PERBANDINGAN METODE ALGORITMA C4.5 DAN NAÏVE BAYES UNTUK MEMPREDIKSI KELULUSAN TEPAT WAKTU MAHASISWA STMIK PALANGKARAYA

PROPOSAL TUGAS AKHIR

Disusun untuk Memenuhi Syarat Penulisan Tugas Akhir pada
Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) Palangkaraya

OLEH :

AULIA RAHMAYANTI

NIM C1555201057

PROGRAM STUDI TEKNIK INFORMATIKA

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER

(STMIK) PALANGKARAYA

2021

PERSETUJUAN

PERBANDINGAN METODE ALGORITMA C4.5 DAN NAÏVE BAYES UNTUK MEMPREDIKSI KELULUSAN TEPAT WAKTU MAHASISWA STMIK PALANGKARAYA

Proposal Tugas Akhir ini telah disetujui untuk Diseminarkan
pada Tanggal 09 Maret 2021

Pembimbing I,	Pembimbing II,
Lili Rusdiana, M.Kom	H. Suratno, S.Kom, M.Si
NIK. 198707282011007	NIK. 196912281997101

Mengetahui
Ketua STMIK Palangkaraya,

Suparno, M.Kom
NIK. 196901041995105

PENGESAHAN

PERBANDINGAN METODE ALGORITMA C4.5 DAN NAÏVE BAYES UNTUK MEMPREDIKSI KELULUSAN TEPAT WAKTU MAHASISWA STMIK PALANGKARAYA

Proposal Tugas Akhir ini telah Diseminarkan, Dinilai dan Disahkan
Oleh Tim Seminar pada Tanggal 13 Maret 2021

Tim Seminar Proposal :

1. Sulistyowati, S.Kom, M.Cs
   Ketua

   .....

2. Lili Rusdiana, M.Kom
   Sekertaris

   .....

3. H. Suratno, S.Kom, M.Si
   Anggota

   .....

DAFTAR ISI

HALAMAN PERSETUJUAN ii

HALAMAN PENGESAHAN iii

DAFTAR ISI iv

DAFTAR TABEL vi

DAFTAR GAMBAR vii

DAFTAR LAMPIRAN viii

BAB I PENDAHULUAN 1

1.1 Latar Belakang Masalah 1

1.2 Perumusan Masalah 3

1.3 Batasan Masalah 3

1.4 Tujuan dan Manfaat 4

1.5 Sistematika Penulisan 5

BAB II LANDASAN TEORI 7

2.1 Tinjauan Pustaka 7

2.2 Data Mining 10

2.2.1 Pengertian Data Mining 10

2.2.2 Pekerjaan Dalam *Data Mining* 11

2.3 Set Data 13

2.4 Klasifikasi 15

2.4.1 Konsep Klasifikasi 15

2.4.2 Pengukuran Kinerja Klasifikasi 15
2.5 Pohon Keputusan 17
2.5.1 Pengertian Pohon Keputusan 17
2.5.2 Model Pohon Keputusan 17
2.6 Algoritma C4.5 18
2.6.1 Pengertian Algoritma C4.5 18
2.6.2 Tahapan Algoritma C4.5 18
2.7 Naïve Bayes 20
2.8 Weka 21
BAB III METODE PENELITIAN 23
3.1 Metode Penelitian 23
3.2 Kerangka Pemikiran 25
3.1.1 Algoritma 26
3.1.2 Simulasi Perhitungan Manual 28
DAFTAR PUSTAKA 32
LAMPIRAN 33

DAFTAR TABEL

Tabel 1 Perbandingan Penelitian 8
Tabel 2 Tipe Atribut 14
Tabel 3 *Class* pada *Confussion Matrix* 15
Tabel 4 Sampel Data *Training* 28
Tabel 5 Kelas Kategorial 28
Tabel 6 Hasil perhitungan *Entropy* dan *Gain* 30
Tabel 7 Contoh Data Testing Untuk Perhitungan *Naïve Bayes* 30
Tabel 8 Hasil perhitungan *Naïve Bayes* 31

DAFTAR GAMBAR

Gambar 1 Posisi *data mining* dalam bisnis cerdas 11
Gambar 2 Tampilan Weka 22
Gambar 3 Kerangka Pemikiran Penelitian 25
Gambar 4 *Flowchart* Algoritma C4.5 26
Gambar 5 *Flowchart Naïve Bayes* 27

