

## IMPLEMENTASI *BUSINESS INTELLIGENCE* UNTUK MENENTUKAN MAHASISWA PENERIMA BEASISWA

<sup>1</sup>Ricky Akbar, <sup>2</sup>Widya Wulandari, <sup>3</sup>Zikrya Hasanah, <sup>4</sup>Hesti Gravina, <sup>5</sup>Yudha Restu Alditya,  
<sup>6</sup>Agum Sadewa

<sup>1,2,3,4,5,6</sup>Program Studi Sistem Informasi, Fakultas Teknologi Informasi Universitas Andalas

Jl. Universitas Andalas, Limau Manis, Pauh 25163 Padang-Sumatera Barat

Email: <sup>1</sup>rickyakbar1984@gmail.com, <sup>2</sup>widyawulandari.wewe@gmail.com, <sup>3</sup>zikrya.hasanah@yahoo.com,  
<sup>4</sup>hestigravina6@gmail.com, <sup>5</sup>ryudhagen96@gmail.com, <sup>6</sup>agumymous63@gmail.com

### ABSTRAK

Seiring dengan perkembangan teknologi informasi membuat segala aspek pekerjaan menjadi lebih mudah. Perkembangan teknologi informasi dapat di implementasikan dalam berbagai aspek pekerjaan, salah satunya ialah pada bidang akademik dalam menentukan mahasiswa-mahasiswa pada sebuah universitas yang berhak menerima beasiswa sesuai dengan kriteria-kriteria yang telah ditentukan. Dalam implementasinya, tentu sulit dalam menentukan siapa saja mahasiswa yang berhak menerima beasiswa tersebut dikarenakan data yang teramat banyak pada setiap angkatan mahasiswa. Maka dari itu pihak akademik pada sebuah universitas ingin mendapatkan informasi mengenai mahasiswa yang tepat yang berhak untuk menerima beasiswa tersebut, namun banyaknya data mahasiswa yang ada membuat proses pengambilan informasi berjalan lambat jika dilakukan dengan *query standard* pada *database*. Dalam menentukan mahasiswa yang berhak menerima beasiswa tentu banyak faktor yang membuat mahasiswa tersebut pantas untuk dipilih karena banyak kriteria yang akan menentukan mahasiswa penerima beasiswa tersebut, seperti nilai indeks prestasi akademik mahasiswa, dan dari tingkat ekonomi mahasiswa. *Business intelligence* merupakan serangkaian proses untuk mendapatkan informasi dengan mudah dari data yang berukuran besar. *Business intelligence* tak lepas dari penggunaan *data warehouse* dan analisis data. Pada penelitian ini dilakukan pembuatan *tree decision* untuk pembuatan calon mahasiswa yang masuk kategori dalam penerimaan beasiswa selanjutnya akan dilakukan pembuatan dalam pengurutan calon mahasiswa yang berhak menerima beasiswa tersebut dengan menggunakan *RapidMiner*. Metode yang akan digunakan dalam penelitian ini adalah *Online Analytical Processing (OLAP)* serta operasi-operasi yang ada di dalamnya. Dengan menggunakan metode OLAP dan *RapidMiner* diharapkan akan mampu mempermudah dalam *data mining* yang menghasilkan pohon keputusan dalam pengambilan keputusan serta menentukan mahasiswa yang berhak menerima beasiswa pada sebuah universitas.

**Kata kunci:** *business intelligence*, *data warehouse*, OLAP, *RapidMiner*.

### A. PENDAHULUAN

Seiring dengan perkembangan teknologi informasi membuat segala aspek pekerjaan menjadi lebih mudah. Perkembangan teknologi informasi juga membuat data yang dihasilkan menjadi semakin melimpah. Informasi yang dihasilkan dari data dan teknologi tersebut menjadi faktor yang penting bagi perkembangan sebuah instansi pemerintah ataupun perusahaan.

