

Memprediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode *Decision Tree* J48 di Universitas Advent Indonesia

¹Yusran Timur Samuel, ²Bern Jonathan, ³Josephine

^{1,2,3} Fakultas Teknologi Informasi, Universitas Advent Indonesia

e-mail: ¹ y.tarihoran@unai.edu, ² 1581009@students.unai.edu, ³ josephinenaibaho@yahoo.com

Abstrak

Setiap perguruan tinggi, pasti menginginkan mahasiswanya lulus tepat waktu. Salah satu alasannya agar akreditasi perguruan tinggi tersebut dapat meningkat. Namun pada kenyataannya ada mahasiswa yang tidak dapat menyelesaikan pendidikannya tepat waktu. Untuk itu, penulis bermaksud untuk meneliti bagaimana membuat sistem yang dapat memprediksi apakah seorang mahasiswa dapat lulus tepat waktu menggunakan metode *decision Tree J48*, agar dapat ditentukan tindakan yang dapat menolong mahasiswa tersebut selesai tepat waktu jika sekiranya mahasiswa tersebut memiliki kecenderungan untuk tidak tamat tepat waktu. Melalui kajian literatur dan wawancara, maka didapatkan 19 atribut yang dapat mempengaruhi ketepatan waktu lulus di Universitas Advent Indonesia. Dalam pengolahan data, penulis menggunakan aplikasi WEKA dengan mengimplementasikan algoritma J48. Dari hasil penelitian diperoleh akurasi sebesar 90,24%, dapat disimpulkan bahwa algoritma ini memiliki tingkat akurasi yang cukup baik untuk melakukan prediksi mahasiswa tamat tepat waktu.

Hasil Penelitian mempunyai kesimpulan bahwa cuti adalah *root* teratas dari atribut mahasiswa tamat tidak tepat waktu. Sementara lulus mata kuliah, lingkungan yang mempengaruhi, pelunasan biaya kuliah, pernah disiplin, kerajinan mengerjakan tugas, bekerja di dalam lingkungan universitas, jurusan sesuai dengan peminatan SMA, jumlah organisasi, dan pengaruh lingkungan adalah atribut pendukung, bisa di gunakan untuk memprediksi.

Kata Kunci : *Data Mining, Decision Tree, Prediksi Tepat Waktu*

Predicting Timely Students Graduation Using the Decision Tree J48 Method at Universitas Advent Indonesia

Abstract

Every universities wants its students to graduate on time. The reason is that the accreditation of the universities may increase. But in reality, there are students who cannot complete their education on time. Therefore, the author wants to examine how to make a system that can predict whether a student can graduate on time using the J48 decision tree method, so that actions can be taken to help the student finishes their studies on time if they tend not to graduate on time. Through the interviews and literature studies, 19 attributes were found that could affect the timeliness of the students graduation at Universitas Advent Indonesia. To process the data, the author uses the WEKA application by implementing the J48 algorithm.

Results of the study has obtained an accuracy of 90,24%, it can be concluded that this algorithm has a fairly good level of accuracy to make predictions of students graduating on time. The results of the study have concluded that academic leave is main reason for the students not to graduate on time. While passing the classes, environmental influences, financial difficulties, discipline problems, perseverance in doing assignments, working as student labor in the university, the linearity of studies in accordance with high school specialization, numbers of students' organizations they followed, and the friendship influences are supporting attributes that can be used to predict.

Keywords: *Data Mining, Decision Tree, Timely Prediction*

1. PENDAHULUAN

Kelulusan tepat waktu mempunyai manfaat baik bagi universitas maupun bagi mahasiswa sendiri. Bagi universitas, mahasiswa yang tamat tepat waktu atau lebih cepat dari target masa studi yang sudah ditetapkan, memiliki nilai yang positif untuk meningkatkan nilai akreditasi dari BAN PT di mana salah satu butir penilaiannya adalah banyaknya mahasiswa yang dapat menyelesaikan program studinya tepat waktu atau lebih cepat dari waktu yang telah ditetapkan. Selain keuntungan bagi universitas, tentunya lulus tepat waktu menjadi harapan mahasiswa yang berkuliah di suatu universitas atau perguruan tinggi karena dapat lebih menghemat biaya perkuliahan, kesempatan mendapat lapangan kerja yang lebih baik, dan keuntungan-keuntungan lainnya. Untuk itu, universitas akan mengusahakan berbagai cara agar para lulusannya dapat tamat tepat waktu. Salah satu usaha agar para lulusan dapat tepat waktu adalah meningkatkan mutu pendidikan di universitas tersebut.

