

INTISARI

KOMPARASI PERFORMA MODEL ALGORITMA *SUPPORT VECTOR MACHINE* DAN *SUPPORT VECTOR MACHINE PARTICLE SWARM OPTIMIZATION* UNTUK KLASIFIKASI *HOAX* BAHASA INDONESIA

Oleh

Sheraton Pawestri
15/378075/PA/16550

Di era teknologi ini informasi mudah didapatkan baik dari internet, surat kabar, atau dari mulut ke mulut. Namun informasi juga bisa disalahgunakan seperti membuat berita atau informasi palsu atau *hoax*. Penelitian tentang klasifikasi *hoax* telah banyak dilakukan. Data yang berbentuk teks menyebabkan data memiliki dimensi yang tinggi dan fitur yang sangat banyak. Salah satu metode klasifikasi data berdimensi tinggi adalah Support Vector Machine (SVM). SVM saja tanpa seleksi fitur akan menghasilkan fitur yang banyak. Fitur yang banyak akan memakan waktu yang lama serta tidak efisien sehingga perlu adanya seleksi fitur. Salah satu metode seleksi fitur adalah *Particle Swarm Optimization* (PSO).

Pada penelitian ini dilakukan komparasi performa klasifikasi *hoax* dengan metode klasifikasi SVM dan SVM+PSO. Data yang diambil bertema pilpres 2019 dan politik. 200 data pilpres 2019 dengan rincian 100 *hoax* dan 100 *non-hoax* dan 20 tema politik dengan rincian 10 *hoax* dan 10 *non-hoax*. Evaluasi klasifikasi berupa akurasi, *precision*, *recall* dan lama waktu *train test* SVM dengan dan tanpa PSO.

Evaluasi penelitian ini, rata-rata akurasi sebesar 0.89545 atau 89.545%, rata-rata *precision* sebesar 92.905% dan rata-rata *recall* sebesar 83.901% untuk SVM. Rata-rata akurasi tertinggi menjadi 90.454%, rata-rata *precision* yaitu 92.238% dan *recall* yaitu 86.238% untuk SVM+PSO. Pada penelitian ini, PSO mampu meningkatkan performa namun kenaikan tidak pesat yaitu sekitar 1%.

Kata kunci: *hoax*, particle swarm optimization, support vector machine, seleksi fitur

ABSTRACT

COMPARISON OF SUPPORT VECTOR MACHINE AND SUPPORT VECTOR MACHINE PARTICLE SWARM OPTIMIZATION FOR HOAX CLASSIFICATION IN INDONESIAN NEWS

by

Sheraton Pawestri
15/378075/PA/16550

In this technological era, information is easily obtained either from the internet, newspapers, or by word of mouth. But information can also be misused such as making hoaxes. Many studies about hoax classification. Text data can have many dimensions and features. One of text classification methods that have been used is Support Vector Machine (SVM). SVM without feature selection will produce many features that take a long time an inefficient for processing classification. So, feature selection is needed. One method of feature selection is Particle Swarm Optimization (PSO).

In this study a comparison of the performance of the hoax classification with the SVM and SVM + PSO classification methods was performed. There are 200 data (100 hoaxes, 100 non-hoaxes) with presidential election 2019 data and 20 data with political theme (10 hoaxes, 10 non-hoaxes). Classification is done twice, SVM and SVM+PSO. Evaluations contain accuracy, precision, recall, and the length of time the SVM train test with and without PSO.

Evaluations contain an average accuracy of 0.89545 or 89.545%, an average precision of 92.905%, and an average recall of 83.901% for SVM. Then the highest average accuracy is 90.454%, the highest average precision is 92.238% and the highest recall is 86.238% for SVM+PSO. In this study PSO was able to improve performance but the increase was not rapid, it was around 1%.

Keywords: hoax, particle swarm optimization, support vector machine, feature selection