ABSTRAK

Klasifikasi teks adalah proses penempatan suatu teks ke suatu kategori atau kelas sesuai dengan karakteristik dari teks tersebut. Dalam tahapannya tiap dokumen menunjuk pada satu kelas tertentu maka dibutuhkan proses untuk menggali informasi dari dokumen tersebut. Klasifikasi digunakan dalam proses mencari set model atau fungsi yang membedakan kelas dari data. Dalam hal ini data yang akan diklasifikasikan adalah publikasi ilmiah berupa paper, jurnal dan pro-ceeding dalam bidang Computer Science (Ilmu Komputer). Dataset yang ada akan dibagi menjadi data training dan data testing. Dalam klasifikasi, dilakukan pembelajaran terhadap data training yang dimana terdapat atribut yang ditetapkan sebagai kelas. Algoritma Support Vector Machine (SVM) menggunakan multiclass One-Againts-One (OVO) memiliki performansi lebih baik dibandingkan dengan algoritma lain seperti One-Againts-All (OVA) SVM dan Multinomial naïve bayes. Untuk meningkatkan performansi klasifikasi menggunakan feature selection Chi-square (x^2) sehingga akurasi meningkat menjadi 72.14%. Perbandingan komposisi menghasilkan akurasi sebesar 73.9 % pada perbandingan komposisi data training sebesar 90% dan data testing sebesar 10%.

Kata Kunci: klasifikasi teks, support vector machine, multiclass, One-Againts-One