

# ANALISIS REKOMENDASI PEMINATAN MENGUNAKAN METODE DECISION TREE DENGAN ALGORITMA C4.5

Tri Sutrisno<sup>1</sup>, Stefanny Claudia<sup>2</sup>

<sup>1,2</sup> Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara,  
Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia  
E-mail: <sup>1</sup>tris@fti.untar.ac.id, <sup>2</sup>stefannyclaudia96@yahoo.com,

## Abstrak

Perancangan yang dibuat adalah Analisis Rekomendasi Peminatan Menggunakan Metode Decision Tree dengan Algoritma C4.5. Aplikasi yang dibuat juga memiliki fitur konsultasi akademik online yang dapat digunakan oleh mahasiswa dan dosen pembimbing akademik. Tujuan pembuatan aplikasi ini adalah untuk menghasilkan hasil analisis peminatan kajian ilmu untuk mahasiswa Teknik Informatika dalam topik skripsinya dan memudahkan dosen Penasehat Akademik mengarahkan kajian ilmu yang akan ditekuni oleh mahasiswa bimbingan akademiknya. Pengujian aplikasi prediksi menggunakan data nilai mahasiswa angkatan 2013 yang sudah lulus. Dari data nilai tersebut dipilih 32 mata kuliah yang paling berpengaruh terhadap peminatan bidang mahasiswa. Berdasarkan pengujian, aplikasi sudah berjalan, dapat digunakan sebagaimana mestinya, dan menghasilkan hasil analisis rekomendasi peminatan. Hasil analisis divalidasi dengan membandingkan kondisi nyata peminatan mahasiswa. Dengan presentase rata-rata akurasi sebesar 72,6227%.

**Kata kunci :** Decision Tree, Algoritma C4.5

## Abstract

The application created are used to analyze which thesis preference subject suits students academic performance based on their academic grades. The application also provide online academic consultations features which students can use for their academic consultations. To find their thesis preference, the application use decision tree method with C4.5 algorithm. Testing prediction system using students data from 2012 to 2015 who have found their thesis preference. The value data used is 32 mandatory courses in the Faculty of Information Technology before thesis preference. The application can run, use and perform well in accordance with the design made. Testing is to compare the accuracy of the selected tree model build from training data and the thesis preference students have selected. The average accuracy percentage of this a 72, 6227%.

**Keywords :** Decision Tree, Algorithm C4.5

## 1. PENDAHULUAN

Proses pembelajaran di perkuliahan memerlukan kurikulum yang mendukung suatu mata kuliah di masing-masing semester. Kurikulum adalah seperangkat rencana dan pengaturan mengenai tujuan, isi, dan bahan pelajaran serta cara yang digunakan sebagai pedoman penyelenggaraan kegiatan pembelajaran untuk mencapai tujuan pendidikan nasional.

Kurikulum Operasional yang dibuat oleh pihak program studi bertujuan untuk mengetahui distribusi dan prasyarat mata kuliah di setiap semester. Dengan adanya Kurikulum Operasional, setiap mahasiswa dapat mengetahui mata kuliah apa saja yang akan diambil di semester tertentu beserta dengan prasyarat mata kuliah yang harus dipenuhi dengan syarat nilai tertentu. Jika terdapat mahasiswa yang ingin mengambil mata kuliah yang memiliki prasyarat

mata kuliah tertentu, maka mata kuliah yang menjadi prasyarat harus dipenuhi di semester sebelumnya dan telah memenuhi syarat nilai tertentu.

Kegiatan konsultasi akademik di Fakultas Teknologi Informasi Universitas Tarumanagara merupakan pertemuan antara mahasiswa dengan dosen penasehat akademik yang dibutuhkan oleh setiap mahasiswa sebelum melakukan pengisian kartu rencana studi (KRS). Kegiatan konsultasi akademik dengan dosen penasehat akademik bertujuan agar mahasiswa dapat berkonsultasi mengenai masalah perkuliahan dan membantu mahasiswa dalam mengambil rencana studi, karena seringkali mahasiswa kesulitan untuk memilih mata kuliah yang akan diambil pada semester berikutnya.