Data instansi pemerintah yang melimpah menyebabkan proses pengambilan informasi yang berakibat pada proses pengambilan keputusan yang lambat. Hal tersebut menjadi sebuah tantangan yang berarti bagi sebuah instansi pemerintah atau perusahaan. Oleh karena itu instansi pemerintah atau perusahaan tersebut dituntut agar dapat mengolah dan mengelola seluruh data yang ada secara cepat dan efisien agar dapat menghasilkan informasi yang berkualitas sehingga mengakibatkan pengambilan keputusan yang tepat.

Data yang melimpah akan menjadi tidak berguna apabila tidak ada pengelolaan yang baik dan juga analisa terhadap aktifitas bisnis merupakan kebutuhan yang tidak bisa ditawar [1].

Kecepatan pengolahan data dan pengambilan informasi yang akan dijadikan untuk pengambilan keputusan pada sebuah instansi pemerintah atau perusahaan, membuat instansi pemerintah dan perusahaan mampu menanggulangi dan menghindari kejadian atau peristiwa yang dapat mengakibatkan kerugian pada instansi pemerintah tersebut.

Pentingnya penggunaan *Business Intelligence (BI)* untuk instansi pemerintah yaitu pada sebuah universitas negeri yang salah satu prosesnya yakni pada proses penentu mahasiswa yang berhak menerima beasiswa. Banyaknya data mahasiswa tersebut membuat pihak yang berwenang pada instansi pemerintah untuk melakukan pengolahan data dengan sangat baik sehingga nantinya dapat menghasilkan sebuah informasi dan pengambilan keputusan yang tepat.

Instansi pemerintah merasa kesulitan dalam mengelola data yang cukup besar bila menggunakan *tools* yang sangat sederhana yang tidak mampu menghasilkan informasi yang tepat dan mengakibatkan kekeliruan dalam pengambilan keputusan.

Instansi pemerintah yang dimaksud ialah Universitas Andalas dimana universitas ini memiliki jumlah mahasiswa aktif kurang lebih 28.000 mahasiswa termasuk dengan program diploma III dan pascasarjana. Bisa dibayangkan berapa banyak data mahasiswa yang dimiliki oleh Universitas Andalas. Dari data tersebut Universitas Andalas ingin mengambil informasi yang penting untuk mengetahui siapa saja mahasiswa yang berhak menerima beasiswa berdasarkan kriteria indeks prestasi akademik dan tingkat ekonomi mahasiswa.

Adapun sejumlah data yang dibutuhkan ialah 10.000 data yang terdiri dari data mahasiswa angkatan 2013 dan data mahasiswa angkatan 2014 dimana pada setiap angkatan akan dipilih beberapa mahasiswa yang berhak untuk menerima beasiswa pada setiap jurusan dengan kriteria indeks prestasi akademik yang memuaskan dan dengan tingkat ekonomi menengah ke bawah yaitu pekerjaan orangtua dari mahasiswa tersebut yang merupakan buruh atau petani dan orangtua mahasiswa yang tidak memiliki pekerjaan.

Maka dalam penelitian ini akan dikaji mengenai analisis data untuk menentukan mahasiswa yang berhak menerima beasiswa dengan menggunakan *tool RapidMiner*.

## B. LANDASAN TEORI

### B.1. *Business Intelligence* (BI)

Istilah kecerdasan (*intelligence*) telah digunakan oleh para peneliti pada kecerdasan buatan (*artificial intelligence*) sejak tahun 1950-an. Namun, hanya pada 1990-an bisnis intelijen menjadi istilah populer pada komunitas-komunitas bisnis dan IT. Pada tahun 2000-an, *business analytics* diperkenalkan untuk mewakili kunci komponen analitis dalam BI [2].

*Business intelligence* adalah sebuah istilah yang menggabungkan arsitektur, *tools*, *database*, *analytical tools*, aplikasi, dan metodologi. Dengan mengaplikasikan teknologi *business intelligence* pada perusahaan, data yang sebelumnya tidak memiliki nilai ekonomis bagi perusahaan dapat diubah menjadi bahan bernilai tinggi untuk menentukan strategi bisnis kedepan [3].

