
Analisis Pola Pembayaran UKT Mahasiswa dengan Clustering Menggunakan K-Means dan DBSCAN

Yori Adi Atma¹⁾*, Harfebi Fryonanda²⁾, Taruma Leo Wijaya³⁾

¹⁾²⁾³⁾Jurusan Teknologi Informasi, Politeknik Negeri Padang, Indonesia

*Corresponding Email: yori@pnp.ac.id

Abstrak

Penelitian ini bertujuan untuk menganalisis data sosial ekonomi mahasiswa baru sebagai dasar dalam penetapan level Uang Kuliah Tunggal (UKT) secara objektif dan berbasis data. Metode yang digunakan adalah teknik unsupervised learning berupa clustering dengan algoritma K-Means dan DBSCAN. Data yang dianalisis berasal dari isian formulir UKT mahasiswa baru Politeknik Negeri Padang tahun ajaran 2024/2025, yang mencakup informasi penghasilan orang tua, jumlah tanggungan, kepemilikan rumah, dan status bantuan sosial. Hasil analisis menunjukkan bahwa K-Means berhasil mengelompokkan mahasiswa ke dalam tiga klaster dengan rata-rata penghasilan dan UKT yang selaras. Klaster dengan penghasilan tertinggi cenderung berada pada level UKT 7 dan 8, sedangkan klaster dengan penghasilan lebih rendah terkonsentrasi pada level UKT menengah. Sementara itu, metode DBSCAN hanya menghasilkan satu klaster valid dengan banyak data teridentifikasi sebagai noise, menunjukkan keterbatasannya dalam menangani dataset kecil dan homogen. Kesimpulan dari penelitian ini adalah bahwa metode K-Means lebih sesuai digunakan sebagai alat bantu dalam mengevaluasi dan meninjau ulang kebijakan UKT secara objektif berdasarkan data sosial ekonomi mahasiswa.

Kata Kunci: UKT, clustering, K-Means, DBSCAN

Abstract

Abstract This study aims to analyze the socioeconomic data of new students as a basis for objectively and data-driven determination of tuition fee levels (UKT). The method used is an unsupervised learning technique in the form of clustering using the K-Means and DBSCAN algorithms. The data analyzed was obtained from the UKT application forms of new students at Politeknik Negeri Padang for the 2024/2025 academic year, which includes information such as parental income, number of dependents, home ownership, and social assistance status. The analysis results show that K-Means successfully grouped students into three clusters with average income and tuition fee levels that were aligned. The cluster with the highest income tended to fall into UKT levels 7 and 8, while clusters with lower income were concentrated in the mid-level UKT categories. Meanwhile, the DBSCAN method produced only one valid cluster, with many data points identified as noise, indicating its limitation in handling small and homogeneous datasets. The conclusion of this study is that the K-Means method is more suitable to be used as a decision-support tool in evaluating and reviewing UKT policies objectively based on students' socioeconomic data.

Keywords: UKT, clustering, K-Means, DBSCAN

PENDAHULUAN

Penentuan besaran Uang Kuliah Tunggal (UKT) pada Perguruan Tinggi Negeri merupakan bagian penting dalam mewujudkan sistem pendidikan tinggi yang adil dan inklusif. Kebijakan UKT yang diatur melalui Permendikbud Nomor 55 Tahun 2013 menetapkan bahwa besaran biaya pendidikan disesuaikan dengan kemampuan ekonomi mahasiswa, bukan dibebankan secara merata (Renaissance, 2024). Dalam implementasinya, proses penentuan level UKT dilakukan berdasarkan data sosial ekonomi yang diisi oleh mahasiswa baru, seperti pekerjaan dan penghasilan orang tua, jumlah tanggungan keluarga, kondisi tempat tinggal, serta bukti dukungan seperti Kartu Indonesia Pintar (KIP) atau SKTM (Ramadha & Nursafitri, 2024). Di Politeknik Negeri Padang, reviewer UKT memiliki peran penting dalam memverifikasi dan menganalisis data tersebut sebelum menetapkan level UKT dari 1 hingga 8.

