PENERAPAN TEKNIK KLASIFIKASI UNTUK KLASIFIKASI PRESTASI AKADEMIK PESERTA DIDIK (STUDI KASUS DI POLITEKNIK LP3I BANDUNG)

Tubagus Riko Rivanthio, S.Kom., M.Kom.
Dosen Program Studi Manajemen Informatika
Politeknik LP3I Bandung
e-mail: riko_rivanthio@yahoo.com

Abstrak : Penelitian ini bertujuan untuk membuat pola untuk klasifikasi prestasi akademik peserta didik berdasarkan variabel yang dimiliki peserta didik. Hasil klasifikasi peserta didik akan menjadi indikator bagi Politeknik LP3I Bandung dalam melakukan pengawasan kepada peserta didik. Metode penelitian yang digunakan adalah CRISP-DM, pemodelan yang digunakan adalah teknik klasifikasi dengan menggunakan algoritma C5.0 yang diolah menggunakan aplikasi SPSS Clementine. Pemodelan menghasilkan 7 pola untuk klasifikasi predikat cumlaude, sangat memuaskan dan memuaskan, serta menghasilkan beberapa informasi yang bermanfaat bagi lembaga sebagai bahan evaluasi dalam meningkatkan kualitas pendidikan.

Kata Kunci : Klasifikasi, Prestasi Akademik, Algoritma C5.0, Data Mining

1. Pendahuluan

1.1. Latar Belakang Masalah

Politeknik LP3I Bandung merupakan lembaga pendidikan tinggi formal jenjang Diploma III (tiga), yang memiliki tujuan untuk menghasilkan lulusan yang profesional, berjiwa mandiri dan berbudi luhur serta berdaya saing dalam penyerapan tenaga kerja di dunia usaha dan industri. Untuk menghasilkan lulusan yang profesional, maka Politeknik LP3I Bandung selalu mengutamakan kualitas dari setiap lulusannya. Tujuan dari Politeknik LP3I Bandung tersebut dapat tercapai jika dilakukan pengawasan terhadap prestasi akademik dari peserta didik. Pengawasan peserta didik pada politeknik LP3I Bandung dilakukan oleh dosen pembimbing akademik. Namun dalam melakukan pengawasannya para dosen pembimbing akademik tidak mempunyai perangkat yang dapat digunakan dalam proses pengawasan. Perangkat tersebut adalah perangkat yang dapat digunakan untuk klasifikasi prestasi akademik peserta didik, sebagai perangkat yang memberikan gambaran kepada Peserta didik tentang hasil akhir prestasi akademik yang akan diperolehnya. Kebutuhan akan perangkat klasifikasi prestasi akademik Peserta didik dapat dipenuhi dengan menggunakan disiplin ilmu data mining. Dimana data mining memiliki satu tugas atau fungsinya adalah klasifikasi, disamping fungsi tugas lainnya yaitu deskripsi, estimasi, prediksi, kluster, dan asosiasi. Klasifikasi dapat dilakukan secara manual dengan menjalankan algoritma yang relevan dengan fungsi ini. Algoritma yang saat ini dikenal luas dalam dunia

data mining adalah C5.0. yang telah dikomputerisasi menjadi salah satu fitur dalam banyak perangkat lunak, salah satunya adalah Clementine. Algoritma C5.0 diharapkan dapat menghasilkan pola klasifikasi prestasi akademik peserta didik, berdasarkan variabel yang dimiliki oleh setiap peserta didik.

1.2. Rumusan Masalah

Dalam latar belakang masalah telah diuraikan bahwa perangkat lunak Clementine akan digunakan untuk menjalankan algoritma C5.0 untuk tujuan klasifikasi prestasi akademik peserta didik. Dalam melakukan tugas klasifikasi prestasi akademik diperlukan data peserta didik sebagai variabel prediktor, dan variabel klasifikasi prestasi akademik sebagai variabel target. Keseluruhan ini nilai variabel ini akan diolah dengan perangkat lunak Clementine dengan menggunakan algoritma C5.0.

Beberapa masalah yang akan muncul bagi pelaksanaan tugas klasifikasi ini adalah:

1. Variabel apa sajakah yang relevan untuk menjadi prediktor pada klasifikasi prestasi akademik peserta didik?
2. Bagaimanakah memodelkan permasalahan klasifikasi prestasi akademik ini dalam perangkat Clementine, sehingga dari variabel prediktor seorang peserta didik, akan dapat diperoleh klasifikasi prestasi akademik peserta didik.
3. Bagaimanakah memanfaatkan model yang dihasilkan bagi pengawasan yang dilakukan oleh Pembimbing Akademik?

