

ABSTRAK

Data mining merupakan suatu eksplorasi dan analisis dari suatu kuantitas data yang besar yang bertujuan untuk menemukan *pattern* yang bermakna. Banyaknya data berdimensi tinggi yang tidak berlabel *class* dalam dunia ini mendorong untuk dilakukannya klasterisasi. Klasterisasi adalah mengelompokkan data ke dalam klaster-klaster tertentu berdasarkan pendekatan kemiripan antar objek dalam suatu klaster dan perbedaannya antar klaster.

Principal Component Analysis Self Organizing Map (PCA-SOM) merupakan salah satu algoritma klasterisasi yang berusaha menerapkan prinsip kerja PCA pada algoritma klasterisasi SOM. PCA-SOM melakukan *mapping* ke neuron-neuron untuk mencari neuron pemenang dengan cara memproyeksikan data ke vektor basis dari neuron itu sendiri.

Tugas akhir ini diimplementasikan menggunakan Matlab 7.6 dan Microsoft Excel 2007. Parameter performansi hasil diukur dari pencapaian nilai akurasi, nilai overall purity dan evaluasi nilai variansi untuk tiap klaster terbentuk.

Kata kunci: data mining, *pattern*, *class*, klasterisasi, PCA-SOM, *mapping*