

PENERAPAN ALGORITMA K-MEANS DAN ELBOW METHOD UNTUK ANALISIS POLA KETERLAMBATAN PEMBAYARAN SANTRI PADA SISTEM KEUANGAN PESANTREN BERBASIS MACHINE LEARNING

Alan¹, Ade Irma Purnamasari², Denni Pratama³, Riri Narasati⁴.

Program Studi Teknik Informatika^{1,2,4}
Program Studi Komputerisasi Akuntansi³

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
alanmaarif14@gmail.com

(*) Corresponding Author : alanmaarif14@gmail.com
Published : 30 Juni 2026

Abstract—The financial management of Pesantren Kebon Kelapa Al Ma'rifah still faces significant challenges related to delays in students' administrative payments, which affect cash flow stability and the effectiveness of financial planning. The limitations of manual systems and the absence of data-driven analysis make it difficult for administrators to map delay patterns and identify student groups at high risk of payment delinquency. This study aims to analyze payment delay patterns using the K-Means Clustering algorithm and determine the optimal number of clusters through the Elbow Method. The dataset consists of student payment transactions from 2025, including due dates, payment dates, bill amounts, payment amounts, and the number of days in arrears. The research process includes collecting administrative and historical transaction data, data preprocessing such as cleaning, date format reconciliation, feature normalization, and delay variable engineering. Subsequently, the Within-Cluster Sum of Squares (WCSS) evaluation was conducted to determine the optimal number of clusters before applying the K-Means model. The analysis results indicate that three clusters provide the most ideal configuration: on-time payers, moderate-delay payers, and high-delay payers. The centroid analysis reveals that the number of delay days is the most dominant factor in cluster formation. These findings demonstrate that machine learning—particularly K-Means—is effective in computationally and objectively mapping student payment behaviors. Beyond providing segmentation insights, this study can assist the pesantren in formulating risk management strategies, establishing payment reminder policies, and enhancing cash flow efficiency. Therefore, the research offers practical contributions to the modernization of pesantren financial systems as well as theoretical contributions to the application of clustering methods in data-driven educational administration analysis.

Keywords: K-Means Clustering; Elbow Method; Machine Learning; Payment Delays; Pesantren Financial System

Abstrak—Pengelolaan keuangan Pesantren Kebon Kelapa Al Ma'rifah masih menghadapi tantangan signifikan terkait keterlambatan pembayaran administrasi santri yang berdampak pada kestabilan arus kas dan efektivitas perencanaan keuangan. Keterbatasan sistem manual dan absennya analisis berbasis data menyebabkan pengelola kesulitan dalam memetakan pola keterlambatan serta mengidentifikasi kelompok santri yang berisiko tinggi menunggak pembayaran. Penelitian ini bertujuan menganalisis pola keterlambatan pembayaran santri dengan menerapkan algoritma *K-Means Clustering* serta menentukan jumlah kluster optimal menggunakan *Elbow Method*. Data yang digunakan merupakan data transaksi pembayaran santri tahun 2025 yang terdiri dari tanggal jatuh tempo, tanggal pembayaran, jumlah tagihan, nominal pembayaran, dan jumlah hari keterlambatan. Proses penelitian meliputi pengumpulan data administratif dan transaksi historis, *praproses* data seperti pembersihan, rekonsiliasi format tanggal, normalisasi fitur, serta rekayasa variabel keterlambatan. Selanjutnya dilakukan evaluasi *Within-Cluster Sum of Squares* (WCSS) untuk menentukan jumlah kluster optimal sebelum model *K-Means* diterapkan. Hasil analisis menunjukkan bahwa tiga kluster merupakan konfigurasi paling ideal, yaitu kelompok pembayar tepat waktu, kelompok dengan keterlambatan sedang, dan kelompok dengan perilaku keterlambatan tinggi. Analisis *centroid* mengungkap variabel keterlambatan hari sebagai penentu paling dominan dalam pembentukan kluster. Temuan ini membuktikan bahwa metode *machine learning*, khususnya *K-Means*, efektif dalam memetakan perilaku pembayaran santri secara *komputasional* dan objektif. Hasil penelitian

ini tidak hanya memberikan gambaran segmentasi perilaku pembayaran, tetapi juga dapat membantu pihak pesantren dalam merumuskan strategi pengelolaan risiko, menetapkan kebijakan pengingat pembayaran, serta meningkatkan efisiensi arus kas. Dengan demikian, penelitian ini memberikan kontribusi praktis bagi modernisasi sistem keuangan pesantren serta kontribusi teoretis bagi penerapan metode *clustering* dalam konteks analisis administrasi pendidikan berbasis data.