2. Landasan Teori

2.1. Pengertian Data Mining

Beberapa pengertian yang data mining yang diperoleh dari berbagai sumber antara lain :

1. “Data mining adalah proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar” (Tan, 2006:2).
2. “Data mining merupakan analisis dari peninjauan kumpulan data untuk menentukan hubungan yang tidak diduga dan meringkas data dengan cara berbeda sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data” (Larose, 2006:2).
3. “Data mining adalah kegiatan yang meliputi pegumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining ini bisa dipakai untuk memperbaiki pengambilan keputusan dimasa depan” (Santosa, 2007:10)
4. “Sebagai Disiplin Ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki” (Susanto, 2010:2)

A. *Cross Industry Standard Process for Data Mining (CRISP-DM)*

Menurut Daniel T. Larose (2006:5), Cross-Industry Standart 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.

Enam fase CRISP-DM (Cross Industry Standard Process for Data Mining) (Larose, 2006:6).

1. Business Understanding Phase (Fase Pemahaman Bisnis)
2. Data Understanding Phase (Fase Pemahaman Data)
3. Data Preparation Phase (Fase Pengolahan Data)
4. Modeling Phase (Fase Pemodelan)
5. Evaluation Phase (Fase Evaluasi)
6. Deployment Phase (Fase Penyebaran)

Gambar 1. Proses Data Mining menurut CRISP-DM

B. Tahapan Penambangan *Data Mining*

Menurut Daniel T. Larose (2006:27) tujuan dari *Data mining* adalah menjadikan data yang bermanfaat dari database yang ada, maka dari itu database perlu menjalani preprocessing, dalam bentuk pembersihan data dan transformasi data. Adapun tahapan tersebut dapat dilihat pada gambar 2. Tahap – tahap tersebut, bersifat interaktif dimana pemakai terlibat langsung atau dengan perantaraan knowledge base.(Larose, 2006), yang diawali dengan Pembersihan data, Integrasi data, Seleksi data, Transformasi data, Proses *mining*, Presentasi pengetahuan

Gambar 2. Tahapan Penambahan Data

C. Definisi Klasifikasi

Classification adalah sebuah model dalam *data mining* dimana, *classifier* dikonstruksi untuk memprediksi *categorical label*, seperti “aman ” atau “beresiko” untuk data aplikasi peminjaman uang; “ya” atau “tidak ” untuk data marketing; atau “*treatment A*”, “*treatment B*” atau “*treatment C*” untuk data medis. Kategori tersebut dapat direpresentasikan dengan nilai yang sesuai dengan kebutuhannya, dimana pengaturan dari nilai tersebut tidak memiliki arti tertentu. (Han, 2006)

Gambar 3 Klasifikasi sebagai pemetaan sebuah himpunan variabel input x ke dalam label kelasnya y

D. Klasifikasi menggunakan *Decision Tree*

Decision Tree Induction adalah pengetahuan dalam bentuk pohon keputusan yang *training tuple* nya telah diberi label kelas. *Decision Tree* bisa disebut juga alat non-parametrik analisis diskriminan, yang dirancang untuk mewakili aturan keputusan dalam bentuk yang disebut pohon biner. Sebuah *Decision Tree* seperti sebuah *flowchart* dengan berstruktur pohon, dimana setiap internal *node* (*non-leaf node*) mendenotasikan sebuah tes pada sebuah variabel, setiap cabang merepresentasikan sebuah hasil dari tes, dan setiap *leaf node* (atau terminal *node*) memegang sebuah label kelas. (han, 2006:291)

E. Algoritma C5.0

Menurut Larose (2006:124), algoritma C5.0 merupakan pengembangan dari algoritma ID3 dan C4.5. Algoritma C5.0 mengkonstruksi pohon keputusan dari sebuah set data *training*, yang berupa kasus-kasus atau record-record (tupel) dalam basis data, dimana tiap record memiliki variabel kontinu atau diskrit atau keduanya.

