

# PENERAPAN DATA MINING SEBAGAI MODEL SELEKSI PENERIMA BEASISWA PENUH (STUDI KASUS: STIE PERBANAS SURABAYA)

Hariadi Yutanto<sup>1</sup>, Nurcholis Setiawan<sup>2</sup>

STIE Perbanas Surabaya

<sup>1</sup>antok@perbanas.ac.id, <sup>2</sup>cholis@perbanas.ac.id

---

## Abstrak

Program beasiswa yang diperuntukkan bagi siswa SMA/SMK/MA telah ada di STIE Perbanas Surabaya sejak tahun 2000. Beberapa tahapan dan kriteria dalam penentuan beasiswa terkadang mengakibatkan sulitnya membuat keputusan dalam menentukan calon penerima beasiswa. Data calon penerima disajikan dalam bentuk table akan tetapi tidak mudah dalam proses membaca table dengan cepat dan tepat ditambah lagi dengan jumlah data yang cukup besar. Tujuan dari penelitian mengumpulkan informasi tentang data calon penerima beasiswa penuh dalam satu table mulai dari hasil kemampuan akademik, non akademik, komitmen, penghasilan orang tua dan rekomendasi. Informasi disajikan dengan cara mengklasifikasikan calon penerima beasiswa penuh menggunakan algoritma C4.5 dengan mencari nilai *Entropy* dan *Gain* dalam membentuk pohon keputusan. Dengan menggunakan system informasi pembuatan pohon keputusan dari data calon mahasiswa menjadi lebih mudah dan cepat sehingga dapat melihat factor apa saja pertimbangan dalam pengambilan keputusan.

**Kata Kunci** : data mining, beasiswa penuh, C4.5, model pendukung keputusan

---

## 1.1 Latar Belakang

STIE Perbanas Surabaya (STIEP) merupakan sebuah Perguruan Tinggi Swasta, yang berdiri sejak 1975. Pada tahun 2000 STIEP memulai membuka jalur pendaftaran beasiswa penuh bagi masyarakat umum untuk lulusan SMU/SMK/MA dan sederajat. Proses seleksi penerimaan beasiswa berada dibawah sebuah kepanitian mahasiswa baru SPMB dilakukan dalam beberapa tahapan dengan system gugur, penentuan hasil akhir adalah pada test psikologi. Data hasil penerimaan setiap tahunnya cukup banyak sehingga apabila tidak dimanfaatkan data tersebut akan menjadi tumpukan data pada gudang data. Data mining merupakan metode untuk menggali dan mencari pengetahuan baru dari data tumpukan data tersebut.

Data Mining merupakan teknologi baru yang sangat berguna untuk menemukan informasi yang sangat penting dari gudang data. Ketersediaan data yang banyak dan kebutuhan akan informasi atau pengetahuan sebagai pendukung pengambilan keputusan untuk membuat solusi bisnis dan dukungan infrastruktur di bidang teknik informatika

merupakan cikal-bakal dari lahirnya teknologi data mining. Sehingga informasi tersebut bisa digunakan sebagai solusi pengambilan keputusan di dunia bisnis, untuk pengembangan bisnis.

Konsep teknologi informasi yang berkaitan dengan data dan informasi adalah bagian dari *data mining*. Konsep data mining adalah sebagai proses atau teknik pemodelan yang mempergunakan analisis dengan variasi data yang banyak untuk mendapatkan pola dan hubungan diantara variasi data tersebut. Kehadiran data mining dilatarbelakangi dengan adanya masalah data explosion atau ledakan data yang dialami oleh banyak organisasi yang telah mengumpulkan data sekian tahun lamanya.

Pada STIEP terdapat beberapa tahapan seleksi beasiswa dari rata-rata 300 pendaftar calon beasiswa dari tahun 2010-2015 STIEP hanya menerima kuota 15 siswa beasiswa penuh dan 5 orang beasiswa Perbanas Jatim. Berikut adalah tahapan seleksi penerimaan beasiswa yaitu

1. Seleksi Administrasi dilakukan melalui pendaftaran online <http://spmb.perbanas.ac.id>, dengan kriteria nilai

2. Seleksi Wawancara data dilaksanakan setelah lolos seleksi administrasi, terdapat 20 pewawancara yang masing-masing memberi rekomendasi 10 orang untuk lolos tahap selanjutnya.
3. Seleksi Psikotest adalah tahapan akhir dari proses penerimaan

Permasalahan yang sering dihadapi adalah pada proses penentuan lolos seleksi wawancara, dibutuhkan waktu yang lama dalam pengambilan keputusan banyak kriteria pertimbangan dalam menentukan lolos atau tidaknya calon penerima beasiswa, penyajian data disajikan dalam bentuk excel dan dilakukan *sorting* data.