Terdapat empat komponen dasar dari *business intelligence* yang saling bersinergi agar sebuah *business intelligence* dapat berfungsi yaitu:

#### 1) *Data warehouse*

*Data warehouse* berperan sebagai sumber data dari *business intelligence*. Sebuah *data warehouse* merupakan koleksi data yang terorientasi pada subyek, tidak mengalami perubahan, serta memiliki rentang waktu yang cukup lebar yang berfungsi dalam mendukung pengambilan keputusan manajemen.

#### 2) *Business analytics*

Secara singkat *business analytical* adalah kumpulan peralatan untuk memanipulasi, menambang dan menganalisa data yang terdapat di dalam *data warehouse*.

#### 3) *Report dan queries*

Termasuk didalamnya segala bentuk pelaporan baik secara statis (tidak berubah) ataupun dinamis sesuai dengan perubahan data dan setiap macam *query* yang ada seperti *drill down*, *multidimensional view*, pengelompokan, dsb.

#### 4) *Data, text dan web mining* serta peralatan matematika level atas dan statistik.

*Data mining* merupakan sebuah proses untuk menemukan relasi atau informasi yang tidak diketahui didalam *database* besar ataupun *data warehouse* dengan menggunakan peralatan *intelligent*. *Text mining* adalah *semiautomated* proses ekstraksi pola dari data tidak terstruktur dalam jumlah yang besar. Perlu diingat bahwa perbedaan *text mining* dengan *data mining* yaitu pada sumber data dari *data mining* adalah data terstruktur yang terdapat didalam *database* [3].

Suatu instansi/perusahaan biasanya memiliki ratusan bahkan ribuan data, data-data tersebut jika akan diperiksa dengan secara manual yaitu sortir satu persatu akan membutuhkan waktu yang lama dan memiliki tingkat kesalahan lebih tinggi dibandingkan dengan *business intelligence*. Karena *business intelligence* pengolahan dan analisis data menjadi lebih teratur dan lebih cepat dengan dibantu *tools* dan aplikasi *data mining*. Jadi dengan adanya *business intelligence*, dapat mempermudah instansi/perusahaan menemukan data yang dapat berguna baik bagi instansi/ perusahaan tersebut maupun bagi masyarakat luas. Data-data yang terkumpul tersebut yang kemudian menjadi *data warehouse*. Data-data yang terdapat pada *data warehouse* akan digunakan untuk melakukan *business analytical*, melakukan analisa terhadap data pada *data warehouse* sesuai kebutuhan atau tujuan instansi/ perusahaan. Selanjutnya dengan aplikasi *data mining* dapat diadaptasi data sesuai dengan *business analytical* yang telah dilakukan.

Pada jurnal implementasi *business intelligence* untuk menentukan mahasiswa penerima beasiswa, hal pertama yang dilakukan yaitu mencari data mahasiswa Universitas Andalas dari tahun 2013 sampai 2014 sebagai *data warehouse*. Kemudian melakukan *business analytical* pada *data warehouse*, yaitu kriteria yang diperlukan untuk menentukan mahasiswa yang cocok sebagai penerima beasiswa. Selanjutnya membuat *report* dan *query* dengan dibantu aplikasi *data mining*, aplikasi *data mining* yang digunakan yaitu *RapidMiner*.

## B.2. RapidMiner

Saat ini *RapidMiner* termasuk salah satu solusi *data mining* dan prediksi analisis yang banyak dipakai. *RapidMiner* mempunyai banyak kelebihan, diantaranya:

- 1) Tersedia secara bebas sebagai aplikasi *open source* tetapi juga menyediakan *commercial licence* yang cocok untuk aplikasi *closed-source*.
- 2) Sudah matang, perkembangan pengguna, komunitas pengembang, biaya pengembangan dengan *RapidMiner* relatif lebih rendah dibandingkan dengan solusi *data mining* lainnya.
- 3) Sangat cocok untuk tujuan penelitian, intuitif, diatur dengan baik dan *user friendly* untuk antarmuka pengguna grafis.
- 4) Dapat mengakses ke banyak *database* dan mampu membaca berbagai format *file*, sejumlah besar operator *built in* untuk *data mining* memiliki lebih dari 250 *learning algoritma* dan banyak praproses [4].

