

PREDIKSI KELULUSAN MAHASISWA MENGUNAKAN ALGORITMA C.45 (Studi Kasus, Sekolah Tinggi Ilmu Administrasi Trinitas Ambon)

Daniel Hartono Renyut^{*1}, Yuyun², Ferdinand³

^{1,2}Universitas Handayani Makassar, ^{1,3}STIA Trinitas Ambon

e-mail : ^{*1}danielrenyut29stmik@gmail.com, ²yuyunwabula@handayani.ac.id,

³ferdinand.ren2018@yahoo.com

Kelulusan mahasiswa merupakan salah satu tolak ukur dalam menjadikan pendidikan lanjutan sebagai strategi dewan. Tujuan dari penelitian ini Untuk mengimplementasikan algoritma C.45 dalam memprediksi Kelulusan Mahasiswa Untuk mengetahui factor-faktor yang mempengaruhi dalam kelulusan mahasiswa tepat waktu dan sebaliknya. Algoritma yang digunakan dalam penelitian ini adalah algoritma c.45 dimana Algoritma C4.5 digunakan membentuk pohon keputusan yang dapat digunakan untuk membentuk pohon keputusan. Hasil dari penelitian ini adalah Mahasiswa dapat di prediksi dan dilakukan pengujian dengan menerapkan teknik data mining untuk melakukan prediksi berdasarkan data training dengan attribube SKS dan Jumlah IPK merupakan faktor yang mempengaruhi ketepatan waktu lulus dan Hasil evaluasi yang dilakukan dengan melakukan pengujian Confusion Matrix dengan data uji sebanyak 404 data mahasiswa dengan pembagian data kelulusan yang di prediksi lulus dengan tepat waktu sebanyak 368 sedangkan mahasiswa yang tidak lulus dengan tepat waktu terdapat 36 data mahasiswa sehingga dengan melakukan pengujian confusion matrix memperoleh hasil akurasi sebesar 85%.

Kata Kunci : Prediksi, Kelulusan, Mahasiswa, Algoritma c.45.

I. PENDAHULUAN

Sekolah Tinggi Ilmu Administrasi Trinitas merupakan perguruan Tinggi Swasta yang mampu berkompetensi dengan perguruan tinggi baik itu lulusan, sumberdaya manusia dan manajemennya. Salah satu unsur terpenting dalam penyelenggaraan perguruan tinggi adalah mahasiswa. Mahasiswa sering disebut sebagai kelompok masyarakat yang memiliki ciri intelektualitas yang lebih luas dibandingkan dengan kelompok usia mereka yang bukan mahasiswa ataupun kelompok usia lain yang dibawah mereka. [1]

Dalam dunia pendidikan mahasiswa merupakan objek utama dari setiap penyelenggaraan pendidikan yang

senantiasa mengedepankan kedisiplinan ilmu yang bermanfaat untuk mahasiswa itu sendiri ataupun orang lain, biasanya banyak hal yang terjadi kepada mahasiswa ketika proses pendidikan itu dilaksanakan, mulai ketidak patuhan akan aturan, ketidak nyamanan dalam belajar, ketidak seriusan dalam belajar, kemampuan dan bakat yang dimiliki yang tidak menunjang pada bidang yang diterimanya, kesenjangan sosial, ketidak mampuan secara financial, dan lain-lain, itu semua menyebabkan terjadinya ketidak konsitenan dalam proses belajar, sehingga dampak yang ditimbulkannya cukup kompleks, mulai dari kesiapan untuk menerima materi yang kurang, banyaknya mahasiswa yang mengulang, drop out dan berhenti sebelum lulus.

Kelulusan mahasiswa merupakan salah satu tolak ukur dalam menjadikan pendidikan lanjutan sebagai strategi dewan. Saat ini sudah biasa dilacak bahwa jumlah mahasiswa baru di suatu perguruan tinggi tidak sebanding dengan jumlah lulusan setiap tahunnya. Dalam organisasi pendidikan lanjutan, mahasiswa merupakan salah satu faktor penting dalam kemajuan program studi. [2]

Lama masa studi mahasiswa dalam menyelesaikan kuliahnya merupakan salah satu bagian penilaian akreditasi bagi perguruan tinggi, namun masih banyak mahasiswa yang tidak dapat menyelesaikan kuliahnya selama 8 semester karena berbagai macam faktor. Sehingga hal itu mempengaruhi ketepatan kelulusan mahasiswa dalam proses penyelenggaraan pendidikan di perguruan tinggi. [3]

Data Mining adalah teknik yang memanfaatkan data dalam jumlah yang besar yang tersedia didalam database untuk menguraikan penemuan pengetahuan. Terdapat banyak metode dalam data mining diantaranya seperti Decision Trees, Bayesian, Artificial Neural Networks, Nearest Neighbor, Support Vector Machines dan lainnya. [4]

