

**Abstrak**

Pandemi virus corona yang merebak menyebabkan berbagai dampak pada kehidupan setiap orang. Keberlangsungan pandemi yang berkepanjangan dan kebijakan sosial distancing yang diterapkan oleh pemerintah membuat masyarakat berkomunikasi serta meluapkan opini melalui sosial media. Twitter adalah salah satu sosial media dengan penambahan pengguna yang cukup masif pada masa pandemi. Beredarnya opini dan informasi di Twitter dapat membuat polemik di masyarakat sehingga menimbulkan kesalahpahaman di masyarakat. Pada akhirnya, hal tersebut akan menumbuhkan berbagai macam emosi public yang dapat mempengaruhi kebijakan yang diambil oleh pemerintah selanjutnya. Untuk dapat mengetahui emosi publik, dilakukan klasifikasi emosi terhadap teks Twitter berbahasa Indonesia yang berkaitan dengan covid menggunakan random forest dengan ekstraksi fitur TF-IDF dan Word2vec. Random forest adalah algoritma *ensemble* dengan *supervised learning* yang secara teori unggul dalam menangani imbalance pada data dan cepat dalam melakukan klasifikasi [1]. Data yang digunakan melalui tahap preprocessing yaitu *case folding*, *punctuation removal*, *stopword removal*, *stemming* dan *tokenization*. Pada penelitian ini, klasifikasi emosi terhadap teks Twitter mengenai Covid-19 yang berbahasa Indonesia berhasil dilakukan dengan nilai performasi sebanyak 0,43 menggunakan TF-IDF dan 0.38 menggunakan Word2vec. Sebagai pembandingan, digunakan juga data dari penelitian sebelumnya [2] yang mendapatkan nilai performasi 0.57 dengan TF-IDF dan 0.60 dengan Word2vec.

**Kata kunci :** random forest, tf-idf, word2vec, klasifikasi emosi, covid-19

---