Kata Kunci: *K-Means Clustering; Elbow Method; Machine Learning; Keterlambatan Pembayaran; Sistem Keuangan Pesantren*

INTRODUCTION

Berbagai penelitian sebelumnya menunjukkan bahwa metode *clustering* khususnya *K-Means* memiliki efektivitas tinggi dalam mengelompokkan data pendidikan dan keuangan siswa, sehingga relevan untuk diterapkan dalam konteks pengelolaan keuangan pesantren. Penelitian yang dilakukan oleh [1] mengelompokkan siswa berdasarkan ketepatan pembayaran SPP pada SMK Di ponegoro 2 Rawalo. Hasil penelitian tersebut menunjukkan bahwa *K-Means* mampu memetakan siswa yang tepat waktu dan yang menunggak secara akurat, sehingga dapat membantu pihak sekolah dalam memperkirakan risiko keterlambatan pembayaran. Temuan ini diperkuat oleh [2] yang menerapkan metode *clustering* untuk menganalisis preferensi penggunaan pembayaran digital di kalangan mahasiswa. Studi ini menemukan bahwa pengelompokan perilaku pembayaran digital memberikan kontribusi signifikan terhadap pemahaman inklusi keuangan mahasiswa, yang relevan untuk mengidentifikasi perilaku pembayaran santri pada lembaga pendidikan lain.

Dalam konteks pengelolaan keuangan institusional [3] menerapkan *K-Means* untuk mengklasifikasikan data Anggaran Pendapatan dan Belanja Daerah (APBD) Meskipun fokus penelitian ini berada pada instansi pemerintahan, metodologi yang digunakan relevan sebagai rujukan dalam pengelolaan keuangan pesantren, terutama dalam segmentasi anggaran dan deteksi pola pembiayaan. Sementara itu, penelitian [4] pada SMKN 1 Stabat menerapkan *K-Means* untuk mengelompokkan jadwal mengajar guru, menunjukkan bahwa algoritma *clustering* ini sangat fleksibel dan dapat digunakan pada berbagai kebutuhan manajemen sekolah. Studi serupa oleh [5] mengelompokkan tingkat prestasi siswa di MTs Miftahul Ulum Bengkak menggunakan *K-Means*, yang menegaskan bahwa teknik *clustering* efektif dalam memahami perilaku atau karakteristik siswa,

baik dalam bidang akademik maupun aspek administratif.

Dalam ranah keuangan, penelitian [6] mengenai penentuan *kolektabilitas* pinjaman menggunakan kombinasi *Deep Learning* dan *K-Means* menunjukkan bahwa *clustering* mampu mendeteksi tingkat risiko kredit secara akurat. Hal ini sangat relevan dengan konteks keterlambatan pembayaran santri karena konsep *kolektabilitas* paralel dengan risiko tunggakan. Selanjutnya [7] menunjukkan bahwa *K-Means* memiliki kemampuan prediktif kuat dalam memodelkan risiko pasar finansial, sehingga memberikan dasar metodologis yang kuat bahwa *K-Means* dapat juga digunakan untuk memprediksi risiko pembayaran pendidikan.

Dari perspektif teoretis, tinjauan yang dilakukan oleh, [8] menyoroti perkembangan model *clustering* di bidang *Educational Data Science*, termasuk isu *fairness* dalam pengelompokan siswa. Hal ini memberikan dasar bahwa penggunaan metode *clustering* seperti *K-Means* harus memperhatikan keadilan dan objektivitas dalam analisis data pendidikan. Lebih lanjut, artikel ulasan oleh [9] dalam *Information Sciences* menjelaskan berbagai varian *K-Means* dan kelebihanannya dalam menangani data besar. Ulasan tersebut menjadi landasan penting bahwa *K-Means* tidak hanya sederhana dan cepat, tetapi juga mampu memberikan performa optimal dalam *klasterisasi* data skala besar seperti data pembayaran santri.

Akhirnya, penelitian oleh [10] yang menerapkan *K-Means* untuk pembagian jurusan siswa di sekolah menengah menunjukkan bahwa algoritma ini efektif dalam mengelompokkan siswa berdasarkan karakteristik tertentu. Meskipun konteksnya akademik, pendekatan tersebut dapat dengan mudah diadaptasi untuk mengelompokkan santri berdasarkan karakteristik pembayaran mereka. Secara keseluruhan, seluruh penelitian ini memperkuat bahwa metode *K-Means* dan teknik *clustering* lain dapat menjadi solusi strategis dalam memahami dan memprediksi perilaku pembayaran pada lembaga pendidikan, termasuk pesantren.