Namun, dalam praktiknya, proses evaluasi ini masih menghadapi beberapa tantangan. Salah satunya adalah potensi subjektivitas dalam menilai apakah data yang diajukan benar-benar merepresentasikan kondisi ekonomi mahasiswa. Selain itu, belum adanya pendekatan berbasis analisis data untuk menemukan pola sosial ekonomi mahasiswa membuat penetapan level UKT rentan terhadap ketidaktepatan sasaran (Retnoningsih & Marom, 2017). Misalnya, mahasiswa yang seharusnya mendapatkan UKT rendah justru ditempatkan pada level yang lebih tinggi karena keterbatasan evaluasi kuantitatif. Di sisi lain, mahasiswa dengan kondisi ekonomi mapan mungkin saja mendapatkan level UKT yang lebih rendah dari seharusnya.

Dalam menghadapi permasalahan tersebut, diperlukan pendekatan berbasis data untuk membantu reviewer dalam mengelompokkan mahasiswa berdasarkan kondisi sosial ekonominya secara objektif. Salah satu pendekatan yang potensial untuk digunakan adalah metode unsupervised learning dalam data mining, khususnya teknik clustering (Atma & Montesori, 2022). Clustering memungkinkan data sosial ekonomi mahasiswa dikelompokkan ke dalam beberapa segmen atau klaster berdasarkan kemiripan nilai antarindividu. Metode yang umum digunakan

antara lain K-Means dan DBSCAN. K-Means bekerja dengan membagi data ke dalam sejumlah kluster berdasarkan jarak centroid (Darmi & Setiawan, 2017), sementara DBSCAN mampu membentuk kluster dengan fleksibel, bahkan untuk data yang tidak terstruktur atau memiliki outlier (Mutiah et al., 2024).

Penelitian ini bertujuan untuk menganalisis data sosial ekonomi mahasiswa baru di Politeknik Negeri Padang menggunakan metode K-Means dan DBSCAN, untuk menemukan pola pengelompokan yang dapat merepresentasikan tingkat kemampuan ekonomi mereka (Adha et al., 2021). Selain itu, hasil clustering akan dibandingkan dengan level UKT yang telah ditetapkan untuk menilai kesesuaiannya. Dengan demikian, penelitian ini diharapkan dapat menjadi kontribusi awal dalam pengembangan sistem pendukung keputusan penentuan level UKT yang lebih transparan dan berbasis data.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksploratif dengan metode *unsupervised learning* berupa *clustering* untuk menganalisis data sosial ekonomi mahasiswa baru. Data yang digunakan bersumber dari formulir isian UKT mahasiswa Politeknik Negeri Padang tahun ajaran 2024/2025, yang telah diverifikasi oleh tim reviewer. Data tersebut mencakup variabel-variabel seperti pekerjaan dan penghasilan orang tua dan wali, jumlah tanggungan, kepemilikan rumah, kondisi rumah, keikutsertaan dalam program bantuan sosial (seperti KIP, PKH, KKS), serta skor dan level UKT yang telah ditetapkan.

Pemilihan sampel dilakukan secara purposif dengan mengambil seluruh data mahasiswa baru yang telah mengisi lengkap formulir UKT dan diverifikasi hingga akhir Juni 2024. Setelah data terkumpul, dilakukan tahap praproses yang meliputi pembersihan data, konversi data kategorik menjadi numerik menggunakan teknik label encoding, dan normalisasi data numerik menggunakan min-max scaling agar setiap atribut memiliki rentang nilai yang setara.

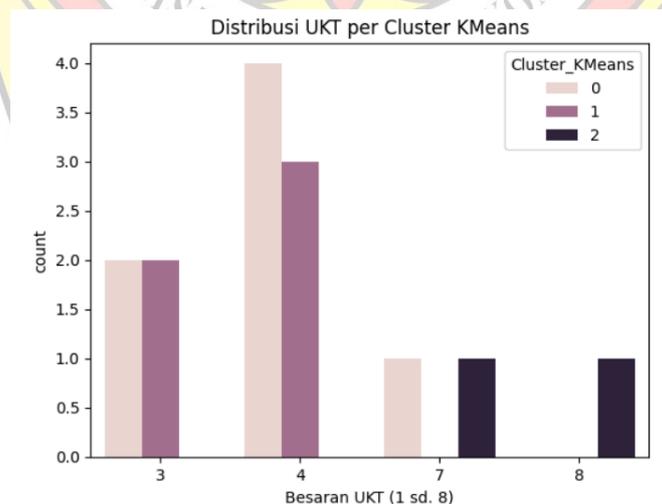
Untuk proses clustering, dua metode digunakan sebagai pembanding, yaitu K-Means dan DBSCAN. K-Means digunakan karena kemampuannya dalam membentuk kluster dengan jumlah yang ditentukan sebelumnya secara cepat dan efisien

(Saputra & Yusuf, 2024). Sementara DBSCAN digunakan karena dapat mendeteksi kluster dengan bentuk tidak beraturan serta mampu mengidentifikasi data yang tidak termasuk ke dalam kluster (Nur et al., 2025). Pemilihan jumlah kluster untuk K-Means dilakukan dengan pendekatan Elbow Method (Lashiyanti et al., 2023), sedangkan DBSCAN ditentukan berdasarkan nilai parameter epsilon dan minimum samples yang divalidasi menggunakan visualisasi k-distance graph (Monko & Kimura, 2023).

