

Analisis Kualitas Pengelompokan Algoritma K-Means di Knime dan Excel untuk PTMT Pasca Vaksinasi Covid-19

Lili Kartikawati

SMK Negeri 4 Yogyakarta, Yogyakarta, Daerah Istimewa Yogyakarta, Indonesia
cintakufortuna@gmail.com

Abstrak: Pembelajaran tatap muka yang melibatkan interaksi diantara teman dan guru sangat sulit digantikan dengan pembelajaran jarak jauh. Vaksinasi pendidik dan tenaga kependidikan menjadi salah satu prioritas negara dalam upaya akselerasi pembelajaran tatap muka. Pendidik dan tenaga kependidikan di satuan pendidikan Kota Yogyakarta telah melaksanakan vaksinasi COVID-19 secara lengkap pada April 2021, sehingga SMK Negeri 4 Yogyakarta wajib menyediakan layanan pembelajaran tatap muka terbatas (PTMT) dan memfasilitasi pembelajaran jarak jauh. Pengelompokan yang berkualitas dibutuhkan untuk membatasi jumlah peserta didik yang mendapatkan layanan PTMT sesuai dengan standar protokol kesehatan. Proses pengelompokan sekumpulan data ke dalam klaster yang memiliki kemiripan ditentukan dengan perhitungan jarak. Penelitian ini melakukan pengelompokan peserta didik dengan aplikasi KNIME dan microsoft excel. Selanjutnya dilakukan analisis perhitungan kualitas pengelompokan dari kedua aplikasi untuk mendapatkan prototype pengelompokan terbaiknya. Perhitungan manual memanfaatkan rumus euclidean distance untuk menghitung jarak antar dokumen dan metode silhouette coefficient untuk menghitung perbedaan karakteristik (perbedaan jarak) diantara *cluster* dan persamaan karakteristik (kedekatan jarak) data-data didalam satu *cluster*. Hasil pengklasteran algoritma K-Means keluaran dari aplikasi KNIME memiliki kualitas 0,3208 (lemah) sedangkan manual excel memiliki kualitas 0,6331 (sedang), sehingga hasil manual excel direkomendasikan untuk membantu kurikulum sekolah mengelompokkan peserta didik yang akan diberikan layanan PTMT.

Kata kunci: algoritma K-Means; vaksinasi COVID-19; PTM Terbatas; Knime.

Quality Analysis of K-Means Algorithm Grouping in Knime and Excel for PTMT Post Covid-19 Vaccination

Abstrac: Face-to-face learning that involves interaction between friends and teachers is very difficult to replace with distance learning. Vaccination of educators and education staff is one of the country's priorities in accelerating face-to-face learning. Educators and education personnel in the Yogyakarta city education unit have implemented a complete COVID-19 vaccination in April 2021, so that SMK Negeri 4 Yogyakarta is required to provide limited face-to-face learning services (PTMT) and facilitate distance learning. Quality grouping is needed to limit the number of students who receive PTMT services. The process of grouping a set of data into clusters that have similarities is determined by calculating the distance. This study grouped students with the KNIME application and Microsoft Excel. Furthermore, the analysis of the calculation of the quality of the grouping of the two applications is carried out to get the best prototype. The manual calculation uses the Euclidean distance formula to calculate the distance between documents and the silhouette coefficient method to calculate the characteristic difference (distance difference) between clusters and the characteristic equation (closeness of distance) of the data in one cluster. The results of the K-Means algorithm clustering output from the KNIME application have a quality of 0.3208 (weak) while the excel manual has a quality of 0.6331 (medium), so the results of the excel manual are recommended to help the school curriculum classify PTMT.

Keywords: K-Means algorithm; vaccination COVID-19; PTM Limited; Knime.

1. Pendahuluan

Surat Keputusan Bersama Menteri Pendidikan dan Kebudayaan, Menteri Agama, Menteri Kesehatan, serta Menteri Dalam Negeri pada Selasa 30 Maret 2021 telah menetapkan dua

agenda besar yaitu Vaksinasi pendidik dan tenaga kependidikan serta kebijakan pembelajaran di masa pandemi COVID-19. Tantangan terbesar pembelajaran di masa pandemi COVID-19 diutarakan Mendikbud adalah kenyataan bahwa

pembelajaran tatap muka yang melibatkan interaksi antar teman dan guru sangat sulit digantikan dengan pembelajaran jarak jauh.

Vaksinasi pendidik dan tenaga kependidikan menjadi salah satu prioritas negara dalam upaya akselerasi pembelajaran tatap muka. (Kementerian Pendidikan, Kebudayaan, 2020). Setelah pendidik dan tenaga kependidikan di satuan pendidikan divaksinasi COVID-19 secara lengkap, pemerintah pusat, pemerintah daerah, kanwil atau kantor Kemenag mewajibkan satuan pendidikan untuk menyediakan layanan pembelajaran tatap muka terbatas (PTMT) dengan tetap menerapkan protokol kesehatan dan pembelajaran jarak jauh.

Vaksinasi COVID-19 bagi seluruh pendidik dan tenaga kependidikan di Kota Yogyakarta telah dilakukan bulan April 2021. Pelaksanaan vaksinasi COVID-19 yang telah berlangsung tersebut menjadikan SMK Negeri 4 Yogyakarta sebagai bagian dari sekolah di bawah naungan Pemerintah Daerah D.I. Yogyakarta berkewajiban untuk menyelenggarakan PTMT. Dua hal penting yang harus dipahami ketika akan membukalayanan PTMT di SMK Negeri 4 Yogyakarta adalah kondisi kelas jaga jarak minimal 1,5 meter dan maksimal 18 peserta didik per kelas, jumlah hari dan jam PTMT. Sarana yang dimiliki sebanyak 36 ruang kelas teori dan 18 ruang bengkel praktik terpadu, untuk mengakomodir 54 rombongan belajar dimana dalam satu rombel terdiri dari 36 peserta didik.

Melihat kondisi sarana dan jumlah peserta didik yang dimiliki sekolah, maka sangat penting bagi sekolah untuk merumuskan strategi yang harus dilakukan agar pelaksanaan PTMT tepat sasaran dan berkualitas. Pemberlakuan PTMT pada seluruh peserta didik dan mata pelajaran sangat tidak memungkinkan, mengingat besarnya jumlah peserta didik dibandingkan sarana, pendidik, alokasi waktu dan persyaratan operasional pembelajaran di masa pandemi COVID-19 yang tidak memenuhi syarat. Pertimbangan tersebut mengerucutkan kebijakan sekolah dalam penyelenggaraan PTMT dengan tidak mengikut sertakan seluruh mata pelajaran yang ada dan tidak mengikut sertakan seluruh peserta didik yang dimiliki.

Kesenjangan belajar dengan PJJ tidak dialami oleh seluruh peserta didik dalam seluruh materi kompetensi matapelajaran yang diselenggarakan tingkat pendidikan. Sekolah harus mampu menambang data untuk mendapatkan informasi peserta didik mana saja yang terkendala pembelajaran jarak jauh (PJJ) dan mata pelajaran apa saja yang belum mencapai target ketuntasan belajar dengan PJJ.

