

ABSTRACT

In the digital age and technological advancements, social media has become an important platform for users to share their thoughts, opinions, and feelings on a variety of topics, including sporting events such as the Asian Cup. So, it is important to understand the views and opinions of the public by analyzing social media data. This analysis classifies them into three categories: positive, negative, and neutral based on a predetermined class. This data analysis technique uses CRISP-DM, which is an industry-standard data mining process that starts with business insights, data understanding, data preparation, modeling, price evaluation, and implementation. The feature is selected using a ranking technique that extends the query so that all data is collected based on a specific class. The number of features needed to improve accuracy. The next step is to apply the classification algorithm technique, with the Naive Bayes technique. The results of classification using the Naive Bayes method in this study have an accuracy level of 92%. After the validation process using K-Fold cross validation, the accuracy value obtained from Naïve Bayes was 92%. Based on the results of the model classification, neutral sentiment dominated with an accuracy result of 64.1%, for positive sentiment had an accuracy result of 33.6%, and for negative sentiment, the correct result was 2.3%. These accurate results show a lot of data that is neutral at the 2024 Asian Cup in Qatar. This shows the number of tweets expressing a neutral stance towards the host of the 2024 Asian Cup in Qatar. Because of this event the participants were not much harmed by the organizing committee or referees and so on. Then the fans were also enthusiastic about participating in the 2024 Qatar Asian Cup until the end of the event because there was no indication of cheating and harming the visiting team in the event.

Keywords: Sentiment Analysis Classification, Asian Cup, Social Media X, Naive Bayes.