

KORELASI ANTARA PROFIL DAN NILAI AKADEMIS SISWA DENGAN MENGGUNAKAN ALGORITMA *K-MEANS*

Surohman¹, Luky Fabrianto*², Faizal Riza³, Novianti M Faizah⁴

¹²³ Program Pasca Sarjana STMIK Nusa Mandiri, Jakarta

⁴ Information Systems Department, Tama Jagakarsa University, Jakarta, Indonesia.

Email: ¹14002280@nusamandiri.ac.id, ²14002269@ nusamandiri.ac.id, ³14002266@ nusamandiri.ac.id

⁴novianti@jagakarsa.ac.id

*Penulis Korespondensi

(Naskah masuk: 08 Januari 2020, diterima untuk diterbitkan: 21 Juli 2021)

Abstrak

Hampir setiap pelajar di Indonesia terdaftar dengan atribut profil yang lengkap, seperti : Nama, Jenis Kelamin, Jenis Tinggal, Alat Transportasi, Usia Orangtua, Pendidikan Orangtua, Pekerjaan Orangtua, Penghasilan Orangtua dan atribut lainnya. Dari data atribut profil tersebut dapat diklasifikasi berdasarkan kedekatan nilai antara atribut yang dimiliki masing-masing siswa. Disisi lain siswa juga memiliki data yang berisi nilai akademis yang juga dapat dibuat klusterisasi.

Data yang dipakai dalam penelitian ini melibatkan 512 instances yang didapat dari sebuah Sekolah Menengah Kejuruan (SMK) di Jakarta. Metode yang pakai untuk klusterisasi menggunakan algoritma *K-Means*. Penelitian ini akan mencari korelasi klusterisasi profil siswa terhadap nilai akademisnya.

Tahapan penelitian diawali dengan persiapan dataset profil dan dataset nilai siswa, atribut dari dataset profil yang dipakai hanya atribut yang dianggap dapat merepresentasikan profil siswa dan keluarganya. Tahap berikutnya adalah mentransformasi data atribut non numerik (kategorik dan interval) menjadi numerik. Dilanjutkan dengan tahap perhitungan jarak antar data dan tahap terakhir mencari pola korelasi antara kluster profil dan kluster nilai akademis yang terbentuk.

Dengan metode *elbow* jumlah kluster yang paling ideal dalam penelitian ini adalah antara 3 dan 4 kluster, dimana nilai *Silhouette Coefficient* tertinggi adalah 0,8103 untuk pengelompokan 3 kluster.

Kata kunci: clustering, *k-Means*, metode elbow, *Silhouette Coefficient*, data mining

CORRELATION BETWEEN PROFILE AND ACADEMIC GRADES OF STUDENTS USING *K-MEANS* ALGORITHM

Abstract

Almost every student in Indonesia is registered with complete profile attributes, such as: Name, Gender, Type of Stay, Transportation Equipment, Parents' Age, Parental Education, Parents' Work, Parents' Earnings and other attributes. From the profile attribute data it can be clustered based on the closeness of the values between the attributes possessed by each student. On the other hand students also have attribute data that contains academic values that can also be clustered.

The data used in this study involved 512 instances obtained from a Vocational High School (SMK) in Jakarta. The method used for clustering is using the *K-Means* algorithm. This research will look for correlation of student profile clustering to its academic value.

The stages of the research began with the preparation of the profile dataset and the student value dataset, the attributes of the profile dataset used were only those attributes that were considered to represent the profiles of students and their families. The next step is to transform non-numeric attribute data (categorical and interval) into numeric. Followed by the stage of calculating the distance between data and the final stage looking for patterns of correlation between profile clusters and academic value clusters that are formed.

With the *elbow* method, the most ideal number of clusters in this study is between 3 and 4 clusters, where the highest *Silhouette Coefficient* value is 0.8103 for grouping 3 clusters.

Keywords: clustering, *k-Means*, metode elbow, *Silhouette Coefficient*, data mining

1. PENDAHULUAN

Di era digitalisasi saat ini, data profil siswa mencakup atribut yang lebih detail dan juga atribut dari keluarga siswa tersebut, seperti jenis tempat tinggal, alat transportasi, email, nomor telepon seluler, rekening bank, pendidikan orangtua, usia orangtua, penghasilan orangtua, Penerima KPS (Kartu Perlindungan Sosial) dan lain sebagainya.

Dalam penelitian ini penulis akan mencari bentuk atau pola korelasi antara profil pribadi siswa terhadap prestasi akademis yang diraih oleh siswa dengan membuat klasterisasi dari data yang ada menggunakan algoritma *K-Means*, tentunya dari hasil penelitian ini penulis berharap dapat mencari pola keterkaitan antara profil pribadi dan keluarga siswa terhadap nilai akademisnya, sehingga membantu para pendidik untuk mencari solusi yang berkaitan dengan peningkatan nilai akademis masing-masing siswa.