Penelitian ini mengelompokkan peserta didik dengan aplikasi KNIME dan microsoft excel. Selanjutnya dilakukan analisis perhitungan kualitas pengelompokkan dari kedua aplikasi untuk mendapatkan prototype pengelompokkan terbaiknya. Perhitungan manual memanfaatkan rumus *euclidean distance* untuk menghitung jarak antar dokumen dan metode *silhouette coefficient* untuk menghitung perbedaan karakteristik (perbedaan jarak) diantara *cluster* dan persamaan karakteristi (kedekatan jarak) data-data didalam satu *cluster* (Hidayati et al., 2021). Penelitian terdahulu untuk menyelesaikan permasalahan pengelompokkan data dengan menerapkan Teknik *K-Means Clustering* diantaranya: Metode *K-Means Clustering* untuk memilih media promosi Penerimaan Mahasiswa Baru pada Politeknik Negeri Tanah Laut oleh Julianto, Permadi, Noviyanti (2017) dengan judul "*Aplikasi Pemilihan Strategi Promosi Penerimaan Mahasiswa Baru Politeknik Negeri Tanah Laut Menggunakan Metode K-Means Clustering*" (Julianto & Permadi, 2017), Gambaran efektivitas pembelajaran E-learning pada masa pandemi COVID-19 dengan memanfaatkan metode K-Means oleh Lesmana, Akbari, Rahman, Gustian (2020) dengan judul "*Penerapan K-Means dalam Efektivitas Pembelajaran E-Learning pada Masa Pandemi COVID-19*" (Lesmana et al., 2020), Data responden yang dikumpulkan dilakukan pengelompokkan menjadi 3 *cluster* "rendah", "sedang" dan "tinggi" sebagai tingkatan kepuasan pembelajaran online pada masa pandemi COVID-19 oleh Sianipar, Siahaan, Siregar, Zer, Hartama (2020) dengan judul "*Penerapan Algoritma K-Mean dalam Menentukan Tingkat Kepuasan Pembelajaran Online pada Masa Pandemi COVID-19*" (R Sianipar et al., 2020), Data tindak kejahatan pencurian yang ada di Polres Asahan dilakukan *clustering* dengan menggunakan algoritma K-Means menjadi 3 kelompok yaitu tinggi, menengah dan rendah menggunakan aplikasi berbasis web yang dikembangkan oleh Damanik, Irianto, Dahriansah (2021) dengan judul "*Penerapan Metode Clustering dengan Algoritma K-Means Tindak Kejahatan Pencurian di Kabupaten Asahan*" (Afni Syahpitri Damanik et al., 2021), Clustering kinerja dosen berdasarkan kualitas pelaksanaan Tridharma (Pendidikan, pengabdian kepada masyarakat dan kegiatan penunjang lainnya) untuk mengontrol keseimbangan kinerja dosen dilakukan oleh Sugiharti, Muslim (2016) berjudul "*On-Line Clustering of Lecturers Performance of Computer Science Department of Semarang State University Using K-Means Algorithm*" (Sugiharti & Muslim,

2016), Memetakan kelompok belajar siswa berdasarkan nilai rapor pada semester I sampai dengan IV dalam 3 Cluster yaitu sangat berprestasi, berprestasi dan kurang berprestasi menggunakan metode Clustering K-Means dilakukan oleh Dewi, Defit, Yunus (2021) berjudul “Akurasi Pemetaan Kelompok Belajar Siswa Menuju Prestasi Menggunakan Metode K-Means (Studi Kasus SMP Pembangunan Laboratorium UNP)” (Dewi et al., 2020).

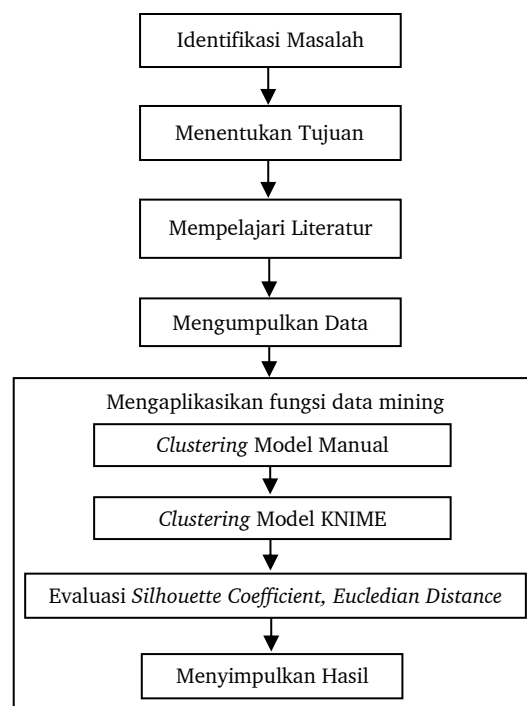
Penelitian diatas sebagian besar memanfaatkan angket, respondent, jejak pendapat untuk dijadikan sumber data pengelompokkan dengan memanfaatkan algoritma *K-Means* sedangkan pada penelitian ini, sumber data adalah data fisik yang bersifat jelas yaitu nilai rapor tengah semester genap 2020/2021 peserta didik X Usaha Perjalanan Wisata 1 di SMK Negeri 4 Yogyakarta untuk dilakukan pengelompokkan menjadi 2 *cluster* tipe pembelajaran yaitu PJJ dan PTMT. Pengelompokkan peserta didik dengan aplikasi KNIME dan microsoft excel, dilanjutkan dengan melakukan analisis perhitungan kualitas pengelompokkan dari kedua aplikasi untuk mendapatkan *prototype* pengelompokkan terbaiknya. Perhitungan manual memanfaatkan rumus *euclidean distance* untuk menghitung jarak antar dokumen dan metode *silhouette coefficient* untuk menghitung perbedaan karakteristik (perbedaan jarak) diantara *cluster* dan persamaan karakteristik (kedekatan jarak) data-data didalam satu *cluster*. Uraian diatas merupakan tindakan penelitian ini yang berjudul “Analisis Kualitas Pengelompokkan Algoritma *K-Means* di *Knime* dan *Excel* Untuk PTMT Pasca Vaksinasi Covid-19”. Tujuan dari penelitian ini diantaranya; 1) Melakukan pengelompokkan (PJJ dan PTMT) peserta didik menggunakan algoritma *K-Means* secara otomatis dengan *Knime* dan secara manual dengan *Excel*, 2) Menghitung kualitas pengelompokkan dengan metode *silhouette coefficient* dari pengelompokkan secara otomatis oleh KNIME dan pengelompokkan secara manual oleh *excel*, 3) Mendapatkan salah satu *prototype* pengelompokkan dengan kualitas terbaik diantara keduanya sehingga dapat mejadi rekomendasi pengelompokkan yang efektif dan efisien bagi kurikulum sekolah sesuai dengan kebutuhan dan sumber daya yang dimiliki.

2. Metode Penelitian

Metode eksperimen menjadi pilihan untuk diterapkan pada penelitian ini. Penggunaan teknik ini mempunyai tujuan mencari dan menemukan sendiri permasalahan yang dihadapi

dengan mengadakan percobaan (Santoso, 2009). Penelitian ini akan melakukan percobaan untuk menguji kualitas pengelompokkan dari hasil pengelompokkan data raport tengah semester genap 2020/2021 kelas X Usaha Perjalanan Wisata 1 di SMK Negeri 4 Yogyakarta. Susunan kerangka kerja penelitian akan membantu kelancaran pelaksanaan yang akan dilakukan. Tahapan-tahapan yang dilakukan dalam penelitian ini diawali dari identifikasi masalah, menentukan tujuan, mempelajari literatur, pengumpulan data, pengolahan data dengan algoritma *K-Means Clustering* metode *K-Means* di aplikasi KNIME, dievaluasi kualitas *clusteringnya* dan disimpulkan hasilnya seperti *flowchart* pada gambar 1.