DAFTAR LAMPIRAN

Lampiran 1 Surat Tugas
Lampiran 2 Kartu Kegiatan Konsultasi Tugas Akhir
Lampiran 3 Surat Tugas Penguji Seminar Proposal
Lampiran 4 Berita Acara Seminar Proposal Tugas Akhir
Lampiran 5 Kartu Kegiatan Seminar Proposal Skripsi
Lampiran 6 Daftar Hadir Seminar Proposal Tugas Akhir

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Prediksi merupakan suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi sangat perlu dilakukan dalam pembelajaran formal supaya dapat memperkecil kemungkinan mahasiswa mendapatkan nilai yang kurang baik maupun menghindari mahasiswa program studi Teknik Informatika tidak lulus dalam suatu mata kuliah. Dengan menggunakan *data mining*, setiap kumpulan atau gudang data dapat memberikan pengetahuan penting. Dalam *data mining* terdapat banyak teknik dalam pengerjaannya, diantaranya algoritma *Naïve Bayes*, algoritma C4.5, jaringan saraf tiruan dan masih banyak lainnya.

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya merupakan organisasi pendidikan yang bergerak di bidang teknologi informasi. Informasi tingkat kelulusan dari mahasiswa STMIK Palangkaraya sangat penting untuk meningkatkan pelayanan yang dapat membuat mahasiswa lulus tepat waktu. Penggunaan *data mining* dapat digunakan sebagai pertimbangan dalam mengambil keputusan lebih lanjut tentang faktor yang mempengaruhi kelulusan khususnya faktor dalam data induk mahasiswa.

Program Sarjana (S1) jurusan Teknik Informatika STMIK Palangkaraya adalah program pendidikan akademik setelah pendidikan menengah, yang memiliki beban studi ditentukan antara 145-149 SKS (Satuan Kredit Semester) yang ditempuh selama 8 (delapan) semester dengan masa studi paling lama 7 tahun akademik atau 14 semester (Pedoman STMIK Palangkaraya, 2020: 19-20). Jika masa studi mahasiswa dapat diketahui lebih dini, maka pihak program studi Teknik Informatika dapat melakukan tindakan-tindakan yang diperlukan supaya mahasiswa dapat lulus tepat waktu.

Berdasarkan masalah tersebut di atas, penulis mencoba melakukan penelitian untuk menganalisa dan membandingkan akurasi dua metode algoritma yang merupakan bagian dari *Top 10 algorithm data mining*, yaitu algoritma C4.5 dan *Naïve Bayes* pada sejumlah *datasets*. Adapun sumber *datasets* adalah data mahasiswa program studi Teknik Informatika (STMIK Palangkaraya) dimana setiap *datasets* memiliki jumlah data (*instances*) dan jumlah atribut (*attributes*) yang berbeda.

Pemilihan penggunaan Algoritma C4.5 dan Naïve Bayes pada penelitian ini didasarkan pada beberapa alasan, yaitu: Selain kedua algoritma tersebut sama-sama mudah diimplementasikan dan sama-sama dapat memberikan hasil yang baik dalam kasus klasifikasi, kedua algoritma tersebut juga mempunyai beberapa keunggulan masing-masing. C4.5 merupakan algoritma klasifikasi pohon keputusan yang efisien dalam menangani atribut bertipe diskret dan numerik (Han, et al., 2012). Algoritma Naive Bayes, (Han, et al., 2012) menjelaskan bahwa algoritma ini hanya membutuhkan satu kali scan data training.

Perbandingan pengukuran akurasi yang diuji akan menghasilkan informasi yang efisien, akurat dan riil. Atas dasar inilah, penulis tertarik untuk mengambil judul tugas akhir. "Perbandingan Metode Algoritma C4.5 dan Naïve Bayes Untuk Memprediksi Kelulusan Tepat Waktu Mahasiswa STMIK Palangkaraya".

1.2 Perumusan Masalah

Berdasarkan masalah yang telah diuraikan pada latar belakang, maka dapat dirumuskan permasalahan dalam penelitian ini yaitu “Bagaimana membandingkan akurasi Algoritma C4.5 dan *Naïve Bayes* untuk memprediksi kelulusan tepat waktu mahasiswa STMIK Palangkaraya?”.

