

5. Kesimpulan

Pada Tugas Akhir ini, algoritma klasifikasi dengan performa yang terbaik adalah SVM dengan nilai akurasi dan F1-measure sebesar 0.846, disusul dengan algoritma NB dengan nilai akurasi dan F1-measure sebesar 0.829. Kedua algoritma dapat dikategorikan memiliki performa yang baik karena nilai akurasi dan F1-measure sama-sama lebih mendekati 1 daripada mendekati 0 [12]. Walaupun karakteristik soal pada dataset beragam, algoritma SVM dan NB masih dapat melakukan klasifikasi dengan baik. Ekstraksi fitur dengan TFPOS-IDF dapat memberikan performansi yang lebih baik dibandingkan TF-IDF pada algoritma SVM. Sementara itu, TF-IDF memiliki performansi yang lebih baik dibandingkan dengan TFPOS-IDF pada algoritma NB. Kemudian, memodifikasi *stopwords* dapat membantu memberikan performansi yang baik untuk kedua algoritma. Selain itu, melakukan *random oversampling* pada data dapat meningkatkan performa yang dihasilkan untuk algoritma SVM dan NB. Penelitian ini masih dapat dikembangkan dari sisi ketersediaan jumlah data yang digunakan dan menggunakan kombinasi Word2Vec dan TFPOS-IDF sebagai metode ekstraksi fitur.