Percobaan akan melakukan pengelompokkan model pembelajaran yang akan diterapkan dengan menggunakan algoritma *K-Means* pada aplikasi *Knime* dan *microsoft excel*. Hasil pengelompokkan dilakukan pengujian kualitas menggunakan metode *silhouette coefficient*. Metode *silhouette coefficient* melakukan pengujian kualitas pengelompokkan dengan menghitung jarak antar dokumen dengan menggunakan rumus *euclidean distance*. Hasil percobaan akan mampu menjawab permasalahan untuk mendapatkan *prototype* pengelompokkan yang berkualitas sehingga sehingga pengelompokkan peserta didik yang efektif dan efisien dapat disusun oleh kurikulum SMK Negeri 4 Yogyakarta sesuai dengan kebutuhan dan sumber daya yang dimiliki.



Gambar 1. Kerangka Kerja Penelitian
Analysis Techniques

Solusi dan penyelesaian masalah diperoleh melalui proses pemaparan gambaran masalah, analisis masalah dan memahami masalah. Permasalahan pelaksanaan PTMT di SMK Negeri 4 Yogyakarta adalah keterbatasan sumber daya dibandingkan dengan jumlah peserta didik dengan persyaratan kondisi kelas jaga jarak minimal 1,5 meter dan maksimal 18 peserta didik per kelas, jumlah hari dan jam PTMT. Sarana yang dimiliki sebanyak 36 ruang kelas teori dan 18 ruang bengkel praktik terpadu untuk mengakomodir 54 rombongan belajar dimana dalam satu rombongan terdiri dari 36 peserta didik.

Capaian penelitian yang menjadi tujuan adalah menentukan *cluster* tipe pembelajaran peserta didik (PJJ, PTMT) berdasarkan sumber data raport tengah semester peserta didik kelas X Usaha Perjalanan Wisata di SMK Negeri 4 Yogyakarta dan mengevaluasi kualitas *clusteringnya* dengan menggunakan metode *silhouette coefficient* dan rumus *euclidean distance* untuk menghitung jarak.

Literatur Clustering dengan Algoritma K-Means, Aplikasi KNIME dan Evaluasi kualitas clustering dengan Silhouette Coefficient

Tahapan *clustering* data menggunakan algoritma K-Means adalah; Pertama, menentukan jumlah *cluster k* yang dipengaruhi oleh kebutuhan sistem. Kebutuhan jumlah *cluster* tipe pembelajaran adalah 2 yaitu PJJ dan PTMT. Kedua, Menentukan nilai titik tengah (*centroid*) sesuai dengan jumlah *cluster* yang telah ditentukan.

$$C_i = \text{Min} + \frac{(i-1) \cdot (\text{max} - \text{min})}{n} + \frac{(\text{max} - \text{min})}{2 \cdot n} \quad (\text{Persamaan 1})$$

dimana :

C_i : *Centroid* dari kelas i
Min : Nilai terkecil dari data kelas kontinyu
Max : Nilai terbesar dari data kelas diskret
 n : Jumlah kelas diskret

Ketiga, alokasikan masing-masing data ke pusat *cluster* terdekat dengan rumus:

$$Be = \frac{(\sqrt{(O_i - T_i)^2})}{(O_j - T_j)^2} \quad (\text{Persamaan 2})$$

dimana :

Be : Jarak data ke pusat *cluster*
 O : Data *record*
 T : Data *centroid*

Keempat, ulangi kembali untuk menentukan pusat *cluster (centroid)* baru dengan rumus:

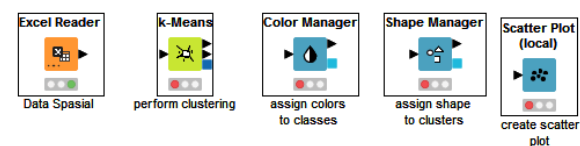
$$C_i = \frac{O_i + \dots + O_n}{\Sigma O} \quad (\text{Persamaan 3})$$

dimana :

C_i : *Centroid* dari kelas i
 O_1 : Nilai data *record* ke 1
 O_n : Nilai data *record* ke n
 ΣO : Jumlah data *record*

Kelima, pusat *cluster* yang baru dilakukan implementasi pengalokasian. Jika pusat *cluster* masih mengalami perubahan maka lakukan proses iterasi langkah ke 3 sampai dengan proses pencarian dihentikan.

Aplikasi KNIME memiliki *platform integrasi* terbaik untuk kebutuhan laporan dan analisis data. Software open source untuk data *analytics*, *reporting*, dan *integration platform* yang bisa menyatukan berbagai komponen dalam machine learning dan *data mining* dengan konsep diagram data. Keuntungan yang diperoleh dengan menggunakan software KNIME adalah memodelkan tiap-tiap tahapan pada *machine learning* dengan berbagai bahasa pemrograman. *Node* yang akan digunakan pada *workflow clustering* yang memanfaatkan KNIME diantaranya adalah *excel reader*, *kmeans*, *color manager*, *shape manager* dan *scatter plot (local)* terlihat pada gambar 2.



Gambar 2. Kebutuhan Node *Clustering*

Tahapan evaluasi kualitas *clustering* dengan *Silhouette Coefficient* dilakukan dengan tahapan sebagai berikut; Pertama, hitung rata-rata jarak objek dengan semua objek lain yang berbeda di dalam satu *cluster*.

$$d(i, C) = \frac{1}{|A|} \sum_{j \in A, j \neq i} d(i, j) \quad (\text{Persamaan 4})$$

Dimana :

$a(i)$: Rata-rata jarak objek i dengan semua objek lain dalam satu *cluster*
 A : Jumlah objek dalam satu *cluster*
 $d(i, j)$: Jarak objek i terhadap j

Kedua, hitung rata-rata jarak objek dengan semua objek lain yang berada pada *cluster* lain, kemudian ambil nilai paling minimum

$$d(i, C) = \frac{1}{|A|} \sum_{j \in C, j \neq i} d(i, j) \quad (\text{Persamaan 5})$$

Dimana :

$d(i, C)$ atau $b(i)$: Rata-rata jarak objek i dengan semua objek lain dalam *cluster* lainnya
 A : Jumlah objek
 $d(i, j)$: Jarak objek i terhadap j
 C : *Cluster* lain
 $b(i) = d(i, C)$

Ketiga, hitung nilai *silhouette coefficient*

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (\text{Persamaan 6})$$

Dimana :

- a(i) : Rata-rata jarak objek i dengan semua objek lain dalam *cluster* yang sama
b(i) : Rata-rata jarak objek i dengan semua objek lain dalam *cluster* lainnya

Nilai hasil *silhouette coefficient* berada dikisaran nilai -1 sampai dengan 1. Nilai mendekati 1 diartikan memiliki pengelompokan data dalam suatu *cluster* yang lebih baik dari pada yang mendekati -1 sebagai pengelompokan yang buruk dalam suatu *cluster* [9].