*RapidMiner* mulai dikembangkan pada tahun 2001 di unit *artificial intelligence* dari Technical University of Dortmund dimana Ralf Klinkenberg, Simon Fischer, dan Ingo Mierswa bersama-sama membuat proyek analisis prediktif dengan nama *Yet Another Learning Environment (YALE)* [5].

## C. PEMBAHASAN DAN HASIL

Pada bab ini akan diuraikan mengenai hasil penelitian serta pembahasan dari hasil penelitian, yaitu bagaimana hasil dari pengolahan data-data mahasiswa yang ada sehingga nanti didapati *output* berupa pohon keputusan untuk menentukan siapa saja yang berhak menerima beasiswa.

### C.1. Langkah Pengolahan Data

Pada aplikasi ini, data masukan yang akan diolah pada sistem yaitu berupa Nomor Induk Mahasiswa (NIM), Pekerjaan orangtua, dan Indeks Prestasi Kumulatif (IPK). Pada sistem ini, data yang diolah berupa data seluruh mahasiswa Universitas Andalas angkatan 2014 sampai dengan angkatan 2015.

### C.2. Kriteria Mahasiswa yang Memperoleh Beasiswa

Adapun kriteria dari mahasiswa yang akan mendapat beasiswa berdasarkan sistem ini yaitu:

- 1) Bukan mahasiswa tahun akhir
- 2) Pekerjaan orangtua ialah sebagai buruh/tani dan tidak memiliki pekerjaan
- 3) Memiliki IPK lebih dari 2,75.

### C.3. Proses Eksekusi Data

Proses eksekusi dari data menggunakan *RapidMiner* yaitu:

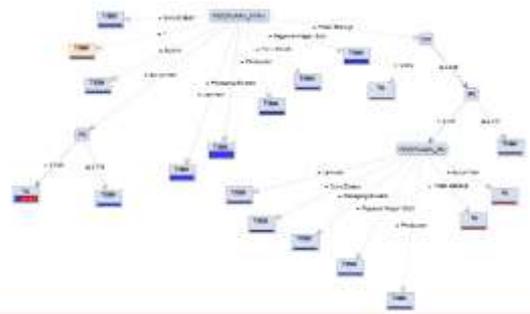
- 1) *Input/read* data yang akan diproses, dimana data disini berupa data *Excel*.
- 2) Kemudian pilih pada *Import Configuration Wizard* data yang akan diproses.
- 3) Kemudian lakukan validasi data

### C.4. Hasil Proses Data

Adapun hasil dari proses data ini yaitu:

#### 1) Pohon Keputusan

Pada tahap awal data dikriteriakan berdasarkan pekerjaan orangtua. Bagi mahasiswa yang pekerjaan orangtuanya sesuai dengan kriteria, maka akan dilanjutkan dengan proses penyeleksian IPK, sehingga terlihat hasil seperti Gambar 1.



