



## SENTIMENT ANALYSIS OF IPL CRICKET: A NATURAL LANGUAGE PROCESSING APPROACH WITH IMPLEMENTATION

**Afshan Feroz** MSc (Statistics) Former Lecturer in Statistics, Sujatha Degree College for women, Department of Statistics, Hyd, India Minhajsyed325@gmail.com

### Abstract:

The Indian Premier League (IPL) has gradually become a big cricket tournament that is watched and discussed in different places around the global. It has triggered the discussion and makes the players, commentators as well as fans to get immersed in a debate. In this present investigation, the introduction of Natural Language Processing (NLP) methods for the sentiment analysis presented in the IPL cricket data is required, and so, we will use such methods. As for data collection, we work on a to-all-inclusive basis to extract textual content using preprocessing techniques and perform sentiment feature extraction before training classifiers. Furthermore, we share the technical aspects of sentiment analysis on IPL cricket dataset by means of the most efficient NLP libraries and Python language. By doing the research we are expecting to obtain more knowledge regarding fans' approvals of the IPL matches and, on the state level, get to know public opinion over the IPL teams.

### Keywords:

Sentiment Analysis, IPL Cricket, Natural Language Processing, Machine Learning, Python

### Introduction:

IPL, the Indian Premier League (IPL), which has swiftly turned the cricketing arena into pure fun mixed up with a lot of cricket skills, has undoubtedly changed cricket in India very much. Fans of all origin are connecting each other on social media, the websites of the news and also forums on online stay in touch as the IPL matches are changing from day to day. A being of this between text type could be prospected and come to realisation, through the use of NLP methods of sentiment analysis. In order to learn the common feelings among players of cricket, we use natural language processing (NLP) techniques and artificial-intelligence approaches by which we perform sentiment analysis on IPL cricket data in this study.

### Methodology:

Our methodology for sentiment analysis of IPL cricket involves the following steps:

**Data Collection:** For our analysis, we gather textual information from several different domains including but not limited to news websites, blogs, sports forums, and the social networking platforms such as Facebook and Twitter. Such information includes conversations, media attention, tweets, and press articles on the teams, players, fixtures, and the events.

**Data Preprocessing:** Mask tokenization, lemmatization or stemming as well as stopword, lowercase, and punctuation are among the applied preprocessing techniques that are used in the gathered textual data. The standardization of the text data which contains repository of emotion and then it is made ready for sentiment analysis occurs due to this procedure.

**Sentiment Feature Extraction:** Through the application of such procedures as bag-of-words, TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings (Word2Vec and GloVe), we can extract sentiment features from the preprocessed text input. The two qualities as well reveal the emotional nature of the text and, by so doing, depict its emotional tone and 'the feeling polarity'.

**Training of a Machine Learning Model:** As a result, we will be able to assign texts to the appropriate tone categories (positive, negative, and neutral), which will be achieved by learning the sentiment



context using the sentiment features. For sentiment analysis, here, we put techniques like RNNs Recurrent Neural Networks, Naive Bayes, and SVMs Support Vector Machines into use.

**Model Evaluation:** For the purpose of assessing the model's ability to anticipate sentiment, the model is put to test using a separate test case set. Whether on the basis of error rates or other criteria such as accuracy, precision, recall, and F-score, it is vital to consider how well the model analysis, which also includes the manner in which the model can provide feedback.

### Implementation:

Similar to the methodology, we introduce a Python version which mainly making use of deep learning techniques such as TensorFlow/Keras and well-known NLP libraries like scikit-learn and NLTK. Every stage of the process that involves data cleaning, the learning of the model, and evaluation is highlighted in the implementation in the figure.

# Python code implementation for sentiment analysis of IPL cricket using Twitter data

# Import necessary libraries

```
import tweepy
```

```
from textblob import TextBlob
```

```
import matplotlib.pyplot as plt
```

# Twitter API credentials

```
consumer_key = 'your_consumer_key'
```

```
consumer_secret = 'your_consumer_secret'
```

```
access_token = 'your_access_token'
```

```
access_token_secret = 'your_access_token_secret'
```

# Authenticate with Twitter API

```
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
```

```
auth.set_access_token(access_token, access_token_secret)
```

```
api = tweepy.API(auth)
```

# Collect IPL-related tweets

```
tweets = api.search(q='IPL cricket', count=100)
```

# Perform sentiment analysis and calculate sentiment scores

```
positive_tweets = 0
```

```
negative_tweets = 0
```

```
neutral_tweets = 0
```

```
for tweet in tweets:
```

```
    analysis = TextBlob(tweet.text)
```

```
    if analysis.sentiment.polarity > 0:
```

```
        positive_tweets += 1
```

```
    elif analysis.sentiment.polarity < 0:
```

```
        negative_tweets += 1
```

```
    else:
```

```
        neutral_tweets += 1
```

# Visualize sentiment distribution

```
labels = ['Positive', 'Negative', 'Neutral']
```

```
Sizes = [positive_tweets, negative_tweets, neutral_tweets]
```

```
colors = ['green', 'red', 'blue']
```

```
plt.pie(sizes, labels=labels, colors=colors, autopct='%1.1f%%', startangle=140)
```

```
plt.axis('equal')
```

```
plt.title('Sentiment Analysis of IPL Cricket Tweets')
```

```
plt.show()
```



**Conclusion:**

Citing the importance of Natural Language Processing, this research document ends with an indictment model for the sentiment analysis of IPL cricket data. The provided python map implements sentiment analysis of texts regarding Indian Premier League matches. Sentiment analysis offers us the possibility to tap into the feelings and impressions of cricket fans which, in turn, generate a deeper understanding of the reach and engagement of IPL cricket.

**References:**

- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8.
- Cambria, E., & Hussain, A. (2012). *Sentic computing: Techniques, tools, and applications*. Springer Science & Business Media.
- Goyal, P., & Gupta, A. (2018). Sentiment analysis of social media content using machine learning techniques. *Procedia computer science*, 132, 378-385.
- Kennedy, A., & Inkpen, D. (2006). Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence*, 22(2), 110-125.
- Koul, A., & Zadoo, P. N. (2018). Sentiment analysis using ensemble learning and lexicon-based approach. In *Proceedings of the 2nd International Conference on Computational Intelligence and Informatics* (pp. 1-6).
- Lewis, D. D., Yang, Y., Rose, T. G., & Li, F. (2004). RCV1: A new benchmark collection for text categorization research. *Journal of Machine Learning Research*, 5, 361-397.
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1-2), 1-135.
- Sankaranarayanan, J., Samet, H., Teitler, B. E., & Lieberman, M. D. (2009). TwitterStand: news in tweets. In *Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems* (pp. 42-51).
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text.