

# Penerapan Teknik Clustering Data Mining untuk Memprediksi Kesesuaian Jurusan Siswa (Studi Kasus SMA PGRI 1 Subang)

Tubagus Riko Rivanthio<sup>1</sup>, Mardhiya Ramdhani<sup>2</sup>, Ahmad Sahi<sup>3</sup>

<sup>1,2,3</sup>Politeknik LP3I Bandung

---

**Keywords:**

*clustering  
data mining  
suitability  
majors  
students*

**ABSTRACT**

SMA PGRI 1 Subang is a private school that has several missions, one of which is the establishment of academic and non-academic achievements. In an effort to achieve the mission, the supervision of student achievement must be conducted. The effort is to provide understanding in the selection of majors in accordance with the interests and talents of students. But in the activity of providing understanding, the school does not yet have a model that can evaluate the interests and talents of students to choose majors. The model can be obtained using student data processing. Data processing can be done using data mining, namely data mining clustering techniques. The technique will produce a model in the selection of majors. This clustering process is the process of grouping similar data based on the similarity of data held by students. The research method used is the CRISP-DM method which has 6 stages consisting of: Business Understanding, Data Understanding, Data Processing, Modeling, Evaluation, and Dissemination. The data that is processed is 620 data consisting of class of students in 2014, 2015, 2016. The results of processing using clustering obtained 6 clusters that have different models for each cluster. The results of this study can be used by schools in recommending courses chosen by students according to students' interests and talents, so students can learn optimally.

---

**Corresponding Author:**

Tubagus Riko Rivanthio,  
Politeknik LP3I Bandung,

Jl. Pahlawan No.59, Sukaluyu, Kec. Cibeunying Kaler, Kota Bandung, Jawa Barat 40123.

Email: [riko\\_rivanthio@yahoo.com](mailto:riko_rivanthio@yahoo.com)

---

**1. PENDAHULUAN**

SMA PGRI 1 Subang merupakan salah satu sekolah swasta yang ada di Kab. Subang yang dinaungi oleh Yayasan PGRI dan didirikan pada tahun 1978. SMA PGRI 1 Subang memiliki visi yaitu Unggul Dalam Akademis Dan Non Akademis. Dalam upaya mencapai visi tersebut SMA PGRI 1 Subang harus melakukan pengawasan terhadap prestasi yang diperoleh oleh siswa. Prestasi akademik siswa dipengaruhi oleh beberapa faktor salah satunya adalah pemilihan jurusan sesuai dengan minat dan bakat siswa.

Minat merupakan faktor yang memiliki peran atau pengaruh yang cukup besar dalam memperoleh prestasi akademik. Dengan adanya minat maka siswa akan belajar dengan bersungguh-sungguh karena sesuai dengan minat dimilikinya [1]. Selain dari faktor minat, faktor bakat pun merupakan faktor yang dapat mempengaruhi prestasi akademik siswa, Hal ini diungkapkan oleh Tolbert dalam Prisca D. Sampe [2] bahwa bakat studi merupakan hal yang terpenting untuk memperoleh prestasi akademik diperlukan bakat studi yang memadai. Bakat adalah kemampuan yang sudah berkembang, terwujud dan terbukti. Bakat studi merupakan kemampuan studi yang sudah dapat dilihat hasilnya.

Namun dalam kenyataannya tidak semua siswa memahami minat dan bakat diri sendiri. Hal ini dipengaruhi oleh usia siswa sma yang masih labil, dimana pada usia tersebut cenderung hanya ikut-ikutan dari temannya atau paksaan dari orang tua tanpa melihat minat dan bakat siswa. Sehingga tidak sedikit siswa yang salah dalam mengambil jurusan yang mengakibatkan siswa akan mengalami kesulitan dalam mengikuti

pelajaran, hal ini akan berdampak terhadap prestasi siswa. Maka peran guru sangat diperlukan dalam memberikan pemahaman dan motivasi kepada siswa agar siswa dapat memahami minat dan bakat yang dimiliki oleh siswa. Sehingga siswa mampu memilih jurusan sesuai dengan minat dan bakat siswa.