Gambar 1. Pohon keputusan

- 2) Pentotalan jumlah mahasiswa yang mendapat beasiswa dilihat berdasarkan pohon keputusan. Pada Gambar 2 dijelaskan secara detail mengenai informasi mahasiswa yang akan menerima beasiswa berdasarkan kriteria yang telah ditentukan. Pada Gambar 1 hanya menampilkan apakah mahasiswa masuk ke dalam kategori atau tidak, namun pada Gambar 2 menampilkan jumlah mahasiswa yang masuk ke dalam kategori melalui penyaringan kriterianya. Apabila mahasiswa masuk dalam kriteria, maka detail informasi mahasiswa yang menerima beasiswa akan dijelaskan secara detail pada Gambar 2.
- 3) Akurasi  
Disini dijelaskan berapa tingkat kedekatan antara nilai prediksi dengan nilai aktual. Seperti yang terlihat Pada Gambar 3.
- 4) Presisi/*confidence*  
Presisi/ *confidence* merupakan proporsi kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya [6]. Pada penelitian ini presisi digunakan untuk menampilkan tingkat ketepatan antara informasi yang diminta dengan jawaban yang diberikan oleh sistem, yang terlihat pada Gambar 4.

```

Tree
PERKULIAHAN_AYAH = ? : Tidak (Tidak=116, Ya=0)
PERKULIAHAN_AYAH = BUKAN: Tidak (Tidak=249, Ya=0)
PERKULIAHAN_AYAH = BUKAN: Tidak (Tidak=249, Ya=0)
PERKULIAHAN_AYAH = BUKAN: Tidak (Tidak=249, Ya=0)
    IIRK > 2.743: Ya (Tidak=565, Ya=1200)
    IIRK > 2.743: Tidak (Tidak=552, Ya=0)
PERKULIAHAN_AYAH = Guru/Dosen: Tidak (Tidak=669, Ya=0)
PERKULIAHAN_AYAH = Lain-lain: Tidak (Tidak=1026, Ya=0)
PERKULIAHAN_AYAH = Pegawai/Swasta: Tidak (Tidak=2169, Ya=0)
PERKULIAHAN_AYAH = Pegawai Sempit: Tidak (Tidak=1156, Ya=0)
PERKULIAHAN_AYAH = Pensiunan: Tidak (Tidak=576, Ya=0)
PERKULIAHAN_AYAH = TNI / POLRI: Tidak (Tidak=202, Ya=0)
PERKULIAHAN_AYAH = Tidak Bekerja
    IIRK > 3.031: Ya (Tidak=0, Ya=4)
    IIRK > 3.035
        IIRK > 2.747
            PERKULIAHAN_IBU = BUKAN: Ya (Tidak=0, Ya=27)
            PERKULIAHAN_IBU = Guru/Dosen: Tidak (Tidak=63, Ya=0)
            PERKULIAHAN_IBU = Lain-lain: Tidak (Tidak=97, Ya=0)
            PERKULIAHAN_IBU = Pegawai/Swasta: Tidak (Tidak=64, Ya=0)
            PERKULIAHAN_IBU = Pegawai Sempit: Tidak (Tidak=23, Ya=0)
            PERKULIAHAN_IBU = Pensiunan: Tidak (Tidak=11, Ya=0)
            PERKULIAHAN_IBU = Tidak Bekerja: Ya (Tidak=0, Ya=73)
            IIRK > 2.747: Tidak (Tidak=122, Ya=0)
    
```

Gambar 2. Pentotalan jumlah

accuracy: 84.5% (+ 4.82% index: 94.5%)

	Real Total	Real Ya	Class prediction
pred Total	788	3	88.87%
pred Ya	676	1342	72.27%
class result	83.75%	89.91%	

Gambar 3. Akurasi hasil

precision: 73.18% (+ 5.67% index: 72.27%) (positive class: Ya)

	Real Total	Real Ya	Class prediction
pred Total	788	3	88.87%
pred Ya	676	1342	72.27%
class result	83.75%	89.91%	

Gambar 4. Presisi hasil

5) *Recall/ sensitivity*

*Recall/ sensitivity* adalah proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar [6]. Pada penelitian ini *recall* digunakan untuk menampilkan tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi, yang terlihat pada Gambar 5.

recall: 98.85% (+ 0.43% index: 99.85%) (positive class: Ya)

	Real Total	Real Ya	Class prediction
pred Total	788	3	88.87%
pred Ya	676	1342	72.27%
class result	83.75%	89.91%	

Gambar 5. Recall

6) *Area Under Curve (AUC)*

*Area Under Curve (AUC)* dihitung untuk mengukur perbedaan performansi metode yang digunakan [7]. AUC terbagi atas tiga jenis, yaitu:

- a. *AUC Optimistic*
- b. *AUC Pessimistic*
- c. *AUC Neutral*

Pada AUC terdapat kurva *Receiver Operating Characteristic (ROC)* yang digunakan untuk mengekspresikan data *confusion matrix*. Garis horizontal mewakili nilai *false positives (FP)* dan garis vertikal mewakili nilai *true positives (TP)*.

