

Affandy Fahrizain. 2020. **Sistem Deteksi Persebaran Sentimen Masyarakat Terhadap Zonasi Sekolah Menggunakan K-Nearest Neighbor (KNN)**. Skripsi ini di bawah bimbingan Drs. Kartono, M.Kom dan Badrus Zaman, S.Kom., M.Cs. Program Studi S1 Sistem Informasi. Fakultas Sains dan Teknologi, Universitas Airlangga.

ABSTRAK

Kebijakan zonasi sekolah menuai banyak pro dan kontra di kalangan masyarakat. Banyaknya data *tweet* terkait zonasi sekolah mampu membentuk sebuah *insight* yang dapat dimanfaatkan untuk mengetahui persebaran sentimen masyarakat serta mengetahui daerah mana saja yang paling terpengaruh oleh zonasi sekolah dan sebaliknya. Pada penelitian ini dilakukan pembangunan sistem persebaran klasifikasi *tweet* dengan algoritma K-Nearest Neighbor (KNN). Sistem yang dibangun telah mampu melakukan *scraping* dan klasifikasi data secara otomatis. Tahapan penelitian yang dilakukan ada 7, yaitu pengumpulan data, *labelling* data, *preprocessing* data, klasifikasi model KNN, evaluasi model KNN, perancangan sistem, dan implementasi sistem. Pengumpulan data dilakukan dengan melakukan pencarian menggunakan kata kunci Zonasi Sekolah. Sebelum proses klasifikasi dilakukan tahap *preprocessing* yang meliputi *case folding*, *tokenizing*, penghapusan *stopwords*, *stemming*, dan pembobotan *term* sehingga mengubah data *tweet* menjadi matriks TF-IDF. Data diklasifikasikan kedalam 3 kelas, yaitu positif, negatif, dan netral dengan menghasilkan akurasi terbaik sebesar 0,79 dengan nilai K sebesar 28.

Kata kunci: klasifikasi, *tweet*, KNN, zonasi sekolah, persebaran sentimen masyarakat

Affandy Fahrizain. 2020. **Detection System for Public Sentiment Distribution of Educational Zoning Using K-Nearest Neighbor (KNN)**. This thesis supervised by Drs. Kartono, M.Kom and Badrus Zaman, S.Kom., M.Cs. Bachelor of Computer Science (Information System). Faculty of Science and Technology, Universitas Airlangga.

ABSTRACT

The school zoning policy has caused many pros and cons to society. The number of tweets related to school zoning can be used to create an insight that can be utilized to find out the distribution of people's sentiments and to identify which areas are most affected by school zoning and vice versa. In this study, the K-Nearest Neighbor (KNN) method is used to build a tweet classification distribution system. The system that built has been successfully done data scraping and classification automatically. There are 7 stages of research, namely data collection, data labeling, data preprocessing, KNN model classification, KNN model evaluation, system design, and system implementation. School zoning is used as a keyword for scraping in data collection. After data gathered, then data is labeled by the crowdsource labeling method. Before data used to build a machine learning model, they need to be preprocessed first. The data are classified into 3 classes namely positive, negative, and neutral. Then evaluation for built of machine learning model shown that model best accuracy is 0.79 with a K value of 28.

Keywords: classification, tweet, KNN, school zoning, distribution of people sentiment