Dalam upaya pemberian pemahaman tentang pemilihan jurusan yang sesuai dengan minat dan bakat siswa. Guru SMA PGRI 1 belum memiliki suatu model yang dapat membantu dalam melakukan evaluasi minat dan bakat siswa untuk memilih jurusan. Model evaluasi yang dibutuhkan adalah model yang dapat memberikan informasi yang efektif dan efisien terkait dengan minat dan bakat siswa SMA PGRI 1 Subang. Informasi tersebut diperoleh dengan menggunakan proses pengolahan data yang dimiliki oleh siswa.

Kebutuhan pemodelan untuk evaluasi dapat dipenuhi dengan menggunakan disiplin ilmu data mining. Menurut Daniel T. Larose [3], Data mining adalah analisis data pengamatan set untuk menemukan hubungan tak terduga dan untuk meringkas data dalam cara-cara baru yang baik dimengerti dan berguna untuk pemilik data. Data mining memiliki satu tugas atau fungsinya adalah klaster, disamping fungsi tugas lainnya yaitu deskripsi, estimasi, prediksi, klasifikasi, dan asosiasi. Pengklasteran mengacu pada pengelompokan catatan, observasi, atau kasus ke dalam kelas dari objek serupa. Sebuah kluster adalah kumpulan dari catatan yang mirip satu sama lain, dan berbeda dengan catatan dalam kluster lain. Pembentukan kluster atau kelompok menggunakan algoritma K-Means, dimana K-Mean adalah algoritma yang menghubungkan setiap cluster dengan sebuah centroid (titik pusat). Tiap titik ditempatkan kedalam cluster dengan centroid yang terdekat [4]. Menurut Bandirali dalam jurnal Siti Nur Arofah [5], K-Means merupakan salah satu metode clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster. Metode ini mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik yang sama dan data yang mempunyai karakteristik yang berbeda di kelompokkan ke dalam cluster yang lain

Algoritma K-Means sering digunakan oleh para peneliti dalam melakukan pengelompokan data seperti menentukan pengelompokan prestasi siswa tinggi, menengah, dan cukup yang dipaparkan dalam jurnal Fitri Larasati Sibuea [6] pada Jurnal Teknologi dan Sistem Informasi tentang pemetaan siswa berprestasi menggunakan metode k-means clustering. Dimana pada jurnal tersebut K-Means mampu mengelompokkan siswa berdasarkan nilai kegiatan ekstrakurikuler, Nilai Tugas, UTS, UAS, absensi dan nilai sikap, menjadi 3 kelompok prestasi siswa. Contoh lainnya adalah menentukan calon penerima beasiswa bidik misi di POLBENG yang dipaparkan oleh Jaroji [7] dalam jurnal INOVTEK POLBENG tentang K-Means untuk menentukan calon penerima beasiswa bidik misi di POLBENG. Dimana pada jurnal tersebut algoritma K-Means mampu mengelompokkan calon penerima beasiswa bidik misi kedalam 4 cluster, yang terdiri dari cluster sangat layak, cluster layak, cluster pertimbangan, dan cluster pertimbangan.

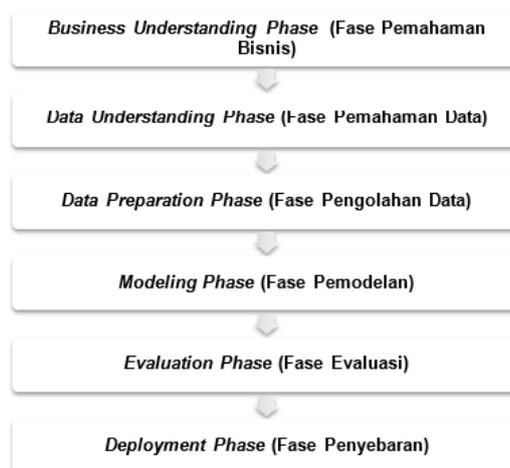
Sesuai pemaparan tentang teori data mining, clustering dan K-Means pada paragraf sebelumnya, bahwa proses clustering data mining merupakan rangkaian kegiatan untuk melakukan pengamatan pada data terdahulu, sehingga diperoleh kelompok-kelompok yang sejenis yang dapat digunakan untuk memprediksi data yang akan datang. Dengan menggunakan teknik klaster pada data mining, diharapkan dapat ditemukan suatu pola atau model dalam pemilihan jurusan oleh siswa SMA PGRI 1 Subang. Oleh karena itu, masalah dalam penelitian ini adalah, bagaimanakah membuat alat untuk pemilihan kesesuaian jurusan menggunakan teknik clustering data mining?.