Penelitian ini menggunakan sumber data capaian tengah semester dari populasi kelas X Usaha Perjalanan Wisata di SMK Negeri 4 Yogyakarta. Sampel berupa nilai dan sampling adalah data nilai pelajaran. Data didapatkan dengan melakukan download dari sistem informasi akademik (SIA) SMK Negeri 4 Yogyakarta menggunakan hak akses wali kelas.

Dalam mengaplikasikan Fungsi Data Mining, diperoleh pola-pola dari sumber data dan hasil *clustering* yang dilakukan dengan menggunakan aplikasi KNIME dan microsoft excel, untuk dapat diidentifikasi hubungan dari kejadian yang terjadi sehingga didapatkan pengetahuan baru yang mendukung rekomendasi pengelompokan PTMT yang berkualitas. Metode *clustering* yang digunakan adalah algoritma *K-Means* dan teknik evaluasi dilakukan menggunakan metode *silhouette coefficient* dan rumus *euclidean distance* untuk menghitung jarak.

Pada tahap kesimpulan, dilakukan perbandingan hasil evaluasi kualitas *Clustering* Algoritma *K-Means* dari dua pengaplikasian *pengclusteran* yang telah dilakukan, yaitu di aplikasi microsoft excel dengan aplikasi KNIME. Nilai *silhouette coefficient* yang paling mendekati angka 1 menjadi rekomendasi model pengelompokan PTMT kepada kurikulum SMK Negeri 4 Yogyakarta.

3. Hasil dan Pembahasan

Penelitian ini telah melakukan kegiatan yaitu preprosesing dan transformasi data, proses pengelompokan dan pengujian kualitas pengelompokan. Sumber data berupa nilai raport tengah semester 2020/2021 kelas X Usaha Perjalanan Wisata 1 yang diunduh dari portal sistem informasi akademik (SIA) SMK Negeri 4 Yogyakarta, dilakukan pembersihan data untuk menghilangkan kolom yang tidak digunakan dalam proses *clustering*. Penelitian *pengclusteran* yang sama dengan menggunakan nilai telah

dilakukan oleh Sri Dewi (Dewi et al., 2020) yaitu menggunakan nilai rapor siswa yang didapat pada Semester I sampai dengan Semester IV dan penelitian yang telah dilakukan oleh Mardalius (Mardalius, 2018) yaitu menggunakan nilai Ujian Nasional SMA jurusan IPA yang meliputi data nilai bahasa indonesia, bahasa inggris, matematika, fisika, biologi, dan kimia. Pada penelitian ini menggunakan sampel peserta didik sejumlah 35 dikarenakan 1 peserta didik nomor presensi 19 pindah sekolah dan nilai 16 Mata Pelajaran yang telah dilakukan proses rerata dari dua nilai pengetahuan dan keterampilan sehingga didapatkan nilai akhir yang disajikan pada tabel 1. Pada penelitian yang dilakukan oleh Sri Dewi (Dewi et al., 2020) dan Mardalius (Mardalius, 2018) baru melakukan *pengclusteran* saja belum menganalisis kualitas *pengclusterannya*, sedangkan pada penelitian ini telah mengimplementasikan pengelompokan dengan menggunakan aplikasi analisis data otomatis Knime dan aplikasi analisis data manual excel serta melakukan analisis kualitas *pengclusterannya* untuk mendapatkan salah satu *prototype* pengelompokan yang terbaik diantara keduanya.

Nilai hasil *silhouette coefficient* berada dikisaran nilai -1 sampai dengan 1. Nilai mendekati 1 diartikan memiliki pengelompokan data dalam suatu *cluster* yang lebih baik dari pada yang mendekati -1 sebagai pengelompokan yang buruk dalam suatu *cluster* (Rousseuw, 1990). Pada hasil pengukuran kualitas didapatkan nilai *Si* dari pengelompokan di Knime sebesar 0,3208 dan *Si* dari pengelompokan manual pada excel sebesar 0,6331. Pengelompokan manual pada excel sebagai pengelompokan yang memiliki kualitas lebih baik dari pada pengelompokan dengan Knime. Tahapan pengelompokan pada kedua *prototype* yang dilakukan pada penelitian ini dipaparkan pada tabel 1.

Proses Clustering Algoritma K-Mean Secara Manual di Excel

Clustering manual pada aplikasi excel ditentukan ada 2 kelompok yaitu *online* dan PTMT. Untuk mempermudah dalam pengolahan maka kelompok *online* digitalisasi kedalam angka 0 dan kelompok PTMT didigitalisasi kedalam angka 1. Tahapan yang dilakukan mulai dari penentuan jumlah *cluster*, menentukan titik *centroid* dengan menggunakan rumus persamaan 1 dan menghitung jarak terdekat menggunakan rumus dipersamaan 2 sehingga didapatkan daftar jarak yang dipetakan dalam tipe pembelajaran pada mata pelajaran pendidikan agama ditampilkan pada tabel 2.

Tabel 1. Data Nilai Raport Tengah Semester X Usaha Perjalanan Wisata 1

| Inisial Nama Siswa | Muatan Nasional | | | Muatan Kewilayahan | | | C1 Dasar Bidang Keahlian | | | C2 Dasar Program Keahlian | |
|--------------------------|------------------|-----|------------|--------------------|-----|----------|--------------------------|-----|-------|---------------------------|-----------------|
| | Pendidikan Agama | .. | B.Ingggris | Seni Budaya | .. | Sikomdig | IPA Terapan | .. | KIP | .. | Bahasa Mandarin |
| | 1 | ... | 6 | 7 | ... | 9 | 10 | ... | 12 | ... | 16 |
| A.I | 81,65 | ... | 84,50 | 82,96 | ... | 88,00 | 85,00 | ... | 84,14 | ... | 87,50 |
| A.N.K. | 80,74 | ... | 82,00 | 87,09 | ... | 83,50 | 82,50 | ... | 80,34 | ... | 85,00 |
| A.N | 84,73 | ... | 89,50 | 85,68 | ... | 87,50 | 83,50 | ... | 86,89 | ... | 90,00 |
| A.R.B. | 83,08 | ... | 88,50 | 85,56 | ... | 87,00 | 85,00 | ... | 85,72 | ... | 83,00 |
| A.E.S. | 82,13 | ... | 83,00 | 93,56 | ... | 85,00 | 85,50 | ... | 82,78 | ... | 86,00 |
| A.H.N. | 82,92 | ... | 88,00 | 89,48 | ... | 87,00 | 78,57 | ... | 83,64 | ... | 87,50 |
| B.O.V. | 79,70 | ... | 83,00 | 85,13 | ... | 87,00 | 83,00 | ... | 80,70 | ... | 87,00 |
| C.A.F. | 83,95 | ... | 83,00 | 90,33 | ... | 85,00 | 83,00 | ... | 84,89 | ... | 85,50 |
| C.S.P. | 81,69 | ... | 89,50 | 95,08 | ... | 87,50 | 86,25 | ... | 83,97 | ... | 86,50 |
| C.V.D. | 82,19 | ... | 85,00 | 93,87 | ... | 83,00 | 86,25 | ... | 81,33 | ... | 86,50 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| T.R. | 79,33 | ... | 86,50 | 92,37 | ... | 86,00 | 84,50 | ... | 80,70 | ... | 83,50 |
| V.R.F. | 82,87 | ... | 88,00 | 92,48 | ... | 89,00 | 87,65 | ... | 84,47 | ... | 90,00 |
| YU | 82,02 | ... | 84,00 | 86,45 | ... | 87,50 | 87,07 | ... | 85,14 | ... | 86,50 |
| Y.A | 82,52 | ... | 86,00 | 95,45 | ... | 83,00 | 85,50 | ... | 81,81 | ... | 87,50 |
| Z.R.D. | 77,40 | ... | 79,50 | 79,89 | ... | 79,50 | 77,50 | ... | 75,00 | ... | 81,50 |

