

Abstract

Finding information from so many documents that will be more and more every day is a difficult task and waste time. To find a needed document from other unnecessary document is done by reading those documents one by one. This can be overcome by doing document mapping in different cluster based on the information inside.

Clustering is done for grouping the data into clusters which the object in the same cluster have high similarity. But there is a problem in clustering process when the data have a high dimension. Therefore, it is needed an algorithm that can overcome this problem. Self Organizing Maps (SOM) is a right algorithm to make a data group with high dimension. SOM is neural network algorithm that work by update the value, so it suits by vector input value. By this way, network that fill with neurons will arrange itself based on some input value in a cluster.

Focus on this final task is implementation of Self Organizing Maps (SOM) algorithm on clustering of news with Indonesian Language. The testing is done by analyzing the effect of dimension and input parameters like epoch and learning rate toward the performances and accuration system. At the and of testing, can be proved that SOM can overcome high dimension data well. The counting of clustering result performance is done based on precision and entropy.

Key word: Clustering, Unsupervised learning, Self Organizing Maps.