

## Evaluasi Pembayaran Keuangan Siswa berdasarkan Penghasilan Wali Siswa menggunakan Metode *Clustering K-Means*

Imam Ahmad Ashari<sup>1,\*</sup>, Iis Setiawan Mangku Negara<sup>1</sup>, R Bagus Bambang Sumantri<sup>2</sup>

<sup>1</sup> Program Studi Teknologi Informasi, Universitas Harapan Bangsa, Indonesia

<sup>2</sup> Program Studi Sistem Informasi, Universitas Harapan Bangsa, Indonesia

\* Correspondence: imamahmadashari@uhb.ac.id

**Copyright:** © 2022 by the authors

Received: 17 Agustus 2022 | Revised: 18 Agustus 2022 | Accepted: 1 Desember 2022 | Published: 20 Desember 2022

### Abstrak

Salah satu masalah yang sering terjadi di administrasi sekolah adalah keterlambatan pembayaran biaya pendidikan. Oleh karena itu perlu adanya evaluasi terhadap proses pembayaran pendidikan agar ke depan proses pembayaran dapat berjalan dengan tertib dan disiplin. Penelitian ini bertujuan untuk membuat sebuah model kluster pengelompokan pembayaran administrasi siswa. Jenis penelitian ini adalah kuantitatif menggunakan metode K-Means *clustering* untuk mengelompokkan data pembayaran berdasarkan 2 variabel, yaitu waktu pembayaran dan penghasilan wali siswa yang dilaksanakan di SD swasta Semarang. Data yang digunakan pada penelitian ini adalah data pembayaran pada tahun ajaran 2019/2020 yang berjumlah 1.933 *record*, meliputi transaksi dari 419 siswa. Penentuan jumlah kluster dihitung menggunakan metode *elbow*, didapatkan kluster terbaik dari data yang digunakan adalah 3 kluster, yaitu kluster 0, 1 dan 2. Hasil temuan kami didapat kluster 2 memiliki presentase terbesar dalam pembayaran administrasi bulanan lebih awal, yaitu 52,5%, presentase tepat waktu paling tinggi ada di kluster 1 yaitu 74,8%, dan presentase telat bayar tertinggi ada di kluster 0 yaitu 28,6%. Hasil analisis menunjukkan bahwa faktor utama keterlambatan pembayaran bukanlah dari penghasilan wali siswa melainkan ada faktor eksternal lainnya, dibuktikan dari presentase telat bayar tertinggi ada di kluster 0, dimana rata-rata penghasilan wali siswa  $\geq 10.000.000$ .

**Kata kunci:** biaya pendidikan; evaluasi pembayaran; metode *clustering*; algoritma k-means; metode *elbow*

### Abstract

*One of the problems that often occurs in school administration is the late payment of tuition fees. Therefore, it is necessary to evaluate the education payment process so that in the future the payment process can run in an orderly and disciplined manner. This study aims to create a cluster model for grouping student administration payments. This type of research is quantitative using the K-Means Clustering method to classify payment data based on 2 variables, namely the time of payment and the income of students' guardians carried out in private elementary schools in Semarang. The data used in this study is payment data for the 2019/2020 academic year, which totals 1,933 records, covering transactions from 419 students. Determining the number of clusters is calculated using the elbow method, the best clusters obtained from the data used are 3 clusters, namely clusters 0, 1 and 2. Our findings show that cluster 2 has the largest percentage of early monthly administration payments, namely 52.5%, the percentage is on time the highest was in cluster 1, namely 74.8%, and the highest percentage of late payments was in cluster 0, namely 28.6%. The results of the analysis show that the main factor for late payments is not the guardian's income but other external factors, as evidenced by the highest percentage of late payments in cluster 0, where the average income of student guardians is  $\geq 10,000,000$ .*



**Keywords:** *cost of education; payment evaluation; clustering method; k-means algorithm; elbow method*

## PENDAHULUAN

Pendidikan sebagai proses peningkatan sebuah kualitas sumber daya manusia berperan penting dalam upaya pembangunan sebuah bangsa dan negara. Pendidikan merupakan proses yang berkelanjutan dan tidak pernah berakhir, sehingga dapat menghasilkan kualitas yang berkesinambungan, yang ditunjukkan pada perwujudan sosok manusia masa depan, dan berakar pada nilai-nilai budaya bangsa serta Pancasila (Sujana, 2019). Melalui sebuah pendidikan yang baik, dapat diperoleh berbagai hal baru sehingga bisa digunakan dalam menciptakan sumber daya manusia yang berkualitas. Biaya pendidikan adalah salah satu faktor pendukung bagi penyelenggara pendidikan, khususnya sekolah atau instansi swasta (Prabowo et al., 2021). Tanpa adanya biaya pendidikan operasional sekolah tidak akan berjalan dan berkembang. Hanya saja keterlambatan pembayaran oleh siswa masih sering terjadi yang menyebabkan pembangunan sekolah tertahan. Seperti yang terjadi pada tahun ajaran yang berjalan sekarang, yaitu 2021/2022 di bulan september masih ada 24 siswa yang terlambat melakukan pembayaran SPP. Oleh karena itu perlu adanya evaluasi untuk memastikan faktor keterlambatan pembayaran tersebut.