Pada algoritma C5.0, pemilihan atribut yang akan diproses menggunakan *information gain*. Secara heuristik akan dipilih atribut yang menghasilkan simpul yang paling bersih (*purest*). Jika dalam cabang suatu *decision tree* anggotanya berasal dari satu kelas maka cabang ini disebut *pure*. Kriteria yang digunakan adalah *information gain*. Jadi dalam memilih atribut untuk memecah obyek dalam beberapa kelas harus kita pilih atribut yang menghasilkan *information gain* paling besar. Ukuran *information gain* digunakan untuk memilih atribut uji pada setiap node di dalam tree. Ukuran ini digunakan untuk memilih atribut atau node pada pohon. Atribut dengan nilai *information gain* tertinggi akan terpilih sebagai parent bagi node selanjutnya. Formula untuk *information gain* adalah (Kantardzic M, 2003:185) :

$$I(s_1, s_2, \dots, s_m) = - \sum_{i=1}^m p_i \log_2(p_i) \quad (1)$$

S adalah sebuah himpunan yang terdiri dari s data sampel. Diketahui atribut class adalah m dimana mendefinisikan kelas-kelas di dalamnya, C_i (for $i= 1, \dots, m$), s_i

adalah jumlah sampel pada S dalam class C_i , untuk mengklasifikasikan sampel yang digunakan maka diperlukan informasi dengan menggunakan aturan seperti diatas (Formula *Information Gain*). Dimana p_i adalah proporsi kelas dalam output seperti pada kelas C_i dan diestimasikan dengan s_i/s . Atribut A memiliki nilai tertentu $\{a_1, a_2, \dots, a_v\}$. Atribut A dapat digunakan pada partisi S ke dalam v subset, $\{S_1, S_2, \dots, S_v\}$, dimana S_j berisi sample pada S yang bernilai a_j pada A . Jika A dipilih sebagai atribut tes (sebagai contoh atribut terbaik untuk split), maka subset ini akan berhubungan pada cabang dari node himpunan S . S_{ij} adalah jumlah sample pada class C_i dalam sebuah subset S_j . Untuk mendapatkan informasi nilai subset dari atribut A tersebut maka digunakan formula (Ernawati, 2006:21)

F. Definisi dari variabel yang digunakan

1. Peserta didik : Pengertian peserta didik menurut ketentuan umum UU RI No.20 Tahun 2003 tentang Sistem Pendidikan Nasional adalah anggota masyarakat yang berusaha mengembangkan potensi diri melalui proses pembelajaran yang tersedia pada jalur, jenjang, dan jenis pendidikan tertentu.
2. Prestasi Akademik : Definisi prestasi akademik menurut adalah bukti peningkatan atau pencapaian yang diperoleh seorang siswa sebagai pernyataan ada tidaknya kemajuan atau keberhasilan dalam program pendidikan (Azwar, 2002).
3. Predikat Kelulusan : Berdasarkan KEPMEN NOMOR 232/U/2000 tentang Pedoman Penyusunan Kurikulum Pendidikan Tinggi Dan Penilaian Hasil Belajar Peserta didik Predikat, Pasal 15 (lima belas) ayat 1 (satu) dan 2 (dua) kelulusan terdiri atas 3 tingkat yaitu : memuaskan, sangat memuaskan, dan dengan pujian, yang dinyatakan pada transkrip akademik. IPK sebagai dasar penentuan predikat kelulusan program sarjana dan program diploma adalah:
 - a. IPK 2,00 - 2,75 : memuaskan;
 - b. IPK 2,76 - 3.50 : sangat memuaskan;
 - c. IPK 3.51 - 4,00 : dengan pujian (cumlaude).
4. Kehadiran : Kehadiran peserta didik di sekolah (*school attendance*) adalah kehadiran dan keikutsertaan peserta didik secara fisik dan mental terhadap aktivitas sekolah pada jam-jam efektif di sekolah. (Imron, 1994).
5. Indeks Prestasi Semester : Berdasarkan PERMENDIKBUD RI Nomor 49 Tahun 2014 Tentang Standar Nasional Pendidikan Tinggi, Pasal 23, bahwa hasil penilaian capaian pembelajaran lulusan di tiap semester dinyatakan dengan indeks prestasi semester (IPS).
6. Indeks Prestasi kumulatif : Berdasarkan PERMENDIKBUD RI Nomor 49 Tahun 2014 Tentang Standar Nasional Pendidikan Tinggi Pasal 23, bahwa hasil penilaian capaian pembelajaran lulusan pada akhir program studi dinyatakan dengan indeks prestasi kumulatif (IPK).
7. Tes Psikologi (*psikotest*) : Tes psikologis (*psikotest*) pada dasarnya adalah alat ukur yang obyektif dan dibakukan (distandarisasikan) atas sampel perilaku tertentu. (Anastasia, 1998)
8. Asal Sekolah : Pada sistem pendidikan Indonesia terdapat dua jenis status sekolah yaitu sekolah negeri dan sekolah swasta. Sekolah negeri dan swasta

memiliki lima perbedaan yaitu kepemilikan, pengelolaan, pembiayaan, sumber daya manusia, sarana dan prasarana (Nurkholis, 204).