Penelitian terkait salahsatunya adalah pengelompokan siswa berdasarkan prestasi nilai akademik untuk mengetahui kemampuan siswa dalam menguasai mata pelajaran (Butarbutar et al., 2017), adapun (Listiyoko et al., 2017) bertujuan meningkatkan nilai rata-rata kelas dengan cara melakukan klasifikasi siswa dengan tujuan pendidik dapat melakukan evaluasi perkembangan siswa, serta ada pula klasterisasi siswa berprestasi dengan melihat nilai akademisnya (Sibuea & Safta, 2017). Secara umum algoritma yang dipakai dengan algoritma pada penelitian ini sama namun pendekatan yang dipakai berbeda tetapi sama-sama mengelompokkan siswa dari nilai akademis dengan aspek lainnya.

Penelitian ini menggunakan hasil dari klasterisasi dengan algoritma *K-means* yang diimplementasikan pada dua dataset (dataset profil dan dataset nilai), hasil dari klasterisasi yang didapat disesuaikan menurut karakteristik masing-masing klaster.

2. METODE PENELITIAN

2.1. Algoritma *K-Means*

Dalam penelitian ini digunakan Algoritma *K-Means* dimana Algoritma *K-Means* dapat membuat kelompok berdasarkan nilai atributnya yang dimasukkan kedalam sejumlah klaster *K* (Younus et al., 2015), flowchart algoritma *K-Means* dapat dilihat seperti pada gambar 1.

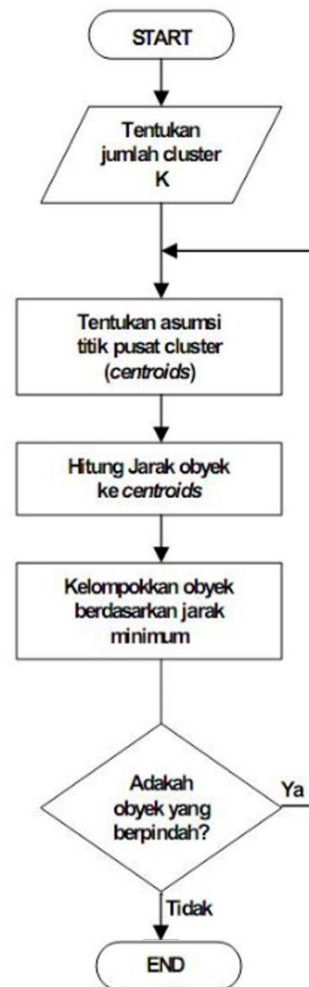
Rumus yang dipakai pada algoritma *K-Means* menggunakan rumus *Euclidean distance*, pada persamaan (1), d_{ij} adalah jarak nilai, x_{ik} merupakan nilai sebuah objek dan x_{jk} adalah nilai dari objek yang lain berikutnya.

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (1)$$

n = Dimensi Data

i = Posisi titik 1

j = Posisi titik 2



Gambar 1. Alur algoritma *K-Means* (Younus et al., 2015)

2.2. Pre-processing

Pada dataset terdapat dua buah jenis atribut yang harus ditransformasi terlebih dahulu, yaitu atribut yang memiliki data kategorik dan data interval, untuk data kategorik dilakukan proses *binning*, dimana nilai dari data non-numerik diubah menjadi data yang bisa dikalkulasi (Murti et al., 2005). Sedangkan untuk data interval, caranya dengan mencari nilai rata-ratanya.

Berikut beberapa contoh data yang dipakai dan transformasinya, data Jenis Tinggal ada 4 kategori, maka kategori pertama bersama orangtua diberi nilai 1000, bersama Wali nilai 0100, kost nilai 0010 dan lainnya 0001. Untuk Jenis Kelamin nilai 1 untuk laki-laki dan nilai 0 untuk perempuan. Sedangkan data interval, misal Penghasilan Orangtua Rp. 1.000.000,- s/d Rp. 2.000.000,- maka data tersebut dirata-ratakan dimana nilai terendah ditambah nilai tertinggi dibagi dua, dalam contoh ini menjadi Rp. 1.500.000,-.

2.3. Scoring (Metode Elbow)

Dalam menentukan nilai K yang akan dibentuk dalam algoritma K -Means yang paling ideal dapat dicari dengan metode *scoring*, atau metode *elbow*, yang mana merupakan suatu metode yang digunakan untuk menentukan nilai K yang akan dipakai ketika menggunakan algoritma K -Means, cara mendapatkan nilai K yang paling ideal adalah dengan cara melihat persentase hasil perbandingan antara jumlah kluster yang akan membentuk siku pada suatu titik, jika digambarkan dengan grafik, nilai sebuah kluster (n) dan nilai kluster berikutnya (m) memiliki titik kenaikan atau penurunan yang paling drastis maka nilai K untuk algoritma K -Means yang paling ideal adalah nilai antara nilai n dan m (Muningsih & Kiswati, 2018).