Seluruh proses analisis dilakukan menggunakan bahasa pemrograman Python di lingkungan Jupyter Notebook dengan bantuan pustaka pandas, scikit-learn, dan matplotlib untuk visualisasi hasil.

HASIL DAN PEMBAHASAN

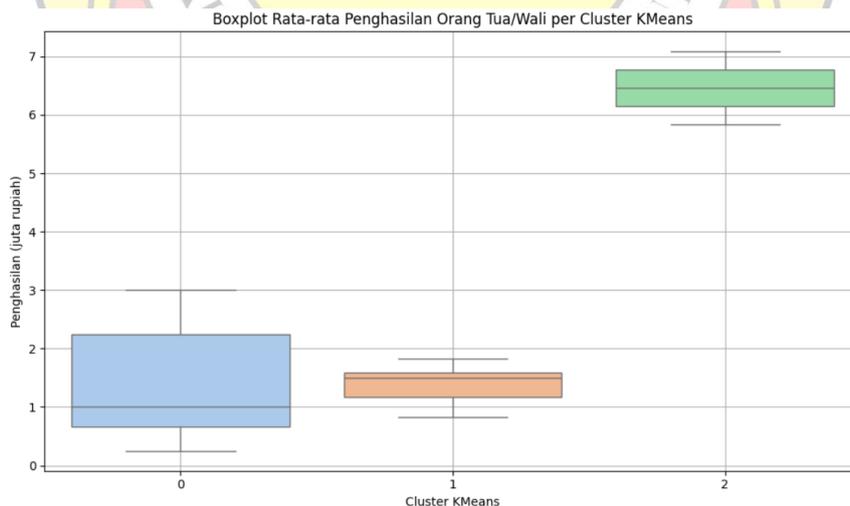
Hasil proses clustering menggunakan algoritma K-Means menghasilkan tiga kelompok mahasiswa berdasarkan kondisi sosial ekonomi mereka. Distribusi UKT pada masing-masing kluster yang ditampilkan pada Gambar 1 menunjukkan adanya keterkaitan yang cukup jelas antara kluster dan besaran UKT. Mahasiswa dalam Cluster 2 didominasi oleh UKT tinggi, yaitu level 7 dan 8, dengan rata-rata UKT mencapai 7.5. Hal ini menunjukkan bahwa kelompok ini kemungkinan terdiri dari mahasiswa dengan latar belakang ekonomi yang tergolong mampu.



Gambar 1. Distribusi UKT per Cluster

Sementara itu, Cluster 0 dan Cluster 1 menunjukkan kecenderungan distribusi UKT yang lebih menengah, dengan rata-rata masing-masing 4.14 dan 3.6. Cluster 0 memiliki penyebaran UKT yang relatif bervariasi, termasuk level 3, 4, dan 7. Sebaliknya, Cluster 1 lebih konsisten berada pada level 3 dan 4. Temuan ini mengindikasikan bahwa meskipun kedua klaster berada di kisaran menengah, terdapat perbedaan dalam karakteristik sosial ekonomi anggotanya. Potensi overlap pada klaster menengah ini juga membuka ruang evaluasi terhadap efektivitas proses verifikasi manual UKT.

Untuk memperkuat interpretasi terhadap hasil clustering, Gambar 2 menyajikan boxplot rata-rata penghasilan orang tua atau wali dari mahasiswa dalam masing-masing klaster. Hasil visualisasi ini menunjukkan bahwa Cluster 2 memiliki rentang penghasilan yang jauh lebih tinggi dibandingkan dua klaster lainnya, dengan median berada di kisaran 6–7 juta rupiah. Sementara itu, Cluster 0 dan Cluster 1 memiliki rentang penghasilan yang lebih rendah, dengan Cluster 1 memiliki median yang lebih kecil dari Cluster 0. Menariknya, sebaran data pada Cluster 0 tampak lebih lebar, yang bisa diartikan bahwa klaster ini mencakup mahasiswa dari latar belakang ekonomi yang lebih beragam, sedangkan Cluster 1 cenderung lebih homogen.