Berdasarkan penelitian Hossain et al. (2019) analisis permasalahan yang dilakukan menggunakan salah satu teknik dari data mining untuk menemukan pola keterlambatan pembayaran siswa berdasarkan variabel waktu pembayaran dan penghasilan wali siswa. Metode yang diusulkan menghitung nilai ambang sebagai pusat K-Means dan berdasarkan nilai ini jumlah *cluster* terbentuk. Pada setiap iterasi K-Means, jika jarak Euclidian antara dua titik lebih kecil atau sama dengan nilai *threshold*, maka kedua titik data tersebut akan berada dalam satu grup yang sama. Jika tidak, metode yang diusulkan akan membuat cluster baru dengan titik data yang berbeda. Data mining adalah proses menemukan struktur data dari kumpulan data yang besar. Dengan proses ini, para pengambil keputusan dapat membuat keputusan tertentu untuk pengembangan lebih lanjut dari masalah dunia nyata, salah satu teknik data mining yang terkenal adalah metode K-Means *Clustering*. Data mining bekerja dengan menemukan pola atau pengetahuan dengan menggunakan beberapa metode atau algoritma, beberapa peran utama data mining yaitu: estimasi, prediksi, klasifikasi, *clustering*, dan asosiasi (Uska et al., 2020). Teknik *clustering* adalah teknik pengelompokan data dengan mendefinisikan kelas dari data pengetahuan tanpa menggunakan label kelas (Rodriguez et al., 2019). Metode *clustering* digunakan untuk menemukan struktur kluster dalam kumpulan data yang dicirikan oleh kesamaan terbesar dalam kluster yang sama dan perbedaan terbesar antara kluster yang berbeda (Sinaga & Yang, 2020). *Clustering* dianggap sebagai pendekatan pemodelan data yang memberikan tinjauan singkat data (Toma, 2021). Di antara banyak algoritma *clustering*, K-Means merupakan algoritma yang banyak digunakan karena memiliki algoritma sederhana dan konvergensi yang cepat (Yuan & Yang, 2019).

Algoritma K-Means adalah metode *clustering* yang paling umum digunakan dalam penelitian (Syakur et al., 2018). Metode ini mempartisi data ke dalam kluster sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu kluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kluster yang lain. Algoritma K-Means merupakan algoritma pembelajaran yang sederhana yang dapat menyelesaikan sebuah permasalahan untuk meminimalkan kesalahan ganda (Nurzahputra et al., 2017). K-Means adalah contoh terkenal dari algoritma *clustering* partisi yang memperoleh partisi K objek melalui perbaikan berulang sehingga jumlah jarak kuadrat antara objek dan pusat kluster diminimalkan (Jothi et al., 2019). Kelemahan algoritma K-Means antara lain penentuan jumlah kluster berdasarkan asumsi dan sangat mengandalkan pemilihan awal centroid untuk mengatasi kelemahan tersebut (Umargono et al., 2020). Untuk

menentukan jumlah kluster atau partisi K yang dipakai pada algoritma K-Means dapat menggunakan metode *elbow* untuk menentukannya (Marutho et al., 2018). Metode *Elbow* merupakan metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah kluster terbaik dengan melihat persentase perbandingan jumlah kluster yang akan membentuk siku pada suatu titik (Nainggolan et al., 2019).

Pada penelitian sebelumnya algoritma K-Means digunakan untuk melakukan analisis pergerakan harga saham perbankan, dimana analisis pergerakan saham bank di klasifikasikan menjadi tiga cluster, yaitu *lower*, *middle*, dan *upper* kluster (Zuhroh et al., 2021). Pada bidang investasi, algoritma K-Means digunakan untuk menganalisis efisiensi investasi, dimana hasil analisis dari algoritma K-Means akan digunakan sebagai patokan investasi masa depan (Li et al., 2020). Pada bidang analisis perilaku, algoritma K-Means digunakan untuk menganalisis kebiasaan perilaku mahasiswa dalam kehidupan sehari-hari dan untuk membuat penilaian yang akurat tentang perilaku konsumen (Yang et al., 2020). Pada studi kasus pembayaran Algoritma K-Means pernah digunakan untuk menganalisis hubungan Pendapatan dengan metode pembayaran pelanggan pada usaha kecil menengah (Marisa et al., 2021). Algoritma K-Means juga pernah di kombinasikan dengan metode *Analytic Hierarchy Process* (AHP) untuk mencari rekomendasi pembiayaan UMKM, algoritma K-Means digunakan untuk pengelompokkan laba bersih. Kombinasi AHP digunakan dalam pengambilan keputusan multikriteria dan cukup baik untuk memecahkan masalah identifikasi pendanaan nasabah yang membutuhkan banyak kriteria (Sukmadewanti et al., 2018).