G. Faktor-faktor yang mempengaruhi Prestasi Akademik

Menurut Kusumaningsih (2009) bahwa faktor yang mempengaruhi prestasi akademik terbagi atas dua tahap, yaitu:

- a. Tahap pengalaman pra perkuliahan dan karakteristik latarbelakang peserta didik.
- b. Tahap pengalaman Perkuliahan

Sedangkan menurut Slameto (2010) ada dua faktor yang mempengaruhi keberhasilan belajar, yaitu :

- a. Faktor internal terdiri dari : Faktor Jasmaniah, Faktor Psikologi yaitu, intelegensi, perhatian, minat, bakat, motivasi, kematangan dan kesiapan, dan Faktor Kelelahan
- b. Faktor Eksternal terdiri dari : Faktor Keluarga, Faktor Sekolah, Faktor Masyarakat,

H. Perangkat Lunak SPSS Clementine

Menurut David Olson (2008), banyak perangkat lunak statistik dan analitik yang tersedia untuk melakukan penggalian data, seperti Enterprise Miner dari SAS, Intelligent Miner dari IBM, PolyAnalist dari Megaputer, Clementine dari SPSS, dan lain sebagainya. SPSS Clementine memungkinkan pengguna untuk melakukan penggalian data dari tahap pembersihan data, mempersiapkan data, melakukan analisis sampai tahap penggunaan model secara cepat dan mudah. Metode yang ditawarkan pun lengkap mulai dari analisis dasar sampai ke tingkat analisis yang kompleks

3. Obyek Dan Metodologi Penelitian

Pada bab ini akan membahas dua pembahasan yaitu obyek penelitian dan metodologi penelitian yang digunakan.

3.1. Obyek Penelitian

Penelitian dilakukan di Politeknik LP3I Bandung, dengan menggunakan data peserta didik angkatan peserta didik 2010 dan 2011 dari semua program studi yang ada di Politeknik LP3I Bandung yaitu Program Studi Administrasi Bisnis, Manajemen Informatika, Akuntansi, Hubungan Masyarakat.

Politeknik LP3I Bandung merupakan salah satu lembaga pendidikan tinggi yang bernaung dibawah bendera LP3I Grup. Kiprahnya dihadirkan sebagai lanjutan program profesi yang telah hadir di dunia pendidikan Indonesia selama kurang lebih 26 tahun. Dengan format "*Vocational based Polytechnic*", maka Politeknik LP3I Bandung memproyeksikan seluruh peserta didiknya dipromosikan untuk magang dan kerja sebelum lulus kuliah sesuai dengan ketentuan yang berlaku.

3.2. Metodologi Penelitian

Adapun tahapan proses *data mining* dengan menggunakan metode CRISP-DM digambarkan pada gambar 4 berikut ini :

Gambar 4. Tahapan Metode Penelitian Menggunakan CRISP-DM

1. **Business Understanding Phase (Fase Pemahaman Bisnis)**

Tujuan dari penelitian ini adalah membuat suatu perangkat yang dapat dijadikan acuan oleh dosen pembimbing akademik dalam melakukan pengawasan prestasi akademik peserta didiknya. Adapun cara yang digunakan untuk membuat perangkat tersebut adalah dengan menggunakan teknik klasifikasi pada *data mining*.

2. **Data Understanding Phase (Fase Pemahaman Data)**

Data yang digunakan pada penelitian ini adalah data akademik dari peserta didik pada angkatan tahun 2010 dan 2011 dari seluruh program studi yang ada di Politeknik LP3I Bandung. Jumlah data peserta didik adalah sebanyak 508 data. Adapun data yang diperoleh terbagi atas dua kategori yang terdiri dari data profil dan data akademik peserta didik.

3. **Data Preparation Phase (Fase Pengolahan Data)**

Langkah selanjutnya adalah fase pengolahan data yang diawali dengan tahap pengumpulan data, menyeleksi dan menganalisa data untuk penentuan variabel yang digunakan, serta melakukan transformasi data

a. **Tahapan Pengumpulan Data**

Data diperoleh melalui database pada sistem informasi akademik Politeknik LP3I Bandung yang terdiri dari : Rata-rata kehadiran peserta didik , Nilai rata-rata *Psikotest* peserta didik , Data Asal Sekolah Peserta didik , Indeks Prestasi Semester Satu peserta didik dan Indeks Prestasi Kumulatif.

b. Pembersihan data

Pembersihan data dilakukan untuk memeriksa data yang tidak konsisten seperti data nilai peserta didik dan data indeks prestasi yang tidak lengkap, atau membuang data yang tidak konsisten seperti data peserta didik yang mengundurkan diri ataupun cuti akademik, data peserta didik yang belum lulus kuliah, data peserta didik yang belum mengikuti *psikotest*.

c. Transformasi data

Algoritma C5.0 membutuhkan *input data* dalam bentuk data numerik. Maka dari itu pada beberapa data kandidat variabel dilakukan proses transformasi kedalam bentuk numerik tertentu.