Tujuan yang ingin dicapai dari penelitian ini adalah memperoleh model atau pola dalam pemilihan kesesuaian jurusan siswa SMA PGRI 1 Subang agar siswa tidak akan mengalami kesulitan dalam proses pembelajaran. Penelitian ini diharapkan dapat berguna bagi SMA PGRI 1 Subang dalam memberikan rekomendasi terhadap jurusan yang akan dipilih oleh siswa sesuai dengan minat dan bakat yang dimiliki. Sedangkan tujuan khusus dari penelitian ini adalah:

1. Diketahuinya variabel yang menentukan kesesuaian jurusan yang akan dipilih oleh siswa.
2. Diperolehnya perangkat atau alat yang dapat digunakan untuk melakukan evaluasi kesesuaian pemilihan jurusan oleh siswa.

## 2. METODE

Cross Industry Standard Process for Data Mining (CRISP-DM) digunakan sebagai standard proses data mining sekaligus sebagai metode penelitian. Menurut Daniel T. Larose (2006:5), Cross-Industry Standard Process for Data Mining (CRISP-DM) yang dikembangkan tahun 1996 oleh analisis dari beberapa industri seperti Daimler Chrysler, SPSS dan NCR. CRISP-DM menyediakan standar proses data mining sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian. Dalam CRISP-DM terdapat enam fase atau tahapan proses. Adapun tahapan proses data mining dengan menggunakan metode CRISP-DM digambarkan pada gambar 1 berikut ini:



Gambar 1. Tahapan Metode Penelitian Menggunakan CRISP-DM

Untuk menjamin keberhasilan, penelitian ini akan dilakukan dengan metode CRISP-DM dengan tahapan penelitian meliputi:

1. Business Understanding Phase (Fase Pemahaman Bisnis)  
Tujuan dari penelitian ini adalah menganalisa kesesuaian pemilihan jurusan oleh siswa dengan menggunakan teknik *data mining* yaitu metode *clustering* (pengelompokkan) berdasarkan variable yang dimiliki oleh siswa
2. Data Understanding Phase (Fase Pemahaman Data)  
Langkah selanjutnya adalah fase pengolahan data, dalam fase ini ada berapa tahapan dalam pengolahan data yaitu
  - a. Pengumpulan data
  - b. Menyeleksi atau pembersihan data untuk penentuan atribut yang akan digunakan, seperti data prestasi, minat dan bakat siswa
3. Data Preparation Phase (Fase Pengolahan Data)  
Melakukan transformasi data yaitu merubah data disesuaikan dengan kebutuhan *tool/software* yang akan digunakan pada tahap fase pemodelan, seperti data bentuk teks menjadi angka
4. Modeling Phase (Fase Pemodelan)  
Melakukan proses clustering untuk menghasilkan kelompok-kelompok data yang sejenis
5. Evaluation Phase (Fase Evaluasi)  
Mengamati dan menganalisa hasil pengelompokan yang diperoleh dari pemodelan, serta implemtasi hasil clustering
6. Deployment Phase (Fase Penyebaran)  
Implementasi dari clustering yang dihasilkan dan pembuatan Laporan

### 3. HASIL DAN PEMBAHASAN

Hasil penelitian pemilihan kesesuaian jurusan siswa SMA PGRI 1 Subang adalah sebagai berikut:

1. Fase Pemahaman Bisnis  
Pada fase ini dilakukan tahapan menentukan tujuan penelitian dan ruang lingkup penelitian, dimana tujuan penelitian adalah untuk memperoleh model atau pola dalam pemilihan jurusan siswa SMA PGRI 1 Subang. Ruang lingkup dalam penelitian ini adalah menggunakan teknik clustering data mining, sedangkan data yang dibutuhkan adalah data prestasi siswa SMA PGRI 1 Subang.
2. Fase Pemahaman Data  
Pada tahap fase pemahaman data adalah memahami data yang diperlukan dalam melakukan pengelompokkan data siswa menggunakan teknik clustering data mining. Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual, maka dari itu data yang diperlukan adalah

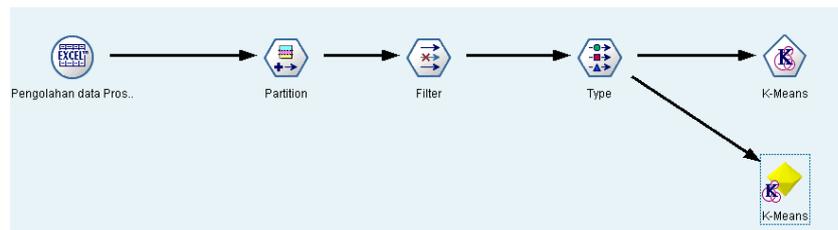
- a. Data nilai Ujian Nasional SMP, hal ini diperlukan untuk melihat data prestasi awal siswa sebelum menempuh pembelajaran di SMA. Data ini digunakan sebagai data bakat siswa
- b. Data nilai Ujian Nasional Berbasis Komputer SMA, data ini diperlukan untuk melihat hasil pembelajaran siswa.
- c. Data peminatan jurusan siswa, data ini diperlukan untuk melihat minat siswa terhadap jurusan yang ada di SMA PGRI 1 Subang, yaitu IPA dan IPS
- d. Data Hasil Psikotest, data hasil psikotest adalah data yang digunakan oleh SMA PGRI 1 Subang dalam penjurangan jurusan siswanya, dimana hasil psikotest ini dihasilkan dari wawancara guru konseling terhadap siswanya dan nilai Ujian Nasional yang dihasilkan oleh siswa.
- e. Data siswa yang digunakan adalah siswa lulusan tahun 2017, 2018, 2019

### 3. Fase Pengolahan Data

Pada tahap pengolahan data adalah melakukan proses seleksi data yaitu menghilangkan data-data yang tidak kumulit atau data yang kosong, karena hal ini akan berpengaruh terhadap hasil clustering, setelah dihilangkan data yang kosong maka selanjutnya adalah penyeragaman data yaitu bentuk atau tipe data diseragamkan, data yang diseragamkan adalah nilai UNBK dikonversi kedalam nilai mutu

