Köppen M. Hybrid intelligence for social networks. Springer Nature; 2017.
4. Yu Y, Duan W, Cao Q. The impact of social and conventional media on firm equity value: a sentiment analysis approach. *Decis Supp Syst.* 2013;55(4):919–26.
5. Pang B, Lee L. Opinion mining and sentiment analysis. *Found Trends Inf Ret.* 2008;2(1–2):1–94.
6. Satapathy R, Cambria E, Hussain A. Sentiment analysis in the bio-medical domain. Springer Nature; 2017.
7. Saif H, He Y, Alani H. Alleviating data sparsity for twitter sentiment analysis; 2012. P. 1–8.
8. Case study for sentiment analysis on twitter. [cited 2015 Jan]. [https://www.researchgate.net/publication/322337829\\_A\\_Case-Study\\_for\\_Sentiment\\_Analysis\\_on\\_Twitter](https://www.researchgate.net/publication/322337829_A_Case-Study_for_Sentiment_Analysis_on_Twitter).
9. Da Silva NFF, Hruschka ER. Tweet sentiment analysis with classifier ensembles. *Decis Supp Syst.* 2014;66:170–79.
10. San Francisco Pride. [cited 2019 Jun 29]. [www.sfproude.org](http://www.sfproude.org).
11. The GLBT Historical Society. [cited 1985]. [www.glbthistory.org](http://www.glbthistory.org).
12. Hindustan Times. [cited 2019 Jun 04]. <http://m.hindustantimes.com>.
13. This is the state of LGBTI rights around the world in 2018. [cited 2018 Jun 14]. <https://www.weforum.org/agenda/2018/06/lgbti-rights-around-the-world-in-2018/>.