Tabel 1. Transformasi IPK

IPK	Predikat Prestasi Akademik
≥ 3.50	Cumlaude
3.00 – 3.49	Sangat Memuaskan
≤ 2.99	Memuaskan

Tabel 2. Transformasi Asal Sekolah

Predikat Prestasi	Nilai Angka
Negeri	1
Swasta	0

4. Modeling Phase (Fase Pemodelan)

Tahap selanjutnya adalah fase pemodelan yaitu metode dan pemilihan algoritma yang akan digunakan pada penggalian *data mining* untuk membuat klasifikasi prestasi akademik peserta didik menggunakan algoritma C5.0 pada aplikasi SPSS Clementine. Adapun tahapan penerapan algoritma C5.0 untuk klasifikasi prestasi akademik peserta didik menggunakan SPSS Clementine adalah sebagai berikut :

1. Menyusun data yang diperoleh dari hasil transformasi kedalam aplikasi *Microsoft Excel*
2. Kemudian data diolah menggunakan *SPSS Clementine* versi 10.1
3. Adapun *decision tree* yang diperoleh dari hasil eksekusi algoritma C5.0 adalah sebagai berikut ini seperti yang ditampilkan pada gambar 5.

Gambar 6 *Decision tree* Klasifikasi Prestasi Akademik Peserta Didik

Pohon keputusan (*decision tree*) dari pemodelan klasifikasi prestasi akademik menggunakan algoritma C5.0 menghasilkan beberapa pola/*ruleset* klasifikasi dan informasi. Adapun pola/*ruleset* klasifikasi dan informasi tersebut adalah sebagai berikut :

- Terdapat tujuh pola/*ruleset* untuk klasifikasi prestasi akademik peserta didik yang diperoleh dari hasil pemodelan (penjelasan lebih detail dapat dilihat pada bab analisa hasil penelitian)
- Selain diperoleh pola/*ruleset* dari klasifikasi prestasi akademik peserta didik, terdapat beberapa informasi lainnya yang berguna bagi Politeknik LP3I Bandung, yaitu :
 - Terdapat 20% peserta didik dari 508 peserta didik yang memperoleh predikat memuaskan, artinya terdapat 20% peserta didik akan mengalami kesulitan dalam mendapatkan pekerjaan.
 - Peserta didik yang memperoleh predikat memuaskan yang berjumlah 102, dari jumlah tersebut terdapat 35 peserta didik (6,8% dari 508) yang memiliki nilai rata-rata *psikotest*nya kurang dari atau sama dengan 3,00. Jumlah peserta didik yang memiliki nilai rata-rata *psikotest*nya kurang dari atau sama dengan 3,00.
 - Terdapat 113 peserta didik (22,2% dari 508) yang memiliki rata-rata kehadiran kurang dari atau sama dengan 12,75.
 - Terdapat 93 peserta didik yang memiliki IP semester satu lebih dari 3,41 dan rata-rata kehadiran lebih dari 13,37, dengan nilai rata-rata *psikotest* lebih dari 3,41, dimana peserta didik tersebut diprediksi akan memperoleh predikat cumlaude, namun dari 93 terdapat 5 peserta didik yang akhirnya memperoleh predikat sangat memuaskan.

5. *Evaluation Phase (Fase Evaluasi)*

Tahap evaluasi adalah tahapan untuk mengamati dan menganalisa hasil klasifikasi yang diperoleh dari pemodelan, serta implementasi dari hasil klasifikasi. Selain dari mengamati hasil klasifikasi, pada tahap ini pun dilakukan analisa dari klasifikasi untuk memperoleh informasi-informasi yang berguna untuk digunakan sebagai bahan evaluasi sistem pendidikan dan proses bisnis yang ada di Politeknik LP3I Bandung.

Pada tahap ini juga dilakukan menganalisa tingkat akurasi dari klasifikasi yang dihasilkan. Analisa tingkat akurasi klasifikasi yang dihasilkan dari pemodelan dapat dilihat dengan menggunakan fasilitas komponen *analysis* pada SPSS Clementine.

6. *Deployment Phase (Fase Penyebaran)*

Pada tahap ini merupakan tahapan membuat laporan hasil kegiatan *data mining*. Laporan akhir mengenai informasi yang diperoleh berdasarkan hasil pengamatan dan analisa hasil klasifikasi. Hasil laporan penelitian tersebut dapat pula dijadikan sebagai acuan untuk penelitian selanjutnya seperti penelitian untuk klasifikasi kesesuaian program studi dan konsentrasi yang dipilih oleh peserta didik dengan kemampuan yang dimiliki oleh peserta didik.