Data mining dan Knowledge Discovery in Databases (KDD) sering kali digunakan secara bergantian dalam menjelaskan proses penggalian informasi dalam basis data yang sangat besar akan tetapi berkaitan satu sama yang lain, skema proses KDD (Han & Kamber, 2006). [5]

Sebelumnya telah dilakukan penelitian oleh (Rahman et al., 2020) dengan penelitian Prediksi Kelulusan Mahasiswa Menggunakan Algoritma C4.5 Penelitian ini bertujuan untuk mengetahui tingkat akurasi algoritma C4.5 dalam prediksi kelulusan mahasiswa, menggunakan metode decision tree algoritma C4.5. Data mahasiswa yang digunakan sebagai atribut adalah IPK, SKS, Umur, Jenis Kelamin. [3]

Sehingga dengan penelitian sebelumnya maka untuk mendapatkan solusi dengan permasalahan yang ada di atas hal ini sangat dibutuhkan agar dapat menentukan strategi yang tepat agar mahasiswa dapat terhindar dari keterlambatan. Dengan prediksi yang akurat dan tepat guna diharapkan penelitian dapat membantu jurusan untuk meningkatkan jumlah lulusan agar memenuhi Standarisasi yang ditetapkan oleh BAN-PT, Penelitian ini berfokus untuk menguji kelayakan prediksi tingkat kelulusan mahasiswa Sekolah Tinggi Ilmu Administrasi Trinitas menggunakan Algoritma C.45 untuk menemukan algoritma dengan tingkat akurasi tertinggi. [6]

II. LANDASAN TEORI

A. Prediksi

Prediksi adalah pemikiran terhadap suatu besaran, misalnya permintaan terhadap satu atau beberapa produk pada periode yang akan datang. Pada hakekatnya prediksi hanya merupakan suatu perkiraan (guess), tetapi dengan menggunakan teknik- teknik tertentu, maka prediksi menjadi lebih dari sekedar perkiraan. Prediksi dapat diartikan perkiraan yang ilmiah (educated guess). Setiap pengambilan keputusan yang menyangkut keadaan dimasa yang akan datang, maka pasti ada peramalan yang melandasi pengambilan keputusan tersebut (Hay's et al., 2017). [4]

Prediksi adalah suatu cara menilai secara metodis sesuatu yang mungkin akan terjadi kemudian dengan bergantung pada data yang diklaim dalam rentang waktu yang lama, sehingga kesalahan (kontras antara sesuatu yang terjadi dan hasil yang diharapkan) dapat dibatasi. (Abidin et al., 2021). [2]

Peramalan adalah seni atau ilmu untuk memperkirakan kejadian di masa depan. Hal ini dapat dilakukan dengan melibatkan pengambilan data historis dan memproyeksikannya ke masa mendatang dengan suatu bentuk model sistematis. Atau bisa juga dengan menggunakan kombinasi model matematis yang disesuaikan dengan pertimbangan yang baik dari seorang manajer. (Rachman, 2018). [7]

B. Algoritma C45

Algoritma C4.5 digunakan untuk membentuk pohon keputusan yang dapat digunakan untuk membentuk pohon keputusan yang ditemukan oleh John Ross Quinlan. Algoritma C4.5 merupakan pengembangan dari algoritma ID3. Berbeda dengan ID3 yang menggunakan information gain dalam algoritma C4.5 pemilihan atribut dilakukan dengan menggunakan Gain Ratio. Gain Ratio digunakan

untuk mengatasi atribut yang memiliki nilai yang sangat bervariasi dan dihitung berdasarkan Split Information. (Budiman et al., 2019). [8]

Algoritma C4.5 adalah algoritma yang sudah cukup terkenal dan merupakan salah satu algoritma dalam Data Mining Salah satu kegunaannya adalah untuk membentuk sebuah pohon keputusan (Decision Tree). Pohon Keputusan sendiri merupakan metode klasifikasi dan prediksi yang digunakan untuk mencari data dan menemukan hubungan yang tersimpan dari variable atau atribut yang digunakan dan sebuah variabel target yang biasa disebut class atau label. Algoritma yang sering digunakan untuk membuat sebuah pohon keputusan adalah ID3, algoritma C4.5, dan CART (Saputra Ginting and Taufiq Luthfi, 2020). [3]

C. Decision Tree

Decision tree (pohon keputusan) merupakan salah satu metode yang populer digunakan dalam pengklasifikasian. Disebut sebagai pohon keputusan karena model yang dihasilkan untuk memprediksi data berupa pohon. Untuk mengklasifikasi sebuah data, setiap attribute dari data tersebut akan diuji melalui serangkaian node yang terdapat pada pohon keputusan dan setelah data sampai pada leaf node, data tersebut akan terklasifikasi sesuai dengan kelas yang terdapat pada leaf node.

