

IMPLEMENTASI ALGORITMA *NAÏVE BAYES* TERHADAP ANALISIS SENTIMEN PADA TIMNAS BULUTANGKIS INDONESIA PADA JEJARING SOSIAL TWITTER

Alpin Anugerah

ABSTRAK

Jejaring sosial bukan hanya berguna sebagai media untuk berkomunikasi yang tergolong optimal, tetapi jejaring sosial ini juga bisa berperan sebagai tempat dalam menampung pendapat ataupun opini dari masyarakat secara meluas. Salah satu jejaring sosial yang kerap dipergunakan dalam menampung seluruh opini masyarakat ini ialah *Twitter*. Pihak yang mempergunakan *Twitter* di tanah air kerap kali menyampaikan opininya terkait dengan seluruh bidang, tanpa terkecuali bidang olahraga, terkhusus berkaitan dengan Timnas Bulutangkis Indonesia. Ditemukan banyak opini yang disampaikan masyarakat terkait dengan Timnas Bulutangkis tanah air, dimana hal ini bisa ditemukan pada jejaring sosial *Twitter*. Dari keadaan tersebut, maka dilaksanakannya penelitian yang berhubungan dengan opini masyarakat perihal Timnas Bulutangkis Indonesia. Salah satu metode yang diterapkan yakni melalui pelaksanaan penganalisisan terhadap sentimen pada Timnas Bulutangkis Indonesia dalam jejaring sosial *Twitter* dengan memanfaatkan metode klasifikasi serta algoritma *Naïve Bayes* guna memperoleh klasifikasi terhadap tweet yang bernilai positif ataupun negatif dari masyarakat terhadap Timnas Bulutangkis Indonesia. Jumlah data yang di crawling sebanyak 3277 data lalu data melalui proses *data duplicate removal* sehingga memperoleh hasil data sebanyak 272 data tweet dan belum terlabelkan. Sebelum proses mengklasifikasikan data yang diperoleh, data harus dilaksanakan pemberian label pada datanya serta pembersihan data terlebih dulu sebelum masuk pada tahapan *text processing*, selanjutnya data akan diberi bobot pada setiap kata dengan *Term Frequency– Inverse Document Frequency* (TF-IDF) yang akan kedepannya kata tersebut akan dijadikan sebagai fitur. Lalu, karena data berlabel positif serta negatif mempunyai jumlah yang jauh berbeda, maka dimanfaatkan metode *Synthetic Minority Oversampling Technique* (SMOTE) guna melaksanakan penyeimbangan terhadap datanya. Tahapan selanjutnya yaitu dilaksanakannya pembagian data yang besarnya yakni 80% 20% dan diklasifikasikan dengan metode *Naive Bayes*. Maka hasil yang diperoleh dari pelaksanaan penelitiannya tersebut ialah didapat bahwa data uji memperoleh *accuracy* yang besaran persentasenya yakni 96%, *precision* dengan besaran persentasenya 100%, *recall* dengan besaran persentasenya 92%.

Kata Kunci: Analisis Sentimen, Klasifikasi, Bulu tangkis, *Naïve Bayes*

IMPLEMENTATION NAÏVE BAYES ALGORITMA ON SENTIMENT ANALYSIS OF THE INDONESIAN BADMINTON NATIONAL TEAM ON TWITTER SOCIAL NETWORKS

Alpin Anugerah

ABSTRACT

Social networks are not only useful as an effective medium of communication, but social networks can also be a place to accommodate opinions or opinions among the general public. One social network that is widely used to accommodate these opinions is Twitter. users Twitter in Indonesia often express their opinions on all fields including sports, especially the Indonesian Badminton National Team. There are various kinds of opinions about the Indonesian Badminton National Team, which can be seen on Twitter. Based on these conditions, research was carried out on public opinion about the Indonesian Badminton National Team. One way to do this is to carry out a sentiment analysis of the Indonesian Badminton National Team on the Twitter using the classification method and the Naïve Bayes to classify positive or negative tweets that people express about the Indonesian Badminton National Team. The amount of data that was crawled was 3,277 data and then the data went through a data duplicate removal so that the results obtained were 272 data tweets and had not been labeled. Prior to the process of classifying the data obtained, the data must be labeled on the data as well as cleaning the data first before entering the text processing, then the data will be weighted for each word with the Term Frequency– Inverse Document Frequency (TF-IDF) which will be said in the future. will be used as a feature. Then, because the data labeled positive and negative have a much different amount, the Synthetic Minority Oversampling Technique (SMOTE) method is used to carry out balancing of the data. The next stage is the implementation of the distribution of data whose magnitude is 80% 20% and is classified using the Naive Bayes method. Then the results obtained from the implementation of this research are that the test data obtained accuracy with a percentage of 96%, precision with a percentage of 100%, recall with a percentage of 92%.

Keywords: *Sentiment Analysis, Classification, Badminton, Naïve Bayes*