Universitas Advent Indonesia adalah Universitas yang memiliki visi menjadi universitas yang unggul di bidang pengajaran, penelitian, dan pengabdian kepada masyarakat, berdasarkan falsafah pendidikan Kristen Advent se-Asia Tenggara pada tahun 2020. Salah satu tekad yang di ambil yaitu secara intelek membatu para mahasiswa untuk memperoleh ketrampilan berkomunikasi serta menyanggulkan mahasiswa membuat keputusan-keputusan berdasarkan fakta. Program pendidikan yang dijalankan untuk memperluas pandangan mahasiswa dan untuk menantang setiap mahasiswa untuk mencapai keunggulan pendidikan.

Pencapaian akademik adalah istilah untuk menunjukkan suatu pencapaian tingkat keberhasilan tentang suatu tujuan karena suatu usaha belajar telah dilakukan oleh seseorang secara optimal. Kinerja akademik seorang mahasiswa dapat dilihat dari hasil proses belajar di perguruan tinggi sehingga pencapaian tingkat keberhasilan akademik mahasiswa dapat dipengaruhi mutu dari pendidikan perguruan tinggi. Dan pendidikan mutu yang tinggi juga dilihat dari presentasi kelulusan yang tepat waktu dari perguruan tersebut.

Mahasiswa yang datang di Universitas Advent Indonesia berasal dari berbagai latar belakang sehingga tidak semua mempunyai kemampuan yang sama dalam menerima proses pendidikan di Universitas Advent Indonesia. Untuk itu perlu di prediksi kecepatan kelulusan seseorang agar dapat diidentifikasi, siapa saja yang memiliki potensi tidak tamat tepat waktu, sehingga dapat diberikan perhatian khusus kepada mahasiswa yang memiliki potensi tidak tamat tepat waktu. Mengenai perhatian yang bagaimana yang harus diberikan kepada mahasiswa yang berpotensi tamat tidak tepat waktu, bukanlah merupakan bagian dari penelitian ini. Penelitian ini hanya berfokus pada bagai mana membangun sebuah sistem yang dapat memprediksi apakah seseorang akan tamat tepat waktu atau tidak

Klasifikasi adalah teknik *data mining* untuk memproses penempatan suatu objek atau konsep ke dalam satu set kategori berdasarkan objek atau konsep yang bersangkutan dan untuk menemukan suatu model atau fungsi yang menggambarkan dan membedakan kelas data atau konsep dengan tujuan dapat menggunakan model untuk membuat prediksi kelas objek di mana kelas labelnya tidak diketahui. Berdasarkan uraian di atas maka penulis tertarik mengadakan penelitian dengan menggunakan metode *data mining* untuk memprediksi apakah seorang mahasiswa yang berada pada tahap awal perkuliahan di Universitas Advent Indonesia akan tamat tepat waktu atau tidak.

2. LANDASAN TEORI

Data Mining

Dalam buku yang ditulis oleh Charu C. Aggarwal menjelas bahwa *data mining* adalah studi tentang mengumpulkan, pembersihan, pengolahan, analisis, dan mendapatkan wawasan yang berguna dari data. *Data mining* juga merupakan sebuah istilah umum yang digunakan untuk menggambarkan berbagai aspek pengolahan data[1]. *Data mining*, sering juga disebut sebagai *Knowledge Discovery in Database* (KDD).

KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [2].