Pada proses konsultasi antara dosen dan mahasiswa yang dilakukan dengan cara tatap muka langsung, cara ini memang efektif tetapi juga memiliki keterbatasan-keterbatasan seperti keterbatasan waktu dan keterbatasan ruang penyimpanan. Mengetahui banyaknya jumlah mahasiswa setiap angkatan, menjadi tidak efisien bagi satu dosen penasehat akademik untuk memberikan rekomendasi rencana studi yang ingin diambil dan sesuai dengan masalah yang dihadapi oleh mahasiswa.

Pada Perguruan tinggi, mahasiswa diharapkan untuk memilih Fakultas, Program Studi, dan Jurusan demi menyelesaikan studi di jenjang perguruan tinggi. Di Universitas Tarumanagara, Fakultas Teknologi Informasi, Program Studi Teknik Informatika, ditawarkan 5 jenis peminatan yang berfungsi sebagai referensi untuk menyelesaikan skripsi.

Dalam pemilihan peminatan yang ditawarkan oleh Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara, banyak mahasiswa yang tidak memilih peminatan yang sesuai dengan kemampuan akademik mereka sehingga mahasiswa mengalami kesulitan dalam pembelajaran hingga proses pembuatan skripsi sesuai dengan peminatan yang dipilih. Minimnya pengetahuan tentang peminatan, membuat peminatan tidak berfungsi dengan baik. Peminatan seharusnya dapat memudahkan mahasiswa dalam penyelesaian studi, karena peminatan bertujuan agar mahasiswa dapat fokus pada pembelajaran konsentrasi spesifik dari peminatan.

Guna membantu memecahkan persoalan yang dihadapi dalam kegiatan konsultasi akademik, dibuat sebuah sistem yang dapat memberikan analisis rekomendasi peminatan mahasiswa menggunakan metode *decision tree* dengan algoritma C4.5 pada aplikasi konsultasi akademik online berbasis *website*.

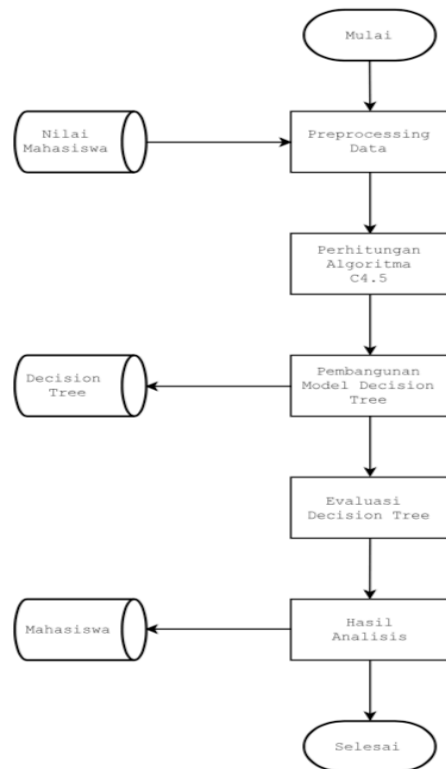
Metode *decision tree* dengan algoritma C4.5 dinilai merupakan metode yang efektif dari *supervised learning*, dimana metode *decision tree* C4.5 merupakan algoritma untuk membangun *decision tree* yang biasa digunakan untuk klasifikasi pada *data mining*. Algoritma C4.5 merupakan algoritma *supervised learning* sehingga algoritma C4.5 membutuhkan *data training* untuk membangun *decision tree* dan membutuhkan *data sample* yang akan digunakan untuk menguji ketepatan dari *tree* yang dibangun.

Aplikasi analisis rekomendasi peminatan menggunakan metode *decision tree* dengan algoritma C4.5, diharapkan mempermudah dosen pembimbing akademik dalam membantu perkembangan studi mahasiswa dan menentukan peminatan yang akan ditempuh berdasarkan data nilai yang mahasiswa peroleh.