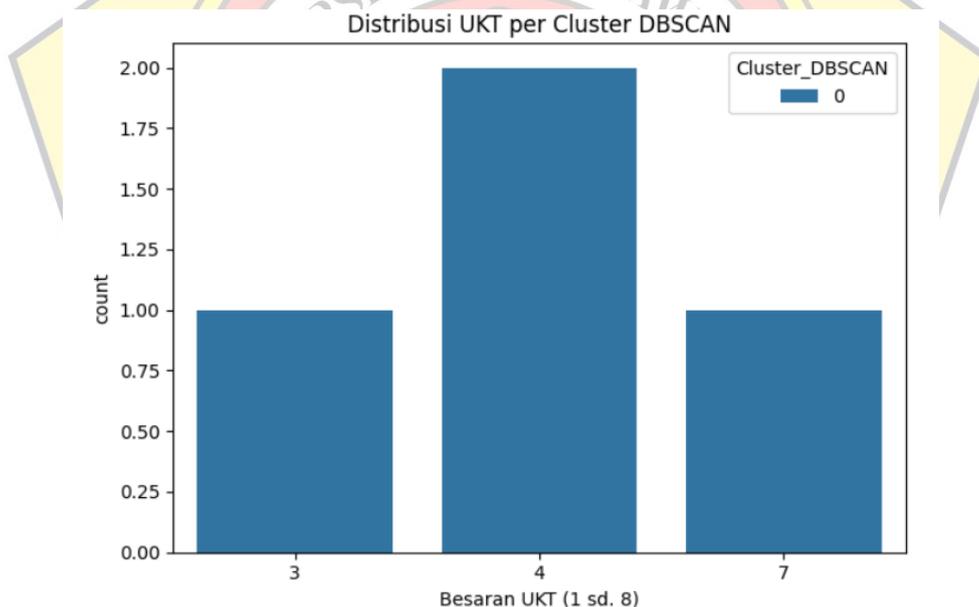


Gambar 2. Boxplot Rata-rata Penghasilan Orang Tua

Secara keseluruhan, hasil clustering K-Means memperlihatkan bahwa pendekatan ini cukup efektif dalam mengelompokkan mahasiswa berdasarkan

kemiripan kondisi sosial ekonomi, yang kemudian tercermin pula dalam level UKT yang diterapkan. Hal ini menunjukkan potensi besar metode ini untuk digunakan sebagai alat bantu dalam proses penentuan atau peninjauan ulang UKT secara objektif dan berbasis data.

Sebagai pembandingan, metode DBSCAN juga diterapkan pada dataset yang sama. Dengan parameter eps sebesar 0.8 dan min_samples 3, hasil clustering yang diperoleh hanya menghasilkan satu kluster valid (Cluster 0) dengan 4 anggota, sedangkan 10 data lainnya ditandai sebagai noise atau outlier (label -1). Distribusi UKT mahasiswa dalam kluster DBSCAN ini ditampilkan pada Gambar 3, yang menunjukkan bahwa mahasiswa tersebar pada UKT level 3, 4, dan 7. Karena hanya terdapat satu kluster valid, maka perhitungan silhouette score tidak dapat dilakukan, sehingga evaluasi kualitas hasil clustering menjadi terbatas.



Gambar 3. Distribusi UKT per Cluster

Hasil dari DBSCAN menunjukkan bahwa metode ini sensitif terhadap nilai parameter eps dan kurang optimal diterapkan pada dataset kecil dan homogen seperti dalam studi ini. Meski memiliki keunggulan dalam mendeteksi outlier, DBSCAN gagal membentuk kluster yang bermakna untuk tujuan pengelompokan UKT. Sebaliknya, K-Means menunjukkan performa yang lebih stabil dan menghasilkan pengelompokan yang lebih jelas serta interpretatif.

Dengan demikian, dalam konteks analisis data sosial ekonomi mahasiswa baru untuk mendukung penetapan level UKT, metode K-Means terbukti lebih sesuai digunakan dibandingkan DBSCAN, baik dari sisi jumlah klaster, sebaran data, maupun kesesuaian dengan kondisi lapangan.