Decision Tree J48

Secara umum Algoritma J48 biasanya digunakan untuk masalah klasifikasi. Sebuah *decision tree* terdiri dari beberapa simpul yaitu *tree's root*, internal *node* dan *leafs*. Konsep entropi digunakan untuk penentuan pada atribut mana sebuah pohon terbagi (*split*). Semakin tinggi *entropy* sebuah sampel, semakin tidak murni sampel tersebut.

Rumus yang digunakan untuk menghitung *entropy* sampel S adalah;

$$Entropy(S) = -p1\log_2p1 - p2\log_2p2 \quad [2]$$

Di mana: p1 adalah proporsi sampel atau grup 1 yang dipasangkan dengan proporsi p2 proporsi grup 2. Maka setelah nilai dari *Entropy* ditemukan atau sudah diperoleh, maka langkah selanjutnya adalah menghitung *gain* untuk mengukur efektifitas suatu atribut yang nantinya digunakan dalam mengklasifikasikan data. Maka dengan itu, *Gain* dapat dihitung dengan menggunakan rumus sebagai berikut:

$$Gain(S, A) = Entropy(s) - \sum_{v \in values(A)} \frac{sv}{S} * Entropy(Sv)$$

Gambar 1: Rumus Gain Entropy [3]

Di mana v adalah suatu nilai yang mungkin untuk atribut A, *values(A)* menyatakan himpunan nilai-nilai yang mungkin untuk atribut A. Sv menyatakan jumlah sampel data untuk nilai v, dan S merupakan jumlah seluruh dari sampel data, sedangkan *entropy(Sv)* merupakan *entropy* untuk sampel-sampel yang memiliki nilai v.

Weka

Weka adalah aplikasi *data mining open source* berbasis Java. Aplikasi ini dikembangkan pertama kali oleh Universitas Waikato di Selandia Baru sebelum menjadi bagian dari Pentaho yang merupakan sebuah Perusahaan aplikasi yang bergerak dalam bidang analisis bisnis, *big data* dan *data mining*. Perusahaan Pentaho didirikan pertama kali pada tahun 2004 di Orlando, Amerika Serikat[4].

ARFF (Attribute-Relation File Format)

ARFF *file* merupakan sebuah teks ASCII yang menggambarkan daftar contoh berbagi set atribut. *File* ARFF dikembangkan oleh sebuah Proyek Belajar Mesin di Departemen Ilmu Komputer dari Universitas Waikato, Selandia Baru untuk digunakan pada perangkat lunak atau aplikasi, yaitu Weka. *File* ARFF memiliki dua bagian yang berbeda. Bagian yang pertama adalah Informasi *Header* yang merupakan informasi data. *Header file* ARFF berisi nama relasi, daftar atribut (kolom dalam data) [4].

Algoritma J48 pada WEKA

J48 merupakan implementasi C4.5 pada WEKA. J48 menangani himpunan data dalam format ARFF, tidak mengandung kode untuk mengkonstruksi pohon keputusan. Kelas ini mereferensi kelas-kelas lain, kebanyakan pada paket Weka. *Classifiers* J48, yang mengerjakan semua proses konstruksi pohon. Adapun kelebihan C4.5 antara lain:

- C4.5 mampu menangani atribut dengan tipe diskrit atau kontinu.
- C4.5 mampu menangani atribut yang kosong (*Missing Value*).

C4.5 telah berkembang menjadi C5. C4.5 merupakan pengembangan dari ID3 (*Iterative Dichotomiser* 3). ID3 merupakan algoritma yang dipergunakan untuk membuat sebuah *decision tree* atau pohon keputusan. Algoritma ID3 ditemukan oleh J. Ross Quinlan, dengan memanfaatkan teori informasi milik Shanon. Idenya adalah membuat pohon dengan percabangan awal yang dapat membagi antara iya dan tidak. Pada WEKA ada pilihan untuk menggunakan ID3 dengan nama yang sama dan atributnya harus bertipe nominal dan tidak boleh terdapat atribut yang kosong [5].