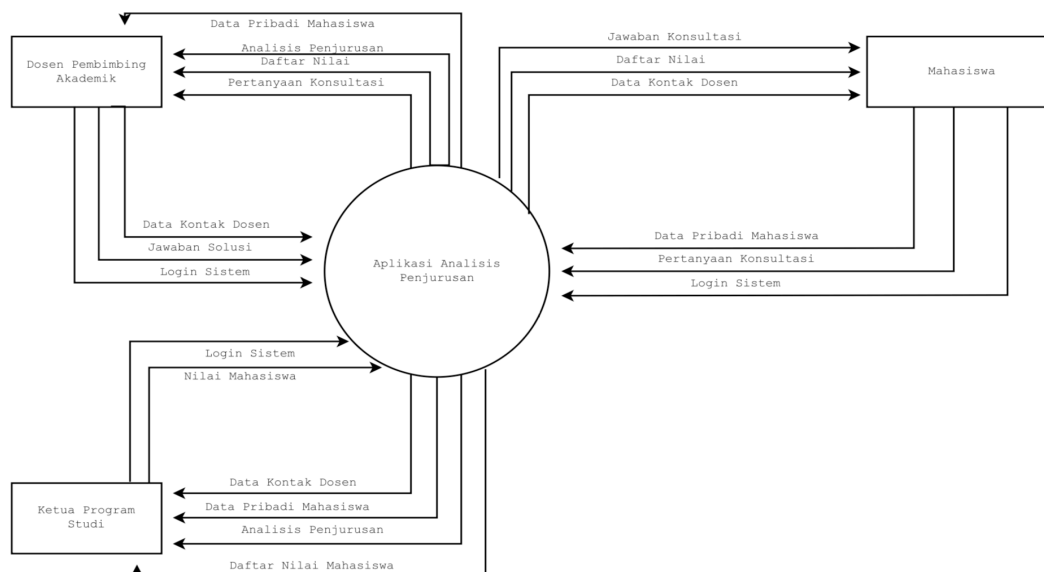
## 2. METODE PENELITIAN

Aplikasi yang dirancang adalah Aplikasi Analisis Rekomendasi Peminatan Menggunakan Metode *Decision Tree* Dengan Algoritma C4.5. Aplikasi ini dirancang berbasis *website* dan akan dibangun menggunakan bahasa pemrograman PHP, dengan basis data dan server menggunakan MySQL yang telah tergabung di dalam XAMPP. Metode perancangan yang akan digunakan dalam pembangunan aplikasi ini adalah metode perancangan SDLC (*System*