SIMPULAN

Simpulan Penelitian ini menunjukkan bahwa metode clustering berbasis K-Means dapat digunakan sebagai pendekatan yang efektif untuk mengelompokkan mahasiswa baru berdasarkan data sosial ekonomi mereka, yang kemudian dapat dijadikan dasar pertimbangan dalam penetapan level Uang Kuliah Tunggal (UKT). Hasil pengelompokan memperlihatkan adanya keterkaitan yang konsisten antara klaster yang terbentuk dengan rata-rata penghasilan orang tua maupun besaran UKT yang dikenakan. Sementara itu, metode DBSCAN yang juga diuji sebagai pembanding menunjukkan keterbatasan dalam membentuk klaster yang valid pada data dengan karakteristik kecil dan homogen, serta menghasilkan proporsi noise yang tinggi. Oleh karena itu, K-Means dinilai lebih cocok digunakan sebagai alat bantu analisis dalam proses evaluasi dan peninjauan kebijakan UKT agar lebih objektif dan berbasis data.

DAFTAR PUSTAKA

- Adha, R., Nurhaliza, N., Sholeha, U., & Mustakim, M. (2021). Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia. *SITEKIN: Jurnal Sains, Teknologi Dan Industri*, 18(2), 206–211. <https://ejournal.uin-suska.ac.id/index.php/sitekin/article/view/12469>
- Atma, Y. A., & Montesori, S. (2022). Analisis Data Mining Untuk Menentukan Profit Perusahaan Menggunakan Metode K-Means. *Device : Journal of Information System, Computer Science and Information Technology*, 3(2), 29–36. <https://doi.org/10.46576/device.v3i2.2699>
- Darmi, Y. D., & Setiawan, A. (2017). Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk. *Jurnal Media Infotama*, 12(2), 148–157. <https://doi.org/10.37676/jmi.v12i2.418>
- Lashiyanti, A. R., Rasyid Munthe, I., Nasution, F. A., & Korespondensi, E. P. (2023). Optimisasi Klasterisasi Nilai Ujian Nasional dengan Pendekatan Algoritma K-Means, Elbow, dan Silhouette. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, 6(1), 14–20.

- Monko, G. J., & Kimura, M. (2023). *Optimized DBSCAN Parameter Selection: Stratified Sampling for Epsilon and Gridsearch for Minimum Samples*. 43-61. <https://doi.org/10.5121/csit.2023.132004>
- Mutiah, S., Hasnataeni, Y., Fitrianto, A., & Jumansyah, L. M. R. D. (2024). *Perbandingan Metode Klastering K-Means dan DBSCAN dalam Identifikasi Kelompok Rumah Tangga Berdasarkan Fasilitas Sosial Ekonomi di Jawa Barat*. 09(September), 247-260.
- Nur, A., Maulidhia, A., Ika, I., Friska, W., Sukarno, I., & Basya, R. (2025). Implementasi Perbandingan Algoritma k-Means dan DB-Scan Pada Beban Listrik Rumah Tangga. *INTEGER: Journal of Information Technology*, 10(1), 85-94.
- Ramadha, N. S., & Nursafitri, W. N. (2024). Analisis Klasifikasi UKT Mahasiswa Berdasarkan Tingkat Penghasilan Orang Tua Menggunakan Algoritma C4.5. *Inventor: Jurnal Inovasi Dan Tren Pendidikan Teknologi Informasi*, 2(1), 1-9. <https://www.ejournal.tsb.ac.id/index.php/inventor/article/view/1248>
- Renaissance, L. E. X. (2024). *Rekonstruksi Kebijakan Hukum Kenaikan Uang Kuliah Tunggal Berdasarkan Tanggung Jawab Konstitusional Negara*. 9(12), 394-422.
- Retnoningsih, Y. D., & Marom, A. (2017). Analisis Kebijakan Penyelenggaraan Pendidikan Berbasis Uang Kuliah Tunggal Bagi Perguruan Tinggi Negeri Fakultas Ilmu Sosial Dan Ilmu Politik Universitas Diponegoro Semarang Jawa Tengah. *Journal of Public Policy and Management Review*, 6(2), 1-14.
- Saputra, A., & Yusuf, R. (2024). Perbandingan Algoritma DBSCAN dan K-MEANS dalam Segmentasi Pelanggan Pengguna Transportasi Publik Transjakarta Menggunakan Metode RFM. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(4), 1346-1361. <https://doi.org/10.57152/malcom.v4i4.1516>