Pada pohon keputusan terdapat 3 jenis node, node – node tersebut adalah:

- a. Root Node
Root Node merupakan node yang letaknya berada diawal tree (diatas). Root Node tidak memiliki input yang berarti tidak ada cabang yang masuk ke node ini. Root Node dapat memiliki output lebih dari satu atau tidak memiliki output sama sekali.
- b. Internal Node
Internal Node merupakan node percabangan, pada node ini hanya terdapat satu input (satu cabang masuk) dan dapat memiliki output satu atau lebih.
- c. Leaf Node
Leaf Node atau terminal node, merupakan node akhir pada decision tree. Node ini hanya memiliki satu input dan tidak memiliki output. Node berperan untuk menunjukkan kelas akhir dari pengklasifikasian. [9]

D. Kelulusan

Kelulusan mahasiswa adalah hal yang penting diperhatikan, karena persentase jumlah kelulusan mempengaruhi penilaian pemerintah serta mempengaruhi status akreditasi program studi tingkat kelulusan dalam suatu perguruan tinggi itu sangat penting bagi setiap mahasiswa ada berepa faktor yang mempengaruhi kelulusan diantaranya biaya perkuliahan IPS semester dari IPS 1 sampai IPS 7, IPK, dan SKS persemester penghasilan orang tua. Adapun indikator yang sering digunakan untuk menentukan seorang mahasiswa bisa lulus adalah IPS banyak perguruan tinggi menggunakan hal ini dan variable yang akan digunakan adalah jenis kelamin IPS 1 – IPS 7 IPK

dan SKS, SKS persemester dan penghasilan rata-rata orang tua(Thaniket & Taufik Luthf, 2020). [8]

E. Faktor Yang Mempengaruhi Kelulusan

Adapun factor yang mempengaruhi tingka kelulusan pada mahasiswa adalah sebagai berikut :

- a. Faktor Kurangnya Motivasi
- b. Problem Pribadi
- c. Faktor Sudah Menikah
- d. Faktor Ekonomi
- e. Kurangnya Motivasi Keluarga

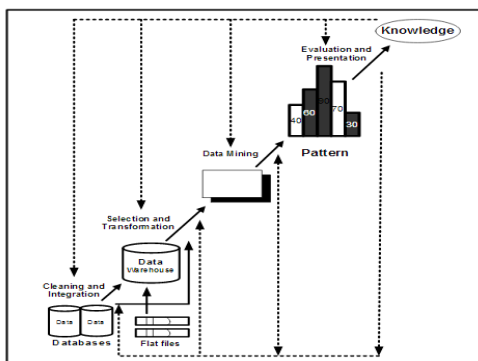
F. Pengenalan Pola, Data Mining dan Machine Learning

Data Mining juga merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data dengan ukuran yang cukup besar melalui serangkaian proses untuk mendapatkan informasi yang berharga dari data tersebut. Data Mining dapat diterapkan pada berbagai bidang yang mempunyai sejumlah data. Menurut Daryl Pregibon bahwa “Data mining adalah campuran dari statistic, kecerdasan buatan, dan riset basis data” yang masih berkembang. Data mining memungkinkan pemakai menemukan pengetahuan dalam data database yang tidak mungkin diketahui keberadaanya oleh pemakai, Istilah lain yang sering digunakan diantaranya knowledge discovery (mining) in databases (KDD). [6]

Machine Learning adalah suatu area dalam artificial intelligence atau kecerdasan buatan yang berhubungan dengan pengembangan teknik-teknik yang bisa diprogramkan dan belajar dari data masa lalu. Pengenalan pola, data mining dan machine learning sering dipakai untuk menyebut sesuatu yang sama. Bidang ini bersinggungan dengan ilmu probabilitas dan statistic kadang juga optimasi.Machine learning menjadi alat analisis dalam data mining.Bagaimna bidang-bidang ini berhubungan bisa dilihat dalam gambar 2.1 Santoso (2007:23).

G.Tahap-Tahap Data Mining

Sebagai suatu rangkaian proses data mining dapat dibagi menjadi beberapa tahap yang diilustrasikan di Gambar 2.2. Tahap-tahap tersebut bersifat interaktif, pemakaian terlibat langsung atau dengan perantaraan knowledge base.