1.3 Batasan Masalah

Melihat objek penelitian yang ada cukup luas serta untuk memudahkan penulis dalam pembahasan, maka penulis membatasi masalah sebagai berikut:

a. Banyaknya variabel yang mempengaruhi kelulusan mahasiswa pada STMIK Palangkaraya, akan tetapi variabel yang digunakan oleh peneliti adalah IPK Semester 1-6 dan Jenis Kelamin.
b. Data yang akan diproses hanya sebagian dari angkatan 2015 dan 2016 program studi Teknik Informatika.
c. *Tool* atau *Software* yang digunakan untuk pengujian metode yaitu Weka.
d. Pembuatan pohon keputusan berdasarkan hasil uji coba semua *dataset*s angkatan 2015 dan 2016 pada Weka.
e. *Dataset*s yang digunakan adalah data lengkap tanpa *missing value*.

1.4 Tujuan dan Manfaat

a. Tujuan

Berdasarkan rumusan masalah yang telah dibahas maka yang menjadi tujuan dari penelitian ini adalah:

1) Menganalisa perbandingan metode Algoritma C4.5 dan *Naïve Bayes* dalam melakukan prediksi kelulusan tepat waktu mahasiswa.
2) Mengetahui algoritma mana yang memiliki akurasi lebih tinggi di antara Algoritma C4.5 dan *Naïve Bayes* dalam melakukan prediksi kelulusan tepat waktu mahasiswa.

b. Manfaat

Berdasarkan tujuan penelitian yang hendak dicapai, maka penelitian ini diharapkan mempunyai manfaat dalam pendidikan baik secara langsung maupun tidak langsung. Adapun manfaat penelitian ini adalah sebagai berikut:

1) Sebagai pengalaman yang berharga dan merupakan upaya untuk menambah wawasan, pengetahuan serta kemampuan untuk menganalisa suatu masalah dan juga sebagai studi perbandingan antara teori–teori yang penulis dapatkan dengan kenyataan yang terjadi di dalam organisasi atau instansi, khususnya di STMIK Palangkaraya.
2) Diharapkan dapat membantu pihak STMIK Palangkaraya dalam pengambilan upaya untuk meningkatkan kelulusan tepat waktu mahasiswanya.
3) Dapat digunakan sebagai referensi dalam melakukan penelitian terhadap metode Algoritma C4.5 dan *Naïve Bayes*.

1.5 Sistematika Penulisan

Sistematika penulisan bertujuan untuk memberikan penjelasan guna mempermudah pembaca dalam memahami isi secara garis besar, maka sistematika penulisan dibagi menjadi dalam lima bab. Setiap bab memiliki hubungan yang saling berkaitan satu sama lainya, sistematika penulisan dapat dijelaskan sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas mengenai latar belakang, perumusan masalah, batasan masalah, tujuan penelitian dan manfaat penelitian, metodologi penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini membahas mengenai teori-teori yang digunakan penulis untuk menyusun skirpsi ini dan perangkat lunak yang digunakan.

BAB III METODE PENELITIAN

Bab ini membahas mengenai metode penelitian yang digunakan, kerangka pemikiran dan simulasi perhitungan metode *data mining* yang digunakan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas hasil dari penelitian yang dibuat secara nyata / hasil eksperimen, pengujian algoritma, analisis, evaluasi dan validasi algoritma.

BAB V KESIMPULAN DAN SARAN

Bab ini membahas mengenai kesimpulan dan saran-saran yang diperlukan untuk pengembangan sistem yang lebih lanjut

DAFTAR PUSTAKA

LAMPIRAN

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Berbagai penelitian yang dilakukan terkait Algoritma C4.5 dan *Naïve Bayes* telah banyak diteliti sebelumnya. Jurnal dan penelitian yang membahas kemiripan teori maupun subjek penelitian dijadikan sebagai acuan dalam penelitian ini. Tabel 1 merupakan penelitian terdahulu yang membahas Algoritma C4.5 dan *Naïve Bayes*.