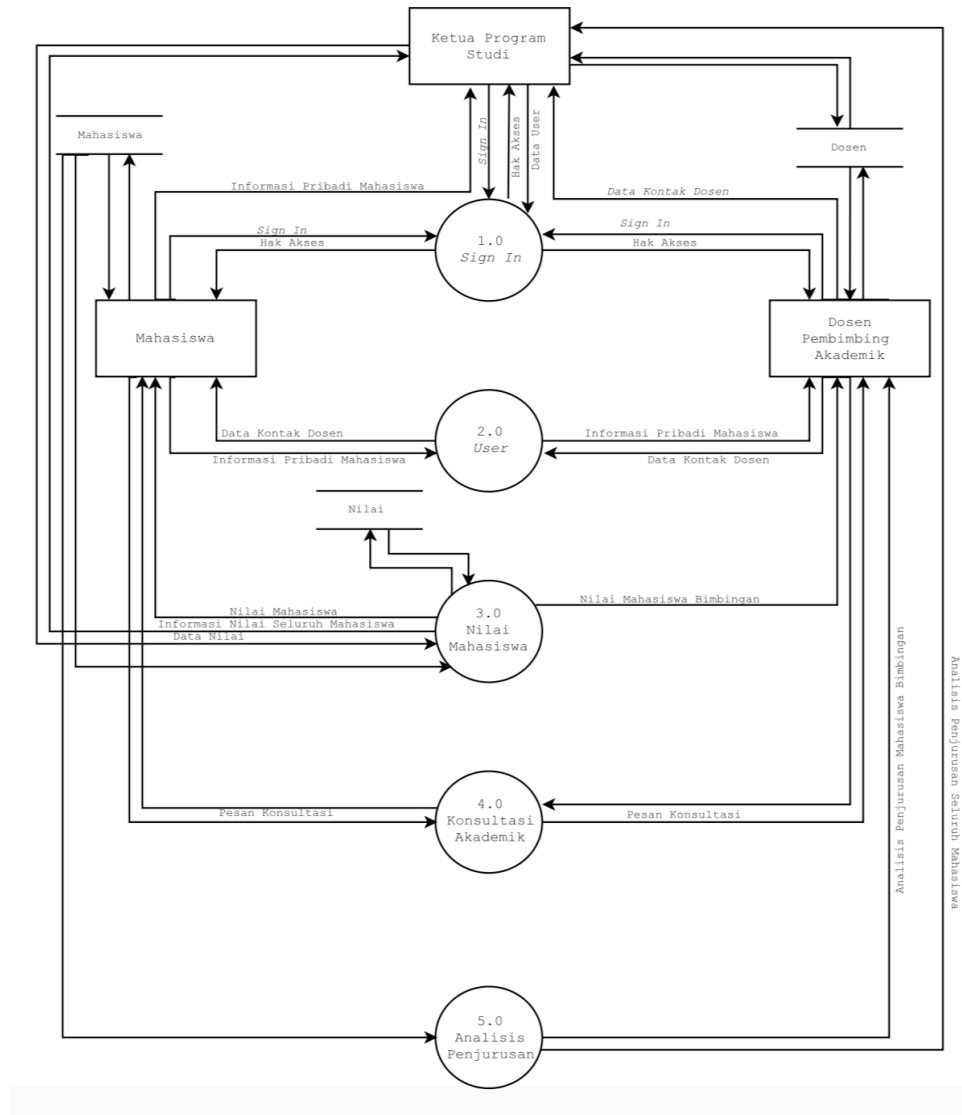
*Development Life Cycle*). Tahapan awal dalam pembuatan aplikasi Analisis Rekomendasi Peminatan Mahasiswa adalah merancang sistem basis data. Rancangan Basis Data. Selanjutnya, dilakukan pembuatan *flowchart* program dan *flowchart* metode *Decision Tree* dengan Algoritma C4.5 berserta dengan *Context Diagram* dan *Data Flow Diagram* level 1 guna membantu penggabungan basis data agar lebih efisien.