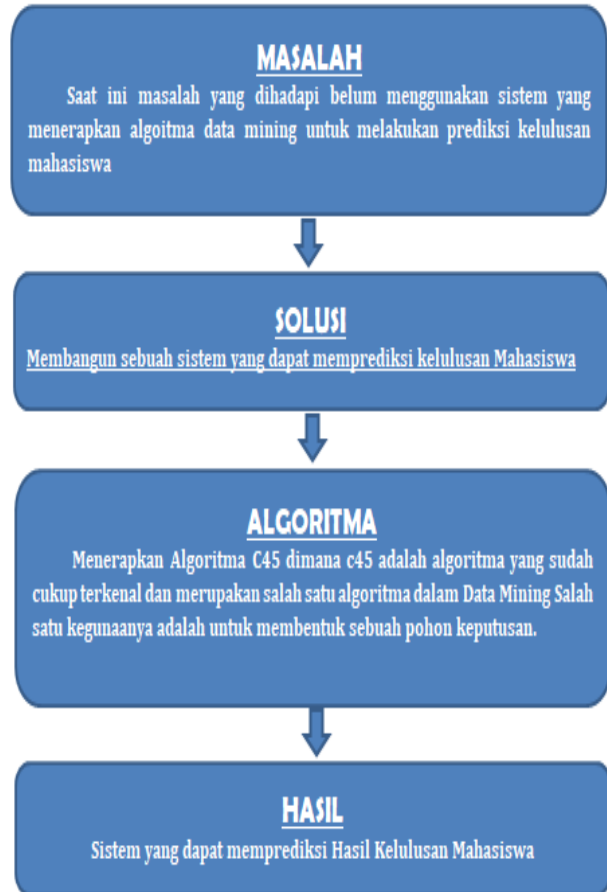
Gambar 2.1 Tahap-tahap Data mining (Sumber : Han,2006).

H.Tahap-tahap data mining ada 6 yaitu :

- a. Pembersihan data (data cleaning)
- b. Integritasi data (data integration)

- c. Seleksi Data (Data Selection)
- d. Transformasi data (Data Transformation)
- e. Proses mining
- f. Evaluasi pola (pattern evaluation)
- g. Presentasi pengetahuan (knowledge presentation)

I. Kerangka Pikir



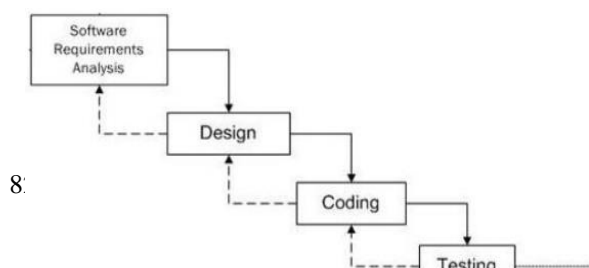
Gambar 2.2 Kerangka Pikir

III. METODE PENELITIAN

A. Jenis Penelitian

Jenis penelitian yang digunakan dalam penelitian ini adalah R&D/Research and Development. Penelitian R & D adalah suatu jenis penelitian yang berfungsi untuk menghasilkan sesuatu yang baru dan dilanjutkan dengan pengujian terhadapnya, kemudian Metode penelitian yang digunakan adalah metode waterfall.

Alasan menggunakan metode ini adalah karena metodewaterfall melakukan pendekatan secara sistematis dan berurutan dalam membangun suatu sistem. Proses metode waterfall yaitu pada pengerjaan darisuatu sistem dilakukan secara berurutan. Sistem yang dihasilkan akan berkualitasbaik, dikarenakan pelaksanaannya secara bertahap sehingga tidak terfokus padatahapan tertentu. Tahapan dari metode waterfall adalah:



Gambar 3.1 Model Waterfall

1. Analisis Kebutuhan
Tahap analisis kebutuhan dilakukan dengan menganalisa kebutuhan user, analisa perangkat lunak dan perangkat keras yang dibutuhkan dalam pengembangan sistem serta kebutuhan lain dalam pembuatan basis data.
2. Desain Sistem
Tahap selanjutnya yaitu mendesain sistem. Tahap ini dibuat sebelum tahap pengkodean. Tujuan dari tahap ini adalah memberikan gambaran tentang apa yang akan dikerjakan dan bagaimana tampilannya. Tahap ini memenuhi semua kebutuhan pengguna sesuai dengan hasil yang dianalisa seperti rancangan tampilan pengembangan sistem ujian online di sekolah, dan membantu mendefinisikan arsitektur sistem secara keseluruhan.
3. Coding (Penulisan Kode Program)
Aktivitas pada tahap ini dilakukan pengkodean sistem. Penulisan kode program merupakan tahap penerjemahan desain sistem yang telah dibuat ke dalam bentuk perintah-perintah yang dimengerti komputer dengan menggunakan bahasa pemrograman. Tahapan ini merupakan tahapan secara nyata dalam mengerjakan suatu sistem. Sistem ini bahasa pemrograman yang dipakai adalah PHP dan database MySQL.
4. Testing Pengujian Program
Pengujian dilakukan untuk memastikan bahwa software yang dibuat telah sesuai dengan desainnya dan semua fungsi dapat dipergunakan dengan baik tanpa ada kesalahan.
5. Operation and Maintenance (Pemeliharaan Program).
Tahap ini merupakan tahap terakhir dalam metode waterfall. Sistem dapat di implementasikan. Pemeliharaan mencakup koreksi dari berbagai error yang tidak ditemukan pada tahap-tahap terdahulu, perbaikan atas implementasi dan pengembangan unit sistem, serta pemeliharaan program. Pemeliharaan sistem dapat dilakukan oleh seorang administrator untuk meningkatkan kualitas sistem agar jauh lebih baik.

