

# **BAB 1**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

SMAN 1 Serang Baru merupakan Sekolah Menengah Atas Negeri yang terletak di Perumahan Kota Serang Baru (KSB) BLOK E, Sukaragam, Kec. Serang Baru, Kab. Bekasi, Prov. Jawa Barat. SMAN 1 Serang Baru berdiri sejak tahun 2007 serta memiliki 2 buah jurusan, yakni Ilmu Pengetahuan Alam (IPA) dan Ilmu Pengetahuan Sosial (IPS). SMAN 1 Serang Baru merupakan satu-satunya sekolah tingkat menengah atas negeri yang berada di Kecamatan Serang Baru, Kabupaten Bekasi. SMAN 1 Serang Baru memiliki predikat AKREDITASI A, serta memiliki total 1115 peserta didik dan 53 tenaga pengajar pada tahun ajaran 2022/2023. Setiap Warga Negara Indonesia berhak mendapatkan pendidikan sebagaimana yang tertuang dalam UUD 1945 pasal 31 ayat 1-2 [1]. Namun, tidak semua Warga Negara Indonesia memiliki kesempatan untuk memperoleh pendidikan yang layak dan berkelanjutan [2]. Maka dari itu, peserta didik yang kurang mampu berhak untuk memperoleh biaya pendidikan serta peserta didik yang berprestasi berhak untuk mendapatkan beasiswa [3].

Pemberian bantuan belajar berupa beasiswa juga kerap kali diberikan kepada siswa SMAN 1 Serang Baru, baik berasal dari program kerjasama antara lembaga swasta ataupun beasiswa yang diberikan oleh pemerintah. Pada prosesnya, pihak penyedia beasiswa akan meminta pihak sekolah dalam menentukan calon penerima beasiswanya, di mana Guru Bimbingan Konseling (BK) yang bertanggung jawab dalam melakukan pemilihan untuk menentukan calon penerima beasiswa. Berdasarkan data penerima beasiswa Sinarmas Land yang diselenggarakan oleh Institut Teknologi dan Sains Bandung pada tahun 2022, sebanyak 10 siswa direkomendasikan untuk mendapatkan beasiswa oleh pihak sekolah dari total 324 siswa yang ada di bangku kelas XII pada tahun ajaran 2021/2022. Pihak sekolah merasa kesulitan karena proses penentuan calon penerima beasiswa yang saat ini dilakukan adalah dengan cara melakukan penelusuran satu-persatu terhadap data siswa yang dirasa unggul dan aktif dalam proses pembelajaran kemudian dilakukan

pemilihan berdasarkan kuota yang ditetapkan oleh pihak penyedia beasiswa, kemudian siswa-siswa yang terpilih akan diwawancarai untuk memvalidasi apakah siswa tersebut telah memenuhi kriteria dari beasiswa yang ditawarkan atau tidak. Dengan banyaknya jumlah siswa yang ada dan terbatasnya kuota calon penerima beasiswa, proses pemilihan yang dilakukan ini memiliki banyak kekurangan seperti membutuhkan waktu yang lama dalam proses pemilihannya serta kemungkinan dilakukannya pemilihan ulang jika siswa yang dipilih sebelumnya tidak memenuhi kriteria dari beasiswa yang ditawarkan. Maka dari itu, diperlukan evaluasi pada proses pemilihan calon penerima beasiswa agar prosesnya menjadi lebih efisien dan tepat sasaran dengan banyaknya kriteria yang dijadikan pertimbangan [4].

Salah satu metode untuk memperoleh informasi dan pengetahuan dari proses pengelolaan data yang berjumlah sangat besar adalah menggunakan *data mining*, dengan salah satu fungsionalitas dari *data mining* yakni *clustering* yang berguna untuk mengelompokkan sejumlah data menjadi kelompok-kelompok data tertentu [5]. Dengan algoritma *Agglomerative Hierarchical Clustering (AHC)*, pembentukan *cluster* didapat berdasarkan hirarki penggabungan satu *cluster* terpisah menjadi gabungan-gabungan *cluster* yang lebih besar berdasarkan kemiripan karakteristik pada data tanpa harus menentukan jumlah *cluster* terlebih dahulu.

Berdasarkan latar belakang yang telah dijelaskan sebelumnya, dibutuhkan suatu metode untuk membantu pihak SMAN 1 Serang Baru dalam proses penentuan calon penerima beasiswa agar lebih efisien dan tepat sasaran. Maka dari itu, akan dilakukan implementasi *data mining* dengan metode *clustering* menggunakan algoritma *Agglomerative Hierarchical Clustering* untuk menghasilkan rekomendasi kelompok calon penerima beasiswa pada sekolah SMAN 1 Serang Baru.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disusun, maka rumusan masalah yang didapat adalah, bagaimana menerapkan data mining dengan metode *clustering* menggunakan algoritma *Agglomerative Hierarchical Clustering* untuk

menghasilkan rekomendasi kelompok calon penerima beasiswa pada sekolah SMAN 1 Serang Baru.

### **1.3 Maksud dan Tujuan**

Maksud dari penelitian ini adalah menerapkan *data mining* dengan metode *clustering* menggunakan algoritma *Agglomerative Hierarchical Clustering* untuk mengelompokkan calon penerima beasiswa pada sekolah SMAN 1 Serang Baru.