Gambar 1 Flowchart Program

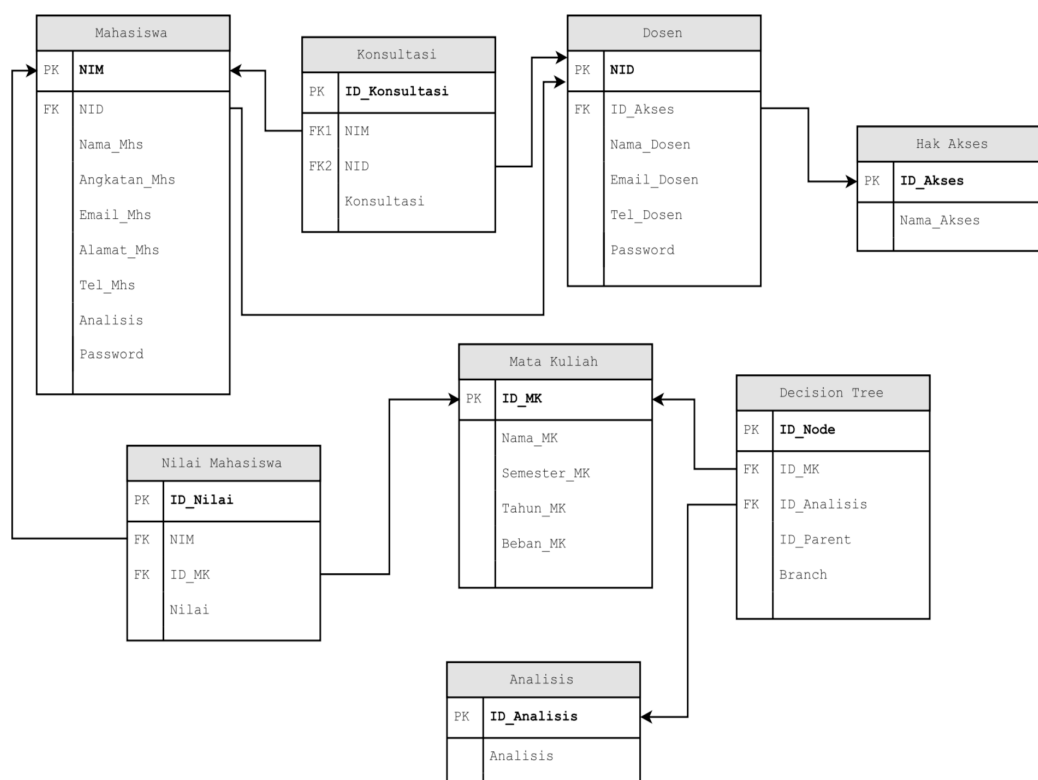


Gambar 2 Context Diagram



Gambar 3 Data Flow Diagram

Lalu dibuat basis data sesuai dengan rancangan menggunakan MySQL yang tersambung pada PHPMyAdmin. Setelah pembuatan basis data, selanjutnya adalah melakukan *preprocessing* terhadap data nilai awal secara manual dengan memilih data mahasiswa yang telah mengambil peminatan dan telah memperoleh nilai minimum “C” pada nilai semester 1 hingga semester 5. Untuk lebih jelasnya berikut diberikan hubungan antar tabel dari basis data yang telah dibuat.



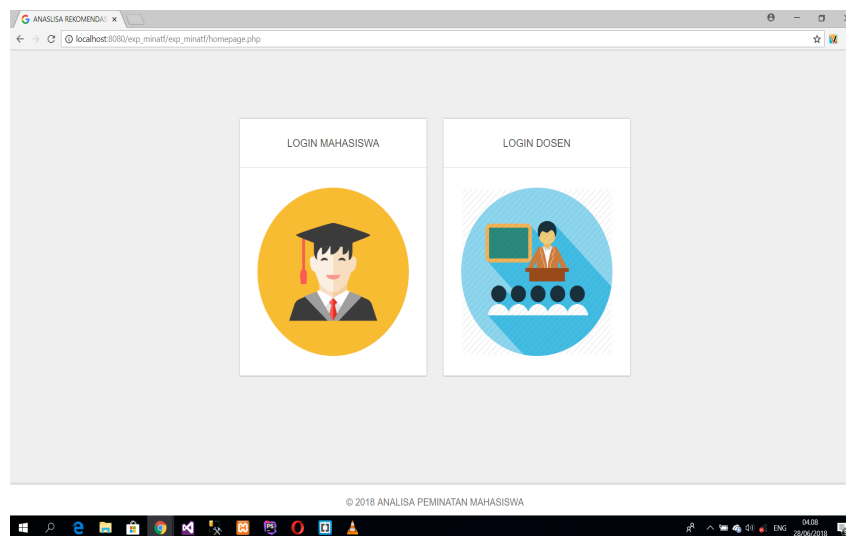
Gambar 4 Hubungan Antar Tabel

Selanjutnya, dilakukan pembuatan kode untuk seluruh fitur aplikasi. Pertama akan dibuat kode untuk login aplikasi, lalu kode untuk *homepage*, tampilan *home* untuk mahasiswa, dosen pembimbing akademik dan ketua program studi. Setelah itu, dilanjutkan dengan pembuatan kode untuk fitur input nilai, input data mahasiswa, dan input data dosen beserta dengan fasilitas *View*, *Edit*, Tambah dan *Delete*. Setelah pembuatan kode diatas selesai, dilanjutkan dengan pembuatan kode untuk fitur utama, yaitu analisis rekomendasi peminatan mahasiswa menggunakan metode *decision tree* dengan algoritma C4.5. Selanjutnya dibuat kode untuk fitur tambahan yaitu fitur konsultasi akademik dimana dosen pembimbing akademik dan mahasiswa dapat bertukar pesan. Seluruh pembuatan kode menggunakan bahasa pemrograman PHP.

### 3. HASIL DAN PEMBAHASAN

Pengujian terhadap program yang telah dibuat dilakukan dengan 2 tahap. Tahap pertama akan dilakukan pengujian terhadap modul, yang bertujuan untuk melihat apakah modul-modul yang telah dibuat dapat berfungsi sesuai dengan rancangan. Tahap kedua akan dilakukan pengujian pada proses perhitungan, bertujuan untuk mengevaluasi proses perhitungan pada program terhadap sejumlah data berhasil mengeluarkan keluaran yang tepat atau tidak.

