## Twitter Sentiment Analysis of India vs Pakistan T20 World Cup Match Using SVM Classifier

[**(**EashanSharma**,**eashansharma31@gmail**.**com**,**](mailto:(EashanSharma%2Ceashansharma31@gmail.com)Aryan gaur**,** [aaryangaur06@gmail.com,](mailto:aaryangaur06@gmail.com) Ms. Shefali Singhal, [shefali.fet@mriu.edu.in](mailto:shefali.fet@mriu.edu.in) **)**

#### Manav Rachna International Institute of Research and Studies

**Abstract:** Feeling investigation or assessment mining is the computational investigation of individuals' viewpoints, perspectives, evaluations, and feelings toward people, elements, occasions, issues, points and their traits. With the huge prominence of online entertainment, individuals have been utilizing this stage to communicate their thoughts, subsequently making it simpler for analysists to concentrate and work on datasets recovered from SNS. It includes extricating tweets from twitter, an illustration of SNS, information purifying and use of an appropriate calculation to get the satisfactory opinion examination. In the ebb and flow research, tweets have been gathered for opinion investigation, for an occasion that happened as of late (India versus Pakistan T20 World Cup ). For this a contextual analysis has been done, utilizing the straight help vector classifier. The accuracy result has been determined while dealing with texts and emojis. The outcomes plainly show the dependence of SVM execution upon input dataset.

**1: Introduction**

With more than 313 multi month to month unique clients and more than 500 million tweets for dependably, twitter has now transformed into a goldmine for affiliations and individuals who have a strong social, political or money related revenue in keeping up and chipping away at their clout and reputation **[4].** Thought evaluation regardless called evaluation mining or feeling AI suggests the use of regular language dealing with, message assessment, computational phonetics, and biometrics to intentionally see, autonomous, survey, and focus on energetic states and dynamic data. Evaluation appraisal is widely applied to voice of the client materials, for example, audits and review reactions, on the web and electronic media, and clinical advantages materials for applications that span from publicizing to client sponsorship to clinical medication **[5].** Twitter is among most comprehensively used electronic media stage to present their perspectives and sentiments while connecting with others online **[6].** Different clients use twitter to pass on their presumptions and considerations on various centers, occasions, various audits on things, associations, and other Twitter clients they are excited about. Hypothesis appraisal on such tweets can be performed by various firms to get a data on friendly classes' point of view on various occasions to foresee individuals' reaction towards their things **[12].** Thusly, doubtlessly knowing customer knowledge from Twitter data is amazingly important for certain applications. Thusly, evidently knowing customer knowledge from Twitter data is incredibly significant for certain applications. The essential arrangement of this paper, using a logical examination of ICC T20 World Cup 2021, is to inspect the adequacy of Support Vector Machine Classifier on two predefined datasets **[7].** In this paper we have talked about the system on the most proficient method to extricate the information and perform opinion investigation on it to acquire the outcomes as Precision, Recall and F-measure in section 3 and section 4 individually. Section 5 talks about the finish of this examination with the future angles examined in segment 6.

### 2: Literature review

We have contemplated, analyzed and consolidated information from some of the most accomplished research papers. like “Sentiment Analysis of Twitter Data: A Survey of Techniques” by Vishal Kharde and Sheetal Sonawane which gives an extensive study on distinctive arrangement methods and KPIs **[1]** orfrom “Sentiment Analysis of Tweets using SVM” by Munir Ahmad, ShabibAftab and Iftikhar Ali which talks about how to utilize SVM on pre-grouped datasets **[8].** From “Comparative Analysis of Sentiment Orientation Using SVM and Naïve BayesTechniques” by Shweta Rana and Archana Singh we figured out how to compute Precision, Recall and F-measure for a classifier and incorporate these learning in the following reseaech **[9].** Research papers such as Barnaghi, P. Ghaffari, and J. G. Breslin, "Text Analysis and Sentiment Polarity on FIFA World Cup 2014 Tweets," in Conference ACM SIGKDD, 2015 and **“**Twitter Sentiment Analysis on Worldwide COVID-19 Outbreaks” by Kamaran H. Manguri and Rebaz N. Ramadhan gave us the idea to perform opinion investigation on a recent and genuine occasion and do a contextual analysis **[10].** From **“**Machine Learning-Based Sentiment Analysis for Twitter Accounts” by Ali hasan and Sania moin we separated SVM classifier into our examination **[11].**

### 3: Methodology Used

The test results were obtained by utilizing twitter reviews. This area is experimentally advantageous in light of the fact that there are large on-line accumulations of such reviews.