Pada penelitian ini algoritma K-Means *Clustering* digunakan dengan tujuan untuk membuat sebuah model kluster pengelompokkan pembayaran administrasi siswa berdasarkan waktu pembayaran dan penghasilan wali siswa. Hasil dari evaluasi akan digunakan sebagai pengambilan kebijakan bagi siswa yang terlambat melakukan pembayaran bulanan. K-Means digunakan untuk mempartisi data yang ada ke dalam bentuk satu atau lebih cluster. Hasil yang diharapkan adalah adanya pengelompokkan data sesuai dengan golongan siswa berdasarkan hubungan penghasilan wali siswa dengan waktu pembayaran administrasi sekolah.

## METODE

Metode yang digunakan pada penelitian ini adalah metode K-Means *Clustering*. Metode K-Means digunakan untuk mengelompokkan data pembayaran berdasarkan 2 variabel, yaitu waktu pembayaran dan penghasilan wali siswa. Data dikelompokkan menjadi 3 kluster, dimana ketentuan jumlah kluster di hitung menggunakan metode *elbow*.

**Tabel 1.** Variabel pengujian

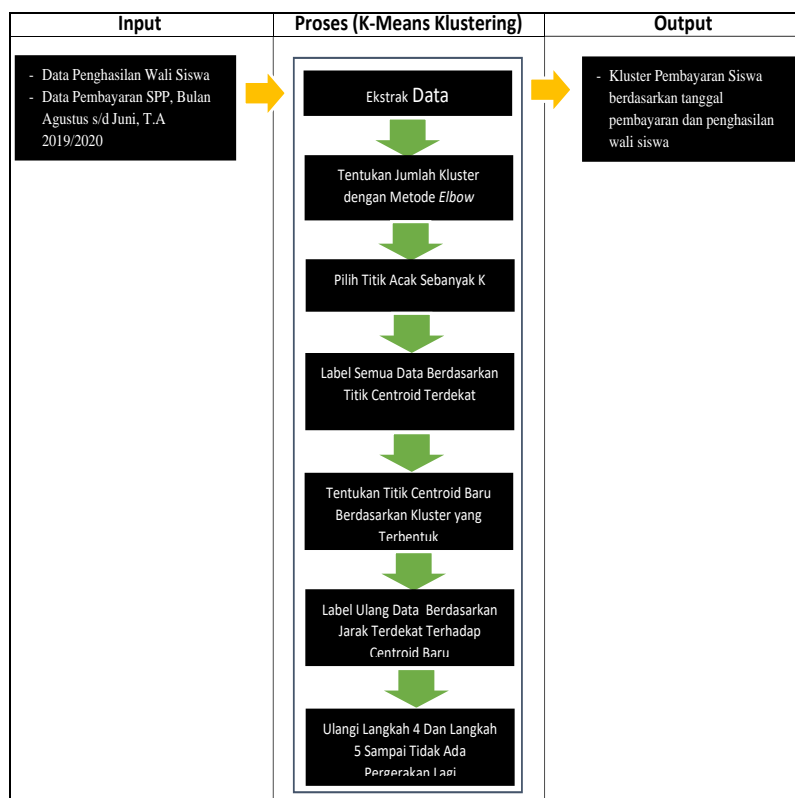
<b>Tanggal Pembayaran</b>	<b>Penghasilan Wali Siswa</b>
08/08/2020	2000000
13/08/2020	3000000
06/08/2020	2500000
12/09/2020	2500000
25/07/2020	2500000
25/07/2020	2500000
03/08/2020	3000000
dst	...

Sumber data yang dipakai dalam penelitian ini adalah data di SD Swasta di Semarang. Data diambil di tahun ajaran aktif 2019/ 2020. Data diambil pada semua bulan, kecuali bulan juli, di karenakan SPP bulan juli di bayarkan pada saat registrasi. Data yang diambil adalah

data penghasilan wali siswa dan data pembayaran SPP. Data dikelompokkan berdasarkan kluster penghasilan wali siswa dan tanggal pembayaran SPP. Data pengelompokkan dianalisis untuk membuat kebijakan terkait dengan teguran atau penundaan pembayaran terhadap wali siswa. Sebelum proses analisis menggunakan metode K-Means *Clustering* data di transformasi terlebih dahulu, beberapa atribut yang tidak diperlukan dihilangkan.

Teknik K-Means *Clustering* untuk memproses data pembayaran SPP siswa dapat dilihat pada gambar 1. Data Variabel Pengujian yang merupakan data relasi dari data pembayaran SPP siswa dan data penghasilan wali murid dapat dilihat pada Tabel 1. Data ini adalah data awal yang akan di proses menggunakan metode K-Means *Clustering*.