Pada tahap pengujian modul ini, dilakukan percobaan terhadap keseluruhan modul yang telah dibuat, yaitu modul *Sign In*, Konsultasi Akademik, Daftar Nilai, Analisis Peminatan, dan *Log Out*. Berikut diberikan tampilan utama program aplikasi analisis rekomendasi peminatan menggunakan metode *decision tree* dengan algoritma C4.5.



Gambar 5 Tampilan Utama

Pengujian terhadap data dilakukan terhadap hasil analisis rekomendasi peminatan mahasiswa. Data yang digunakan adalah data nilai 32 mata kuliah mahasiswa angkatan 2012-2014. Pada *data training* dan *data testing*, data yang diambil adalah data mahasiswa yang telah mengambil peminatan tanpa melihat nilai fitur yang digunakan dan jumlah distribusi kelas. Proporsi *data training* dan *data testing* dibagi dengan ratio 70:30 untuk menemukan hasil analisis yang terbaik.

Pengujian akurasi terhadap proporsi data dilakukan untuk mengetahui apakah sistem dapat berjalan sesuai dengan konsep yang diterapkan. Aplikasi Analisis Rekomendasi Peminatan menggunakan Metode *Decision Tree* dengan Algoritma C4.5 memiliki fungsi utama yaitu untuk menunjukkan hasil analisis yang dilakukan dengan beberapa tahap perhitungan. Perhitungan yang dilakukan dimulai dari menghitung entropi, *information gain*, *split information*, dan *gain ratio*. Perhitungan dilakukan menggunakan data nilai mahasiswa semester 1 hingga semester 5, sebanyak 32 mata kuliah. Setelah perhitungan selesai dilakukan, akan dibuat pohon keputusan yang bertujuan sebagai panduan dalam menunjukkan hasil analisis.

Pada pengujian ini dilakukan percobaan 1, percobaan 2, percobaan 3, percobaan 4, dan percobaan 5. Percobaan 1 adalah percobaan menggunakan 177 data nilai mahasiswa angkatan 2012 dan 2014 sebagai *data training* dan 53 data nilai mahasiswa 2013 sebagai *data testing*. Untuk *data training* digunakan, 108 nilai mahasiswa angkatan 2012 dan 69 nilai mahasiswa angkatan 2014. Dari percobaan ini didapatkan hasil presentase akurasi sebesar 83,0189%.

Pada percobaan 2, digunakan 148 data nilai mahasiswa angkatan 2012-2014. *Data training* menggunakan 114 data nilai mahasiswa dengan pembagian 90 data nilai mahasiswa angkatan 2013 dan 24 data nilai mahasiswa 2014. Untuk, *data testing* menggunakan 34 data nilai mahasiswa angkatan 2013, sehingga didapatkan hasil presentase akurasi sebesar 73,5294% dari percobaan 2.

Pada percobaan 3, digunakan 150 data nilai mahasiswa angkatan 2012-2014. *Data training* menggunakan 150 data nilai mahasiswa dengan pembagian 50 data nilai mahasiswa angkatan 2012, 50 data nilai mahasiswa angkatan 2013, dan 50 data nilai mahasiswa 2014. Sedangkan *data testing* menggunakan 15 data nilai mahasiswa angkatan 2012, 15 data nilai mahasiswa angkatan 2013, dan 15 data nilai mahasiswa angkatan 2014. Dari percobaan ini didapatkan hasil presentasi akurasi sebesar 66,6667%.

Pada percobaan 4, digunakan 195 data nilai mahasiswa angkatan 2012-2015, dengan pembagian 60 data nilai mahasiswa angkatan 2012, 50 data nilai mahasiswa angkatan 2013, 50 data nilai mahasiswa angkatan 2014, dan 35 data nilai mahasiswa angkatan 2015 sebagai *data training*. *Data Testing* menggunakan 40 data nilai mahasiswa angkatan 2013 dan 18 data nilai mahasiswa angkatan 2014. Dari percobaan ini didapatkan hasil presentase akurasi sebesar 69,3103%.