Tabel 1 Perbandingan Penelitian

No	Penulis / Tahun	Topik Penelitian	Metode	Hasil	Perbedaan
1	Choirul Anam, Harry Budi Santoso/ 2018	Perbandingan Kinerja Algoritma C4.5 dan *Naive Bayes* untuk Klasifikasi Penerima Beasiswa	Algoritma C4.5 dan *Naïve Bayes*	Hasil penelitian menunjukkan tingkat akurasi model algoritma C4.5 sebesar 96.40% lebih baik dari tingkat akurasi model algoritma *Naive Bayes* sebesar 95.11%, sedangkan waktu proses dari kedua model algoritma yang diteliti menunjukkan hasil 0 s.	Penelitian sebelumya untuk klasifikasi penerima beasiswa. Sedangkan penulis menggunakan untuk memprediksi kelulusan tepat waktu mahasiswa STMIK Palangkaraya.
2	Syamsul Bahri, Dwi Marisa Midyanti, Rahmi Hidayati / 2018	Perbandingan Algoritma *Naive Bayes* dan C4.5 Untuk Klasifikasi Penyakit Anak	Algoritma C4.5 dan *Naïve Bayes*	hasil perbandingan menunjukkan bahwa algoritma terbaik dengan akurasi sebesar 90.00% yaitu Decision Tree C4.5. Sementara algoritma *Naive Bayes* memperoleh tingkat akurasi sebesar 89.58%	Penelitian sebelumya Untuk klasifikasi penyakit anak. Sedangkan penulis menggunakan untuk memprediksi kelulusan tepat waktu mahasiswa STMIK Palangkaraya.
3	Nurani, Afif/ 2019	Perbandingan Kinerja Algoritma *Naive Bayes* dan C4.5 Untuk Klasifikasi Harga Pangan	Algoritma C4.5 dan *Naïve Bayes*	Hasil perbandingan menunjukkan bahwa metode C45 memiliki tingkat akurasi tertinggi 65% dibandingkan algoritma *Naïve Bayes* dan *Nearest Neighbour* yaitu mencapai 60%	Penelitian sebelumya Untuk klasifikasi harga pangan. Sedangkan penulis menggunakan untuk memprediksi kelulusan tepat waktu mahasiswa STMIK Palangkaraya.
4	Norzam Yahya, Arief	Komparasi Kinerja Algoritma C.45 Dan *Naïve Bayes*	Algoritma C4.5 dan *Naïve Bayes*	Hasil tertinggi pada pengujian dengan data training sebanyak 2133 dan data testing sebanyak 533 yaitu algoritma C45 memiliki nilai	Penelitian sebelumya untuk memprediksi kegiatan penerimaan mahasiswa baru. Sedangkan penulis menggunakan

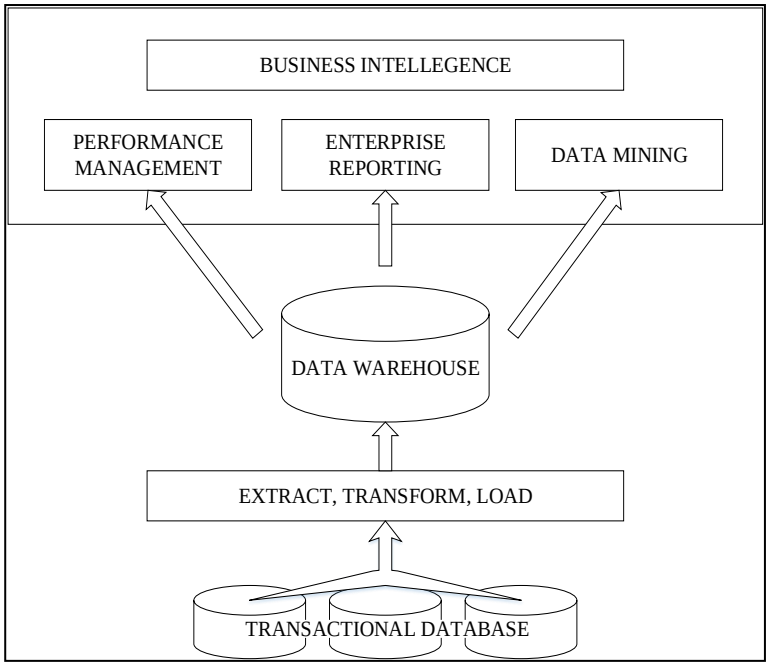
No	Penulis / Tahun	Topik Penelitian	Metode	Hasil	Perbedaan
	Jananto / 2019	Untuk Prediksi Kegiatan Penerimaan Mahasiswa Baru		akurasi lebih tinggi sebesar 88.74% dibandingkan algoritma *Naive Bayes* yang memiliki akurasi sebesar 87.24%.	untuk memprediksi kelulusan tepat waktu mahasiswa STMIK Palangkaraya.
5	Eka Fitriani/ 2020	Perbandingan Algoritma C4.5 Dan *Naïve Bayes* Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan	Algoritma C4.5 dan *Naïve Bayes*	Hasil penelitian menunjukkan algoritma C4.5 menghasilkan nilai akurasi sebesar 91,25% dan nilai AUC sebesar 0,930 dengan tingkat diagnosa *Excellent Classification* sedangkan algoritma *Naïve Bayes* menghasilkan nilai akurasi sebesar 87,11% dan nilai AUC sebesar 0,923 dengan tingkat diagnosa *Excellent Classification*. Sehingga didapat kesimpulan algoritma C4.5 merupakan algoritma yang baik untuk diterapkan pada kelayakan PKH	Penelitian sebelumya untuk menentukan kelayakan penerima bantuan program keluarga harapan. Sedangkan penulis menggunakan untuk memprediksi kelulusan tepat waktu mahasiswa STMIK Palangkaraya.