2.4. Silhouette Coefficient

Untuk mengetahui kualitas dari kluster yang terbentuk, atau seberapa baik suatu objek diletakkan dalam klusternya, dapat digunakan *Silhouette Coefficient*. Metode ini menggabungkan metode *cohesion* dan *separation* (Handoyo et al., 2014). Langkah-langkahnya sebagai berikut:

1. Menghitung rata-rata jarak dari suatu dokumen (i) terhadap semua dokumen lainnya dalam kluster yang sama

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \quad (2)$$

Dimana j adalah dokumen lain dalam kluster yang sama A , dan $d(i, j)$ adalah jarak antara i dan j .

2. Menghitung rata-rata jarak dari dokumen (i) tersebut dengan semua dokumen yang ada di kluster lain, lalu diambil nilai yang paling kecil diantaranya

$$d(i, C) = \frac{1}{|A|} \sum_{j \in C} d(i, j) \quad (3)$$

Dimana $d(i, C)$ adalah jarak rata-rata dokumen i dengan semua objek yang ada pada kluster lain C yang mana $A \neq C$.

$$b(i) = \min_{C \neq A} d(i, C) \quad (4)$$

3. Nilai Silhouette Coefficient adalah :

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (5)$$

3. HASIL DN PEMBAHASAN

Penelitian yang dilakukan melalui beberapa tahap, tahap pertama adalah mempersiapkan data yang akan diolah, kedua melakukan transformasi data yang berjenis non numerik (kategorik dan interval)

menjadi data numerik, ketiga melakukan klasterisasi terhadap data profil siswa dan klasterisasi terhadap data nilai siswa, selanjutnya menganalisa masing-masing kluster yang terbentuk dan terakhir menarik kesimpulan yang berupa korelasi kluster profil siswa terhadap kluster nilai siswa.

3.1. Persiapan Data

Data yang digunakan pada penelitian ini adalah data siswa Tahun Ajaran 2018-2019 dan data nilai Ujian Negara Tahun 2019 dari salahsatu SMK yang berada di Jakarta Utara, ada 512 siswa yang telah mengikuti Ujian Negara Tahun 2019 dan data profilnya cukup lengkap untuk di kalkulasi.

Selanjutnya mengunggah data Profil Siswa dan data Nilai Siswa ke *Google Colab* dan mengecek ada tidaknya *missing value* pada data yang akan dikalkulasi.

Data Profil Siswa

Nama	0
JK	0
Usia	0
Jenis Tinggal	0
Alat Transportasi	0
Usia Ayah	0
Pendidikan ayah	0
Pekerjaan ayah	0
Usia Ibu	0
Pendidikan ibu	0
Pekerjaan ibu	0
Penghasilan Ortu Rata2	0

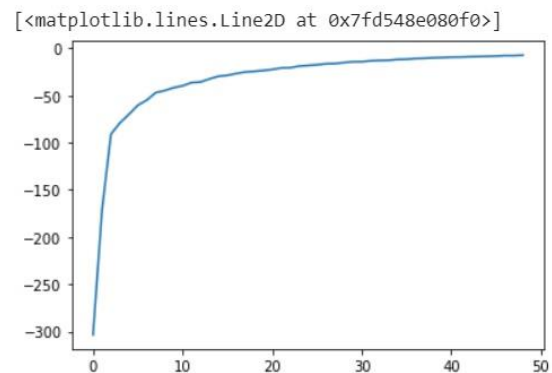
Data Nilai Siswa

Nama	0
BIN	0
ING	0
MAT	0
KMP	0

3.2. Mencari nilai K yang ideal (*Scoring*)

Untuk menentukan nilai K yang paling cocok untuk algoritma K -Means adalah dengan membuat *scoring*, nilai K digunakan untuk menentukan kluster yang akan dibentuk.

Didapat grafik seperti pada gambar 2.



Gambar 2. Grafik hasil *scoring*

Lakukan yang paling signifikan terletak di dekat antara sumbu $x=3$ dan $x=4$, sehingga nilai K yang tepat untuk algoritma K -Means yang akan dijalankan $K=3$ dan $K=4$.

3.3. Menghitung Silhouette Coefficient

Sebelum menjalankan algoritma K -Means dipastikan dahulu nilai K yang paling ideal untuk penelitian ini, karena pada penelitian ini lakukan yang paling signifikan berada diantara $K=3$ dan $K=4$.

Setelah dilakukan perhitungan nilai *Silhouette Coefficient* atau $s(i)$ didapat, seperti pada Tabel. 1 dibawah ini.

