

ABSTRACT

Twitter makes the development of communications and information technology to be fast and easy to access, allowing users to hold and share any information freely. With Twitter every user can convey and share whatever happens either in the form of tweet information or news that is retweet through certain news media without knowing whether the news tweet is true, relevant or irrelevant.

Based on the problem, we need a system of determining the relevance of news using Density Based Spatial Clustering Application with Noise (DBSCAN) and K-Nearest Neighbor (KNN) algorithms to find out relevant news and irrelevant news. Implementation using PHP programming language.

Determination of News Relevance resulted 2 clusters with data testing using K-Nearest Neighbor (KNN) and test data using Clustering Applications with Noise (DBSCAN) to 3 data collection showed that the bigger the relevant data rate of a bigger news.

Keywords: Twitter, News Relevance, Clustering, DBSCAN, Classification, KNN

ABSTRAK

Twitter membuat perkembangan teknologi komunikasi dan informasi menjadi cepat dan mudah untuk diakses sehingga memungkinkan pengguna untuk berpendapat dan berbagi informasi apapun secara bebas. Dengan Twitter setiap pengguna dapat menyampaikan dan berbagi apapun yang terjadi baik berupa tweet informasi maupun berita yang di retweet melalui media berita tertentu tanpa mengetahui apakah tweet berita tersebut benar adanya, relevan atau tidak relevan.

Berdasarkan permasalahan tersebut maka diperlukan sebuah sistem penentuan relevansi berita menggunakan algoritma *Density Based Spatial Clustering Applications with Noise* (DBSCAN) dan *K-Nearest Neighbor* (KNN) untuk mengetahui berita yang relevan dan berita tidak relevan. Implementasi menggunakan bahasa pemrograman PHP.

Penentuan Relevansi berita menghasilkan 2 cluster dengan data testing menggunakan *K-Nearest Neighbor* (KNN) dan data uji menggunakan *Clustering Applications with Noise* (DBSCAN) terhadap 3 data collection menunjukkan bahwa semakin besar data tingkat relevan suatu berita semakin besar.

Kata kunci : Twitter, Relevansi Berita, *Clustering*, DBSCAN, Klasifikasi, KNN