B. Bahan Dan Alat

Kebutuhan Fungsional adalah kebutuhan-kebutuhan yang memiliki keterkaitan langsung dengan sistem, kebutuhan fungsional dari sistem ini meliputi: input, proses dan output. Kebutuhan fungsional yang harus ada dalam sistem yang akan dibangun adalah :

- a. Sistem harus dapat mempermudah Admin atau oengguna dalam mengambil keputusan secara efisien dan efektif
- b. Sistem harus dapat menampung data yang dibutuhkan untuk memprediksi Kelulusan Mahasiswa
- c. Rekapian Laporan

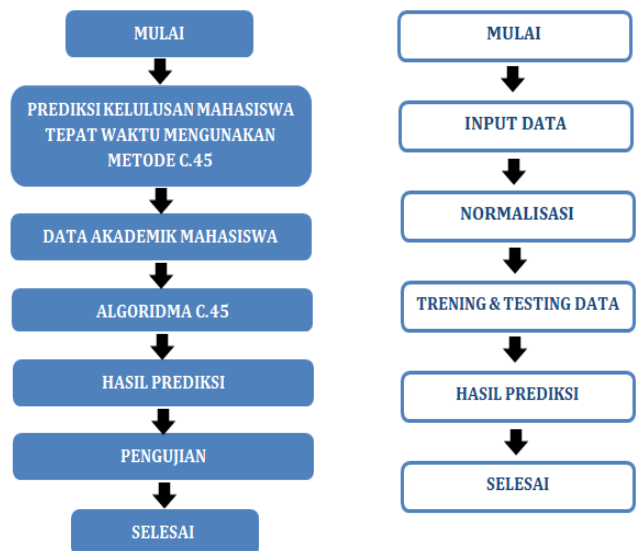
Kebutuhan Non Fungsional :

- a. Kebutuhan perangkat keras (Hardware)
- b. Kebutuhan perangkat Lunak (Software)

C. Populasi Dan Sampel

- a. Populasi
Populasi adalah wilayah generalisasi yang terdiri atas objek atau subjek yang mempunyai kualitas dan karakteristik tertentu yang telah ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya, Populasi dalam penelitian ini yaitu seluruh mahasiswa Sekolah Tinggi Ilmu Administrasi (STIA) Trinitas Ambon.
- b. Sampel
Sampel adalah bagian yang dapat mewakili populasi yang memiliki karakteristik yang sama dan dianggap bisa mewakili populasi.

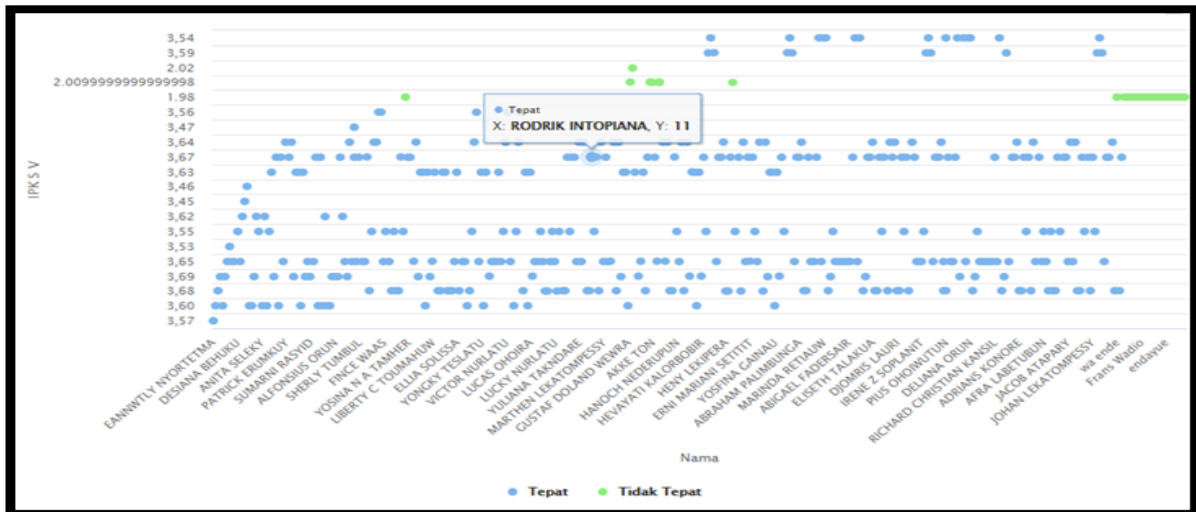
D. Desain Penelitian



Gambar 3.2 Desain Penelitian

E. Alur Penelitian

1. Input Data
2. Normalisasi
3. Data Training dan Data Testing
Data yang digunakan adalah data mahasiswa dengan jumlah data yaitu 404 data dengan jumlah data kelulusan sebanyak 368 dan data belum lulus sebanyak 36 data mahasiswa. Berikut ini merupakan sebaran data kelulusan Mahasiswa.