2.2 Data Mining

2.2.1 Pengertian Data Mining

Menurut Tan dalam (Prasetyo, 2012) "*data mining* adalah proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan". Istilah data mining kadang disebut juga *knowledge discovery*.

Salah satu teknik yang dibuat dalam *data mining* adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan. Kebutuhan untuk prediksi juga dapat memanfaatkan teknik ini. Dalam *data mining*, pengelompokkan data juga bisa dilakukan. Tujuannya adalah agar kita dapat mengetahui pola universal data-data yang ada. Anomali data transaksi juga dapat perlu dideteksi untuk dapat mengetahui tindak lanjut berikutnya yang dapat diambil. Semua hal tersebut bertujuan mendukung kegiatan operasional perusahaan sehingga tujuan akhir perusahaan diharapkan dapat tercapai.



Gambar 1 Posisi *data mining* dalam bisnis cerdas
Sumber: (Prasetyo, 2012)

2.2.2 Pekerjaan Dalam *Data Mining*

Menurut (Prasetyo, 2012), Pekerjaan yang berkaitan dengan data mining dapat dibagi menjadi empat kelompok, yaitu:

a. Model Prediksi (*prediction modelling*)

Model prediksi berkaitan dengan pembuatan sebuah model yang dapat melakukan pemetaan dari setiap himpunan variabel ke setiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan baru yang didapat. Ada dua jenis model prediksi, yaitu klasifikasi dan regresi. Klasifikasi digunakan untuk variabel target diskret, sedangkan regresi untuk variabel target kontinu.

b. Analisis Kelompok (*cluster analysis*)

Analisis kelompok melakukan pengelompokkan data-data ke dalam sejumlah kelompok (*cluster*) berdasarkan kesamaan karakteristik masing-masing data pada kelompok-kelompok yang ada. Data-data yang masuk dalam batas kesamaan dengan kelompoknya akan bergabung dalam kelompok yang berbeda jika keluar dari batas kesamaan dengan kelompok tersebut.

c. Analisis Asosiasi (*association analysis*)

Analisis asosiasi digunakan untuk menemukan pola yang menggambarkan kekuatan hubungan fitur dalam data. Pola yang ditemukan biasanya mempresentasikan bentuk aturan implikasi atau subset fitur. Tujuannya adalah untuk menemukan pola yang menarik dengan cara yang efisien.

d. Deteksi Anomali (*anomaly detection*)

Pekerjaan deteksi anomali berkaitan dengan pengamatan seebuah data dari sejumlah data yang secara signifkan mempunyai karakteristik yang berbeda dari sisa data yang lain. Data-data dari karakteristiknya menyimpang (berbeda) dari data yang lain disebut *outlier*. Algoritma deteksi anomali yang baik harus mempunyai laju deteksi yang tinggi dan laju eror yang rendah. Deteksi anomali dapat diterapkan pada sistem jaringan untuk mengetahui pola data yang memasuki jaringan sehingga penyusupan bisa ditemukan jika pola kerja data yang datang berbeda. Perilaku kondisi cuaca yang mengalami anomali juga dapat dideteksi dengan algoritma ini.