Dari latar belakang permasalahan tersebut peneliti mencoba mengembangkan suatu system informasi untuk penyajian data dari hasil lolos seleksi wawancara yang berisi factor-faktor apa saja dari rekomendasi pewawancara sebagai pendukung keputusan. Penelitian ini menerapkan algoritma C4.5 dalam sebuah sistem informasi yang dapat mengetahui pola pengambilan keputusan penerimaan beasiswa. Dari pola yang dihasilkan akan didapatkan sebuah informasi atau *knowledge*.

### 1.2 Rumusan Masalah

Dengan mengacu pada latar belakang masalah diatas, maka peneliti akan membahas permasalahan tentang “Bagaimana memanfaatkan algoritma C.45 *Data Mining* untuk model penentuan pengambilan keputusan calon penerima beasiswa penuh di STIEP”

### 1.3 Batasan Masalah

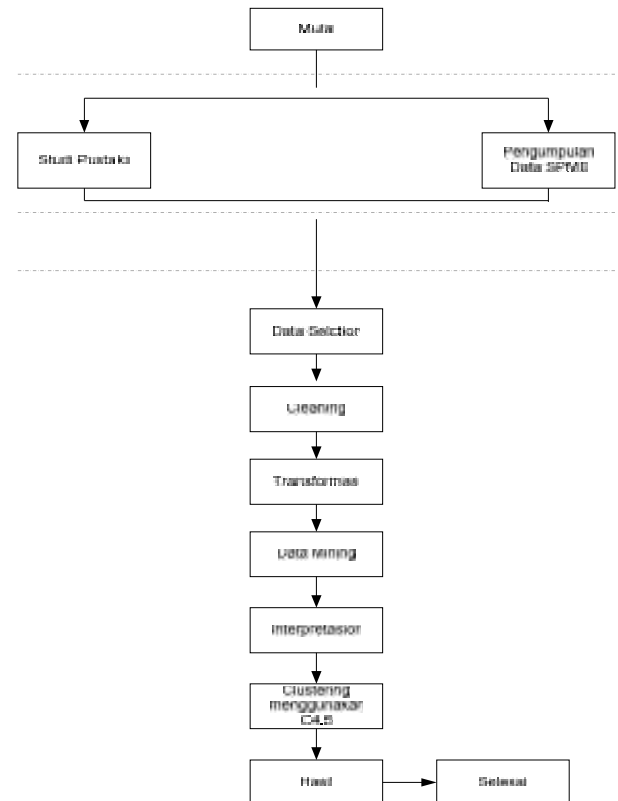
Untuk menghindari pembahasan yang meluas, maka penulis hanya membatasi pembahasan permasalahan hanya pada Penerapan data mining untuk memprediksi calon penerimaan beasiswa penuh dari hasil lolos seleksi wawancara untuk rekomendasi masuk pada tahap psikotest bagi calon mahasiswa baru di STIE Perbanas Surabaya dengan menggunakan Algoritma C4.5 mengikuti tahapan *Knowledge Discovery in Database* (KDD).

### 1.4 Tujuan

Kegiatan penelitian ini bertujuan untuk menerapkan teknik data mining dalam menampilkan informasi untuk pengambilan keputusan mengenai calon penerima beasiswa penuh di STIE Perbanas Surabaya. Dan dengan

penerapan data mining ini diharapkan dapat memberikan pengetahuan yang sebelumnya tersembunyi didalam data sehingga menjadi informasi berharga, berupa informasi yang dapat membantu memprediksi dan mengambil keputusan untuk calon penerima beasiswa penuh di STIE Perbanas.

## 1.5 Metodologi Penelitian



Gambar 1 Tahapan Metode Penelitian

1. Studi Pustaka  
Yaitu mencari informasi dari beberapa buku-buku referensi, serta situs di internet yang dapat membantu memperoleh pengetahuan dan pokok permasalahan yang berhubungan dengan materi penulisan ini.
2. Pengumpulan Data SPMB  
Mengumpulkan data penerima hasil seleksi wawancara yang diambil dari database Sistem Informasi Penerimaan Mahasiswa Baru, data yang diambil untuk dilakukan proses mining.
3. Metode Analisis Data  
Metode Analisis Data Adapun untuk menganalisis data dalam penerapan data mining ini menggunakan tahapan *Knowledge Discovery in Databases* (KDD) yang terdiri dari beberapa

tahapan, yaitu *data selection, preprocessing, transformation, data mining, dan evaluation.*

### 2.1 Data Selection

Dalam data spmb yang lolos administrasi yang digunakan sebagai data wawancara tersaji dalam bentuk excel dalam bentuk tabel dengan atribut dan record. Atribut tersebut memiliki nilai yang dinamakan sebagai *instance.*

Gambar 2. Data Selection

### 2.2 Data Praproses

Data praproses diambil dari data hasil wawancara yang telah dikelompokkan berdasarkan kesimpulan / rekomendasi untuk lolos psikotest.