Jumlah Kluster	$s(i)$
4	0.789459349
3	0.803833866

Nilai K yang dihasilkan dari *scoring* data UNBK siswa dipakai juga pada klusterisasi data profil siswa, sehingga data profil dan data nilai memiliki jumlah kluster yang sama untuk melihat keterkaitannya.

3.4. Menjalankan Algoritma K -Means

Setelah mendapatkan nilai K maka kluster yang dibentuk sudah diketahui, dalam hal ini nilai $K=3$, selanjutnya menjalankan algoritma K -Means untuk masing-masing data (profil siswa dan nilai siswa).

Pada tahap awal penentuan nilai centroid (nilai tengah dalam sebuah kluster) diambil secara random sesuai dengan banyaknya kluster yang didapat dari hasil *scoring*, dalam hal ini 3 kluster. Setelah iterasi pertama dilakukan akan didapat nilai centroid yang baru yang akan menjadi nilai tengah dari tiap-tiap kluster, lalu mengukur jarak setiap objek terhadap nilai centroid dengan menggunakan rumus *Euclidean distance*, iterasi dilakukan hingga nilai centroid tidak mengalami perubahan.

Langkah selanjutnya menambahkan kolom kluster pada data profil siswa dan nilai siswa, maka dalam data profil siswa dan data nilai siswa dilengkapi dengan atribut baru yaitu Kluster.

Data Profil Siswa

Nama
JK
Usia
Jenis Tinggal
Alat Transportasi
Usia Ayah
Pendidikan ayah
Pekerjaan ayah
Usia Ibu

Pendidikan ibu
Pekerjaan ibu
Penghasilan Ortu min
Penghasilan Ortu Rata2
Kluster

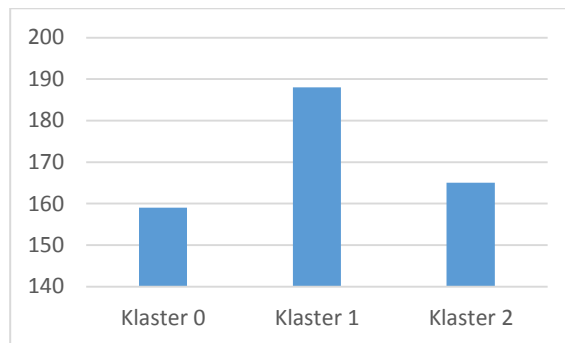
Data Nilai Siswa

Nama
BIN
ING
MAT
KMP
Kluster

Setelah masing-masing dataset (Profil dan nilai) mempunyai klusterisasi selanjutnya kita dapat menarik fakta-fakta yang terdapat dalam dataset pada penelitian ini.

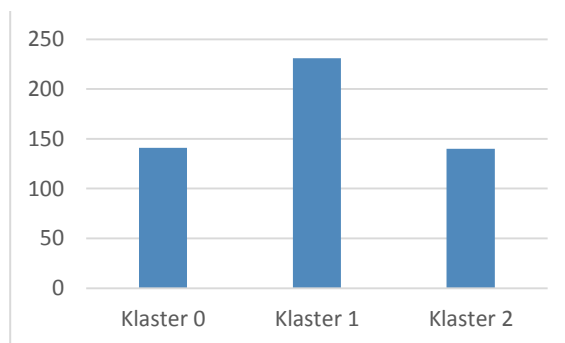
Histogram dari distribusi jumlah siswa tiap-tiap kluster ada pada gambar 3. untuk Profil Siswa dan gambar 4. untuk Nilai Siswa.

Pada klusterisasi profil siswa didapatkan jumlah siswa pada kluster 1 paling banyak, berjumlah 188 siswa, kluster 0 berjumlah 159 siswa dan terakhir kluster 2 berjumlah 165 siswa.



Gambar 3. Histogram distribusi kluster Profil Siswa

Klusterisasi nilai siswa terdiri dari jumlah siswa pada kluster 1 paling banyak, berjumlah 231 siswa, kluster 0 berjumlah 141 siswa dan terakhir kluster 2 berjumlah 140 siswa.



Gambar 4. Histogram distribusi kluster Nilai Siswa

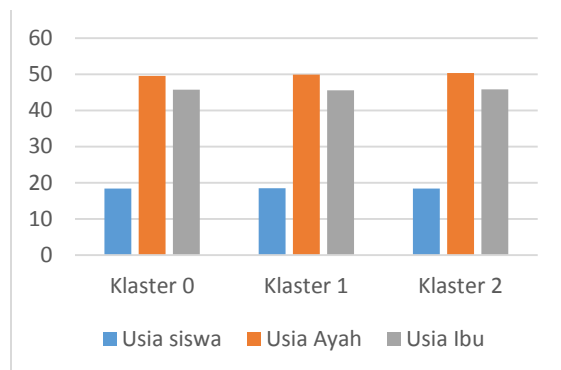
Dari klasterisasi Profil dan Nilai siswa bila dibuat matriks maka hasil klasterisasi dari ke dua dataset tersebut dapat dilihat seperti Tabel 2.