Delta = 0,10
Kelas (n) = PTMT/Online
Max = 87,10
Min = 75,00
n = 2
Toleransi Error = 1,21
C1 = 78,03
C2 = 84,075

Tabel 2. Daftar Jarak antar peserta didik Cluster Tipe Pembelajaran di Mata Pelajaran PAI

| Siswa | Muatan Nasional | | | | | | |
|--------|-----------------------------------|-------|------|-------|-------|------|---|
| | Pendidikan Agama dan Budi Pekerti | | | | | | |
| | 78,03 | 84,08 | I | 79,63 | 82,80 | II | |
| A.I | 81,65 | 3,63 | 2,42 | 0 | 2,02 | 1,15 | 0 |
| A.N.K. | 80,74 | 2,71 | 3,33 | 1 | 1,11 | 2,06 | 1 |
| A.N | 84,73 | 6,70 | 0,65 | 0 | 5,09 | 1,92 | 0 |
| A.R.B. | 83,08 | 5,05 | 1,00 | 0 | 3,44 | 0,27 | 0 |
| A.E.S. | 82,13 | 4,10 | 1,95 | 0 | 2,49 | 0,68 | 0 |
| A.H.N. | 82,92 | 4,89 | 1,16 | 0 | 3,28 | 0,11 | 0 |
| B.O.V. | 79,70 | 1,68 | 4,37 | 1 | 0,07 | 3,10 | 1 |
| C.A.F. | 83,95 | 5,93 | 0,12 | 0 | 4,32 | 1,15 | 0 |
| C.S.P. | 81,69 | 3,66 | 2,38 | 0 | 2,06 | 1,11 | 0 |
| C.V.D. | 82,19 | 4,16 | 1,88 | 0 | 2,56 | 0,61 | 0 |
| D.K.D. | 80,18 | 2,15 | 3,90 | 1 | 0,54 | 2,63 | 1 |
| D.E.A. | 75,00 | 3,03 | 9,07 | 1 | 4,63 | 7,80 | 1 |

| Siswa | Muatan Nasional | | | | | | |
|---------|-----------------------------------|-------|------|-------|-------|------|---|
| | Pendidikan Agama dan Budi Pekerti | | | | | | |
| | 78,03 | 84,08 | I | 79,63 | 82,80 | II | |
| D.N.R. | 82,02 | 3,99 | 2,06 | 0 | 2,38 | 0,79 | 0 |
| H.I.P. | 78,98 | 0,95 | 5,10 | 1 | 0,66 | 3,83 | 1 |
| I.S.Q. | 81,78 | 3,75 | 2,30 | 0 | 2,14 | 1,03 | 0 |
| M.M.R. | 80,19 | 2,16 | 3,88 | 1 | 0,56 | 2,61 | 1 |
| M.F.M. | 82,72 | 4,69 | 1,36 | 0 | 3,08 | 0,09 | 0 |
| M.S.D. | 80,20 | 2,18 | 3,87 | 1 | 0,57 | 2,60 | 1 |
| N. R.W. | 80,73 | 2,70 | 3,35 | 1 | 1,09 | 2,08 | 1 |
| N. A.T | 82,93 | 4,90 | 1,15 | 0 | 3,29 | 0,12 | 0 |
| N.S.A | 79,80 | 1,77 | 4,27 | 1 | 0,17 | 3,00 | 1 |
| N.I.R. | 81,94 | 3,91 | 2,13 | 0 | 2,31 | 0,86 | 0 |
| N.N.A. | 83,80 | 5,77 | 0,27 | 0 | 4,17 | 1,00 | 0 |
| P.P.N. | 81,28 | 3,25 | 2,80 | 0 | 1,64 | 1,53 | 0 |
| Q.K.S. | 80,85 | 2,82 | 3,22 | 1 | 1,22 | 1,95 | 1 |
| R.R.R. | 79,90 | 1,88 | 4,17 | 1 | 0,27 | 2,90 | 1 |
| R.D.U. | 79,74 | 1,71 | 4,33 | 1 | 0,11 | 3,06 | 1 |
| R.S. | 80,95 | 2,93 | 3,12 | 1 | 1,32 | 1,85 | 1 |
| S.R.D. | 87,10 | 9,07 | 3,03 | 0 | 7,47 | 4,30 | 0 |
| T.S.A. | 80,44 | 2,41 | 3,63 | 1 | 0,81 | 2,36 | 1 |
| T.R. | 79,33 | 1,30 | 4,75 | 1 | 0,31 | 3,48 | 1 |
| V.R.F. | 82,87 | 4,84 | 1,21 | 0 | 3,23 | 0,06 | 0 |
| Y.I. | 82,02 | 3,99 | 2,06 | 0 | 2,38 | 0,79 | 0 |
| Y.A | 82,52 | 4,49 | 1,56 | 0 | 2,88 | 0,29 | 0 |
| Z.R.D. | 77,40 | 0,63 | 6,67 | 1 | 2,23 | 5,40 | 1 |

Pada tabel 2 terlihat bahwa terjadi iterasi kedua dikarenakan rerata-centroid yang dihasilkan pada iterasi pertama lebih besar dari nilai toleransi error (1,21). Centroid baru ditentukan berdasarkan rumus persamaan 3. Iterasi ke 2 telah memenuhi sehingga berhenti di iterasi ke 2 dengan detail perhitungan rerata-centroid terlihat pada tabel 3.

Tabel 3. Rerata, Centroid, Rerata Centroid pada Iterasi 1 & 2 Clustering Tipe Pembelajaran

| Kelas | ITERASI 1 | | | ITERASI 2 | | |
|--|-----------|----------|-------------------|------------------------------|----------|-------------------|
| | Rerata | Centroid | Rerata - Centroid | Rerata | Centroid | Rerata - Centroid |
| 1 (PTMT) | 79,63 | 78,03 | 1,61 | 79,63 | 79,63 | 0,00 |
| 0 (ONLINE) | 82,80 | 84,08 | 1,27 | 82,80 | 82,80 | 0,00 |
| TOTAL | | | 2,88 | TOTAL | | 0,00 |
| Oleh karena TOLERANSI error = 1,21 maka 2,88 > 1,21 sehingga lakukan ITERASI Kedua | | | | Hasil Iterasi 2 sdh memenuhi | | |

Proses seperti di atas diberlakukan kepada seluruh mata pelajaran yang berjumlah 16, kemudian hasil tersebut dilakukan perhitungan *silhouette coefficient* yang memanfaatkan persamaan 4 sampai dengan persamaan 6 pada aplikasi excel yang hasilnya terlihat pada tabel 4.