Gambar 2. Data Praproses

### 2.3 Pengelompokan Data

Pengelompokan data prestasi akademik dan non akademik dibagi menjadi 4 kelas yaitu kurang, cukup, baik dan baik sekali

Tabel 1 Prestasi dan Non Akademik

|             |    |
|-------------|----|
| Kurang      | =1 |
| Cukup       | =2 |
| Baik        | =3 |
| Baik Sekali | >3 |

Pengelompokan data komitmen kepada STIE Perbanas dibagi menjadi 3, yaitu Tidak Yakin, Ragu-ragu dan Yakin

Tabel 2 Komitmen Kepada STIEP

|             |     |
|-------------|-----|
| Tidak Yakin | =1  |
| Ragu-Ragu   | =2  |
| Yakin       | >=3 |

Pengelompokan data Minat masuk STIE Perbanas jika gagal BP dibagi menjadi 2, yaitu Tidak Minat dan Minat.

Tabel 3 Minat

|             |      |
|-------------|------|
| Tidak Minat | =1&2 |
| Minat       | =3&4 |

Pengelompokan kesediaan ditempatkan di Prodi Diploma jika gagal BP (Beasiswa Penuh) dibagi menjadi 2, yaitu Tidak Bersedia dan Bersedia

Tabel 5 Kesediaan

|                |    |
|----------------|----|
| Tidak Bersedia | =1 |
| Bersedia       | =2 |

Pengelompokan pendapatan orang tua dibagi menjadi 2, yaitu Ya dan Tidak

Tabel 6 Pendapatan Ortu

|           |       |
|-----------|-------|
| <2,5 Juta | Ya    |
| <2,5 Juta | Tidak |

### 2.4 Perhitungan entropi dan Gain

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (1)$$

Di mana:  
 S = himpunan kelas  
 A = atribut  
 n = jumlah nilai atribut A  
 S<sub>i</sub> = himpunan kelas ke-n  
 |S<sub>i</sub>| = jumlah kasus dalam S<sub>i</sub>  
 |S| = jumlah kasus dalam S

$$Entropy(S) = \sum_{i=1}^k - p_i * \log_2 p_i \quad (2)$$

Di mana:  
 k = himpunan kelas  
 p = jumlah kasus S / total kasus S

Entropy Total

$$= \left(-\frac{68}{121} * \log_2\left(\frac{68}{121}\right)\right) + \left(\frac{53}{121} * \log_2\left(\frac{53}{121}\right)\right)$$

Entropy Prestasi Akademik (Baik)

$$= \left(-\frac{3}{17} * \log_2\left(\frac{3}{17}\right)\right) + \left(\frac{14}{17} * \log_2\left(\frac{14}{17}\right)\right)$$

Entropy Prestasi Akademik (Sangat Baik)

$$= \left(-\frac{10}{37} * \log_2\left(\frac{10}{37}\right)\right) + \left(\frac{27}{37} * \log_2\left(\frac{27}{37}\right)\right)$$



- Santosa Budi, 2007, "*Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*", Yogyakarta : Graha Ilmu.
- Anwar Arifin, Shaufiah, Dwi Dawam, 2012, "*Analisis dan Implementasi Metode Single Layer Perceptron Pada Data Mining Penerimaan Siswa Baru Jalur JPPAN*", pp. 1–7.
- Ardhi Fauzan, 2009, "*Analisis dan Implementasi Algoritma Cruise dengan Metode ID pada Penentuan Seleksi Penerimaan Mahasiswa Baru di Perguruan Tinggi melalui JPPAN*", pp. 1–7.
- Fadli,A. 2011. "*Aplikasi Data Mining. Ilmu Komputer.Com*",  
<http://ilmukomputer.org/2011/03/14/aplikasi-data-minning/> (diakses 20 Oktober 2016)
- Kusrini, & Lutfi, E.T. 2009. "*Algoritma data mining*". Penerbit ANDI : Yogyakarta
- Santosa, B., 2007, *Data Mining : "Teknik Pemanfaatan Data Untuk Keperluan Bisnis, Teori dan Aplikasi"*, Graha Ilmu, Yogyakarta.
- Mujib Ridwan, Hadi Suyono & M.Saros, "*Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier*" *jurnaleeccis*
- Larose, & Daniel T. 2005. "*Discovering knowledge in data: an introduction to data mining*". USA: John Wiley and Sons