3. METODE PENELITIAN

Desain Penelitian

Penelitian dilakukan dengan menggunakan data primer (kuesioner *online*) menggunakan *google docs*, yaitu data yang dikumpulkan selama penelitian mengenai fakta-fakta dari objek yang diteliti. Subyek yang diteliti adalah mahasiswa yang tamat pada tahun ajaran 2012- 2017. Responden akan mengisi kuesioner berdasarkan atribut yang didapatkan dari literatur maupun wawancara dengan ahli psikolog.

Kemudian data responden akan diolah dengan aplikasi WEKA menggunakan metode *Decision Tree* untuk mengimplementasikan algoritma J48 pada program Data Mining. Gambar berikut ini merupakan blok diagram dari tahap-tahap penelitian.



Gambar 1 Blok Diagram Penelitian

Menetapkan Atribut

Berdasarkan studi literatur dan wawancara kepada berbagai pihak maka ditetapkan beberapa atribut yaitu faktor-faktor yang mempengaruhi cepat lambatnya seorang mahasiswa dalam menyelesaikan studi. Adapun faktor-faktor tersebut dibagi ke dalam dua kelompok besar yaitu faktor eksternal dan faktor internal. Faktor Eksternal meliputi Problem kemampuan ekonomi keluarga dan faktor internal meliputi :

- 1) Problem inteligensi
- 2) Problem kepribadian

- 3) Problem sifat
- 4) Problem adaptasi (menyesuaikan diri)
- 5) Problem motivasi
- 6) Problem minat sekolah [6].

Berdasarkan faktor-faktor di atas maka penulis membuat atribut untuk Memprediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode *Decision Tree* J48 Di Universitas Advent Indonesia adalah sebagai berikut:

1. **Jenis Kelamin.** Berdasarkan data yang dimiliki oleh Pembantu Rektor Bidang Kemahasiswaan Jenis Kelamin yang rata-rata telah tamat adalah mahasiswa yang berjenis kelamin pria.
2. **Agama.** Lingkungan Mayoritas di Universitas Advent Indonesia adalah Advent, menurut pendapat Pembantu Rektor Bidang Kemahasiswaan Lingkungan Mayoritas ini mempengaruhi pergaulan dalam perkuliahan. Dari pendapat tersebut Penulis membuat atribut pilihan apakah agama adalah Advent atau Non-Advent.
3. **Jenjang Pendidikan.** Berdasarkan Pendapat Pembantu Rektor Bidang Kemahasiswaan Jenjang Pendidikan yang diambil juga mempengaruhi seseorang lulus tepat waktu atau tidak, karena jenjang pendidikan D3 melakukan lebih banyak praktek daripada belajar di kelas dibandingkan dengan jenjang Pendidikan S1 dan S2.
4. **Jurusan Peminatan Saat SMA/K.** Jurusan peminatan saat SMA/K, menurut ahli psikologi yang penulis wawancara, juga mempengaruhi akan menjadi apa seseorang saat dia berkuliah. Banyak sekolah menengah kejuruan juga yang memang mempersiapkan siswanya untuk siap di pekerjaan. Dan juga Peminatan IPA cenderung lebih memahami konsep untuk menggunakan logika.