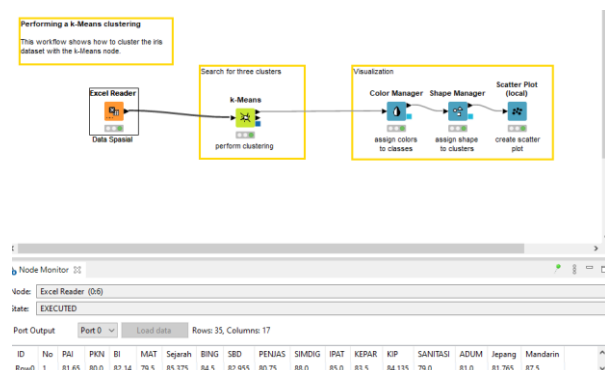
Tabel 4. Hasil Euclidian Distance dan Silhouette Coefficient- Clustering Manual Excel

| Nama Siswa | d(i,centroid_1) | d(i,centroid_0) | Label Manual | a(i) | | b(i) | S(i) |
|------------|-----------------|-----------------|--------------|--------|--------|------|------|
| | | | | d(i,0) | d(i,1) | | |
| A.I | 3,33 | 2,64 | 0 | 0,482 | 1,03 | 0,53 | |
| A.N.K. | 2,10 | 3,89 | 1 | 1,508 | 0,32 | 0,79 | |
| A.N | 4,66 | 2,45 | 0 | 0,446 | 2,36 | 0,81 | |
| A.R.B. | 3,20 | 2,87 | 0 | 0,591 | 0,9 | 0,34 | |
| A.E.S. | 3,72 | 3,38 | 0 | 1,000 | 1,42 | 0,29 | |
| A.H.N. | 3,40 | 3,00 | 0 | 0,676 | 1,1 | 0,38 | |
| B.O.V. | 2,40 | 3,45 | 1 | 1,067 | 0,31 | 0,71 | |
| C.A.F. | 4,32 | 1,88 | 0 | 0,618 | 2,02 | 0,69 | |
| C.S.P. | 5,19 | 1,88 | 0 | 0,620 | 2,89 | 0,79 | |
| C.V.D. | 3,58 | 2,41 | 0 | 0,446 | 1,28 | 0,65 | |
| D.K.D. | 3,67 | 2,54 | 0 | 0,458 | 1,37 | 0,67 | |
| D.E.A. | 1,86 | 6,40 | 1 | 4,020 | 0,44 | 0,89 | |
| D.N.R. | 5,11 | 1,58 | 0 | 0,816 | 2,81 | 0,71 | |

| Nama Siswa | d(i,centroid_1) | d(i,centroid_0) | Label Manual | a(i) | | b(i) | S(i) |
|-------------------------------------|-----------------|-----------------|--------------|--------|--------|--------|------|
| | | | | d(i,0) | d(i,1) | | |
| H.I.P. | 2,52 | 3,19 | 1 | 0,833 | 0,35 | 0,58 | |
| I.S.Q. | 4,48 | 2,81 | 0 | 0,555 | 2,18 | 0,75 | |
| M.M.R. | 2,61 | 3,10 | 1 | 0,757 | 0,4 | 0,47 | |
| M.F.M. | 4,57 | 1,50 | 0 | 0,885 | 2,27 | 0,61 | |
| M.S.D. | 3,89 | 2,71 | 0 | 0,506 | 1,59 | 0,68 | |
| N. | 4,80 | 2,35 | 0 | 0,453 | 2,5 | 0,82 | |
| R.W. | 1,90 | 4,88 | 1 | 2,494 | 0,41 | 0,84 | |
| N.S.A | 2,31 | 6,82 | 1 | 4,437 | 0,3 | 0,93 | |
| N.I.R. | 4,34 | 2,06 | 0 | 0,543 | 2,04 | 0,73 | |
| N.N.A. | 5,51 | 1,57 | 0 | 0,824 | 3,21 | 0,74 | |
| P.P.N. | 3,26 | 2,63 | 0 | 0,477 | 0,96 | 0,50 | |
| Q.K.S. | 3,39 | 2,42 | 0 | 0,445 | 1,09 | 0,59 | |
| R.R.R. | 2,23 | 3,59 | 1 | 1,202 | 0,3 | 0,75 | |
| R.D.U. | 3,32 | 3,28 | 0 | 0,907 | 1,02 | 0,11 | |
| R.S. | 1,93 | 4,22 | 1 | 1,839 | 0,39 | 0,79 | |
| S.R.D. | 4,83 | 1,95 | 0 | 0,587 | 2,53 | 0,77 | |
| T.S.A. | 2,66 | 3,41 | 1 | 1,026 | 0,44 | 0,57 | |
| T.R. | 3,14 | 3,01 | 0 | 0,686 | 0,84 | 0,18 | |
| V.R.F. | 6,30 | 1,55 | 0 | 0,837 | 4 | 0,79 | |
| Y.I. | 3,10 | 3,14 | 1 | 0,790 | 0,8 | 0,01 | |
| Y.A | 4,20 | 2,36 | 0 | 0,452 | 1,9 | 0,76 | |
| Z.R.D. | 1,98 | 6,96 | 1 | 4,574 | 0,36 | 0,92 | |
| Nilai Rerata Silhouette Coefficient | | | | | | 0,6331 | |

Proses Clustering Algoritma K-Mean Menggunakan Aplikasi KNIME

Proses implementasi clustering yang dilakukan pada KNIME dengan memanfaatkan *node-node* yang tersusun pada *workflow*, selanjutnya data nilai raport peserta didik selama tengah semester yang berbentuk excel dimasukan kedalam *node excel reader* dan tekan tombol *run* pada aplikasi KNIME sehingga tampilan pada aplikasi KNIME dapat dilihat pada gambar 3.



Gambar 3. Workflow dan Proses Clustering di KNIME

Pada gambar di atas, *node excel reader* menggunakan data siswa siswi yang ada, lalu menggunakan *node k-means* dari *knime* yang sudah tersedia, dilanjutkan adanya *visualization*

yaitu ada 3 *node color manager*, *shape manager* dan juga *scatter plot*. Hasil *clustering* yang dilakukan oleh *KNIME* dapat dilihat *visualisainya* pada gambar 4.