Tabel 2. Matriks Hasil Klasterisasi (Profil dan Nilai Siswa)

		Nilai		
		Klaster 0	Klaster 1	Klaster 2
Profil	Klaster 0	43	80	36
	Klaster 1	59	83	46
	Klaster 2	39	68	58

Dari Tabel 2. terdapat 184 dari 512 siswa (36%) yang bersesuaian letak klasternya, dimana seorang siswa yang ada di klaster 0 pada klasterisasi Nilai juga berada di klaster 0 pada klasterisasi Profil, dan seterusnya.

Dengan klasterisasi yang didapat kita bisa mendapatkan fakta tentang pengelompokan dari profil dan nilai siswa. Gambar 5. menggambarkan rata-rata usia siswa yang hampir sama pada setiap klaster, yaitu 18 tahun, demikian pula dengan usia ayah dan ibu, rata-rata usia ayah ditiap klaster adalah 50 tahun dan ibu 45 tahun.

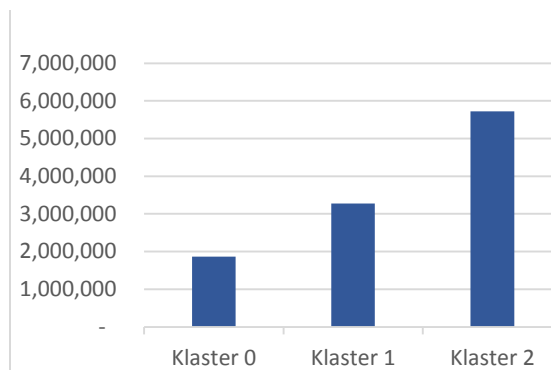


Gambar 5. Rata-rata usai siswa dan orangtuanya

Penghasilan per bulan orangtua yang dipakai dalam penelitian ini adalah penghasilan ayah ditambah dengan penghasilan ibu, penghasilan orangtua pada data ini menggunakan range seperti Rp. 1.000.000 - Rp. 1.999.999, Rp. 2.000.000 - Rp. 4.999.999, dan seterusnya, maka kami mengambil nilai tengah dari range tersebut. Namun penghasilan kedua orangtua tersebut masih dirata-ratakan lagi dengan penghasilan orangtua yang menjadi anggota masing-masing klaster. Didapatkan pada klaster 0 rata-rata penghasilan orangtuanya Rp. 1.864.000, klaster 1 Rp. 3.272.000 dan klaster 2 Rp. 5.724.000. Gambar 6. merupakan grafik dari rata-rata penghasilan orangtua masing-masing klaster.

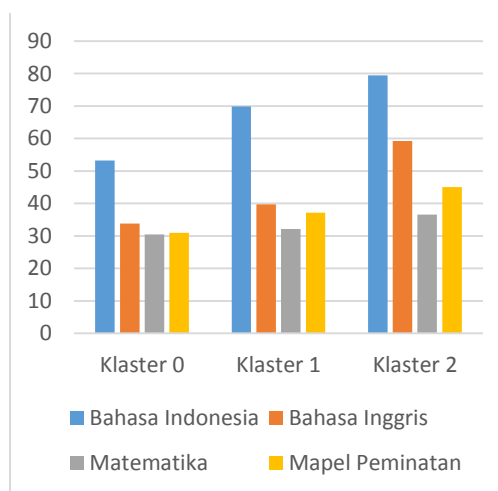
Penghasilan per bulan orangtua yang dipakai dalam penelitian ini adalah penghasilan ayah ditambah dengan penghasilan ibu, penghasilan orangtua pada data ini menggunakan range seperti Rp. 1.000.000 - Rp. 1.999.999, Rp. 2.000.000 - Rp.

4.999.999, dan seterusnya, maka kami mengambil nilai tengah dari range tersebut. Namun penghasilan kedua orangtua tersebut masih dirata-ratakan lagi dengan penghasilan orangtua yang menjadi anggota masing-masing klaster. Didapatkan pada klaster 0 rata-rata penghasilan orangtuanya Rp. 1.864.000, klaster 1 Rp. 3.272.000 dan klaster 2 Rp. 5.724.000. Gambar 6. merupakan grafik dari rata-rata penghasilan orangtua masing-masing klaster.