Berdasarkan gambar 1, penelitian ini dimulai dari menyiapkan data input, data yang dipakai adalah data penghasilan wali siswa dan data pembayaran SPP. Data input akan di proses menggunakan metode K-Means klustering, dimana terdapat 7 (tujuh) proses tahapan yang di lalui, yaitu proses ekstrak data, menentukan jumlah kluster dengan metode *elbow*, memilih titik acak sebanyak K, memberi label ke semua data berdasarkan titik centroid terdekat, menentukan titik centroid baru berdasarkan kluster yang terbentuk, melakukan labeling ulang data berdasarkan jarak terdekat terhadap centroid baru dan yang terakhir adalah mengulangi langkah 4 dan 5 sampai tidak ada pergerakan lagi. Dari proses tersebut akan di dapatkan output yaitu kluster pembayaran siswa berdasarkan tanggal pembayaran dan penghasilan wali siswa.



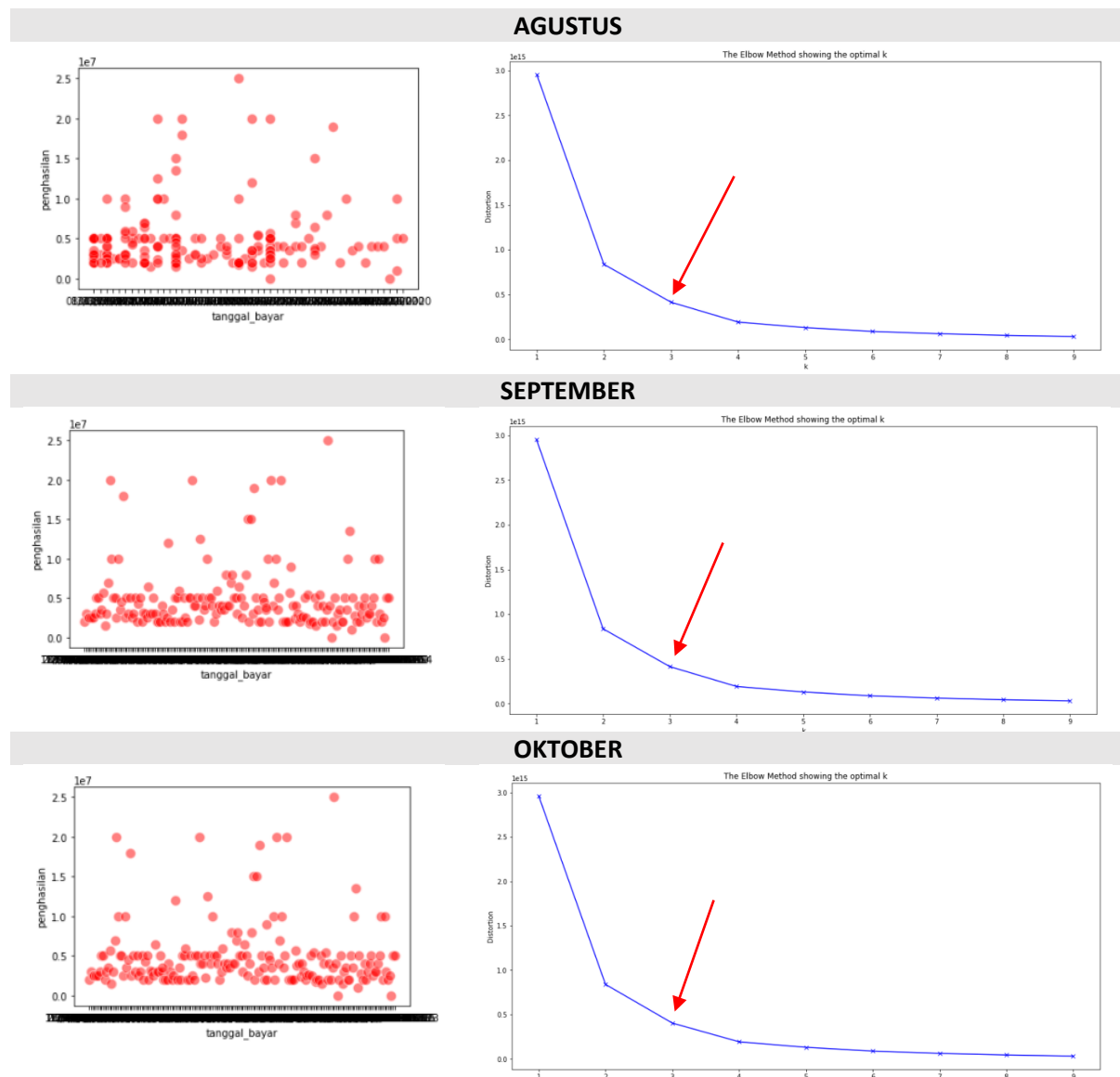
Gambar 1. Desain penelitian

## HASIL DAN PEMBAHASAN

### Hasil

Jumlah kluster yang akan digunakan dalam penelitian ditentukan menggunakan metode *elbow*. Metode *elbow* telah terbukti berhasil mengoptimalkan estimasi jumlah kluster yang dipakai pada metode K-Means *Clustering* (Sammouda & El-zaart, 2021). Data uji yang

digunakan adalah data uji pada bulan Agustus, September dan Oktober. Hasil dari pengujian penentuan jumlah kluster menggunakan metode *elbow* dapat dilihat pada gambar 2.