| Row ID | D PENJAS | D SIMDIG | D IPAT | D KEPAR | D KIP | D SANITASI | D ADUM | D Jepang | D Mandarin | S Cluster |
|--------|----------|----------|--------|---------|--------|------------|--------|----------|------------|-----------|
| Row0 | 80.75 | 88 | 85 | 83.5 | 84.135 | 79 | 81 | 81.765 | 87.5 | cluster_1 |
| Row1 | 80 | 83.5 | 82.5 | 75.5 | 80.335 | 75 | 81 | 81.19 | 85 | cluster_1 |
| Row2 | 82.75 | 87.5 | 83.5 | 83.165 | 86.885 | 87 | 79 | 80.07 | 90 | cluster_0 |
| Row3 | 80 | 87 | 85 | 78.5 | 85.72 | 75 | 82.5 | 81.405 | 83 | cluster_1 |
| Row4 | 80 | 85 | 85.5 | 77.165 | 82.78 | 75 | 82 | 89.5 | 86 | cluster_1 |
| Row5 | 80 | 87 | 78.57 | 88.335 | 83.64 | 77 | 81 | 77.835 | 87.5 | cluster_1 |
| Row6 | 80 | 87 | 83 | 80.665 | 80.695 | 75 | 82 | 78.645 | 87 | cluster_1 |
| Row7 | 80 | 85 | 83 | 81.165 | 84.885 | 89 | 82 | 80.595 | 85.5 | cluster_0 |
| Row8 | 80 | 87.5 | 86.25 | 86.165 | 83.97 | 83.5 | 86 | 77.12 | 86.5 | cluster_0 |
| Row9 | 82.5 | 83 | 86.25 | 80 | 81.33 | 75 | 82 | 79.19 | 86.5 | cluster_1 |
| Row10 | 82.5 | 85.5 | 86 | 79.835 | 81.805 | 75 | 82 | 77.5 | 86.5 | cluster_1 |
| Row11 | 80 | 83 | 76.785 | 76 | 81.805 | 75 | 79 | 75 | 82.5 | cluster_1 |
| Row12 | 83.25 | 88 | 87.645 | 78.75 | 82.805 | 89 | 81.5 | 81.98 | 86.5 | cluster_0 |
| Row13 | 80 | 85 | 82.5 | 82.665 | 80.97 | 79 | 80 | 79.02 | 86 | cluster_1 |
| Row14 | 81.25 | 87.5 | 86.715 | 76.5 | 81.805 | 81.5 | 81.5 | 86.905 | 88.5 | cluster_0 |
| Row15 | 80 | 85 | 81.5 | 78.375 | 83.14 | 75 | 81 | 78.765 | 85.5 | cluster_1 |
| Row16 | 82.5 | 88.5 | 85.285 | 79 | 82.47 | 89 | 80.5 | 80.355 | 86 | cluster_0 |
| Row17 | 82.5 | 85 | 85.43 | 76.5 | 82.805 | 75 | 82.5 | 82.5 | 86.5 | cluster_1 |
| Row18 | 80 | 87 | 85.855 | 81.835 | 87.97 | 92 | 82 | 85.645 | 84.5 | cluster_0 |
| Row19 | 80.75 | 82.5 | 80.75 | 75.835 | 76.28 | 75 | 79 | 76.335 | 83.5 | cluster_1 |
| Row20 | 80 | 81.5 | 78 | 82.415 | 75.165 | 75 | 79 | 64.02 | 80 | cluster_1 |
| Row21 | 82 | 86 | 81.93 | 82.585 | 84.47 | 81.5 | 82.5 | 86 | 88.5 | cluster_0 |
| Row22 | 84.5 | 86.5 | 84.215 | 82.75 | 85.64 | 89 | 82 | 85.69 | 89 | cluster_0 |
| Row23 | 80.75 | 88 | 83.5 | 78.75 | 82.305 | 75 | 82.5 | 78.715 | 87.5 | cluster_1 |

Gambar 4. Visualisasi *Pengclustering* di *KNIME*

Hasil *pengclustering* yang telah dilakukan *KNIME* selanjutnya dilakukan perhitungan nilai *silhouette coefficient* di aplikasi *excel* untuk melihat kualitas hasil *clustering* sehingga hasilnya terlihat pada tabel 5.

Tabel 5. Hasil *Euclidian Distance* dan *Silhouette Coefficient* - *clustering* di *Knime*

| Nama Siswa | $d(i, \text{centroid}_1)$ | $d(i, \text{centroid}_0)$ | Label Manual | $a(i)$ | | $S(i)$ |
|-------------------------------------|---------------------------|---------------------------|--------------|----------|----------|--------|
| | | | | $d(i,0)$ | $d(i,1)$ | |
| A.I | 1,41 | 1,60 | 1 | 0,75 | 0,63 | 0,16 |
| A.N.K. | 2,59 | 0,84 | 0 | 1,08 | 1,35 | 0,20 |
| A.N | 2,21 | 2,46 | 1 | 1,03 | 0,53 | 0,48 |
| A.R.B. | 1,02 | 2,12 | 1 | 0,86 | 0,39 | 0,55 |
| A.E.S. | 1,55 | 1,47 | 0 | 0,77 | 0,73 | 0,04 |
| A.H.N. | 1,93 | 1,93 | 0 | 0,79 | 0,42 | 0,47 |
| B.O.V. | 2,35 | 1,51 | 0 | 0,76 | 0,7 | 0,08 |
| C.A.F. | 1,82 | 2,05 | 1 | 0,84 | 0,39 | 0,54 |
| C.S.P. | 1,52 | 3,41 | 1 | 1,62 | 1,23 | 0,24 |
| C.V.D. | 1,81 | 1,79 | 0 | 0,76 | 0,5 | 0,33 |
| D.K.D. | 1,27 | 1,98 | 1 | 0,81 | 0,4 | 0,50 |
| D.E.A. | 5,15 | 3,28 | 0 | 1,53 | 1,13 | 0,26 |
| D.N.R. | 1,58 | 2,82 | 1 | 1,22 | 0,77 | 0,37 |
| H.I.P. | 2,38 | 0,50 | 0 | 1,38 | 1,69 | 0,19 |
| I.S.Q. | 1,42 | 2,14 | 1 | 0,87 | 0,39 | 0,56 |
| M.M.R. | 1,75 | 0,81 | 0 | 1,10 | 1,38 | 0,20 |
| M.F.M. | 0,84 | 1,89 | 1 | 0,78 | 0,44 | 0,44 |
| M.S.D. | 0,81 | 1,48 | 1 | 0,76 | 0,73 | 0,05 |
| N.R.W. | 1,76 | 2,21 | 1 | 0,90 | 0,41 | 0,55 |
| N. A.T | 4,31 | 2,27 | 0 | 0,93 | 0,44 | 0,53 |
| N.S.A | 5,90 | 3,86 | 0 | 1,98 | 1,67 | 0,16 |
| N.I.R. | 1,06 | 1,89 | 1 | 0,78 | 0,44 | 0,44 |
| N.N.A. | 1,01 | 1,92 | 1 | 0,78 | 0,42 | 0,46 |
| P.P.N. | 1,04 | 1,35 | 1 | 0,80 | 0,83 | 0,04 |
| Q.K.S. | 1,78 | 1,25 | 0 | 0,83 | 0,94 | 0,12 |
| R.R.R. | 1,98 | 1,26 | 0 | 0,83 | 0,93 | 0,11 |
| R.D.U. | 3,31 | 2,76 | 0 | 1,19 | 0,73 | 0,38 |
| R.S. | 3,58 | 1,67 | 0 | 0,75 | 0,59 | 0,22 |
| S.R.D. | 1,05 | 2,15 | 1 | 0,88 | 0,39 | 0,56 |
| T.S.A. | 2,46 | 1,11 | 0 | 0,90 | 1,07 | 0,16 |
| T.R. | 2,16 | 1,87 | 0 | 0,77 | 0,45 | 0,41 |
| V.R.F. | 1,47 | 3,51 | 1 | 1,70 | 1,32 | 0,22 |
| Y.I. | 1,21 | 2,08 | 1 | 0,85 | 0,39 | 0,55 |
| Y.A | 1,73 | 2,37 | 1 | 0,98 | 0,48 | 0,51 |
| Z.R.D. | 5,77 | 3,72 | 0 | 1,86 | 1,54 | 0,18 |
| Nilai Rerata Silhouette Coefficient | | | | | | 0,3208 |

4. Simpulan dan Saran

Permasalahan tentang *clustering* yang diselesaikan pada penelitian ini adalah menemukan model yang tepat untuk melakukan pengelompokan PTMT sehingga layak untuk direkomendasikan kepada kurikulum SMK Negeri 4 Yogyakarta. Hasil evaluasi yang dilakukan pada kedua model yang telah dilakukan dengan perbandingan terlihat pada tabel 6.

