

INTISARI

Perbandingan Relevansi Hasil : Sistem Rekomendasi K-Means Clustering dan Sistem Rekomendasi K-Nearest Neighbor Classifier pada Kasus Distribusi Obat

Oleh

Mohammad Daffa Gashandy
20/455449/PA/19664

Penelitian ini mencoba meningkatkan efektivitas penjualan obat dengan mengusulkan sistem rekomendasi berbasis pembelajaran mesin. Dua algoritma yang diusulkan, *k-means clustering* dan *k-nearest neighbor* (*k*-NN) *classifier*, digunakan untuk mengatasi permasalahan pemilihan algoritma dalam pengembangan sistem rekomendasi. Meskipun keduanya dianggap baik, penelitian ini mengeksplorasi perbedaan karakteristik dan metode evaluasi mereka, dengan tujuan mengembangkan sistem rekomendasi yang optimal.

Tahap implementasi sistem rekomendasi *hybrid filtering* melibatkan langkah-langkah seperti *preprocessing* dan pelabelan data dari Food and Drug Administration (FDA), serta data cuaca dan usia provinsi. Dua metode implementasi yang berbeda digunakan: *k-means clustering* dengan modifikasi *principal component analysis* (PCA) dan *k*-NN dengan modifikasi *gradient boosting*. Setelah model memberikan hasil rekomendasi, dilakukan pembobotan berdasarkan kesamaan label cuaca, interseksi interval usia, dan nomor riwayat obat dengan provinsi. Hasilnya, produk dengan bobot tertinggi direkomendasikan.

Pengujian relevansi hasil rekomendasi dilakukan dengan melibatkan perusahaan obat-obatan, DEXA Group. Evaluasi dilakukan oleh *salesman* dan *data scientist*. Sistem rekomendasi *k-means clustering* cenderung memberikan hasil terbatas karena faktor pemilihan klaster, sementara *k*-NN memberikan hasil yang lebih beragam. Namun, survei penilaian menunjukkan bahwa sistem *k-means clustering* dinilai lebih relevan oleh *salesman* (80%) dan *data scientist* (70%) dibandingkan dengan *k*-NN (70% dan 50%). Penelitian ini diharapkan memberikan panduan untuk pengembangan strategi distribusi obat melalui penerapan sistem rekomendasi yang lebih optimal

Kata-kata kunci : Sistem Rekomendasi, Hybrid Filtering, K-Means Clustering, K-Nearest Neighbor, Gradient Boosting, Principal Component Analysis, Preprocessing, Obat, Salesman

ABSTRACT

Comparison of the Relevance of Result: K-Means Clustering Recommendation System and K-Nearest Neighbor Classifier Recommendation System in the Case of Drug Distribution.

By

Mohammad Daffa Gashandy
20/455449/PA/19664

This research aims to enhance the effectiveness of drug sales by proposing a machine learning-based recommendation system. Two proposed algorithms, k-means clustering and k-nearest neighbor (k-NN) classifier, are employed to address the algorithm selection challenge in recommendation system development. Although both are considered effective, the research explores their characteristic differences and evaluation methods with the goal of developing an optimal recommendation system.

The implementation phase of the hybrid filtering recommendation system involves steps such as preprocessing and labeling data from the Food and Drug Administration (FDA), as well as weather and age data for provinces. Two different implementation methods are used: k-means clustering with modified principal component analysis (PCA) and k-NN with modified gradient boosting. After the model provides recommendation results, weighting is performed based on the similarity of weather labels, age interval intersections, and drug history numbers with provinces. As a result, products with the highest weights are recommended.

Relevance testing of the recommendation results involves the pharmaceutical company DEXA Group. Evaluation is conducted by salesmen and data scientists. The k-means clustering recommendation system tends to yield limited results due to cluster selection factors, while k-NN provides more diverse results. However, survey assessments indicate that the k-means clustering system is deemed more relevant by salesmen (80%) and data scientists (70%) compared to k-NN (70% and 50%). This research is expected to provide guidance for the development of drug distribution strategies through the implementation of a more optimal recommendation system.

Keywords: Recommendation System, Hybrid Filtering, K-Means Clustering, K-Nearest Neighbor, Gradient Boosting, Principal Component Analysis, Preprocessing, Drugs, Salesman.