### : Dataset

Phantom buster’s twitter web crawler was used to extract randomly chosen tweets that contained text, emoticons and gestures created using parenthesis and salutations. This area is useful in light of the fact that there are huge online assortment of twitter posts from various users worldwide and their opinions can be summarized into positive, negative and neutral statements. Two datasets which contained 300 tweets each were used.

|  |  |
| --- | --- |
| Dataset1 #indvspak | |
| Tweet Date | Content |
| Sun Oct 24 | Under Virat Kohli, India is making new histories.Champion Trophy 2017:India lost to Pakistan in finals by 180 runs. Cricket T20 World Cup 2021: India lost to Pakistan in inaugural by 10 wickets. Lekin gyan inse lo Diwali par !#IndvsPak |
| Sun Oct 24 | Dear neighbour, It's hard to fix a broken TV, even for us. Break something else  today.#IndvsPak #ThirteenZero #SixZero. |
| Sun Oct 24 | #IndvsPak Did you miss MS Dhoni. |
| Sun Oct 24 | On this anxious morning, one little counsel. It is generally really smart to watch cricket by fending off governmental issues, disdain and presumption. Partake in the day, praise your success not your adversaries rout. Treat it as a game not war. #indvspak |
| Sun Oct 24 | Wow what a beautiful moment between Kohli and Rizwan! #IndvsPak |

Table 3.1 Dataset 1 #indvspak

|  |  |
| --- | --- |
| Dataset 2 #shami | |
| Tweet Date | Content |
| Mon Oct 25 | A very horrific campaign against Muhammad Shami. It is disgraceful to abuse him as a Muslim. Stop this nonsense. #Shami #StandWithShami. |
| Wed Oct 27 | Most of the trolls who abused Mohammed Shami were from Pakistan but Indian Liberals blamed Indian majority for mistreating fellow minority guy.  This is how fake narrative originates which pushes Indians into a guilt trip. |
| Mon Oct 25 | Assamese News channel @PragNews needs to know how much cash #Shami took from Pakistan. Calls him a Pakistani specialist in the Indian group. |
| Mon Oct 19 | Mohammad #shami we are all with you These individuals are loaded up with disdain since no one gives them any adoration. Pardon them. |
| Mon Oct 25 | Cricketers are standing up for #Shami while Bollywood still lacks the spine to stand up for SRK. |

Table 3.2 Dataset 2 #shami

### : Linear Support Vector Machine

Straight Support vector Machine turned out to be perhaps the most undeniable learning strategy for dealing with the backslide and gathering issues. Support vector machines address an expansion to nonlinear models of the summed up picture calculation created by Vladimir Vapnik. The SVM calculation depends on the measurable learning hypothesis and the VapnikChervonenkis (VC) aspect presented by Vladimir Vapnik and Alexey Chervonenkis. Support vector machines (SVM) are a gathering of directed learning strategies that can be applied to order or relapse. A couple of techniques were concocted and examined due to centrality of the SVM advancement issue which are talked about in **[3].** Straight Support Vector Machine works on immense online dataset which are taken from the web based objections and become well known considering its applications in text course of action, word-sense disambiguation. Backing Vector Machines (SVM) are another legitimate learning method that can be seen as another technique for get ready arranging classifiers contemplating polynomial cutoff points, widened premise limits, brain systems. Backing Vector Machines use a hyper-plane disengaging plane to make a classifier. For issues that can't be clearly isolated in the space of information, this machine offers a probability to observe an answer by making an improvement of the main information space

into a high layered part space, where an ideal segregating hyper plane can be found.