4. Analisa Hasil Penelitian

Terdapat tiga pembahasan yang akan dibahas pada bab ini, yaitu analisa dan implementasi dari klasifikasi serta informasi yang diperoleh dari pemodelan klasifikasi prestasi akademik menggunakan *data mining*, sedangkan pembahasan yang ketiga adalah analisa tingkat akurasi pemodelan klasifikasi.

4.1. Analisa pola/*ruleset* untuk klasifikasi dan informasi yang diperoleh dari pemodelan decision tree

Berdasarkan pemodelan decision tree dari pemodelan klasifikasi prestasi akademik menggunakan data mining, diperoleh beberapa analisa dari pola/*ruleset* untuk klasifikasi prestasi akademik dan beberapa informasi yang diperoleh dari hasil pemodelan. adapun hasil analisisnya adalah sebagai berikut :

1. Terdapat tujuh pola/*ruleset* untuk klasifikasi prestasi akademik peserta didik yang diperoleh dari hasil pemodelan, yaitu terdiri dari :
 - a. Jika IP semester satu $> 3,41$ dan rata-rata kehadiran $> 13,37$ kali, serta nilai rata-rata hasil *psikotest* $> 3,41$, maka peserta didik memiliki peluang untuk memperoleh predikat cumlode.
 - b. Jika IP semester satu $>$ dari $3,41$ dan rata-rata kehadiran $> 13,37$ kali, serta jika nilai rata-rata *psikotest* $\leq 3,41$, maka peserta didik memperoleh predikat sangat memuaskan.
 - c. Jika IP semester $>$ dari $3,41$ dan rata-rata $\leq 13,37$ kali, maka peserta didik memiliki peluang untuk memperoleh predikat sangat memuaskan
 - d. Jika IP semester satu $\leq 3,41$, dan jika nilai rata-rata *psikotest* $> 3,00$, serta rata-rata kehadiran $> 12,75$ maka peserta didik memiliki peluang untuk memperoleh predikat sangat memuaskan.

- e. Jika rata-rata kehadiran $\leq 12,75$, namun jika memperoleh IP semester satu diantara 3,08 sampai 3,41, dan jika memiliki nilai rata-rata *psikotest* $> 3,00$, maka peserta didik masih memiliki peluang untuk memperoleh predikat sangat memuaskan.
 - f. Jika IP semester satu $\leq 3,41$, dan jika nilai rata-rata *psikotest* $\leq 3,00$, maka peserta didik akan memiliki peluang memperoleh predikat memuaskan.
 - g. Jika nilai rata-rata *psikotest* $> 3,0$, namun IP semester satu $\leq 3,08$ dan rata-rata kehadiran $\leq 12,75$, maka peserta didik hanya memiliki peluang memperoleh predikat memuaskan.
3. Diperlukan suatu evaluasi sistem pembelajaran pada peserta didik agar jumlah peserta didik yang memperoleh predikat memuaskan dapat diminimalkan sekecil mungkin, agar peserta didik dapat dengan mudah mendapatkan pekerjaan sesuai yang diharapkan.
 4. Sistem penerimaan peserta didik di Politeknik LP3I Bandung tidak melakukan ujian saringan masuk terhadap calon peserta didik, sehingga terdapat peserta didik yang memiliki nilai *psikotest* rendah, dimana peserta didik yang memiliki nilai *psikotest* yang rendah kesulitan dalam memperoleh predikat cumlaude atau sangat memuaskan. Maka dari itu diperlukan penambahan dosen psikologi, agar rasio dosen psikologi dan peserta didik berimbang, sehingga peserta didik yang memiliki nilai *psikotest* rendah akan mendapatkan perhatian khusus dari dosen psikologi
 5. Pengawasan kehadiran peserta didik secara aturan dilakukan oleh dosen pembimbing akademik, namun pada prakteknya dosen pembimbing akademik kurang memiliki peran dalam melakukan pengawasan, sehingga dari 508 peserta didik yang diamati, terdapat 113 peserta didik (22%) yang memiliki rata-rata kehadiran kurang dari 12,75. Selain kurangnya pengawasan dari dosen pembimbing akademik, peran dari aturan minimal 80% untuk kehadiran peserta didik tidak berjalan semestinya, sehingga terdapat mahasiswa yang memiliki kehadiran dibawah 80% dapat mengikuti ujian.
 6. Indeks prestasi (IP) semester satu memiliki peran cukup penting dalam menentukan klasifikasi prestasi akademik, karena semakin besar nilai indeks prestasi pada semester satu, maka semakin mudah peserta didik merencanakan strategi pembelajaran untuk semester berikutnya. Namun, hal tersebut perlu didukung dengan kehadiran dan nilai *psikotest* peserta didik, sedangkan atribut atau variabel asal sekolah tidak memiliki pengaruh sama sekali terhadap klasifikasi prestasi akademik peserta didik.
 7. Berdasarkan pemaparan sebelumnya yaitu terdapat 5 peserta didik yang diprediksi akan memperoleh predikat cumlaude, namun kenyataannya kelima peserta didik tersebut tidak memperoleh predikat cumlaude. Hal ini disebabkan oleh kurangnya konsistensi motivasi belajar peserta didik, sehingga diperlukan perhatian khusus oleh dosen pembimbing akademik agar selalu memberikan motivasi kepada peserta didik bimbinganya.