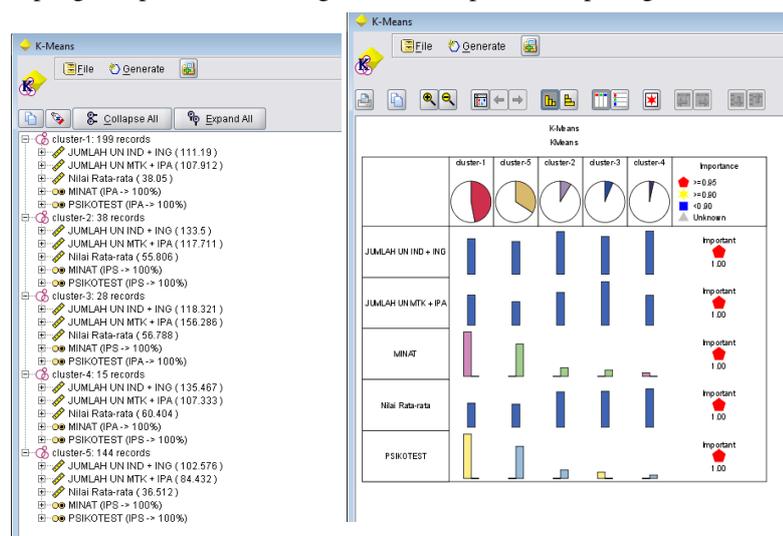
### 4. Fase Pemodelan

Fase Pemodelan adalah tahap yang melibatkan algoritma K-Means dalam melakukan clustering data mining, algoritma K-Means akan menghasilkan pengelompokkan siswa berdasarkan data nilai UN, data UNBK, data minat, dan hasil psikotest. Alat yang digunakan untuk melakukan pemodelan adalah menggunakan SPSS Clementine, hal ini diperlukan agar dalam melakukan pengelompokkan diperoleh data yang cukup akurat dan efektif. Adapun proses pemodelan menggunakan SPSS Clementine adalah sebagai berikut:



Gambar 2. Poses pemodelan Clustering K-Means menggunakan SPSS Clementine

Hasil pemetaan pengelompokkan clustering K-Means dapat dilihat pada gambar 3 sebagai berikut:



Gambar 3. Kelompok yang dihasilkan Clusteting K-Means

Pengelompokkan data yang dihasilkan adalah 5 kelompok adapun uraiannya adalah sebagai berikut:

<p><b>cluster-1 : 199 Records</b></p> <p>* JUMLAH UN IND + ING * Mean = 111.19 * Standard Deviation = 19.583</p> <p>* JUMLAH UN MTK + IPA * Mean = 107.912 * Standard Deviation = 31.669</p> <p>* Nilai Rata-rata * Mean = 38.05 * Standard Deviation = 10.902</p> <p>* MINAT</p> <p>* IPA (100%) * IPA 100% IPS 0%</p> <p>* PSIKOTEST</p> <p>* IPA (100%) * IPA 100% IPS 0%</p>	<p><b>cluster-2 : 38 Records</b></p> <p>* JUMLAH UN IND + ING * Mean = 133.5 * Standard Deviation = 20.675</p> <p>* JUMLAH UN MTK + IPA * Mean = 117.711 * Standard Deviation = 20.114</p> <p>* Nilai Rata-rata * Mean = 55.806 * Standard Deviation = 6.488</p> <p>* MINAT</p> <p>* IPS (100%) * IPA 0% IPS 100%</p> <p>* PSIKOTEST</p> <p>* IPS (100%) * IPA 0% IPS 100%</p>	<p><b>cluster-3 : 28 Records</b></p> <p>* JUMLAH UN IND + ING * Mean = 118.321 * Standard Deviation = 27.374</p> <p>* JUMLAH UN MTK + IPA * Mean = 156.286 * Standard Deviation = 28.43</p> <p>* Nilai Rata-rata * Mean = 56.788 * Standard Deviation = 13.411</p> <p>* MINAT</p> <p>* IPS (100%) * IPA 0% IPS 100%</p> <p>* PSIKOTEST</p> <p>* IPA (100%) * IPA 100% IPS 0%</p>
<p><b>cluster-4 : 15 Records</b></p> <p>* JUMLAH UN IND + ING * Mean = 135.467 * Standard Deviation = 19.896</p> <p>* JUMLAH UN MTK + IPA * Mean = 107.333 * Standard Deviation = 10.154</p> <p>* Nilai Rata-rata * Mean = 60.404 * Standard Deviation = 8.664</p> <p>* MINAT</p> <p>* IPA (100%) * IPA 100% IPS 0%</p> <p>* PSIKOTEST</p> <p>* IPS (100%) * IPA 0% IPS 100%</p>	<p><b>cluster-5 : 144 Records</b></p> <p>* JUMLAH UN IND + ING * Mean = 102.576 * Standard Deviation = 17.037</p> <p>* JUMLAH UN MTK + IPA * Mean = 84.432 * Standard Deviation = 19.724</p> <p>* Nilai Rata-rata * Mean = 36.512 * Standard Deviation = 6.412</p> <p>* MINAT</p> <p>* IPS (100%) * IPA 0% IPS 100%</p> <p>* PSIKOTEST</p> <p>* IPS (100%) * IPA 0% IPS 100%</p>	

