

## ABSTRACT

The POI recommendation system involves large volumes of data, resulting in scalability issues that not only result in increased processing time but also result in reduced accuracy. In addition to accuracy, the type of data used also greatly affects the results of the recommendations. In the recommendation system, there are several types of *rating* namely implicit (binary) and explicit (scalar). Binary *ratings* will result in lower accuracy when it is not handled with the right approach. Some studies related to binary *ratings* still obtain low accuracy, which affects the results of recommendations.

Various techniques have been developed in the site recommendation system to address the problem. Collaborative filtering is a technique that has high accuracy in location recommendation systems, especially user-based with similarity cosine capable of handling datasets with binary *ratings*. In this research, we developed a user-based collaborative filtering approach combined with optimized K-means clustering to improve the accuracy of the POI recommendation system. The dataset used is the dataset with implicit feedback, i.e. Gowalla and Foursquare. Determination of K value in K-Means were optimized using Davies-Bouldin Index (DBI) method. DBI measurement approach is to maximize the distance of inter cluster and minimize the distance of intra cluster. Then the dataset in the cluster was based on latitude and longitude. With the clustering results obtained, a recommendation system was built using user-based collaborative filtering. Similarity of users was calculated using cosine similarity.

This research was developed using Python programming language and some other supporting software. From the experimental results showed that the optimization of the K value using the Davies-Bouldin Index (DBI) method, resulting in better clustering with many variation value of K. The K-means + UBCF method produces precision 0.086 for Gowalla 1 and 0.041 datasets for Foursquare datasets with N=20. Recommendation quality was represented with f-measure value 0.072 for Gowalla 1 dataset and 0.036 for Foursquare with N=20. Then the

results were compared to DBSCAN + UBCF, and the proposed method had better accuracy of 1% increase in precision value. This proves that K-Means + UBCF can handle datasets with implicit feedback and improve system precision.

**Keywords:** POI recommendation system, K-Means, Davies-Bouldin Index method, collaborative filtering

## INTISARI

Sistem rekomendasi POI melibatkan volume data yang besar, sehingga muncul masalah skalabilitas yang tidak hanya menghasilkan peningkatan waktu pemrosesan namun juga menyebabkan berkurangnya akurasi. Selain akurasi, tipe data yang digunakan juga mempengaruhi hasil rekomendasi. Dalam sistem rekomendasi, terdapat beberapa jenis *rating* yaitu implisit (biner) dan eksplisit (skalar). *Rating* biner akan menghasilkan akurasi yang rendah ketika tidak ditangani dengan yang tepat. Beberapa penelitian terkait *rating* biner masih memperoleh akurasi yang rendah.

Berbagai teknik telah dikembangkan dalam sistem rekomendasi lokasi untuk menangani permasalahan tersebut. *Collaborative filtering* adalah teknik yang memiliki akurasi yang tinggi dalam sistem rekomendasi lokasi, khususnya *user-based* dengan *cosine similarity* mampu menangani *dataset* dengan *rating* biner. Dalam penelitian ini, dikembangkan pendekatan *user-based collaborative filtering* yang dikombinasi dengan K-means *clustering* yang telah dioptimasi untuk meningkatkan akurasi sistem rekomendasi POI. Penentuan nilai K pada K-Means dioptimasi menggunakan metode *Davies-Bouldin Index* (DBI). Pendekatan pengukuran DBI yaitu memaksimalkan jarak *intercluster* serta meminimalkan jarak *intracluster*. Kemudian *dataset* di-*cluster* berdasarkan *latitude* dan *longitude*. Hasil *clustering* dari *dataset* Gowalla dan Foursquare yang di dapatkan, dibangun sistem rekomendasi menggunakan *user-based collaborative filtering*. Similaritas pengguna di hitung menggunakan *cosine similarity*.

Penelitian ini dikembangkan menggunakan bahasa pemrograman Python dan beberapa *software* pendukung lainnya. Dari hasil eksperimen menunjukkan bahwa optimasi nilai K menggunakan metode Davies-Bouldin Index (DBI), menghasilkan *clustering* yang lebih baik dengan nilai K yang bervariasi dan dapat memaksimalkan hasil *clustering*. Metode K-means+UBCF menghasilkan presisi 0.086 untuk *dataset* Gowalla 1 dan 0.041 untuk *dataset* Foursquare untuk N=20. Kualitas rekomendasi dapat dilihat dari nilai *f-measure* sebesar 0.72 pada *dataset* Gowalla 1 dan 0.036 pada *dataset* Foursquare untuk N=20. Kemudian hasil penelitian dibandingkan dengan DBSCAN+UBCF, dan metode yang diusulkan mempunyai akurasi lebih baik yaitu

meningkat 1% untuk nilai presisi. Hal ini membuktikan bahwa K-Means +UBCF dapat menangani dataset dengan *implicit feedback* dan meningkatkan presisi sistem.

**Kata kunci:** Sistem Rekomendasi POI, K-means, Metode *Davies-Bouldin Index*, *Collaborative Filtering*