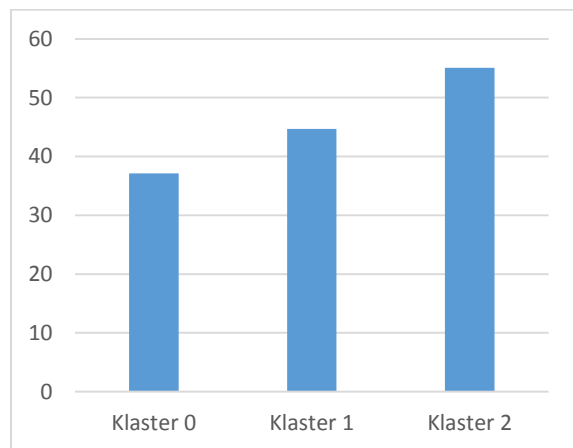
Gambar 6. Rata-rata penghasilan orangtua

Dari data nilai siswa dapat dilihat bahwa mata pelajaran Bahasa Indonesia menjadi nilai yang tertinggi untuk tiap-tiap klaster, disusul Bahasa Inggris. Rata-rata nilai Bahasa Indonesia pada klaster 0 adalah 53,2, sedangkan pada klaster 2 nilainya 79,2, rata-rata nilai ujian Bahasa Inggris klaster 2 adalah 59,2, diikuti dengan klaster 1. Gambar 7. dibawah menampilkan detail rata-rata nilai semua mata pelajaran Ujian Negara dari semua klaster yang terbentuk.



Gambar 7. Rata-rata nilai UNBK

Gambar 8. menunjukan bahwa klaster 2 memiliki Total rata-rata nilai UNBK tertinggi pada penelitian ini dengan nilai 55,1 diikuti klaster 1 dengan nilai 44,7 dan terakhir klaster 0 dengan nilai 37,1.



Gambar 8. Total rata-rata nilai UNBK

Pada dataset profil siswa terdapat *cell* kosong yang cukup banyak pada atribut Jenis Pekerjaan ayah maupun ibu oleh karena itu kami masukan pada kategori Jenis Pekerjaan Lainnya. Gambar 9 dan 10. Grafik distribusi Jenis Pekerjaan ayah dan ibu.

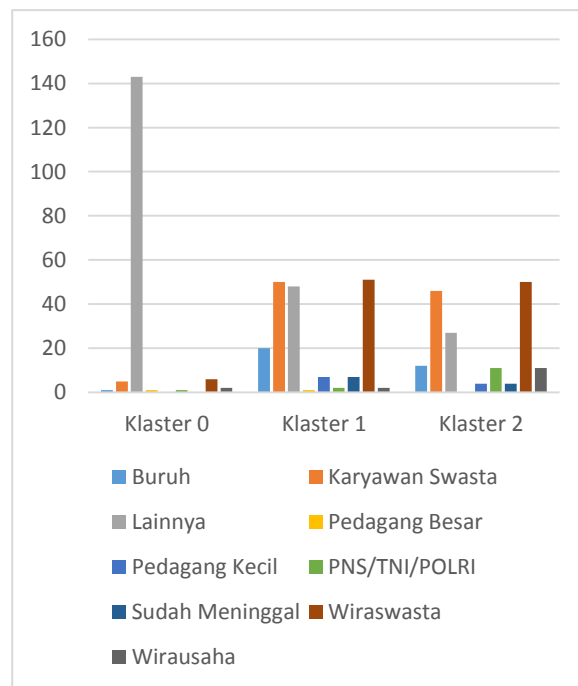
Dalam hal tingkat pendidikan ayah dan ibu juga terdapat banyak kolom yang dikosongkan sehingga kami masukan kedalam kategori tingkat pendidikan lainnya. Gambar 11. dan 12. merupakan tingkat pendidikan yang telah diklasterisasi, kluster 1 dan 2 pada tingkat pendidikan ayah dan ibu relative sama, dimana mayoritas pendidikan orangtua siswa yang berada dalam kluster 1 dan 2 mayoritas SMA/ sederajat. Mayoritas siswa tinggal bersama orangtuanya seperti gambar 13. dibawah, 495 siswa atau 97% dari seluruh siswa tinggal bersama orangtuanya sehingga tidak perlu diklasterisasi.

Alat transportasi siswa yang biasa digunakan kesekolah bisa dilihat pada gambar 14., pada kluster 2 semua siswa menggunakan sepeda motor untuk bersekolah, sedangkan pada kluster 1 kebanyakan siswa menggunakan angkutan umum.