## 5. Fase Evaluasi

Pada tahap ini melakukan evaluasi hasil dari cluster diperoleh oleh algoritma K-Means, yaitu

- a. Cluster pertama adalah kelompok yang memiliki minat jurusan IPA, yang memiliki data hasil psikotest adalah IPA, dengan nilai rata-rata UNBK adalah 38.05 dengan standar deviasi 10,9, memiliki nilai mean jumlah nilai UN SMP Bahasa Indonesia dan Bahasa Inggris adalah 111.19 dengan standar deviasi 19,5, dan memiliki nilai mean jumlah nilai UN SMP Matematika dan IPA yaitu 107.9 dengan standar deviasi =31,6. Maka kesimpulannya adalah jika memiliki minat IPA yang tidak didukung oleh nilai bakat yaitu nilai UN SMP, maka hasil nilai UNBK kurang maksimal, sehingga seharusnya siswa diarahkan ke jurusan IPS, karena nilai UN SMP non eksak (B.Indonesia dan B.Inggris) memiliki nilai yang lebih tinggi dibanding dengan nilai eksak (IPA).

Sehingga dapat diperoleh algoritmanya adalah, Jika minat jurusan siswa adalah IPA dan jumlah nilai UN Matematika SMP ditambah nilai UN IPA SMP kurang dari 107.9 dan Jumlah nilai UN B.Indonesia dan nilai UN B.Inggris kurang dari 119.19, maka siswa diarahkan ke jurusan IPS atau diberikan motivasi agar siswa belajar lebih giat untuk memperoleh nilai UNBK yang maksimal.

- b. Cluster kedua adalah kelompok yang memiliki minat jurusan IPS, yang memiliki data hasil psikotest adalah IPS, dengan nilai rata-rata UNBK adalah 55.80 dengan standar deviasi 6.48, memiliki nilai mean jumlah nilai UN SMP Bahasa Indonesia dan Bahasa Inggris adalah 133.5 dengan standar deviasi 20,6, dan memiliki nilai mean jumlah nilai UN SMP Matematika dan IPA yaitu 117.7.dengan standar deviasi = 20,1 Maka kesimpulannya adalah jika memiliki minat IPS yang didukung oleh nilai bakat yaitu nilai UN SMP, maka hasil nilai UNBK cukup maksimal, sehingga siswa sesuai dengan arahan yaitu jurusan IPS, karena nilai UN SMP non eksak (B.Indonesia dan B.Inggris) memiliki nilai yang lebih tinggi dibanding dengan nilai eksak (IPA).

Sehingga dapat diperoleh algoritmanya adalah, Jika minat jurusan siswa adalah IPS dan jumlah nilai UN Matematika SMP ditambah nilai UN IPA SMP kurang dari 107.9 dan Jumlah nilai UN B.Indonesia dan nilai UN B.Inggris kurang dari 119.19, maka siswa diarahkan ke jurusan IPS atau diberikan motivasi agar siswa belajar lebih giat untuk memperoleh nilai UNBK yang maksimal.

- c. Cluster Ketiga adalah kelompok yang memiliki minat jurusan IPS, yang memiliki data hasil psikotest adalah IPA, dengan nilai rata-rata UNBK adalah 56.80 dengan standar deviasi 13.41, memiliki nilai mean jumlah nilai UN SMP Bahasa Indonesia dan Bahasa Inggris adalah 118.3 dengan standar deviasi 27,3, dan memiliki nilai mean jumlah nilai UN SMP Matematika dan IPA yaitu 156.2.dengan standar deviasi = 13,4 Maka kesimpulannya adalah pada kluster ini siswa awalnya tidak memiliki kepercayaan dalam memilih jurusan IPA, namun hasil dari psikotest yang didukung oleh nilai bakat yaitu nilai UN SMP, siswa diarahkan ke jurusan IPA sehingga hasil nilai UNBK cukup maksimal, karena nilai UN SMP eksak (matematika dan IPA) memiliki nilai yang lebih tinggi dibanding dengan nilai non eksak (B.Indonesia dan B.Inggris).