MATERIALS AND METHODS

Prosedur penelitian merupakan rangkaian langkah sistematis yang disusun untuk memastikan bahwa penelitian dapat dilaksanakan secara terarah, terukur, dan sesuai dengan tujuan yang telah ditetapkan. Dalam studi ini, setiap tahapan dirancang untuk menggambarkan alur kerja mulai dari identifikasi masalah hingga interpretasi hasil akhir, sehingga menghasilkan proses penelitian yang komprehensif dan dapat dipertanggungjawabkan secara ilmiah. merangkum seluruh tahap penelitian mulai dari pengenalan permasalahan, pengumpulan data, pengolahan data, penerapan metode analisis, hingga penyusunan kesimpulan. Untuk memberikan pemahaman yang lebih mendalam mengenai alur penelitian, uraian berikut menjelaskan fungsi dan tujuan dari masing-masing tahap secara rinci.



Gambar 1. Alur Penelitian

Alur penelitian tersebut dimulai dari tahap mulai, yang kemudian dilanjutkan dengan identifikasi masalah untuk memahami permasalahan utama terkait pembayaran santri, khususnya yang berkaitan dengan pola keterlambatan. Setelah itu dilakukan pengumpulan data pembayaran santri sebagai sumber utama penelitian. Data yang telah dikumpulkan kemudian masuk ke tahap pra-proses data, yang meliputi proses *cleaning* (pembersihan data dari kesalahan atau data kosong), *transformasi* (penyesuaian format data), *normalisasi* (penyeragaman skala data), serta *rekayasa fitur* untuk menghasilkan variabel yang lebih representatif.

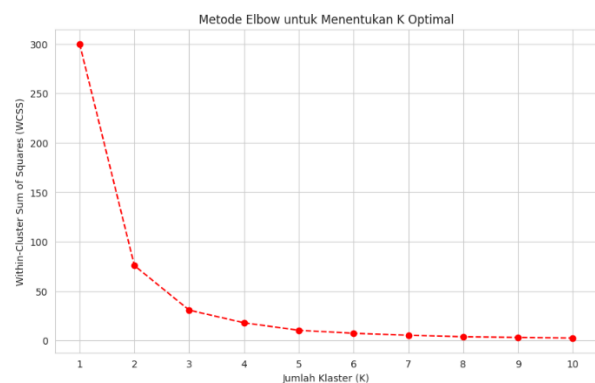
Selanjutnya, dilakukan penentuan jumlah kluster menggunakan metode *Elbow* untuk

mengetahui jumlah kelompok (cluster) yang paling optimal. Berdasarkan hasil tersebut, ditentukan nilai k optimal yang akan digunakan dalam proses clustering. Tahap berikutnya adalah penerapan algoritma K-Means untuk mengelompokkan data pembayaran santri berdasarkan kemiripan karakteristiknya. Hasil dari proses clustering kemudian dievaluasi pada tahap evaluasi hasil kluster untuk memastikan kualitas dan keakuratan pengelompokan.

Setelah itu, dilakukan visualisasi dan analisis centroid untuk memahami karakteristik masing-masing kluster yang terbentuk. Berdasarkan hasil tersebut, peneliti melakukan interpretasi pola keterlambatan pembayaran guna mengidentifikasi kecenderungan atau pola tertentu dalam setiap kelompok. Tahap akhir adalah pembahasan dan implikasi, yang berisi analisis mendalam serta rekomendasi berdasarkan hasil penelitian, hingga akhirnya penelitian mencapai tahap selesai.

RESULTS AND DISCUSSION

Penentuan jumlah kluster optimal dilakukan menggunakan metode *Elbow* dengan menghitung nilai *Within-Cluster Sum of Squares* (WCSS) untuk berbagai nilai k . Nilai WCSS menunjukkan tingkat kekompakan kluster: semakin kecil nilainya, semakin baik pengelompokan. Grafik *Elbow* digunakan untuk mengidentifikasi titik “siku” (*elbow point*), yaitu titik di mana penurunan WCSS



Gambar 2 Penentuan jumlah kluster optimal Elbow

Berdasarkan grafik berikut, titik siku terlihat pada nilai $k = 3$, sehingga jumlah kluster optimal untuk analisis keterlambatan pembayaran santri adalah tiga kluster.

Selain menggunakan Elbow Method, penelitian ini juga melakukan evaluasi tambahan menggunakan Silhouette Score, Davies-Bouldin

Index (DBI), dan Gap Statistic untuk meningkatkan validitas hasil klusterisasi.

Tabel 1 Hasil Evaluasi Penentuan Jumlah Klaster

Jumlah Klaster (K)	Silhouette Score	DBI	Gap Statistic
2	0.5213	0.8123	0.3456
3	0.6124	0.7012	0.4567
4	0.5801	0.7456	0.4102

Berdasarkan tabel di atas, nilai Silhouette Score tertinggi diperoleh pada $K = 3$ sebesar 0,6124, yang menunjukkan kualitas pemisahan klaster yang baik. Nilai Davies-Bouldin Index (DBI) terendah juga diperoleh pada $K = 3$ yaitu sebesar 0,7012, yang menandakan tingkat kemiripan intra-klaster yang rendah dan separasi antar klaster yang baik. Selain itu, nilai Gap Statistic tertinggi juga terdapat pada $K = 3$ sebesar 0,4567. Dengan demikian, berdasarkan ketiga metode evaluasi tersebut, jumlah klaster optimal adalah $K = 3$, dan hasil ini konsisten dengan metode Elbow.