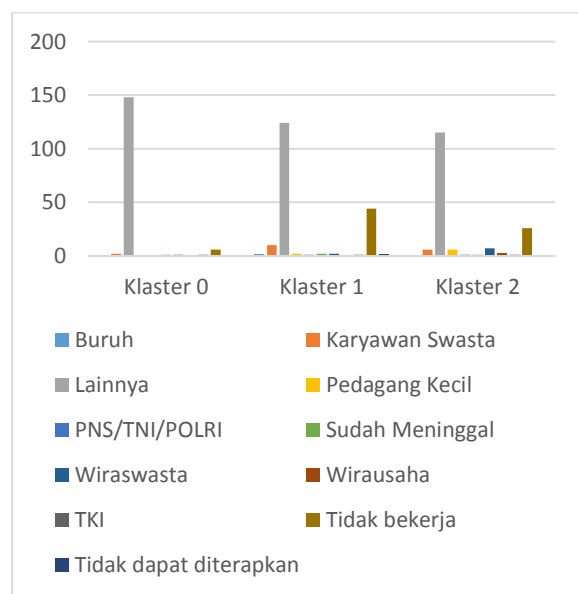
Dari beberapa gambar dan tabel diatas dapat ditarik kesimpulan sebagai berikut :

1. Kluster 0 terdiri dari keluarga dengan rata-rata kondisi ekonomi minimal dan jenjang pendidikan orangtua yang cukup rendah, hal tersebut juga berbanding lurus dengan nilai siswa, dimana nilai para siswa dalam kluster ini juga terendah dalam penelitian ini.
2. Kluster 1 menggambarkan rata-rata kondisi ekonomi yang cukup, pendidikan orangtua juga baik dan pekerjaan ayah kebanyakan berwiraswasta dan karyawan swasta.

3. Kluster 2 merupakan golongan siswa dengan mayoritas pekerjaan ayah adalah karyawan swasta dan wiraswasta, alat transportasi yang digunakan oleh siswa adalah sepeda motor serta kondisi ekonomi keluarga paling tinggi dan nilai ujian negaranya tertinggi dalam penelitian ini.



Gambar 9. Jenis Pekerjaan Ayah

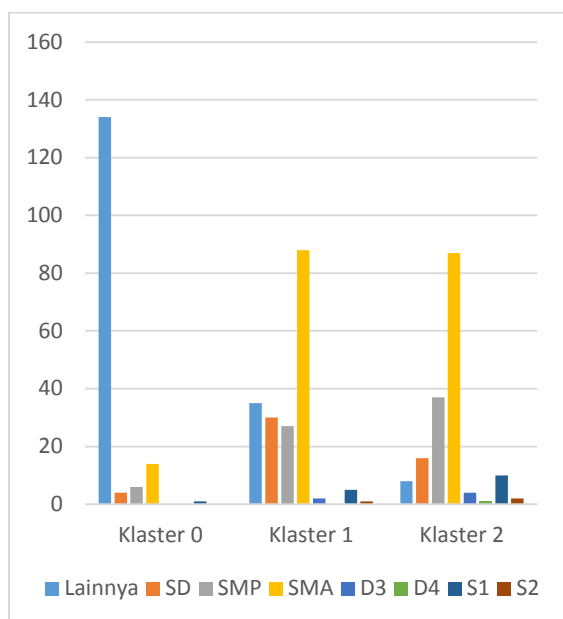


Gambar 10. Jenis Pekerjaan Ibu

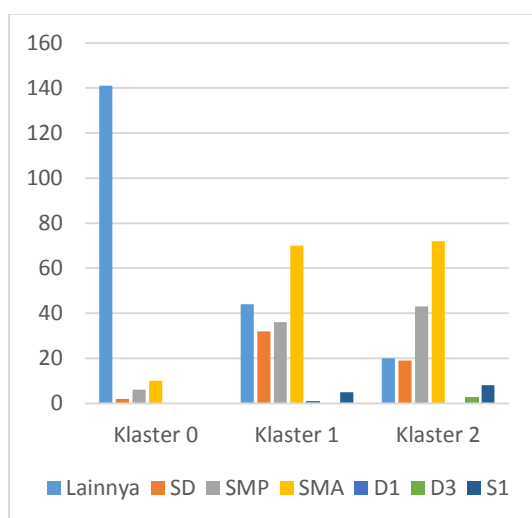
4. Pada penelitian ini menunjukkan kecenderungan semakin mapan kondisi ekonomi sebuah keluarga maka kemungkinan nilai akademis siswa semakin

baik, jika ada siswa dengan kondisi ekonomi yang mapan tapi nilai akademisnya rendah maka bisa disarankan untuk mengikuti Bimbingan Belajar sebagai tambahan, pada penelitian ini terdapat 39 siswa dengan kondisi demikian.