Gambar 3.3 Grafik Sebaran Data

4. Prediksi

Metode yang digunakan untuk memprediksi kelulusan mahasiswa adalah algoritma c45 dengan menampilkan hasil perbandingan tingkat Akurasi

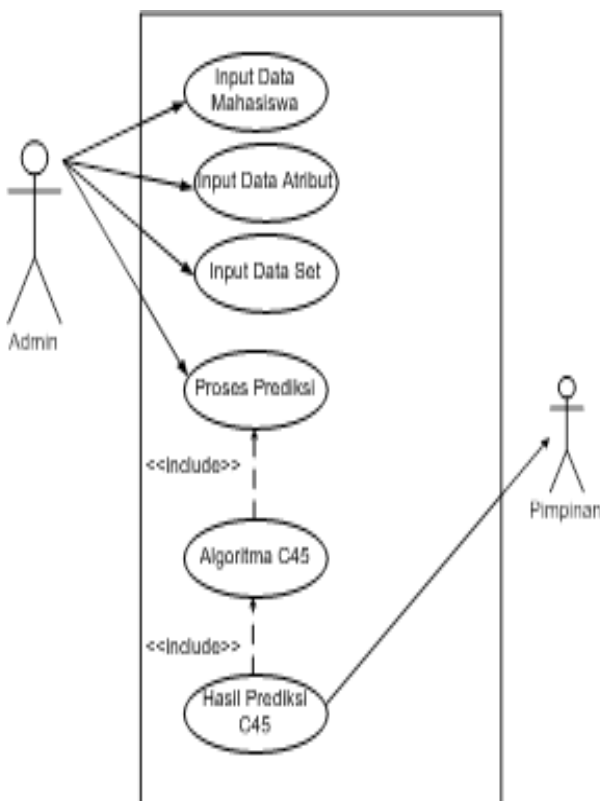
Data sekunder dilakukan dengan melakukan pencarian sebanyak mungkin literatur, baik dari buku, jurnal maupun internet. Selain itu mempelajari literatur yang berkaitan dengan data mining, klasifikasi, Metode Moving Average.

F. Desain Sistem Informasi

Sistem yang diusulkan sebagai solusi pemecahan masalah dengan perancangan sebagai berikut :

H. Teknik Pengumpulan Data

1. Observasi
Pengumpulan data dengan melakukan pengamatan secara langsung terhadap objek penelitian, dengan mencatat hal-hal penting yang berhubungan dengan judul, sehingga diperoleh data yang lengkap dan akurat.
2. Wawancara
Pengumpulan data dengan komunikasi dan wawancara secara langsung dengan pihak akademik dari Sekolah Tinggi Ilmu Administrasi (STIA) Trinitas Maluku.
3. Studi Pustaka
Pengumpulan data dengan menggunakan atau mengumpulkan sumber-sumber tertulis, dengan cara membaca.



Gambar 3.6 Use Case Diagram Sistem yang diusulkan

I. Teknik Pengumpulan Data

1. Observasi
Pengumpulan data dengan melakukan pengamatan secara langsung terhadap objek penelitian, dengan mencatat hal-hal penting yang berhubungan dengan judul, sehingga diperoleh data yang lengkap dan akurat.
2. Wawancara
Pengumpulan data dengan komunikasi dan wawancara secara langsung dengan pihak akademik dari Sekolah Tinggi Ilmu Administrasi (STIA) Trinitas Maluku.
3. Studi Pustaka
Pengumpulan data dengan menggunakan atau mengumpulkan sumber-sumber tertulis, dengan cara membaca, mempelajari dan mencatat hal-hal penting yang berhubungan dengan masalah yang sedang dibahas guna memperoleh gambaran secara teoritis yang dapat menunjang pada penyusunan ini

G. Sumber Data

1. Data Primer
Data yang digunakan merupakan data mahasiswa yang bersumber dari Sekolah Tinggi Ilmu Administrasi (STIA) Trinitas Ambon (MALUKU)
2. Data Sekunder

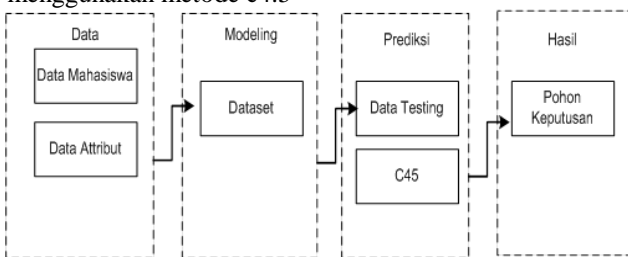
IV. HASIL DAN PEMBAHASAN

A. Pemodelan Sistem

Berdasarkan Analisis sistem yang sedang berjalan maka penulis memberikan satu solusi pemecahan masalah dengan membangun aplikasi Prediksi Kelulusan Mahasiswa Menggunakan Algoritma C 45.