5. **Jurusan Peminatan saat SMA/K yang sesuai dengan Jurusan saat Kuliah.** Ahli Psikologi yang penulis wawancara juga mengatakan kesesuaian jurusan peminatan saat SMA/K juga mempengaruhi mahasiswa karena saat peminatan saat SMA/K sesuai dengan jurusan yang di ambil saat kuliah, mahasiswa lebih mudah untuk menekuni bidang yang ia sudah pelajari juga saat SMA/K.
6. **Pernah Pindah jurusan Saat Kuliah.** Pindah jurusan dapat mengakibatkan terlambatnya seseorang menyelesaikan perkuliahannya. Karena harus mengambil mata kuliah tambahan untuk menyesuaikan dengan jurusan baru yang di ambilnya. [7].
7. **Berkeluarga.** Ketika suami atau istri ada di samping dan memberikan dukungan harusnya ini tidak membuat kuliah itu lamban, atau jika sudah punya anak, punya figur orang tua yang gigih adalah kebanggaan anak dan tentunya menjadi panutan buat dia dimasa depan, walau memang tidak dipungkiri pasti ada kendal, karena harus pandai membagi dan menggunakan waktu sehari-hari dengan baik [7].
8. **Bekerja.** Kuliah sambil bekerja, sambil kerja memang berat, tapi kalau memang keadaan menghendaki itu mestinya dijadikan sebagai sebuah motivasi dan dorongan untuk cepat selesai kuliah. Di berbagai belahan negara dan benua ini berapa banyak orang-orang yang berhasil kuliahnya, malah ada yang mencapai gelar Doktor, mereka tidak menjadikan kuliah sambil kerja itu sebagai suatu beban, tapi motivasi dan meningkatkan optimisme. Dan di berbagai kalangan mahasiswa dan orang muda sekarang malah berusaha mencari kesibukan untuk mendapatkan tambahan uang di sela-sela kuliahnya [7]. Dalam hal bekerja penulis membuat 2 kelompok bekerja yaitu bekerja di luar UNAI dan juga bekerja sebagai *student labor* di UNAI. Karena rata-rata yang penulis amati mahasiswa D3 dan juga S1 lebih cenderung untuk bekerja sebagai *student labor* di UNAI dan rata-rata mahasiswa S2 bekerja di luar UNAI.
9. **Banyaknya Organisasi.** Organisasi yang terlalu aktif dan mempengaruhi mahasiswa lulus tepat waktu cenderung orang yang terlalu aktif dalam berorganisasi dan juga banyak terlalu sibuk untuk mengurus tugas kuliah dan lebih suka untuk berorganisasi [7].
10. **Mengerjakan Tugas Tepat Waktu.** Tepat waktu atau tidak dapat mengindikasikan apakah seseorang pemalas atau tidak. Bila seseorang mengerjakan tugas tidak tepat waktu maka seseorang itu memiliki sifat malas yang menjadi penyebab keterlambatan dalam menyelesaikan

