

ABSTRAK

Ahmad Zaki Baharudin

(170411100054)

Pembimbing I : Devie Rosa Anamisa, S.Kom., M.Kom.

Pembimbing II : Dr. Yeni Kustiyahningsih, S.Kom., M.Kom.

Sejak awal 2019 beberapa negara telah memberitakan virus Covid-19, dimana Virus Covid-19 ini telah memberikan dampak baik negatif maupun positif seluruh dunia dan meningkatnya angka kematian sehingga WHO mencanangkan beberapa program yang dapat membantu memutus rantai penyebaran virus Covid-19, salah satunya adalah program Vaksinasi. Di Indonesia program vaksinasi ini telah memberikan banyak opini masyarakat baik positif maupun negatif mengenai vaksinasi Covid-19. Oleh karena itu pada penelitian ini mencoba merancang suatu sistem mengenai analisis sentimen terhadap vaksinasi Covid-19. Dalam merancang analisis ini melakukan suatu metode, Dalam penelitian ini digunakan istilah frequency-inverse.. document frequency (TF-IDF) yang digunakan untuk pembobotan term (Term Weighting) convert dari tipe data string ke binary sehingga dapat menghasilkan output yang sesuai dengan spesifikasi dari user. Namun pada metode ini membutuhkan waktu komputasi yang cukup lama ketika data dalam jumlah yang besar, Oleh karena itu pada penelitian ini dilakukan penggabungan metode yaitu metode Term Frequency – Inverse Document Frequency (TF-IDF) dan Support Vector Machine (SVM). Dimana metode Support Vector Machine (SVM) adalah salah satu metode yang memiliki waktu yang cukup cepat dalam mengklasifikasi data dalam jumlah yang besar. Tahap evaluasi dilakukan menggunakan *K-Fold Cross Validation* dengan jumlah *10-fold*. Hasil percobaan dilakukan menggunakan data *testing* dengan label *positif* sebanyak 71 data dan label *negative* 192 data. Dari hasil percobaan mendapatkan nilai persentase rata-rata dari *F1-Score* 94.58% dan *classification accuracy* sebesar 95.85%. Model terbaik di dapatkan pada percobaan evaluasi menggunakan *10 fold cross validation* dengan nilai persentase *F1-Score* terbaik didapatkan pada *fold* ke 2 mencapai 98.08% sedangkan nilai persentase akurasi terbaik didapatkan pada *fold* ke 2 mencapai 98.47%.

Kata Kunci : *Analisis Sentimen, Vaksinasi, Covid-19, metode, TF-IDF, SVM, K-Fold Cross Validation*

ABSTRACT

Ahmad Zaki Baharudin

(170411100054)

Pembimbing I : Devie Rosa Anamisa, S.Kom., M.Kom.

Pembimbing II : Dr. Yeni Kustiyahningsih, S.Kom., M.Kom.

Since the beginning of 2019 several countries have reported the Covid-19 virus, where the Covid-19 Virus has had both negative and positive impacts throughout the world and increased mortality rates, so WHO has launched several programs that can help break the chain of the spread of the Covid-19 virus, one of which is Vaccination programme. In Indonesia, this vaccination program has given a lot of public opinion, both positive and negative, regarding the Covid-19 vaccination. Therefore, in this study, we try to design a system regarding sentiment analysis on the Covid-19 vaccination. In designing this analysis, a method is used. In this study, the term frequency-inverse is used. Document frequency (TF-IDF) is used for term weighting (Term Weighting) convert from string data type to binary so that it can produce output that matches the specifications of the users. However, this method requires a long computational time when the data is large. Therefore, in this study, a combination of methods was carried out, namely the Term Frequency - Inverse Document Frequency (TF-IDF) and Support Vector Machine (SVM) methods. Where the Support Vector Machine (SVM) method is one method that has a fairly fast time in classifying large amounts of data. The evaluation stage is carried out using K-Fold Cross Validation with a 10-fold amount. The results of the experiment were carried out using testing data with a positive label of 71 data and a negative label of 192 data. From the experimental results, the average percentage value of the F1-Score is 94.58% and the classification accuracy is 95.85%. The best model was obtained in the evaluation experiment using 10 fold cross validation with the best F1-Score percentage value obtained in the 2nd fold reaching 98.08% while the best accuracy percentage value obtained in the 2nd fold reaching 98.47%.

Keyword : *Sentiment Analysis, Vaccination, Covid-19, method, TF-IDF, SVM, K-Fold Cross Validation*