B. Pengolahan Data

Tahap pengolahan data awal ini yaitu di adopsi dari data kelulusan mahasiswa dengan menetapkan atribut sebagai variabel keputusan dan dataset yang masih berisi informasi pendukung perlu dilkakukan pengolahan dengan membersihkan data informasi pendukung tersebut. Transformasi data pada masing-masing dataset dilakukan ke dalam bentuk yang dibutuhkan pada penelitian ini yaitu menggunakan metode c4.5



Gambar 4.2 Desain Sistem

C. Data

Pengumpulan data dalam penelitian ini menggunakan data kelulusan mahasiswa pada Sekolah tinggi Ilmu Administrasi Trinitas Ambon, kemudian menentukan data yang akan digunakan sebagai atribut untuk melakukan perhitungan sehingga menghasilkan dataset.

D. Modeling

Setelah Melalui proses pengumpulan data kelulusan mahasiswa dan telah melakukan penentuan atribut pada tahap modelling ini menentukan dataset untuk diproses pada tahap selanjutnya.

Jumlah data set yang digunakan dalam penelitian ini dari tahun 2010- 2016 sejumlah 404 data mahasiswa dengan nilai IP yang dihitung berdasarkan nilai IP Semester 1 sampai dengan Nilai IP Semester 5.

Tabel 4.1 Sampel Dataset Tidak Lulus

	A	B	C	D	E	F	G	H	I
1	Nama	Jenis Kelamin	IPK S I	IPK S II	IPK S III	IPK S IV	IPK S V	Jumlah SKS	Target
2	Indang	Pria	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
3	hasan	Wanita	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
4	haze	Wanita	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
5	watina wabula	Wanita	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
6	verdiau	Pria	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
7	fieya	Pria	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
8	endayue	Pria	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
9	la monfi	Pria	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
10	mayang	Wanita	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
11	maweru	Pria	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
12	wa liaya	Wanita	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
13	zabira	Wanita	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
14	watima	Wanita	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
15	pradi	Pria	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
16	elly taukan	Pria	2,01	1,19	1,96	2,21	1,98	100	Tidak Tepat
17									

Tabel 4.1 merupakan tabel sampel dataset untuk mahasiswa yang tidak lulus tepat waktu karena ada yang tidak terpenuhi, untuk jumlah SKS yang terpenuhi dan mencukupi untuk lulus tepat waktu pihak Kampus memberikan Acuan atau standard kelulusan. Untuk standard Kelulusan yaitu berjumlah 144 SKS, SKS dengan jumlah 138 Mencukupi untuk Ujian Skripsi untuk SKS dengan jumlah < 138 tidak mencukupi untuk mengikuti ujian skripsi.

E. Prediksi

Setelah melakukan proses pengolahan dataset kemudian dilakukan tahap prediksi dengan menggunakan data testing yang akan di hitung dan di proses dengan data training yang telah di input pada dataset.

Proses prediksi dilakukan dengan menggunakan algoritma c4.5. Data testing dilakukan dengan pemrograman Web di input untuk mealkukan prediksi dan akan menghasilkan pohon keputusan hasil prediksi.

F. Hasil

Pada bagian ini menampilkan hasil prediksi kelulusan mahasiswa dengan menampilkan hasil perhitungan serta menampilkan pohon keputusan.

G. Hasil Pengujian Confision Matrix

Untuk mengetahui kinerja algoritma C4.5 dalam melakukan klasifikasi terhadap kelas yang telah ditentukan dalam penelitian, untuk menghitung nilai akurasi dilakukan dengan menggunakan confusion matrix, penelitian ini perhitungan nilai akurasinya menggunakan 404 dataset mahasiswa menghasilkan nilai akurasi sebagai berikut :

Tabel 4.2 Tabel Pengujian Confusion Matrix

n = 404	Aktual Positif (+)	Aktual Negatif (-)
Prediksi Positif (1)	TP 343	FP 30
Prediksi Negatif (0)	FN 25	TN 6
	368	36

Maka dapat dihitung nilai accuracy, precision, dan recall seperti yang akan diuraikan berikut ini :

Accuracy mennggambarkan seberapa akurat model dalam mengklasifikasikan dengan benar :

$$\begin{aligned}
 \text{Accuracy} &= (TP+TN) / (TP+FP+FN+TN) \\
 &= (343 + 6) / (343+30+25+6) \\
 &= 349 / 404 \\
 &= 0,849 \\
 &= 0,85 * 100\% = 85\%
 \end{aligned}$$

Precision menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang diberikan oleh model.

$$\begin{aligned}
 \text{Precision} &= (TP) / (TP + FP) \\
 &= 343 / (343+30) \\
 &= 343 / 373 \\
 &= 0.919 \\
 &= 0.92 * 100\% = 92\%
 \end{aligned}$$

Recall atau sensitivity: menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi.