studi, karena mereka akan malas belajar, malas mengikuti perkuliahan, malas mengurus nilai mereka, malas mengurus perkuliahan mereka, dan itu menjadi penyebab mengapa mereka terlambat dalam menyelesaikan studinya[6]. Menurut ahli Psikologi yang penulis wawancara, untuk melihat bagaimana sifat seseorang rajin atau tidak adalah dengan melihat bagaimana ia mengumpulkan tugas dengan tepat waktu atau tidak. Karena bila mahasiswa tersebut rajin maka cenderung mahasiswa lebih memiliki keinginan untuk menyelesaikan studinya. Untuk kategori mengerjakan tugas tepat waktu penulis menggunakan skala Likert dengan 5 pilihan opsi: Selalu, Sering, Kadang-kadang, Jarang, Tidak Pernah.

11. **Mudah Terpengaruh dengan lingkungan baru.** Faktor terpenting ketika memasuki kehidupan baru dan lingkup baru adalah penyesuaian diri atau adaptasi dengan lingkungan yang baru. Karena setiap lingkungan mempunyai keadaan yang berbeda-beda. Begitu pula menjadi seorang mahasiswa, dia juga harus bisa beradaptasi dengan lingkungan yang baru. Beradaptasi dengan kampus, adaptasi dengan akademik, adaptasi dengan proses perkuliahan, adaptasi dengan sistem kampus dan adaptasi dengan segala hal yang berhubungan dengan proses pendidikan di kampus. Karena sistem yang ada di pendidikan kuliah, berbeda dengan sistem perkuliahan sewaktu SMA [6].
12. **Pernah Di Disiplin.** Skorsing atau disiplin dari perkuliahan juga mempengaruhi lama tamat mahasiswa atau apakah akan tamat tepat waktu atau tidak, karena menurut pendapat Pembantu Rektor Bidang kemahasiswaan rata-rata mahasiswa yang terkena disiplin atau skorsing pasti akan tamat tidak tepat waktu.
13. **Pernah Cuti Kuliah.** Pernah cuti kuliah juga mempengaruhi mahasiswa untuk tamat tepat waktu atau tidak tepat waktu. Bila pernah beberapa kali cuti kuliah, mahasiswa pasti akan tamat tidak tepat waktu.
14. **Dominan tinggal di Asrama.** Menurut pendapat Pembantu Rektor bidang kemahasiswaan ada sebagian kecil mahasiswa yang tamat tidak tepat waktu yang tinggal *outside* karena dengan peraturan-peraturan yang di buat di asrama akan membuat mahasiswa lebih rajin, dan juga lebih ada keinginan untuk berkuliah karena jarak gedung perkuliahan dan juga asrama yang cukup dekat dan kecenderungan untuk mengikuti peraturan perkuliahan lebih tinggi.
15. **Berasal dari lingkungan yang tidak mementingkan kuliah.** Menurut Pendapat Ahli Psikologi yang penulis wawancara ada sebagian orang yang bila berkuliah saja sudah cukup memuaskan bagi dirinya dan adalah suatu kebanggaan bagi keluarga bila salah satu anggota keluarganya bisa berkuliah. Hal ini juga bisa mempengaruhi tepat waktu atau tidaknya seseorang berkuliah, karena bisa membuat suatu pola pikir untuk berkuliah saja, dan tidak usah memikirkan waktu tamatnya telat atau tidak.
16. **Orang Tua pegawai Mission.** Menurut Pembantu Rektor bidang kemahasiswaan mahasiswa yang orang tuanya adalah pegawai *Mission* akan mendapatkan bantuan biaya kuliah 60%. Hal ini berpengaruh terhadap tamat tepat waktu karena faktor keuangan juga mempengaruhi hal tersebut.
17. **Lunas Pembayaran saat berkuliah.** Menurut Pembantu Rektor bidang kemahasiswaan, mahasiswa yang pembayarannya tidak lunas tidak boleh ikut dalam ujian tengah atau akhir semester, hal ini dapat mengakibatkan gagalnya mahasiswa tersebut karena tidak bisa mengikuti ujian dan akibatnya tersebut harus mengulang mengambil mata kuliah tersebut. Hal ini tentunya mempengaruhi waktu tempuh perkuliahannya.
18. **Tidak Memenuhi Standar Kelulusan Mata Kuliah Saat Tingkat I.** Bila Mahasiswa tidak memenuhi standar kelulusan mata kuliah pada tingkat I, maka mahasiswa tersebut harus mengambil ulang mata kuliah tersebut kembali. Ini akan mengakibatkan tertundanya waktu untuk menamatkan perkuliahannya.
19. **IPK tingkat I.** Informasi ini digunakan untuk memprediksi apakah seorang mahasiswa akan tamat tepat waktu atau tidak. Pembagian IPK yang penulis buat terdiri dari 3 kategori yaitu IPK <2.00, 2.00-3.25, dan 3.25>4,00. [8]

Beberapa faktor yang perlu dipertimbangkan dalam penggunaan aplikasi ini adalah :

1. Menentukan algoritma pelatihan yang digunakan dalam hal ini penulis menggunakan algoritma J48.
2. Mengatur opsi data yang akan digunakan sebagai *data training* atau *training set*. Dalam hal ini penulis meneliti dengan menggunakan *Use training set* yaitu menggunakan keseluruhan data untuk proses *training*, serta *percentage split* 80:20 yaitu menggunakan 80% untuk *training* dan 20% untuk *testing*. Dan yang terakhir dengan menggunakan *supplied test*.
3. Menentukan atribut apa saja yang akan digunakan pada proses pelatihan. Dalam hal ini semua atribut digunakan untuk proses ini.

Dalam penelitian ini, penulis melakukan ujicoba terhadap tiga metode pengujian *training dan testing set*, yaitu *use training set*, *percentage split* dan metoda *supplied test*.