Sehingga dapat diperoleh algoritmanya adalah, Jika minat jurusan siswa adalah IPS dan jumlah nilai UN Matematika SMP ditambah nilai UN IPA SMP lebih besar dari Jumlah nilai UN B.Indonesia dan nilai UN B.Inggris, maka siswa diarahkan ke jurusan IPA.

- d. Cluster Keempat adalah kelompok yang memiliki minat jurusan IPA, yang memiliki data hasil psikotest adalah IPS, dengan nilai rata-rata UNBK adalah 60.40 dengan standar deviasi 8.6, memiliki nilai mean jumlah nilai UN SMP Bahasa Indonesia dan Bahasa Inggris adalah 135.4 dengan standar deviasi 19,8, dan memiliki nilai mean jumlah nilai UN SMP Matematika dan IPA yaitu 107.3.dengan standar deviasi = 10,1 Maka kesimpulannya adalah pada kluster ini siswa memilih jurusan IPA, namun hasil dari psikotest yang didukung oleh nilai bakat yaitu nilai UN SMP, siswa diarahkan ke jurusan IPS sehingga hasil nilai UNBK cukup maksimal, karena nilai UN SMP non eksak (B.Indonesia dan B.Inggris) memiliki nilai yang lebih tinggi dibanding dengan nilai eksak (IPA).

Sehingga dapat diperoleh algoritmanya adalah, Jika minat jurusan siswa adalah IPA dan jumlah nilai UN Matematika SMP ditambah nilai UN IPA SMP lebih kecil dari Jumlah nilai UN B.Indonesia dan nilai UN B.Inggris, maka siswa diarahkan ke jurusan IPS

- e. Cluster Kelima adalah kelompok yang memiliki minat jurusan IPS, yang memiliki data hasil psikotest adalah IPS, dengan nilai rata-rata UNBK adalah 36.51 dengan standar deviasi 6.4, memiliki nilai mean jumlah nilai UN SMP Bahasa Indonesia dan Bahasa Inggris adalah 102.5.dengan standar deviasi 17,0, dan memiliki nilai mean jumlah nilai UN SMP Matematika dan IPA yaitu 84.4 .dengan standar deviasi = 19,7 Maka kesimpulannya adalah pada kluster ini siswa memilih jurusan IPS dan hasil Psikotest adalah IPS juga, namun karena nilai bakat siswa kurang yaitu nilai UN SMP, maka nilai UNBK Siswa pun kurang, sehingga diperlukan motivasi yang lebih kepada siswa pada kluster ini.

Sehingga dapat diperoleh algoritmanya adalah, Jika minat jurusan siswa adalah IPA dan jumlah nilai UN Matematika SMP ditambah nilai UN IPA SMP kurang dari 102.5 dan Jumlah nilai UN B.Indonesia dan nilai UN B.Inggris kurang dari 119.19, maka siswa diarahkan ke jurusan IPS atau diberikan motivasi agar siswa belajar lebih giat untuk memperoleh nilai UNBK yang maksimal.

#### 6. Fase Penyebaran

Pada fase penyebaran adalah tahap pembuatan program dalam pemenuhan teknologi tepat guna, adapun algoritma pemograman diperoleh dari hasil analisa pada tahap evaluasi, adapun algoritmanya adalah sebagai berikut:

- a. Jika minat jurusan siswa adalah IPA dan jumlah nilai UN Matematika SMP ditambah nilai UN IPA SMP lebih besar dari Jumlah nilai UN B.Indonesia dan nilai UN B.Inggris, maka siswa diarahkan ke jurusan IPA
- b. Jika minat jurusan siswa adalah IPS dan jumlah nilai UN Matematika SMP ditambah nilai UN IPA SMP lebih besar dari Jumlah nilai UN B.Indonesia dan nilai UN B.Inggris, maka siswa diarahkan ke jurusan IPA