**Gambar 2.** Hasil pengujian penentuan jumlah kluster

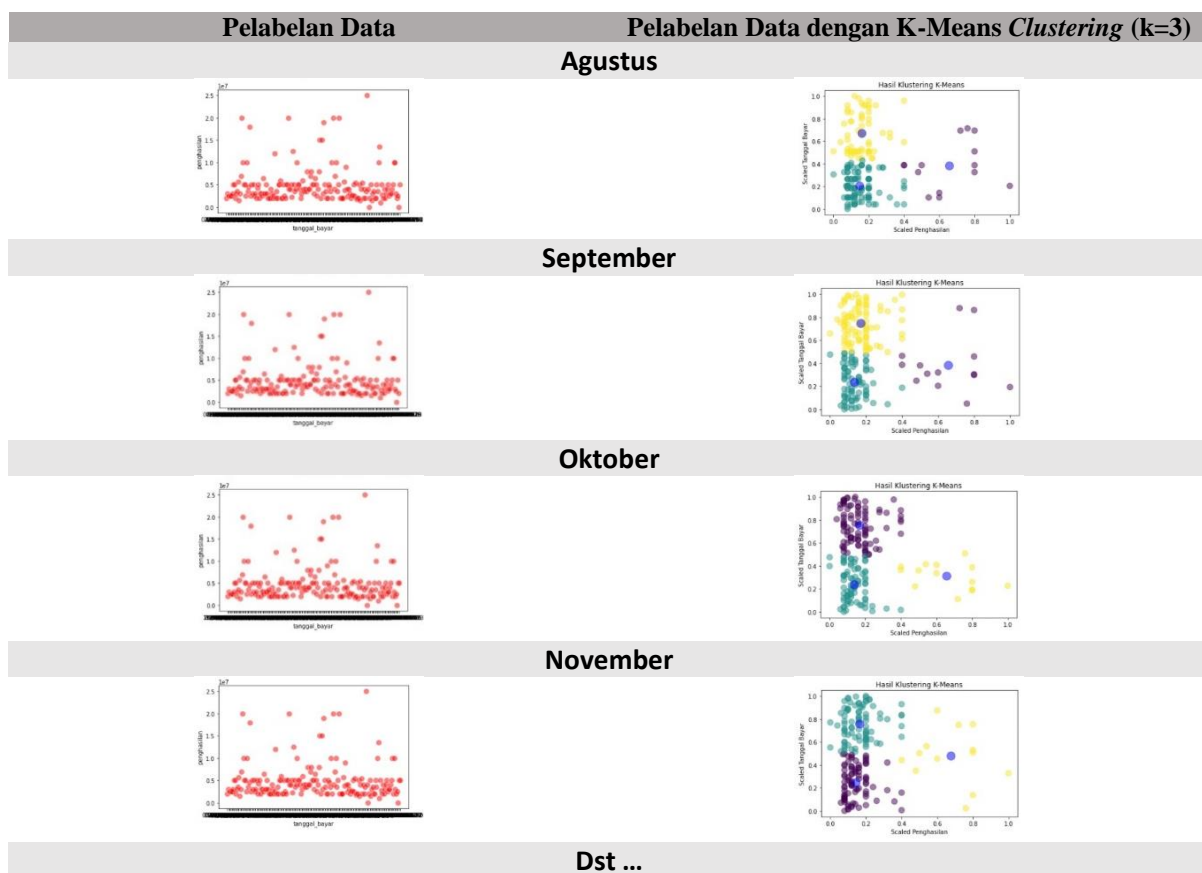
Pada gambar 2 diperlihatkan bahwa titik siku pada pengujian ketiga data mengarah ke kluster 3. Dari pengujian *elbow* di tetapkan bahwa jumlah kluster yang dipakai adalah 3 kluster. Setelah menentukan jumlah kluster menggunakan metode *elbow* langkah selanjutnya adalah membuat klusterisasi data pembayaran dan penghasilan wali siswa menggunakan metode K-Means. Data yang diambil dari database sistem pembayaran direlasikan antara data pembayaran SPP dengan data penghasilan wali siswa. Relasi data dibuat menggunakan *id\_siswa*. Data relasi antara pembayaran SPP dengan data penghasilan wali siswa dapat dilihat pada tabel 2.

Pada tabel 2 dapat dilihat ada beberapa variatif tanggal pembayaran di bulan Agustus. Adapun kebijakan dari sekolah batas keterlambatan pembayaran adalah di tanggal 10. Dari contoh sampel data di atas ada 6 siswa yang telat melakukan pembayaran, dimana 2 siswa di antaranya memiliki wali siswa yang memiliki penghasilan tinggi. Langkah selanjutnya setelah data di relasikan adalah mengimplementasikan metode K-Means *Clustering*. Hasil

dari implementasi metode K-Means *Clustering* pada data pembayaran SPP dan penghasilan wali siswa dapat dilihat pada gambar 3.