Pada metode *use training set*, semua data digunakan untuk proses *training* dan testing, hasil uji coba dapat dilihat pada gambar 4 berikut ini:

```

=== Summary ===
Correctly Classified Instances      190      92.233 %
Incorrectly Classified Instances    16       7.767 %
    
```

Gambar 4 Hasil Akurasi Semua Data Digunakan untuk Proses *Training* dan *Testing*

Hasil pengolahan seperti gambar 4 di atas, terdapat 190 data yang di prediksi benar dan 16 data yang di prediksi salah. Maka dapat disimpulkan bahwa metode ini memiliki nilai akurasi sebesar 92,233%.

Bila menggunakan metode yang kedua yaitu *Percentage Split*, 80% data digunakan sebagai *data training* dan 20% data sebagai *data testing* hasilnya adalah seperti gambar 5 berikut :

```

=== Summary ===
Correctly Classified Instances      37      90.2439 %
Incorrectly Classified Instances     4       9.7561 %
    
```

Gambar 5 Hasil Menggunakan *percentage split* 80:20

Dari hasil yang terlihat pada Gambar 5 di atas terdapat 41 data yang di *testing* dan dari data tersebut 37 data diprediksi benar sesuai dengan hasil yang ada pada data, dan 4 data di prediksi salah. Dengan demikian tingkat akurasi untuk *Cross-Validation* 80% adalah 90,24%.

```

@relation 'Data Training'
@attribute 'Jenis Kelamin' {Laki-laki,Perempuan}
@attribute 'Jenjang' {D3,S1,S2,Mens}
@attribute 'Jurusan Sma' {IPA,IPS,Kejuruan}
@attribute 'Jurusan Sesuai' {Ya,Tidak}
@attribute 'Pindah Jurusan' {Ya,Tidak}
@attribute 'Berkeluarga' {Tidak,Ya}
@attribute 'Bekerja Luar' {Ya,Tidak}
@attribute 'Bekerja di Dalam' {Ya,Tidak}
@attribute 'Organisasi numeric'
@attribute 'Mengerjakan Tugas' {Sering,Selalu,Kadang-kadang,Jarang}
@attribute 'Terpengaruh' {Tidak,Ya}
@attribute 'Disiplin' {Tidak,Ya}
@attribute 'Cuti' {Tidak,Ya}
@attribute 'Lingkungan Teman' {Ya,Tidak}
@attribute 'Lingkungan' {Ya,Tidak}
@attribute 'Anak Mission' {Ya,Tidak}
@attribute 'Lunas' {Ya,Tidak}
@attribute 'Lulus Matkul' {Ya,Tidak}
@attribute 'Ipk' {3.25-4.00,2.00-3.25,<2.00}
@attribute 'Dominan Tinggal' {Ya,Tidak}
@attribute 'Agama' {Advent,Non-Advent}
@attribute 'Tamat Tepat' {Ya,Tidak}@data

Laki-laki,D3,IPA,Ya,Ya,Tidak,Ya,Ya,1,Sering,Tidak,Tidak,Tidak,Ya,Ya,Ya,Ya,3.25-4.00,Ya,Advent,?
Laki-laki,S1,IPA,Tidak,Tidak,Tidak,Tidak,Ya,1,Selalu,Tidak,Tidak,Tidak,Ya,Tidak,Tidak,Ya,Tidak,3.25-4.00,Ya,Advent,?
Laki-laki,S1,IPA,Ya,Tidak,Tidak,Ya,Ya,3,Sering,Tidak,Tidak,Ya,Ya,Tidak,Tidak,Tidak,3.25-4.00,Tidak,Non-Advent,?
Perempuan,S1,IPA,Tidak,Tidak,Tidak,Tidak,Ya,2,Sering,Ya,Tidak,Tidak,Ya,Ya,Tidak,Ya,Tidak,3.25-4.00,Ya,Advent,?
Laki-laki,S1,IPA,Tidak,Tidak,Tidak,Ya,Ya,5,Selalu,Tidak,Tidak,Tidak,Ya,Ya,Tidak,Tidak,Tidak,2.00-3.25,Tidak,Advent,?