*AUC optimistic* adalah perhitungan pengukuran perbedaan performansi metode yang

digunakan untuk menghasilkan nilai yang sesuai dengan kriteria yang ditentukan. Seperti terlihat pada Gambar 6.



Gambar 6. AUC optimistic

*AUC pessimistic* adalah perhitungan pengukuran perbedaan performansi metode yang digunakan untuk menghasilkan nilai yang tidak sesuai dengan kriteria yang ditentukan. Seperti terlihat pada Gambar 7.



Gambar 7. AUC pessimistic

*AUC neutral* adalah perhitungan pengukuran perbedaan performansi metode yang digunakan untuk menghasilkan perbandingan nilai yang sesuai dengan kriteria dan nilai yang tidak sesuai dengan kriteria yang ditentukan. Seperti terlihat pada Gambar 8.



Gambar 8. AUC neutral

#### D. KESIMPULAN

Berdasarkan keseluruhan hasil penelitian yang telah dilakukan, dapat diperoleh kesimpulan sebagai berikut:

- 1) Proses penentuan mahasiswa penerima beasiswa pada Universitas Andalas dilakukan dengan mengeksekusi data mahasiswa yang telah ditentukan kriterianya, yaitu berdasarkan indeks prestasi mahasiswa dan pekerjaan orangtua sebagai buruh/ petani dan orangtua mahasiswa yang tidak memiliki pekerjaan.
- 2) Proses eksekusi data tersebut kemudian di klasifikasikan oleh *tool RapidMiner* berdasarkan kriteria-kriteria yang telah ditentukan tersebut. Berikut adalah bentuk informasi yang merupakan *output* dari pemrosesan data dengan menggunakan *tool RapidMiner*, yaitu:
  - a. *Tree decision* (pohon keputusan) dalam berbagai tampilan bentuk pohon keputusan dimana *user* dapat memilih sesuai dengan ketertarikannya.
  - b. Pentotalan jumlah mahasiswa yang mendapat beasiswa dilihat berdasarkan pohon keputusan.
  - c. Akurasi yaitu berapa tingkat kedekatan antara nilai prediksi dengan nilai aktual.
  - d. Presisi/ *confidence* merupakan proporsi kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya
  - e. *Recall* merupakan tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi.
  - f. *Area Under Curve* (AUC), yang terbagi atas ; AUC *optimistic*, AUC *pessimistic*, dan AUC *neutral*.

- g. Universitas Andalas dapat lebih mudah menganalisa data mahasiswa yang berhak menerima beasiswa dengan bantuan *tool RapidMiner* yang telah mengelola data tersebut dalam klasifikasi kriteria-kriteria yang ditentukan.

#### REFERENSI

- [1] M. S. Anam, Pengembangan Perangkat Lunak Berbasis OLAP untuk Informasi Penjualan (Studi Kasus PT. Semen Gresik (Persero) Tbk.), Matematika, Institut Teknologi Sepuluh Nopember Surabaya, 2007.
- [2] *Business Intelligence and Analytics: From Big Data to Big Impact*, vol. MIS Quarterly Vol. 36 No. 4, pp. pp. 1165-1188, December 2012.
- [3] P. L. Setiawan, *Rancang Bangun Aplikasi Business Intelligence Berbasis Arsitektur Aplikasi Akuntansi Accurate*, Vols. JUISI, Vol. 01, No. 02, Agustus 2015 .
- [4] P. D. A. Pamungkas, "Jurnal Informastika," *Menentukan Kemungkinan Masuknya Calon Mahasiswa Baru Pada Sebuah Perguruan Tinggi Swasta Menggunakan Teknik Klasifikasi Pohon Keputusan Dengan Aplikasi Rapidminer 5.1*, vol. Vol 15 No 1, Juni 2015.
- [5] *Introducing RapidMiner Academia for Students, Professors and Researchers* .:
- [6] PowersD, "Evaluation: From Precision, Recall, and F-Measure to ROOC, Infomedness, Markedness & Correlation," *Journal of Machine Learning Technologies*, pp. 37-63, 2011.
- [7] F. Gorunescu, *Data Mining Concepts, Model and Techniques*, Berling: Springer, 2011.