**Tabel 2.** Data relasi pembayaran spp dengan penghasilan wali siswa

No	ID Siswa	Bulan Pembayaran	Tanggal Pembayaran	Penghasilan Wali Siswa
1	3	Agustus	08/08/2020	2000000
2	4	Agustus	13/08/2020	3000000
3	5	Agustus	06/08/2020	2500000
4	6	Agustus	12/09/2020	2500000
5	7	Agustus	25/07/2020	2500000
6	8	Agustus	25/07/2020	2500000
7	13	Agustus	03/08/2020	3000000
8	16	Agustus	29/07/2020	5000000
9	18	Agustus	16/07/2020	5000000
10	20	Agustus	06/08/2020	3000000
dst	...	...	...	...



**Gambar 3.** Hasil implementasi metode k-means *clustering*

Hasil pada gambar 3 merupakan implementasi dari Metode K-Means *Clustering*, dimana hasil dari metode K-Means *Clustering* adalah pelabelan. Kolom pelabelan data menunjukkan hasil pelabelan dari data mentah yang di inputkan, sehingga semua label berwarna merah, sedangkan hasil pelabelan data dengan 3 kluster K-Means Clustering

menghasilkan 3 warna, yaitu biru, kuning dan ungu, dapat dilihat pada kolom Pelabelan Data dengan K-Means Clustering ( $k=3$ ).

Hasil implementasi metode K-Means Clustering terhadap data pembayaran siswa dan penghasilan wali siswa dapat dilihat pada tabel 2. Setiap baris menunjukkan nama bulan, jumlah data yang di proses dan total jumlah data yang masing – masing berada pada kluster 0, 1, dan 2. Tabel 2 menunjukkan hasil klusterisasi menggunakan Metode K-Means Clustering. Langkah selanjutnya adalah menganalisis karakter dari masing-masing kluster.

**Tabel 2.** Hasil klusterisasi menggunakan metode k-means *clustering*

No	Bulan	Jumlah Data	Kluster 0	Kluster 1	Kluster 2
1	Agustus	183	14	108	61
2	September	183	14	108	61
3	Oktober	184	93	77	14
4	November	184	86	85	13
5	Desember	181	79	14	88
6	Januari	175	12	80	83
7	Februari	171	78	83	10
8	Maret	169	83	76	10
9	April	169	81	78	10
10	Mei	168	75	83	10
11	Juni	166	78	72	16

Hasil analisis menunjukkan bahwa kluster 0 adalah data wali siswa dengan penghasilan  $\geq 10.000.000$ , sedangkan kluster 1 dan 2 penghasilan wali siswa antara 1.000.000 s/d 10.000.000. Setelah menganalisis penghasilan wali siswa langkah selanjutnya adalah menganalisis waktu pembayaran. Sementara itu, tabel 3 memperlihatkan presentase siswa yang melakukan pembayaran lebih awal adalah kluster 2, dimana pendapatan wali siswa antara 1.000.000 s/d 10.000.000, pembayaran tepat waktu tertinggi ada di kluster 1 yaitu 74.8%, sedangkan kluster 0 dimana wali siswa dengan pendapatan  $\geq 10.000.000$  memiliki presentase telat bayar paling tinggi yaitu 28.6%. Dari analisis tersebut maka disimpulkan bahwa besar penghasilan wali siswa bukanlah faktor utama ketepatan waktu pembayaran siswa melainkan ada penyebab dari faktor eksternal lainnya.

**Tabel 3.** Hasil presentase analisis waktu pembayaran

Ketrangan Pembayaran	Kluster		
	0	1	2
Lebih awal	3,6%	15,2%	52,5%
Tepat waktu	67,9%	74,8%	31,1%
Telat bayar	28,6%	10,0%	16,4%
Total Presentase	100,0%	100,0%	100,0%

Pengujian hasil pada penelitian ini menggunakan Confusion Matrix yang dapat menghasilkan nilai *accuracy*, *recall*, *precision* dan *f1-score*. Pengujian menggunakan data pembayaran pada bulan agustus yang disajikan pada gambar 4. Berdasarkan gambar 4, nilai akurasi (*accuracy*) yang didapatkan oleh model klusterisasi dengan metode K-Means *clustering* pada penelitian ini adalah 97%. Terdapat 4 istilah *sebagai* representasi hasil proses klasifikasi pada *confusion matrix*. Keempat istilah tersebut adalah *True Positive* (TP) adalah *True positif* yang diklasifikasikan positif, *True Negative* (TN) adalah *True negatif* yang diklasifikasikan negatif, *False Positive* (FP) adalah *False positif* yang diklasifikasikan positif

dan *False Negative* (FN) adalah False negatif yang diklasifikasikan negatif. Empat kombinasi nilai prediksi dan nilai aktual yang diperoleh dapat dilihat pada gambar 5.

	precision	recall	f1-score	support
0	0.96	1.00	0.98	163
1	1.00	0.70	0.82	20
accuracy			0.97	183
macro avg	0.98	0.85	0.90	183
weighted avg	0.97	0.97	0.96	183