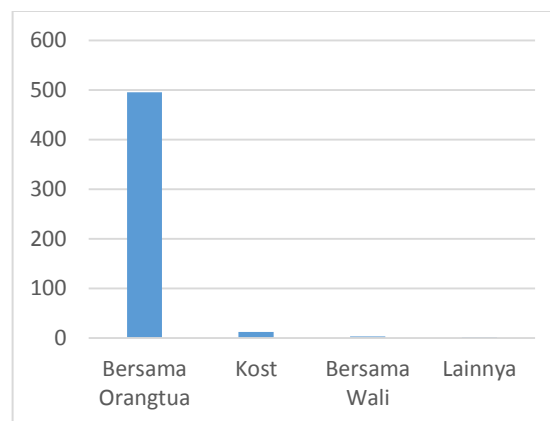
5. Dengan hasil dari penelitian ini juga bisa dijadikan pertimbangan pihak sekolah ketika akan melakukan pembagian kelas bisa memakai klasterisasi profil atau nilai akademis siswa sehingga prosentase siswa dalam setiap kelas sama porsinya.



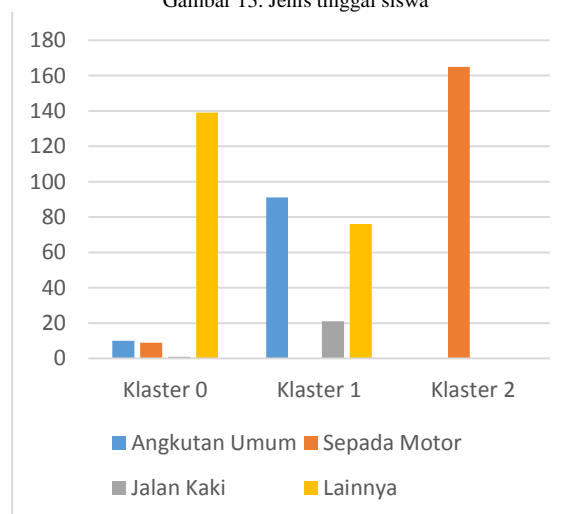
Gambar 11. Tingkat pendidikan ayah



Gambar 12. Tingkat pendidikan ibu



Gambar 13. Jenis tinggal siswa



Gambar 14. Alat transportasi siswa

4. KESIMPULAN DAN SARAN

Pengelompokan siswa berdasarkan aspek sosial dengan pendekatan klasterisasi *k-Means* dapat digunakan. Dalam penelitian ini ada 36% atau 184 dari total 512 siswa yang letak klaster profil dan nilai akademisnya bersesuaian. 43 siswa dengan nilai rendah dan ekonomi keluarga yang minim dan 83 siswa dengan nilai akademis yang tertinggi serta kondisi ekonomi keluarga yang berkecukupan.

Saran untuk penelitian selanjutnya dapat dilakukan dengan pendekatan algoritma atau metode lain, dapat juga menggunakan nilai raport untuk penelitian serupa, dan akan lebih baik jika orangtua mengisi data profil dengan lengkap dan sesuai pilihan yang ada.

DAFTAR PUSTAKA

- BUTARBUTAR, N., WINDARTO, A. P., HARTAMA, D., & SOLIKHUN, S. 2017. Komparasi Kinerja Algoritma Fuzzy C-Means Dan K-Means Dalam Pengelompokan Data

- Siswa Berdasarkan Prestasi Nilai Akademik Siswa. *Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika)*.
<https://doi.org/10.30645/Jurasik.V1i1.8>
- HANDOYO, R., RUMANI, R., & NASUTION, S. M. 2014. Perbandingan Metode Clustering Menggunakan Metode Single Linkage Dan K-Means Pada Pengelompokan Dokumen. *JSM STMIK Mikroskil*.
- LISTIYOKO, L., WATI, R., FAHRUDIN, A., & MUHAMMADIYAH BANTEN, S. 2017. Klasifikasi Siswa Untuk Meningkatkan Nilai Rata-Rata Kelas Menggunakan Metode Data Mining. *Seminar Nasional Teknologi Dan Rekayasa*.
- MUNINGSIH, E., & KISWATI, S. 2018. Sistem Aplikasi Berbasis Optimasi Metode Elbow Untuk Penentuan Clustering Pelanggan. *Joutica*. <https://doi.org/10.30736/Jti.V3i1.196>
- MURTI, D. H., SUCIATI, N., & NANJAYA, D. J. 2005. Clustering Data Non-Numerik Dengan Pendekatan Algoritma K-Means Dan Hamming Distance Studi Kasus Biro Jodoh. *JUTI: Jurnal Ilmiah Teknologi Informasi*.
<https://doi.org/10.12962/J24068535.V4i1.A245>
- SIBUEA, M. L., & SAFTA, A. 2017. Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering. *JURTEKSI*.
<https://doi.org/10.33330/Jurteksi.V4i1.28>
- YOUNUS, Z. S., MOHAMAD, D., SABA, T., ALKAWAZ, M. H., REHMAN, A., AL-RODHAAN, M., & AL-DHELAAN, A. 2015. Content-Based Image Retrieval Using Pso And K-Means Clustering Algorithm. *Arabian Journal Of Geosciences*.
<https://doi.org/10.1007/S12517-014-1584-7>