```

Gambar 6 File ARFF untuk *supplied test*.

Pada metode yang ketiga yaitu *supplied test*, terlebih dahulu dibuat data dengan bentuk ARFF dengan mengikuti format penulisan data ARFF file, yaitu *header* atributnya diletakkan pada bagian atas yaitu jenis kelamin, jenjang, jurusan SMA, jurusan sesuai, pindah jurusan, berkeluarga, bekerja di luar, bekerja di dalam, jumlah organisasi, mengerjakan tugas, terpengaruh, disiplin, cuti, lingkungan teman, lingkungan mempengaruhi, anak *mission*, lunas tepat waktu atau tidak, lulus mata kuliah, IPK, dominan tinggal, agama. Sedangkan data yang ingin di prediksi diletakkan di bagian bawah. Kemudian atribut yang ingin diprediksi yaitu tamat tepat waktu, diisi dengan karakter tanda tanya (?). Gambar 6 di atas adalah tampilan dari ARFF file yang digunakan untuk metode *supplied test*.

Hasil dari pengolahan data dengan menggunakan metode *supplied test* dapat dilihat pada Gambar 7 berikut ini.

```

=== Predictions on test set ===

inst#      actual  predicted error prediction
1          1:?      1:Ya      0.81
2          1:?      1:Ya      0.935
3          1:?      2:Tidak   1
4          1:?      1:Ya      0.935
5          1:?      1:Ya      0.935

```

Gambar 5 Hasil Menggunakan *supplied test*

Dari gambar di atas dapat dilihat bahwa instan pertama, kedua, keempat, dan kelima terprediksi Ya dan instan ke 3 terprediksi Tidak, dengan *error prediction* di atas 80%.

5. Kesimpulan

Memprediksi mahasiswa UNAI apakah akan tamat tepat waktu sesuai dengan waktu normal masa perkuliahan, dapat dilakukan dengan menggunakan teknik *data mining* algoritma klasifikasi *decision tree*

J48. Dari hasil penelitian yang dilakukan dari 19 atribut yang digunakan menghasilkan kesimpulan bahwa atribut *cuti* adalah *root* teratas, artinya atribut ini berperan paling besar untuk menentukan prediksi.

Dari hasil uji coba terlihat bahwa menggunakan semua data sebagai *training set* menghasilkan akurasi 92,233%. Sementara dengan menggunakan *cross validation* 80% di dapati hasil akurasi 90,24%. Dengan demikian dapat di simpulkan bawah algoritma *decision tree* J48 dapat digunakan untuk memprediksi kelulusan mahasiswa tepat waktu di Universitas Advent Indonesia.

6. Referensi

- [1] A. Kurniawati, Analisis Problem-Problem Keterlambatan Penyelesaian Studi Mahasiswa di Stain Salatiga, pp. 58-90, 2011.
- [2] C. Aggarwal, Data Mining, New York: IBM T.J Watson Research Center, 2015.
- [3] G. F. Mandias, Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Di Universitas Klabat Dengan Metode Klasifikasi, p. 353, 2015.
- [4] H. S. & M. S. E. M. Ridwan, Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier., p. 59–64, 2013.
- [5] M. H. L. R. J. F. Frista Yulianora, Analisi Perbandingan Algoritma Decision Tree J48 Dan Naive Bayes Dalam Mengklasifikasikan Pola Penyakit, p. 29, 2014.
- [6] R. Hamid, Penyebab Keterlambatan Mahasiswa Menyelesaikan Studi S1 di IAIN Antasari Banjarmasin, pp. 57-64, 2015.
- [7] S. Fitri, Perbandingan Kinerja Algoritma Klasifikasi Naive Bayesian, Lazy-ibk, Zero-R, dan Decision tree-J48, pp. 33-37, 2014.
- [8] "Weka 3: Data Mining Software in Java," The University of Waikato, [Online]. Available: <http://www.cs.waikato.ac.nz/ml/weka/>. [Accessed 04 December 2017].