Gambar 4. Hasil pengujian menggunakan *confusion matrix*

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	<b>TP</b> <b>14</b>	<b>FP</b> <b>0</b>
	0 (Negative)	<b>FN</b> <b>6</b>	<b>TN</b> <b>75</b>

Gambar 5. Nilai prediksi

### Pembahasan

Data penelitian merupakan data pembayaran SD Swasta di Semarang pada tahun ajaran 2019/2020. Data pembayaran diambil dari bulan agustus sampai dengan mei, dengan jumlah keseluruhan data adalah 1.933 *record*/transaksi. Variabel yang digunakan sebagai bahan penelitian adalah variabel waktu pembayaran dan penghasilan wali siswa. Sebelum dilakukan pemrosesan menggunakan metode K-Means *clustering* langkah pertama adalah ekstrak data, agar data mudah diolah. Data yang diambil adalah data tanggal pembayaran dan data pendapatan wali siswa. Selanjutnya data yang sudah sudah di ekstrak ditentukan jumlah kluster yang kiranya sesuai untuk pelabelan. Penentuan jumlah kluster dilakukan menggunakan metode *elbow*. Dari pengujian penentuan jumlah kluster didapat jumlah kluster yang paling optimal pada data yang digunakan adalah berjumlah 3 kluster.

Setelah dilakukan analisis menggunakan metode K-Means *Clustering* didapat beberapa karakteristik dari pelabelan, yaitu presentase siswa yang melakukan pembayaran lebih awal adalah kluster 2, dimana pendapatan wali siswa antara 1.000.000 s/d 10.000.000, pembayaran tepat waktu tertinggi ada di kluster 1 yaitu 74,8%, sedangkan kluster 0 dimana wali siswa dengan pendapatan  $\geq$  10.000.000 memiliki presentase telat bayar paling tinggi yaitu 28,6%.

Hasil pengujian menggunakan *confusion matrix* menunjukkan akurasi sebesar 97%. Nilai precision negatif 90% dan positif 100%. Nilai recall negatif 100% dan positif 70%. Nilai f1-score negatif 98% dan positif 82%. Jika dibandingkan dengan beberapa metode yang digunakan pada penelitian (Sari et al., 2020), maka presentase hasil pengujian menunjukkan nilai *precision*, *recall* dan *f1-score* sudah cukup baik. Klusterisasi pembayaran administrasi siswa menggunakan metode K-Means *Clustering* merupakan objek penelitian baru yang dilakukan. Dimana penelitian sebelumnya memperoleh hasil prediksi keterlambatan



pembayaran sekolah menggunakan metode K-Nearest Neighbor berhasil dilakukan dengan hasil akurasi 86% (Abdullah et al, 2019).

## SIMPULAN

Analisis hubungan antara penghasilan wali siswa dengan waktu pembayaran sekolah menggunakan metode K-Means *Clustering* berhasil dibuat. Model kluster pengelompokan pembayaran administrasi siswa dikelompokkan ke dalam 3 kluster, yaitu kluster 0,1 dan 2. Analisis hasil menunjukkan bahwa besarnya penghasilan wali siswa tidak begitu berpengaruh terhadap waktu pembayaran siswa. Hasil temuan pada kluster 0 memperlihatkan bahwa masih banyak wali siswa yang berpenghasilan tinggi yang melakukan keterlambatan pembayaran. Dari hasil analisis disimpulkan bahwa keterlambatan pembayaran di sekolah tidak selalu dipengaruhi oleh penghasilan wali siswa, melainkan ada faktor eksternal lainnya, oleh karena itu sekolah perlu mencari faktor lain yang kiranya mempengaruhi keterlambatan pembayaran ini sehingga proses pembayaran dapat berjalan dengan baik. Hasil evaluasi ini dapat digunakan pihak sekolah untuk menindak lanjuti dan meminimalisir keterlambatan pembayaran siswa.