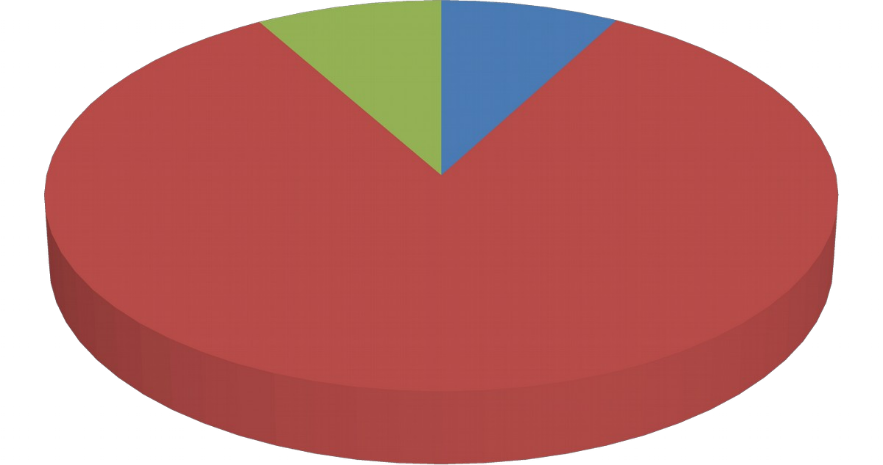
### : India vs Pakistan Case Study

An example was taken to perform sentiment analysis of tweets example being, ICC T20 World Cup. Like any other festival, cricket match between India and Pakistan is celebrated in a grand manner where almost every person from all age groups and from both sides of the border are glued to their screens. In the recent confrontation of the two teams in ICC T20 Men’s World Cup, 2021, social media was hyped for the match and rooting for india but after its defeat, team India suffered a great backlash on all SNS platforms. All the frustration was ranted out on Indian skipper Mohammad Shami for his bad performance.

### : During Match Proceedings

During the showdown among India and Pakistan a tremendous pool of remarks and sentiments were being overwhelmed via online media stages. Fans were applauding group india and dissing Pakistan via online media. There was a ton of certainty towards India's success as a result of the historical backdrop of conflicts where India had won on numerous occasions A separation of a portion of the remarks of this occasion to break down the opinion of individuals towards this occasion. Our examination dependent on the remarks during the match is displayed in Fig 1.

# Sentiment Polarity During Match



Average Positive; 8.39% Average Negative; 8.05%

Average Neutral; 83.56%

Fig1.Sentiment Polarity during India vs Pakistan

### : After Match Proceedings

India experienced a merciless misfortune as Pakistan dominated the game without losing a solitary wicket. After the finish of the match an Indian player named Mohammad shami confronted a ton of reaction from online media. He was manhandled on his religion, blamed for being in plot with Pakistan and a lot more charges. He had been a subject to this treatment a no of times before likewise because of his religion. After every one of the derisive remarks individuals showed up to openly endorse shami flooding the web with remarks with hashtags, for example, #shami and #istandwithshami numerous cricketers and superstars additionally upheld shami saying that anybody can have an off day we ought not hold it against them and misuse them. A portion of the remarks was extricated and was played out to decide general visibility of this occasion. The after effects of the investigation is displayed in Fig 2

# Sentiment Polarity After Match

100

90

80

70

60

50

40

30

20

10

0

Average Neutral Average Negative Average Positive

Fig: 2 Sentiment Polarity after the India vs a Pakistan match

### 4: Results

The exhibition of Sentiment classifier SVM was determined for both the predefined datasets. From the graphs given above it is clearly visible that after the match team India was bashed brutally, especially Shami which resulted in the increase of negative polarity from the people **[1].** The increase in negative polarity is approximately 5%. The Performance was determined based on Precision, Recall and F-Measure. The outcomes determined are shown in table 3 and table 4.

#### : Precision

Precision is a metric that evaluates the quantity of right certain expectations made. Accuracy, in this manner, computes the exactness for the minority class. It is determined as the proportion of accurately anticipated positive models separated by the absolute number of positive models that were anticipated.

Precision = TruePositives/(TruePositives + FalsePositives)

#### : Recall

Recall is a metric that assesses the amount of right specific figures made from all certain assumptions that could have been made. Not in any way like accuracy that primary comments on the right certain assumptions out of each certain estimate, audit offers a hint of missed positive assumptions. Consequently, review thinks about of the consideration of the positive class.

Recall = TruePositives/(TruePositives + FalseNegatives)

#### : F-measure

Request accuracy is extensively used in light of the fact that it is one single measure used to summarize model execution. F-Measure gives an approach to uniting both exactness and survey into a lone measure that gets the two properties. Alone, neither Precision or Recall recaps the whole story. We can have sublime accuracy with shocking review, then again, horrendous accuracy with exceptional review. F-measure gives an approach to discussing the two concerns with a single score. Whenever accuracy and review not entirely set in stone for a twofold or multiclass portrayal issue, the two scores can be joined into the assessment of the F-Measure. The standard F not entirely set in stone as keeps:

F-Measure = (2 \* Precision \* Recall)/(Precision + Recall)

|  |  |  |  |
| --- | --- | --- | --- |
| Performance For Dataset 1 #indvspak | | | |
| Classification | Precision | Recall | F-Measure |
| Negative | 37.50% | 6.82% | 11.55 |
| Neutral | 22.23% | 84.84% | 35.23 |
| Positive | 33.34% | 7.07% | 11.67 |

Table 4.1 Performance of SVM for dataset 1

|  |  |  |  |
| --- | --- | --- | --- |
| Performance For Dataset 2 #Shami | | | |
| Classification | Precision | Recall | F-Measure |
| Negative | 28.33% | 12.78% | 17.62 |
| Neutral | 32.43% | 78.68% | 45.92 |
| Positive | 37.21% | 11.12% | 17.12 |

Table 4.2 Performance of SVM for Dataset 2

### 5: Conclusion

In this paper, the investigation of the exhibition of Support Vector Machine (SVM) for opinion examination was conducted . For execution investigation of SVM, two pre-grouped datasets of tweets , first dataset comprised of tweets in regards to India versus Pakistan Match and second dataset was about the Indian Player Mohammad shami. An internet based disarray framework device is utilized for execution investigation and examination. Results are estimated as far as Precision, Recall and f-measure. The outcomes plainly show the reliance of SVM execution upon input dataset. The presentation reliance of SVM and other AI procedures ought to be investigated further by utilizing enormous and diverse datasets. For near examination the consequences of this paper can be utilized as pattern. Besides it ought to likewise be examined that for grouping reason, which AI calculation performs better on which kind of dataset and what may be the reasons? This can lead the analysts to the further developed renditions of AI calculations for grouping reason**.**

### 6: Future aspects

Sentiment Analysis made a significant advancement few years prior by sorting emojis as certain, negative or impartial. However, the ultimate objective is yet to be reached. The time to expand the perspective and begin ordering extremity with simply certain, negative and nonpartisan and go past and make subcategories like mockery, provoking, provocative, irate, and so forth. This region will require an extremely through and profound examination and formation of many new models, equations, classifiers, and so forth yet we trust that with headway in AI, sentiment analysis will likewise turn out to be a lot of straightforward and careful.

**References**.

* + 1. Kharde, V. and Sonawane, P., 2016. Sentiment analysis of twitter data: a survey of techniques. *arXiv preprint arXiv:1601.06971*.
    2. Sahayak, V., Shete, V. and Pathan, A., 2015. Sentiment analysis on twitter data. *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, *2*(1), pp.178-183..
    3. Pak, A. and Paroubek, P., 2010, May. Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*.
    4. Shalev-Shwartz, S., Singer, Y., Srebro, N. and Cotter, A., 2011. Pegasos: Primal estimated sub-gradient solver for svm. *Mathematical programming*, *127*(1), pp.3-30..
    5. Jianqiang, Z. and Xiaolin, G., 2017. Comparison research on text pre-processing methods on twitter sentiment analysis. *IEEE Access*, *5*, pp.2870-2879.
    6. Pang, B. and Lee, L., 2008. Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, *2*(1–2), pp.1-135..
    7. Bifet, A. and Frank, E., 2010, October. Sentiment knowledge discovery in twitter streaming data. In *International conference on discovery science* (pp. 1-15). Springer, Berlin, Heidelberg**.**
    8. Ahmad, M., Aftab, S. and Ali, I., 2017. Sentiment analysis of tweets using svm. *Int. J. Comput. Appl*, *177*(5), pp.25-29.
    9. Rana, S. and Singh, A., 2016, October. Comparative analysis of sentiment orientation using SVM and Naive Bayes techniques. In *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)* (pp. 106- 111). IEEE.
    10. Manguri, K.H., Ramadhan, R.N. and Amin, P.R.M., 2020. Twitter sentiment analysis on worldwide COVID-19 outbreaks. *Kurdistan Journal of Applied Research*, pp.54-65.
    11. Hasan, A., Moin, S., Karim, A. and Shamshirband, S., 2018. Machine learning-based sentiment analysis for twitter accounts. *Mathematical and Computational Applications*, *23*(1), p.11.
    12. Sharma, E., Gaur, A. and Singhal, S., 2022. In-Depth Opinion Mining and Sentiment Analysis of Twitter Data. In *Proceedings of International Conference on Data Science and Applications* (pp. 37-51). Springer, Singapore.