Tabel 6. Perbandingan *Centroid* dan *Silhouette Coefficient* antara Manual dan KNIME

| Cluster | Manual di Excel | | KNIME | |
|--------------------|-----------------|-----------|---------------|-----------|
| | Cluster 0 | Cluster 1 | Cluster 0 | Cluster 1 |
| PAI | 82,80 | 79,63 | 83,03 | 80,59 |
| PKN | 83,50 | 80,50 | 81,05 | 80,48 |
| BI | 82,60 | 80,55 | 82,37 | 81,76 |
| MAT | 84,88 | 79,63 | 83,27 | 81,96 |
| Sejarah | 86,25 | 79,25 | 87,34 | 84,49 |
| BING | 87,38 | 82,13 | 86,55 | 84,67 |
| SBD | 92,25 | 84,74 | 91,83 | 88,69 |
| PENJAS | 82,93 | 80,44 | 81,66 | 80,60 |
| SIMDIG | 86,63 | 81,88 | 87,36 | 85,02 |
| IPAT | 84,93 | 79,50 | 85,23 | 82,35 |
| KEPAR | 83,74 | 77,67 | 81,11 | 79,12 |
| KIP | 84,73 | 78,24 | 84,50 | 81,09 |
| Sanitasi | 89,22 | 76,60 | 87,50 | 76,33 |
| ADUM | 84,25 | 80,75 | 82,23 | 81,10 |
| Jepang | 81,42 | 73,71 | 82,69 | 78,91 |
| Mandarin | 87,02 | 83,05 | 87,36 | 85,04 |
| Rerata S(i) | 0,6331 | | 0,3208 | |
| Kualitas | Sedang | | Lemah | |

Nilai kualitas dari keluaran *clustering* dengan KNIME diakibatkan oleh karena KNIME dalam menentukan *centroid* secara random sedangkan pada manual excel dilakukan secara detail dengan memasukan nilai tertinggi dan terendah pada rumus baku penentuan *centroid* algoritma *K-Means*. Dari hasil perhitungan kualitas menggunakan metode *silhouette coefficient* pada kedua model tersebut, maka didapatkan model yang berkualitas untuk direkomendasikan kepada kurikulum SMK

Negeri 4 Yogyakarta adalah model penentuan *clustering* secara manual di excel yang menghasilkan nilai *S(i)* sebesar 0,6331 yang memiliki kualitas sedang.

Penelitian yang telah dilakukan ini memiliki potensi untuk dikembangkan sehingga dari model yang telah direkomendasikan untuk pengelompokan tersebut dapat diimplementasikan. Hasil *clustering* selanjutnya perlu dilakukan pembacaan pola-pola yang terbentuk agar mampu menghasilkan pengetahuan baru yang lainnya sehingga dapat melakukan otomatisasi penjadwalan pada perencanaan PTMT sesuai dengan sumber daya yang dimiliki dan persyaratan pelaksanaan PTMT pasca vaksinasi COVID-19. Untuk dapat memberikan penjadwalan PTMT secara otomatis maka dibutuhkan: 1) *Clustering* diberlakukan kepada seluruh kelas yang ada di SMK Negeri 4 Yogyakarta; 2) Menghubungkan keterkaitan pola label pengelompokan peserta didik dengan mata pelajaran yang terbentuk; 3) Mendata durasi waktu yang akan diberlakukan pada PTMT Pasca Vaksinasi COVID-19; 4) Menghitung secara matematis sehingga didapatkan durasi waktu tiap mata pelajaran dan rumpun mata pelajaran yang akan diberlakukan PTMT Pasca Vaksinasi COVID-19.

Daftar Pustaka

- Afni Syahpitri Damanik, N., Prodi Sistem Informasi, M., Royal, S., Sistem Informasi, P., & Manajemen Informasi, P. (2021). Penerapan metode clustering dengan algoritma k-means tindak kejahatan pencurian di kabupaten asahan. *Journal of Computer (Online)*, 1(1), 7–14. <https://jurnal.stmikroyal.ac.id/index.php/j-com/article/view/1065>
- Dewi, S., Defit, S., & Yunus, Y. (2020). Akurasi Pemetaan Kelompok Belajar Siswa Menuju Prestasi Menggunakan Metode K-Means (Studi Kasus SMP Pembangunan Laboratorium UNP). *Jurnal Sistim Informasi Dan Teknologi*. <https://doi.org/10.37034/jsisfotek.v3i1.98>
- Hidayati, R., Zubair, A., Pratama, A. H., & Indana, L. (2021). Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means Clustering. *Techno.Com*, 20(2), 186–197. <https://doi.org/10.33633/tc.v20i2.4556>
- Julianto, V., & Permadi, J. (2017). Aplikasi pemilihan strategi promosi penerimaan mahasiswa baru politeknik negara tanah laut menggunakan metode k-means clustering. *Jurnal Ilmiah Informatika*, 2(1), 99–104.

- <https://doi.org/10.5281/JIMI.V1i2.37>
Kementerian Pendidikan, dan Kebudayaan.
(2020). *Penyesuaian kebijakan pembelajaran di masa pandemi COVID-19*.
- Lesmana, S., Fauzia Akbari, A., Yulia Rahman, E., Gustian, D., Informasi Jl Raya Cibatun, S., Kaler, C., Cisaat, K., Regency, S., & Barat, J. (2020). Penerapan k-means dalam efektivitas pembelajaran e-learning pada masa pandemi covid-19. *Seminar Nasional Informatika*, 2020 (1), 100–110. <http://www.jurnal.upnyk.ac.id/index.php/semnasif/article/view/4090>
- Mardalius, M. (2018). *Implementasi Algoritma K-Means Clustering Untuk Menentukan Kelas Kelompok Bimbingan Belajar Tambahan (Studi Kasus : Siswa SMA Negeri 1 Ranah Pesisir)*.
<https://doi.org/10.31219/osf.io/6mec3>
- R Sianipar, K. D., Wanti Siahaan, S., Siregar, M., Fikrul Ilmi Zer, P. R., Studi Teknik Informatika, P., Tunas Bangsa Pematangsiantar Jl Jend Sudirman Blok, S. A., Pematangsiantar, K., & Utara, S. (2020). Penerapan algoritma k-means dalam menentukan tingkat kepuasan pembelajaran online pada masa pandemi covid-19. *Jurnal Teknologi Informasi*, 4(1), 101–105. <http://jurnal.una.ac.id/index.php/jurti/article/view/1258>
- Rousseuw, L. K. and P. J. (1990). *Finding Groups in Data* (J. W. & Sons (ed.)).
- Sugiharti, E., & Muslim, A. (2016). On-line clustering of lecturers performance of computer science department of semarang state university using k-means algorithm. *Journal of Theoretical and Applied Information Technology*, 10(1). www.jatit.org