## REFERENSI

- Abdullah, R. W., Kusriani, K., & Luthfii, E. T. (2019). Prediksi Keterlambatan Pembayaran Spp Sekolah Dengan Metode K-Nearest Neighbor (Studi Kasus Smk Al-Islam Surakarta). *Informasi Interaktif*, 4(3), 160-164.
- Hossain, M. Z., Akhtar, M. N., Ahmad, R. B., & Rahman, M. (2019). A dynamic K-means clustering for data mining. *Indonesian Journal of Electrical Engineering and Computer Science*, 13(2), 521–526. <https://doi.org/10.11591/ijeecs.v13.i2.pp521-526>
- Jothi, R., Mohanty, S. K., & Ojha, A. (2019). DK-means: a deterministic K-means clustering algorithm for gene expression analysis. *Pattern Analysis and Applications*, 22(2), 649–667. <https://doi.org/10.1007/s10044-017-0673-0>
- Li, L., Wang, J., & Li, X. (2020). Efficiency Analysis of Machine Learning Intelligent Investment Based on K-Means Algorithm. *IEEE Access*, 8, 147463–147470. <https://doi.org/10.1109/ACCESS.2020.3011366>
- Marisa, F., Ahmad, S. S. S., Yusof, Z. I. M., Akhriza, T. M., Purnomowati, W., & Pandey, R. K. (2021). The Analyze of Relationship between Revenue and Customer Payment Methods in Small Medium Enterprise Based on Clustering K-Means. In *Journal of Physics: Conference Series*, 1908(1), 1-9. IOP Publishing. <https://doi.org/10.1088/1742-6596/1908/1/012021>
- Marutho, D., Hendra Handaka, S., Wijaya, E., & Muljono. (2018). The Determination of Cluster Number at k-Mean Using Elbow Method and Purity Evaluation on Headline News. *Proceedings - 2018 International Seminar on Application for Technology of Information and Communication: Creative Technology for Human Life, ISEmantic 2018*, 533–538. <https://doi.org/10.1109/ISEMANTIC.2018.8549751>
- Nainggolan, R., Perangin-Angin, R., Simarmata, E., & Tarigan, A. F. (2019). Improved the Performance of the K-Means Cluster Using the Sum of Squared Error (SSE) optimized by using the Elbow Method. *Journal of Physics: Conference Series*, 1361(1). <https://doi.org/10.1088/1742-6596/1361/1/012015>
- Nurzahputra, A., Muslim, M. A., & Khusniati, M. (2017). Penerapan Algoritma K-Means Untuk Clustering Penilaian Dosen Berdasarkan Indeks Kepuasan Mahasiswa. *Techno.Com*, 16(1), 17–24. <https://doi.org/10.33633/tc.v16i1.1284>
- Prabowo, C., Nul, L., Cep, H., Rohmat, L., & Rizki, A. (2021). *JURNAL DATA SCIENCE & INFORMATIKA ( JDSI ) Teknik Klasifikasi Pembayaran SPP Berdasarkan Tingkat Ketepatan Pembayaran*. 1(1), 1–5.

- Rodriguez, M. Z., Comin, C. H., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. da F., & Rodrigues, F. A. (2019). Clustering algorithms: A comparative approach. In *PLoS ONE*, 14(1), 1-34. <https://doi.org/10.1371/journal.pone.0210236>
- Sammouda, R., & El-Zaart, A. (2021). An optimized approach for prostate image segmentation using K-means clustering algorithm with elbow method. *Computational Intelligence and Neuroscience*, 2021, 1-13. <https://doi.org/10.1155/2021/4553832>
- Sari, V., Firdausi, F., & Azhar, Y. (2020). Perbandingan Prediksi Kualitas Kopi Arabika dengan Menggunakan Algoritma SGD, Random Forest dan Naive Bayes. *Edumatic: Jurnal Pendidikan Informatika*, 4(2), 1-9. <https://doi.org/10.29408/edumatic.v4i2.2202>
- Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, 8, 80716-80727. <https://doi.org/10.1109/ACCESS.2020.2988796>
- Sujana, I. W. C. (2019). Fungsi dan Tujuan Pendidikan Indonesia. *Adi Widya: Jurnal Pendidikan Dasar*, 4(1), 29-39. <https://doi.org/10.25078/aw.v4i1.927>
- Sukmadewanti, I., Arifudin, R., & Sugiharti, E. (2018). Use of K-Means Clustering and Analytical Methods Hierarchy Process in Determining the Type of MSME Financing in Semarang City. *Scientific Journal of Informatics*, 5(2), 148-158. <https://doi.org/10.15294/sji.v5i2.16221>
- Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering*, 336(1), 1-6. <https://doi.org/10.1088/1757-899X/336/1/012017>
- Toma, G. A. (2021). Performance analysis of state of art clustering approaches in data mining. *Int J Inf Technol (IJIT)*, 7(3), 40-44.
- Umargono, E., Suseno, J. E., & Gunawan, S. V. (2020). K-means clustering optimization using the elbow method and early centroid determination based on mean and median formula. In *The 2nd International Seminar on Science and Technology (ISSTEC 2019)* (pp. 121-129). Atlantis Press. <https://doi.org/10.2991/assehr.k.201010.019>
- Uska, M., Wirasasmita, R., Usuluddin, U., & Arianti, B. (2020). Evaluation of Rapidminer-Application in Data Mining Learning using PeRSIVA Model. *Edumatic: Jurnal Pendidikan Informatika*, 4(2), 164-171. <https://doi.org/10.29408/edumatic.v4i2.2688>
- Yang, X., Wu, S., Xia, H., Li, Y., & Li, X. (2020). Campus Economic Analysis Based on K-Means Clustering and Hotspot Mining. *Review of Educational Theory*, 3(2), 42-50. <https://doi.org/10.30564/ret.v3i2.1837>
- Yuan, C., & Yang, H. (2019). Research on K-Value Selection Method of K-Means Clustering Algorithm. *J*, 2(2), 226-235. <https://doi.org/10.3390/j2020016>
- Zuhroh, I., Rofik, M., & Echchabi, A. (2021). Banking stock price movement and macroeconomic indicators: k-means clustering approach. *Cogent Business and Management*, 8(1), 1-10. <https://doi.org/10.1080/23311975.2021.1980247>