

ABSTRAK

Bimasena, Matius Ivan (2025), Studi Perbandingan Algoritma Euclidean, Manhattan dan Chebysev Distance Untuk Optimasi Metode K-Means Clustering Pada Pengelompokan Data Penerimaan Mahasiswa Baru ITB STIKOM Bali. Tesis, Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha

Tesis ini sudah disetujui dan diperiksa oleh Pembimbing I : Dr. I Gede Aris Gunadi, S.Si. M.Kom. dan Pembimbing II : Dr. I Made Agus Wirawan, S.Kom., M.Cs.

Kata-kata kunci: Euclidean Distance, Manhattan Distance, Chebysev Distance, K-Means Clustering, ITB STIKOM Bali, Sekolah, Bali

Institut Teknologi dan Bisnis (ITB) STIKOM Bali adalah salah satu perguruan tinggi swasta di Bali yang telah berdiri sejak tahun 2002. Setiap tahun, ITB STIKOM Bali membuka penerimaan mahasiswa baru dari lulusan SMA/SMK dan sederajat. Tim promosi kampus aktif melakukan promosi di berbagai daerah yang berpotensi mendatangkan calon mahasiswa. Namun dalam proses promosi masih adanya kekurangan dalam menentukan tujuan promosi kampus ke daerah yang berpotensi. Peneliti melakukan pengelompokan data calon mahasiswa baru untuk tahun ajaran 2022 dan 2023 sebanyak 2.689 data, digunakan metode K-means Clustering dengan menentukan kluster terbaik menggunakan Metode Elbow. Dari hasil metode elbow terbentuk 3 kluster optimal. Setelah itu penulis membandingkan tiga metode perhitungan jarak, yaitu Euclidean, Manhattan, dan Chebyshev Distance dengan 3 kluster. Dari hasil evaluasi metode perhitungan jarak, terbukti bahwa Euclidean Distance mempunyai hasil evaluasi yang paling baik yaitu menunjukkan nilai DBI dan *Average Within Centroid Distance* dari perhitungan jarak terkecil dibandingkan Manhattan Distance dan Chebyshev Distance dengan nilai evaluasi *Davies Bouldin Index* 1.202 dan *Average Within Centroid Distance* 16.082. Manhattan Distance cocok digunakan ketika data memiliki fitur yang dipengaruhi oleh perubahan dalam satu dimensi (misalnya, hanya bergerak horizontal atau vertikal dalam satu ruang dimensi). Ini cocok untuk data yang memiliki struktur kotak-kotak atau grid sedangkan Chebyshev Distance baik digunakan ketika pengelompokan (clustering) atau pembelajaran mesin, Chebyshev Distance dapat digunakan jika variabel yang paling dominan atau ekstrem di suatu dimensi lebih kritis untuk diperhitungkan dalam analisis. Setelah ditentukan metode jarak terbaik yaitu Euclidean Distance, penulis melakukan klastering dengan pengelompokan data penerimaan mahasiswa baru berdasarkan program studi tinggi peminat, sedang peminat dan rendah peminat.

ABSTRACT

Bimasena, Matius Ivan (2025), *A Comparative Study of Euclidean, Manhattan, and Chebyshev Distance Algorithms for Optimizing the K-Means Clustering Method in Grouping New Student Admission Data at ITB STIKOM Bali*. Thesis, Computer Science, Graduate Program, Ganesha University of Education.

This thesis has been approved and examined by Supervisor I: Dr. I Gede Aris Gunadi, S.Si., M.Kom., and Supervisor II: Dr. I Made Agus Wirawan, S.Kom., M.Cs.

Keywords : Euclidean Distance, Manhattan Distance, Chebysev Distance, K-Means Clustering, ITB STIKOM Bali, School, Bali

Institut Teknologi dan Bisnis (ITB) STIKOM Bali is a private university in Bali established in 2002. Every year, ITB STIKOM Bali opens admissions for new students graduating from high schools or equivalent institutions. The university's promotion team actively conducts campaigns in various regions with potential to attract prospective students. However, challenges remain in effectively identifying target areas for promotional efforts. This research focuses on clustering data of prospective students for the 2022 and 2023 academic years, totaling 2.689 entries. The K-Means Clustering method was employed to group the data, with the optimal number of clusters determined using the Elbow Method. The analysis identified three optimal clusters. Subsequently, the study compared three distance calculation methods Euclidean, Manhattan, and Chebyshev distances using these three clusters. The evaluation revealed that the Euclidean Distance method produced the best results, showing the lowest Davies-Bouldin Index (DBI) score and Average Within Centroid Distance compared to Manhattan and Chebyshev distances. Specifically, the Euclidean Distance yielded a DBI score of 1.202 and an Average Within Centroid Distance of 16.082. Manhattan Distance is appropriate for data in which features are influenced by variations along a single dimension—such as movement restricted to horizontal or vertical directions within a dimensional space. This distance metric is particularly suitable for data exhibiting a grid or lattice structure. On the other hand, Chebyshev Distance is well-suited for clustering or machine learning tasks where the most dominant or extreme variable in a particular dimension plays a critical role in the analysis. After establishing Euclidean Distance as the most effective method, clustering was performed based on prospective students' interest levels in various study programs, categorized as high, medium, or low interest.