Dari maksud yang telah dipaparkan, tujuan yang hendak dicapai dari penelitian ini yaitu menghasilkan rekomendasi kelompok calon penerima beasiswa untuk memudahkan pihak SMAN 1 Serang Baru dalam proses penentuan calon penerima beasiswa.

### **1.4 Batasan Masalah**

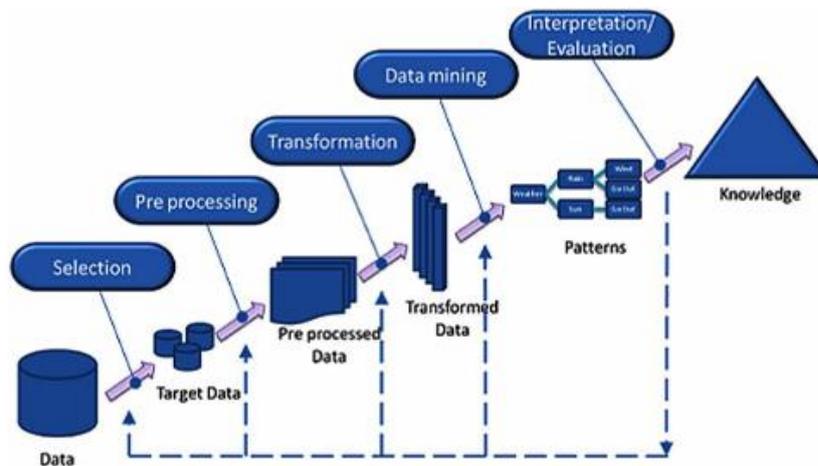
Adapun batasan masalah dari penelitian ini adalah :

1. Data yang digunakan merupakan data siswa kelas XI pada tahun ajaran 2022/2023 atau siswa kelas XII pada tahun ajaran 2023/2024.
2. Kriteria yang dipakai dalam penentuan calon penerima beasiswa adalah kriteria yang kerap kali digunakan oleh pihak penyedia beasiswa pada umumnya, seperti status ekonomi, keaktifan organisasi, jumlah prestasi akademik dan non akademik, serta nilai mata pelajaran Matematika, Bahasa Indonesia, dan Bahasa Inggris.
3. Algoritma *clustering* yang digunakan adalah algoritma *Agglomerative Hierarchical Clustering*.
4. Proses *data mining* menggunakan bahasa pemrograman python.
5. Data masukan berupa *file* berformat *.xlsx*.
6. Informasi hasil kelompok yang terbentuk akan ditampilkan dalam bentuk *website*.

### **1.5 Metodologi Penelitian**

Metode yang digunakan dalam penelitian ini adalah *Knowledge Discovery in Databases (KDD)*. *KDD* adalah proses dalam metode data mining untuk mengekstraksi apa yang dianggap sebagai suatu pengetahuan sesuai dengan

spesifikasi ukuran dan batasan menggunakan *database* dengan melakukan proses *preprocessing*, *sub sampling*, dan transformasi *database* yang diperlukan [6]. Adapun tahapan dari proses *KDD* dapat dilihat pada Gambar 1.1.



Gambar 1.1 Knowledge Discovery in Databases

Berikut adalah penjelasan dari masing-masing tahapan :

### 1. Selection

Tahap ini terdiri dari pembuatan atau pengumpulan data set yang akan diolah untuk dilakukan penemuan pengetahuan (*knowledge discovery*).

### 2. Pre processing

Tahap ini terdiri dari proses pembersihan data mencakup membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data.

### 3. Transformation

Pada tahap ini, terjadi proses transformasi pada data yang telah dipilih, sehingga data tersebut dapat dipastikan telah sesuai untuk proses *data mining*. Proses ini merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam *database*.

#### 4. *Data Mining*

Tahap ini merupakan proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *KDD* secara keseluruhan.

#### 5. *Interpretation / Evaluation*

Tahap ini merupakan bagian dari proses *KDD* yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya. Kemudian, pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

### **1.6 Sistematika Penulisan**

Sistematika penulisan yang akan digunakan, disusun untuk memberikan gambaran umum tentang penelitian yang dikerjakan. Sistematika penulisan dalam penelitian tugas akhir ini adalah sebagai berikut:

#### **BAB 1 PENDAHULUAN**

Bab ini terdiri dari latar belakang masalah penelitian, rumusan masalah, maksud dan tujuan penelitian, manfaat, batasan masalah, metodologi penelitian, dan sistematika penulisan.

#### **BAB 2 TINJAUAN PUSTAKA**

Bab ini berisikan tentang tinjauan umum tempat dilaksanakannya penelitian dan landasan teori yang mendasari judul penelitian.

#### **BAB 3 ANALISIS DAN PERANCANGAN**

Bab ini berisikan tentang analisis permasalahan untuk mengetahui masalah apa yang timbul kemudian mencoba memecahkan masalah tersebut dengan melakukan komparasi dari dua buah *cluster* yang telah dihasilkan.

#### BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisikan hasil implementasi dari proses data mining yang telah dilakukan serta dilakukan pengujian untuk mendapatkan hasil penelitian.

#### BAB 5 KESIMPULAN DAN SARAN

Bab ini berisikan kesimpulan dari penelitian yang sudah dilaksanakan. Adapun saran dari penelitian ini adalah sebagai bahan evaluasi untuk penelitian selanjutnya.