4.2. Tingkat Akurasi klasifikasi yang Diperoleh

Tingkat akurasi klasifikasi prestasi akademik yang dihasilkan dapat dilihat pada gambar 7 yang diperoleh menggunakan fasilitas komponen analysis pada perangkat lunak SPSS Clementine.

Gambar 7. Tingkat akurasi klasifikasi yang diperoleh

4.3. Implementasi dari Pemodelan Klasifikasi Prestasi Akademik Peserta didik

Adapun implementasi dari hasil pemodelan klasifikasi prestasi akademik peserta didik adalah sebagai berikut :

1. Pola/ruleset untuk klasifikasi prestasi akademik yang dihasilkan dari decision tree dapat digunakan oleh peserta didik atau pun pembimbing akademik untuk memprediksi prestasi akademik peserta didik, yaitu dengan cara dibuatkan aplikasi atau program tentang pola/ruleset untuk klasifikasi prestasi akademik. Aplikasi tersebut dapat dintegrasikan dengan sistem informasi yang dimiliki oleh Politeknik LP3I Bandung, sehingga pembimbing akademik atau pun peserta didik dapat melihat informasi tentang pola/ruleset untuk klasifikasi prestasi akademik peserta didik.
2. Pembuatan laporan tentang data jumlah dan identitas peserta didik yang memiliki nilai psikotest kurang dari atau sama dengan 3,00 dibuatkan laporan secara tertulis dan diberikan khusus kepada dosen psikologi dan Wakil Direktur III Bagian Pemasaran dan Kemahasiswaan, agar dosen psikologi dapat membina peserta didik tersebut, sedangkan untuk Wakil Direktur III diharapkan dapat mengevaluasi sistem penerimaan peserta didik.
3. Pembuatan laporan tentang data jumlah kehadiran peserta didik dan data jumlah peserta didik yang memperoleh predikat cumlaude, memuaskan dan

sangat memuaskan dibuat secara tertulis dan diberikan kepada Wakil Direktur I Bidang Pendidikan, agar Wakil Direktur I dan jajarannya melakukan evaluasi pada sistem pembelajaran peserta didik.

5. Kesimpulan Dan Saran

5.1. Kesimpulan

Dari hasil penelitian yang telah dilakukan pada bab sebelumnya, maka dapat disimpulkan bahwa :

1. Variabel yang relevan untuk menjadi prediktor pada klasifikasi prestasi akademik peserta didik adalah variabel nilai rata-rata *psikotest* peserta didik, variabel rata-rata kehadiran peserta didik, dan variabel nilai indeks prestasi semester satu. Variabel indeks prestasi akademik memiliki peran yang dominan dalam melakukan klasifikasi prestasi akademik.
2. Pemodelan klasifikasi prestasi akademik menggunakan perangkat Clementine terdiri dari tujuh tahapan yaitu :
 - a. Menyusun data yang diperoleh dari hasil tahapan transformasi ke dalam *Microsoft Excel*.
 - b. Memasukkan komponen *Excel* pada *pallette Sources*, komponen tersebut digunakan untuk mengambil data yang telah disusun pada *Microsoft Excel*.
 - c. Memilih komponen *Filter* pada *pallette Fields Ops*, komponen tersebut digunakan untuk memilih variabel yang diperlukan dari data yang telah disusun.
 - d. Melakukan pemilihan variabel yang menjadi prediktor dan variabel yang menjadi target, pemilihan tersebut dilakukan pada komponen *Type*.
 - e. Melakukan pemodelan klasifikasi menggunakan algoritma C5.0, dengan memasukan komponen algoritma C5.0.
 - f. Mengeksekusi algoritma C5.0 agar menghasilkan *decision tree* klasifikasi prestasi akademik peserta didik
 - g. Melakukan analisa terhadap *decision tree* untuk memperoleh pola/*ruleset* klasifikasi prestasi akademik peserta didik.
3. Model yang dihasilkan dapat dimanfaatkan sebagai acuan pembuatan sistem informasi klasifikasi prestasi akademik peserta didik di Politeknik LP3I Bandung. Sistem informasi tersebut dapat digunakan untuk mengawasi prestasi akademik prestasi peserta didik oleh pembimbing akademik.