- c. Jika minat jurusan siswa adalah IPS dan Jumlah nilai UN B.Indonesia dan nilai UN B.Ingggris lebih besar dari jumlah nilai UN Matematika SMP ditambah nilai UN IPA SMP maka siswa diarahkan ke jurusan IPS
- d. Jika minat jurusan siswa adalah IPA dan jumlah nilai UN Matematika SMP ditambah nilai UN IPA SMP lebih kecil dari Jumlah nilai UN B.Indonesia dan nilai UN B.Ingggris, maka siswa diarahkan ke jurusan IPS

#### 4. PENUTUP

Hasil penelitian yang telah dilakukan dapat diperoleh beberapa kesimpulan yaitu:

1. Diketahui empat variabel yang dapat digunakan dalam menentukan kesesuaian jurusan di SMA PGRI 1 Subang yaitu nilai ujian nasional SMP, nilai ujian nasional SMA, minat siswa, hasil psikotest
2. Hasil clustering data mining diperoleh 5 cluster yang memiliki rule set dan karakteristik yang berbeda dari cluster satu dengan cluster yang lainnya.
3. Rule set yang dihasilkan pada setiap cluster dapat digunakan dalam pembuatan algoritma program pemilihan kesesuaian jurusan di SMA PGRI 1 Subang

#### UCAPAN TERIMA KASIH

1. Direktorat Riset dan Pengabdian Masyarakat Direktorat Jenderal Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi dan Perguruan Tinggi, yang telah membiaya penelitian ini sesuai dengan Nomor SP DIPA-042.06.1.401516/2018, tanggal 5 Desember 2018 dan Nomor Kontrak Penelitian 2797/L4/PP/2019, tanggal 19 Maret 2019.
2. Direktur Politeknik LP3I Bandung beserta manajemen yang telah memberikan kesempatan waktu dan tempatnya kepada penulis dalam menyelesaikan penelitian program Hibah PDP.
3. Kepala Sekolah SMA PGRI 1 Subang beserta jajarannya yang telah memperkenankan siswa SMA PGRI 1 sebagai objek penelitian penulis.
4. LPPM Politeknik LP3I Bandung yang memberikan bimbingan kepada penulis sehingga penulis dapat menyelesaikan penelitian program Hibah PDP.
5. LPPM UNINDRA yang telah menerima paper hasil penulis, agar dapat dipublikasikan dalam jurnal Faktor Exacta.

#### DAFTAR PUSTAKA

- [1] S. Mustamin, H., Sulasteri, "Faktor-Faktor Yang Mempengaruhi Prestasi Belajar Mahasiswa Jurusan Pendidikan Matematika Fakultas Tarbiyah Dan Keguruan Uin Alauddin Makassar," *J. Mat. dan Pembelajaran*, vol. 1, no. 1, pp. 151–177, 2010.
- [2] P. D. Sampe, "Faktor-faktor yang mempengaruhi pengambilan keputusan untuk pemilihan jurusan di perguruan tinggi pada siswa SMA Kota Ambon," Universitas Kristen Satya Wacana, 2014.
- [3] D. T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining*. 2005.
- [4] F. A. Hermawati, *Data Mining*. Surabaya: Andi Publisher, 2013.
- [5] S. N. Arofah and F. Marisa, "Penerapan Data Mining untuk Mengetahui Minat Siswa pada Pelajaran Matematika menggunakan Metode K-Means Clustering," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 3, no. 2, 2018.
- [6] M. L. Sibuea and A. Safta, "Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering," *JURTEKSI*, vol. 4, no. 1, pp. 85–92, 2017.
- [7] J. Jaroji, D. Danuri, and F. P. Putra, "K-MEANS UNTUK MENENTUKAN CALON PENERIMA BEASISWA BIDIK MISI DI POLBENG," *INOVTEK Polbeng - Seri Inform.*, vol. 1, no. 1, p. 87, 2016.