$$\begin{aligned}
 \text{Recall} &= TP / (TP + FN) \\
 &= 343 / (343+25) \\
 &= 343 / 368 \\
 &= 0.932 * 100\% = 93,2\%
 \end{aligned}$$

Dari hasil pengujian confusion matri dengan dataset mahasiswa yang di uji berjumlah 404 data mahasiswa kelulusan memperoleh hasil akurasi 85% yang di uraikan dalam tabel berikut

Tabel 4.3 Hasil Confusion Matrix

Accuracy	Precision	Recall
85.00%	92%	93.2%

V. KESIMPULAN DAN SARAN

Berdasarkan Hasil Penelitian yang telah dilakukan oleh peneliti telah dapat dilakukan prediksi dengan memanfaatkan algoritma c4.5 untuk memprediksi kepatan waktu lulus mahasiswa sehingga dapat disimpulkan:

1. Mahasiswa dapat di prediksi dan dilakukan pengujian dengan menerapkan teknik data mining dengan menggunakan algoritma c4.5 untuk melakukan prediksi berdasarkan data training dengan attribute SKS dan Jumlah IPK merupakan faktor yang mempengaruhi ketepatan waktu lulus

2. Hasil evaluasi yang dilakukan dengan melakukan pengujian Confusion Matrix dengan data uji sebanyak 404 data mahasiswa dengan pembagian data kelulusan mahasiswa yang di prediksi lulus dengan tepat waktu sebanyak 368 sedangkan mahasiswa yang tidak lulus dengan tepat waktu terdapat 36 data mahasiswa sehingga dengan melakukan pengujian confusion matrix memperoleh hasil akurasi sebesar 85%.

Saran yang dapat berdasarkan kesimpulan yang telah dijabarkan, maka saran dari penelitian ini adalah:

1. Penelitian ini dapat dikembangkan dengan menambah jumlah inputan data training atau dataset karena jumlah data training atau dataser member kontribusi untuk mempengaruhi nilai akurasi, kemudian pemilihan atribut yang akan digunakan dalam melakukan prediksi juga sangat berpengaruh.

2. Untuk penelitian selanjutnya dengan mengembangkan penelitian lebihlanjt untuk membandingkan hasil prediksi, melakukan perbandingan pada algoritma lain agar dapat menghasilkan prediksi yang lebih baik

DAFTAR PUSTAKA

- [1] L. A. R. Hakim, A. A. Rizal, and D. Ratnasari, "Aplikasi Prediksi Kelulusan Mahasiswa Berbasis K-Nearest Neighbor (K-NN)," *JTIM J. Teknol. Inf. Dan Multimed.*, vol. 1, no. 1, pp. 30–36, 2019.
- [2] N. A. Abidin, M. Assidiq, and A. Qaslim, "SISTEM PREDIKSI KELULUSAN MAHASISWA DENGAN METODE BACKPROPAGATION NEURAL NETWORK," *J. Ilm. Maju*, vol. 4, no. 2, pp. 1–5, 2021.
- [3] A. F. A. Rahman, "Prediksi Kelulusan Mahasiswa Menggunakan Algoritma C4. 5 (Studi Kasus Di Universitas Peradaban)," *Indones. J. Informatics Res.*, vol. 1, no. 2, pp. 70–77, 2020.
- [4] R. naufal Hay's and R. Adrean, "Sistem Informasi Inventory Berdasarkan Prediksi Data Penjualan Barang Menggunakan Metode Single Moving Average Pada CV. Agung Youanda," *ProTekInfo (Pengembangan Ris. dan Obs. Tek. Inform.*, vol. 4, pp. 29–33, 2017.
- [5] A. S. Sunge, "Prediksi Kompetensi Karyawan Menggunakan Algoritma C4. 5 (Studi Kasus: PT Hankook Tire Indonesia)," *Semin. Nas. Teknol. Inf. dan Komun. 2018 (SENTIKA 2018)*, vol. 2018, pp. 23–24, 2018.
- [6] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, "Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes: Systematic Review," *Fakt. Exacta*, vol. 13, no. 1, pp. 35–43, 2020.
- [7] R. Rachman, "Penerapan metode moving average dan exponential smoothing pada peramalan produksi industri garment," *J. Inform.*, vol. 5, no. 2, pp. 211–220, 2018.
- [8] F. Aldi and A. A. Rahma, "University Student Satisfaction Analysis on Academic Services by Using Decision Tree C4. 5 Algorithm (Case Study: Universitas Putra Indonesia 'YPTK' Padang)," in *Journal of Physics: Conference Series*, 2019, vol. 1339, no. 1, p. 12051.
- [9] A. Nurkholis and I. S. Sitanggang, "Optimalisasi model prediksi kesesuaian lahan kelapa sawit menggunakan algoritme pohon keputusan spasial," *J. Teknol. Dan Sist. Komput.*, vol. 8, no. 3, pp. 192–200, 2020.