CONCLUSION

penelitian ini menunjukkan bahwa penerapan metode *machine learning*, khususnya algoritma *K-Means Clustering* yang didukung oleh *Elbow Method*, mampu memetakan pola keterlambatan pembayaran santri secara lebih akurat dan sistematis. Pendekatan analitis ini memberikan gambaran perilaku pembayaran santri secara objektif dan dapat menjadi dasar implementasi strategi manajemen keuangan pesantren yang lebih efektif. Dengan demikian, penelitian ini menegaskan bahwa teknologi data *mining* dapat berperan penting dalam meningkatkan kualitas pengelolaan keuangan lembaga pendidikan.

Efektivitas *K-Means* dalam Pemetaan Perilaku Pembayaran Penelitian membuktikan bahwa algoritma *K-Means Clustering* mampu mengelompokkan santri berdasarkan pola keterlambatan pembayaran secara komputasional, objektif, dan sistematis. Proses clustering berhasil membagi santri ke dalam beberapa kelompok perilaku pembayaran yang representatif sehingga mempermudah pihak pesantren dalam memahami tingkat risiko keterlambatan berdasarkan data historis.

Penentuan Jumlah Klaster Optimal Menggunakan *Elbow Method* Metode *Elbow Method* menghasilkan jumlah klaster optimal sebanyak tiga klaster, yaitu: kelompok santri yang membayar tepat waktu, kelompok dengan

keterlambatan sedang, dan kelompok dengan keterlambatan tinggi. Konfigurasi ini terbukti paling ideal berdasarkan nilai WCSS dan analisis centroid, serta menjadi dasar segmentasi perilaku pembayaran santri yang lebih akurat dan informatif.

Kontribusi Praktis terhadap Sistem Keuangan Pesantren Hasil analisis menunjukkan bahwa variabel jumlah hari keterlambatan merupakan faktor dominan dalam pembentukan klaster. Temuan ini memberikan kontribusi praktis bagi pengelola pesantren dalam merumuskan strategi manajemen keuangan, seperti sistem pengingat pembayaran, pemetaan risiko tunggakan, hingga rancangan kebijakan administrasi yang lebih adaptif. Dengan demikian, penelitian ini mendukung modernisasi sistem keuangan pesantren menuju pengelolaan berbasis data dan teknologi.

REFERENCE

- [1] S. Soimah and E. Susilowati, "Penerapan Metode K-Means Cluster untuk Mengklasifikasikan Pembayaran SPP Siswa," *Jurnal Pembelajaran dan Matematika SIGMA*, 2023.
- [2] M. A. Khoiri, R. B. Kotjopradyudi, and A. Y. Eskaluspa, "Analisis Cluster Terhadap Pengelompokan Preferensi Pengguna Pembayaran Digital," *MALCOM*, 2025.
- [3] M. R. Widiyanto, Zakaria, and Irvan, "Penerapan Data Mining untuk Klusterisasi Data APBD Menggunakan K-Means," *Jurnal JTIK*, 2025.
- [4] S. Sutiono, R. F. Wijaya, and M. S. Novelan, "Analisis Metode K-Means Clustering untuk Pengelompokan Jadwal Mengajar Guru," *INTECOMS*, 2023.
- [5] S. Maghfiroh and Z. Fatah, "Analisis Data Mining dengan Algoritma K-Means untuk Menentukan Siswa Berprestasi," *JAMASTIKA*, 2025.
- [6] F. Yenila, H. Marfalino, and S. Defit, "Model Analisis Machine Learning dalam Penentuan Kolektabilitas," *JST*, 2024.
- [7] J. Xu, K. Xu, Y. Wang, Q. Shen, and R. Li, "A K-means Algorithm for Financial Market Risk Forecasting," *arXiv*, 2024.
- [8] T. L. Quy, G. Friege, and E. Ntoutsis, "A Review of Clustering Models in Educational Data Science Towards Fairness-Aware Learning," *arXiv*, 2023.
- [9] A. M. Ikotun, L. Ezugwu, B. Abualigah, B. Abuhaija, and J. Heming, "K-Means Clustering Algorithms: A Comprehensive Review," *Information Sciences*, 2023.

- [10] Q. A. A'yuniyah and M. Reza, "Penerapan Data Mining Dengan Menggunakan Algoritma K-Means untuk Pembagian Jurusan SMA," *JoSYC*, 2024.