Pada percobaan 5, digunakan 170 data nilai mahasiswa angkatan 2012-2015 sebagai *data training*. Pembagian *data training* adalah 60 data nilai mahasiswa angkatan 2012, 76 data nilai mahasiswa angkatan 2013, 34 data nilai mahasiswa angkatan 2015. Sedangkan *data testing* menggunakan 34 data nilai mahasiswa angkatan 2012 dan 17 data nilai mahasiswa angkatan 2014. Dari percobaan ini didapatkan hasil presentase akurasi sebesar 70,5882%.

Dari 5 percobaan yang telah dilakukan didapatkan presentase rata-rata akurasi validasi hasil analisis rekomendasi peminatan mahasiswa adalah sebesar 72,6227%. Untuk lebih jelasnya diberikan tabel pengujian berikut.

Tabel 1 Pengujian Akurasi Terhadap Proporsi Data

Percobaan	Data Training	Data Testing	Benar	Salah	Akurasi
Percobaan 1	177 data 2012-2014	53 data 2012-2014	44	9	83,0189%
Percobaan 2	114 data 2012-2014	34 data 2012-2014	25	9	73,5294%
Percobaan 3	150 data 2012-2014	45 data 2012-2014	30	15	66,6667%
Percobaan 4	195 data 2012-2014	58 data 2012-2014	46	12	69,3103%
Percobaan 5	170 data 2012-2014	51 data 2012-2014	36	15	70,5882%
Rata-rata akurasi					72,6227%

Pada percobaan kedua hingga kelima dilakukan percobaan dengan proporsi jumlah kelas *data training* tidak seimbang, dan hasil analisis menunjukkan hasil akurasi yang berbeda-beda. Hal yang mempengaruhi hasil analisis adalah proporsi kelas *data training* dan model tree yang terbentuk dari *data training*. Presentase rata-rata akurasi dari kelima percobaan sebesar 72,6227%.

#### 4. KESIMPULAN

Kesimpulan yang didapat berdasarkan pembuatan dan pengujian dari aplikasi analisis rekomendasi peminatan menggunakan metode decision tree dengan algoritma C4.5 adalah sebagai berikut :

1. Pada percobaan 1 menghasilkan akurasi sebesar 83,0189% , pada percobaan 2 menghasilkan akurasi sebesar 73,5294%, pada percobaan 3 menghasilkan akurasi sebesar 66,6667%, pada percobaan 4 menghasilkan akurasi sebesar 69,3103% dan pada percobaan 5 menghasilkan akurasi sebesar 70,5882%. Hasil rata-rata akurasi dari seluruh percobaan 1 hingga percobaan 5 memiliki akurasi sebesar 72,6227%.
2. Hasil yang didapatkan dari percobaan yang membandingkan akurasi proporsi jumlah kelas yang sama dengan proporsi jumlah kelas yang berbeda menunjukkan bahwa proporsi jumlah kelas dapat mempengaruhi akurasi dari model tree yang dihasilkan. Model tree yang dihasilkan dengan akurasi tinggi tidak menjadikan model tree tersebut menjadi model yang terbaik jika digunakan untuk data yang baru.

**DAFTAR PUSTAKA**

- [1] Hand, David; Mannila, Heikki. dan Smith, Padhraic. *Principles of data mining*, (London: A Bradford book The MIT Press, 2001).
- [2] Han, Jiawei dan Kimber, Micheline. *Data Mining Concepts and Techniques*. 2nd Edition. San Francisco: Morgan kaufmann Publishers, 2006.
- [3] Karaolis, Minas A.; Moutiris, Joseph A.; Hadjipanayi Demetra.; And Pattichis, Constantinos S. *Assesment Of The Risk Factors Of Coronary Heart Events Based On Data Mining With Decision Trees*, (Crypus: IEEE, 2010).
- [4] Kendall, Kenneth E. dan Kendall, Julie E. *Systems Analysis and Design*, 8<sup>th</sup> Edition, (Upper Saddle River: Prentice Hall, 2006).
- [5] Larose, Daniel T. *Discovering Knowledge in Data: An Introduction to Data Mining*, (Hoboken: John Willey & Sons. Inc, 2005).
- [6] Valacich, Joseph S.; George, Joey F. and Hoffer, Jeffrey A. *Essentials of Systems Analysis and Design*, 4<sup>th</sup> Edition, (Upper Saddle River: Prentice Hall, 2009).