5.2. Saran

Adapun saran untuk pengembangan lebih lanjut dari penelitian ini adalah:

1. Pengembangan lebih lanjut dari penelitian
 - a. Melakukan ekplorasi data yaitu seperti melakukan pemisahan data peserta didik antara peserta didik angkatan tahun 2010 dan angkatan tahun 2011, atau pemisahan data berdasarkan jurusan atau program studi.

- b. Penambahan variabel yang digunakan sebagai prediktor dalam memprediksi prestasi akademik, misalnya variabel non akademis yang relevan, yaitu seperti variabel finansial orang tua peserta didik atau variabel jarak tempuh peserta didik dari rumah ke kampus.
2. Pengembangan hasil penelitian bagi Politeknik LP3I Bandung
 - a. Diperlukan suatu evaluasi sistem pembelajaran pada peserta didik agar jumlah peserta didik yang memperoleh predikat memuaskan dapat diminimalkan sekecil mungkin, sehingga tidak ada peserta didik yang memperoleh predikat memuaskan, agar peserta didik dapat dengan mudah mendapatkan pekerjaan sesuai yang diharapkan.
 - b. Diperlukan suatu perhatian terhadap sistem penerimaan peserta didik baru, agar peserta didik yang mempunyai nilai psikotest yang rendah mempunyai perhatian khusus dari manajemen.
 - c. Pengawasan kehadiran peserta didik oleh pembimbing akademik dilakukan sesuai dengan aturan yang berlaku pada Politeknik LP3I, sehingga jumlah peserta didik yang memiliki kehadiran dibawah 80% dapat diminimalkan.

Daftar Pustaka

- [1] Anastasia, Anne, dkk, Tes Psikologi, Edisi Bahasa Indonesia, PT.Prenhalindo, Jakarta, 1998
- [2] Azwar, S, Tes prestasi: Fungsi pengembangan pengukuran prestasi belajar Yogyakarta: Pustaka Pelajar, 2002
- [3] Ernawati, Iin. Algoritma C5.0 Dan K-Nearest Neighbor. Skripsi. Bogor : Institut Pertanian Bogor, 2008
- [4] Han, Jiawei, et. all, Data Mining : Concept and Techniques, Morgan Kaufmann Publisher, USA, 2006
- [5] Imron, Ali, Manajemen Peserta Didik / Jilid I, Proyek OPF IKIP Malang, Malang, 1994
- [6] Kantardzic, M. *Data Mining : Concepts, models, method and algorithms*. USA : Jhon Wiley & Sons., 2003.
- [7] KEPMEN NOMOR 232/U/2000 tentang Pedoman Penyusunan Kurikulum Pendidikan Tinggi Dan Penilaian Hasil Belajar Peserta didik Predikat
- [8] Kusumaningsih, Y, Faktor-Faktor yang Mempengaruhi Prestasi Belajar. Tersedia dalam (<http://lontar.ui.ac.id>). (diakses 5 Januari 2015).
- [9] Larose, Daniel. *Discovering Knowledge in Data : An Introduction to data Mining*, 2006
- [10] Nurkholis, *Manajemen Berbasis Sekolah*, Teori dan Praktek, Rosda, Bandung, 2004
- [11] Olson, David, dkk, Pengantar Ilmu Penggalan Data Bisnis, Salemba Empat, Jakarta, 2008

- [12] Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 49 Tahun 2014 Tentang Standar Nasional Pendidikan Tinggi
- [13] Slameto. Belajar dan Faktor-Faktor yang Mempengaruhinya, Rineka Cipta, Jakarta, 2010
- [14] Susanto, Sani, dkk, Pengantar Data Mining Menggali Pengetahuan dari Bongkahan Data, Andi, Yogyakarta, 2010.
- [15] Santosa, Budi. Data Mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis, Graha Ilmu, Yogyakarta, 2007
- [16] Tan, Pang-Ning, et.all, *Introduction to Data Mining*. Pearson Education, 2006.
- [17] Undang-Undang RI No.20 Tahun 2003 tentang Sistem Pendidikan Nasional.