

Analisis Teks pada Tweet Berbahasa Indonesia untuk Mendeteksi Pro Kontra Vaksinasi Menggunakan Pendekatan Stance Detection dan Topic Modeling = Text Analytics on Indonesian Tweets to Detect Pro vs Anti Vaccination Using Stance Detection and Topic Modeling

Nia Dwi Rahayuningtyas, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20511035&lokasi=lokal>

Abstrak

Keraguan dan penolakan orang tua terhadap vaksinasi meningkat secara global. Maraknya penyebaran isu vaksinasi melalui media sosial mengarahkan persepsi publik pada keraguan terhadap vaksin yang berujung pada penurunan cakupan imunisasi dan tidak tercapainya target IDL di Indonesia. Pada media sosial Twitter terdapat dua kelompok, yaitu kelompok pro-vaksin yang mendukung vaksinasi dan anti-vaksin yang menolak vaksinasi.

Penelitian ini bertujuan untuk mengidentifikasi apakah sebuah Tweet memiliki kecenderungan ke arah pro atau anti-vaksin dan untuk mengeksplorasi topik-topik terkait pro-vaksin dan anti-vaksin. Dataset diambil dari Twitter dengan kata kunci "vaksin" dan "imunisasi" lebih dari 9.000 data Tweet antara 11 Agustus sampai 10 September 2019. Anotasi dilakukan dalam 3 langkah berturut-turut dengan tiga pasangan label yaitu RELEVANT/IRRELEVANT, SUBJECTIVE/NEUTRAL, dan PRO/ANTI. Tiga eksperimen yaitu pemilihan fitur, algoritma, dan *pipeline* klasifikasi dilakukan untuk mendapatkan model *stance detection* terbaik yaitu nilai rata-rata *micro* tertinggi dari *precision*, *recall*, dan *f1-score*.

Fitur terpilih adalah kombinasi 3 fitur

teks *Count* + *Unigram* + *Bigram* dengan algoritma *Logistic Regression* dan *pipeline Two-stage Classification* (*f1-score* = 80,5%). Algoritma terpilih pada pembentukan *topic modeling* adalah NMF dan LDA masing-masing untuk korpus pro-vaksin dan anti-vaksin dengan nilai koherensi sebesar 0.999.

Topik-topik anti-vaksin meliputi kritik terhadap fatwa halal MUI untuk Vaksin MR, kandungan babi pada Vaksin Meningitis Haji, komersialisasi vaksin, vaksin palsu, KIPI dan bahaya vaksin, vaksin sebagai alat konspirasi dan agenda Yahudi, tuntutan vaksin halal, dan seterusnya. Sedangkan topik-topik pro-vaksin lebih bersifat homogen yaitu mengenai manfaat dan pentingnya imunisasi, aturan pemberian vaksin, dan kampanye dalam bentuk publisitas kegiatan imunisasi, dan anjuran vaksin.

Parents hesitancy and refusal toward immunization was rising globally. The rise of the issue of vaccination through social media directs the public's perception of vaccine hesitancy that lead to a reduction in immunization coverage and the unfulfilled IDL target in Indonesia. There are two groups: pro-vaccine that support vaccines and anti-vaccine that refuse vaccines for various reasons that expressed in tweets on Twitter.

This research aims to identify whether a tweet has a tendency to support, or oppose immunization or vaccines and exploring the topic of pro-vaccine and anti-vaccine corpus. The dataset was taken from Twitter with the keywords "vaksin" and "imunisasi" of more than 9,000 tweets at 11 August until 10 September 2019. Annotation was carried out in 3 consecutive steps with three couple label namely RELEVANT vs IRRELEVANT, SUBJECTIVE vs NEUTRAL, and PRO vs ANTI.

Three experiments, namely the selection of features, algorithms, and pipeline were carried out to get the best model of stance detection which has the highest micro average precision, recall, and f1-scores. The selected feature is combination of **Count** + **Unigram** + **Bigram** features with **Logistic Regression** and **pipeline** **Two-stage Classification** (**f1-score** = 80,5%).

The selected topic modeling algorithms are NMF and LDA for the corpus pro-vaccine and anti-vaccine with coherence score 0.999. Anti-vaccine topics include criticism of the halal MUI fatwa for MR vaccine, pork gelatine in the Hajj Meningitis Vaccine, vaccines for business fields, fake vaccines, KIPI and vaccine hazards, vaccines as part of conspiracy and Jewish agenda, demands for halal vaccines, etc. Whereas pro-vaccine topics are more homogeneous, namely the benefits and importance of immunization, vaccine administration rules, and campaigns with publicity of immunization activities and vaccine recommendations.