2.3 Set Data

Set data (*dataset*) dapat dipandang sebagai kumpulan objek data. Nama lain yang sering digunakan adalah *record, point, event, observation, case*, atau bahkan data. Sementara objek data digambarkan dengan sejumlah atribut yang menangkap (*capture*) karakter dasar objek data, contohnya tinggi badan yang memberikan nilai kuantitatif tinggi badan seseorang, waktu yang menangkap saat sebuah peristiwa terjadi. Atribut terkadang juga diisebut variabel, karakteristik, medan (*field*), fitur, atau dimensi (Prasetyo, 2012).

Atribut adalah sifat atau properti atau karakteristik objek data yang nilainya bisa bermacam-macam dari satu objek ke objek yang lain, dari satu waktu ke waktu yang lain. Ada empat sifat penting yang dimiliki atribut secara umum, yaitu:

a. *Distinctness*, = dan ≠
b. *Order*, <, ≤, > dan ≥
c. *Addition*, + dan –
d. *Multiplication*, * dan /

Dari keempat sifat tersebut dapat diturunkan empat tipe atribut, yaitu nominal, ordinal, interval dan rasio. Tabel 2 menginformasikan penjelasan ketiga tipe atribut dan kaitannya dengan sifat-sifat diatas.

Tabel 2 Tipe Atribut

Tipe Atribut		**Penjelasan**	**Contoh**
Kategoris (Kualitatif)	Nominal	Nilai atribut bertipe nominal memberikan nilai berupa nama. Dengan nama inilah sebuah atribut membedakan dirinya pada data yang satu dengan yang lain (=, ≠).	Kode pos, nomor KTP, nomor induk mahasiswa, jenis kelamin
	Ordinal	Nilai atribut bertipe ordinal mempunyai nilai berupa nama yang mempunyai arti informasi terurut (<, ≤, >, ≥).	Tingkat kelulusan {cumlaude, sangat memuaskan, memuaskan}.
Numerik (Kuantitatif)	Interval	Nilai atribut dimana perbedaan diantara dua nilai mempunyai makna yang berarti (+, -)	Tanggal, suhu (dalam Celcius atau Fahrenheit)
	Rasio	Nilai atribut dimana perbedaan di antara dua nilai dan rasio dua nilai mempunyai makna yang berarti (*, /)	Suhu (dalam kelvin), umur panjang, tinggi.

Sumber: (Prasetyo, 2012)

2.4 Klasifikasi

2.4.1 Konsep Klasifikasi

Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Dalam klasifkasi ada dua pekerjaan utama yang dilakukan, yaitu (1) pembangunan model sebagai prototipe untuk disimpan sebagai memori dan (2) penggunaan model tersebut untuk melakukan pengenalan/klasifikasi/prediksi pada suatu objek data lain agar diketahui di kelas mana objek data tersebut dalam model yang sudah disimpannya (Prasetyo, 2012).

2.4.2 Pengukuran Kinerja Klasifikasi

Dataset hanya memiliki dua kelas, yaitu *class* positif dan *class* negatif, maka dapat dibuatkan Tabel 3 (Wati, 2016).

Tabel 3 *Class* pada *Confussion Matrix*

Classification		*Predicted Class*	
		Class = Yes	*Class* = No
Observed Class	*Class* = Yes	*a true positive* – (TP)	*b false negative* – (FN)
	Class = No	*c false positive* – (FP)	*d true negative* – (TN)

Keterangan:

True Positive (TP) : Proporsi positif dalam dataset yang diklasifikasikan positif

True Negative (TP) : Proporsi negatif dalam dataset yang diklasifikasikan negatif

False Positive (FP) : Proporsi negatif dalam dataset yang diklasifikasikan positif

False Negative (FN) : Proporsi negatif dalam dataset yang diklasifikasikan negatif

Berikut adalah persamaan model *Confussion Matrix*:

a. Nilai *Accuracy* adalah proporsi jumlah prediksi yang benar. Dapat dihitung dengan menggunakan persamaan:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

b. *Specificity* digunakan untuk membandingkan proporsi TN terhadap tupel negatif, yang dihitung dengan menggunakan persamaan:

$$Specificity = \frac{TN}{TN + FP} \quad (2)$$

c. *Negative Predictive Value* (NPV) adalah proporsi kasus dengan hasil *diagnose negative*, yang dihitung dengan menggunakan persamaan:

$$NPV = \frac{TN}{TN + FN} \quad (3)$$

2.5 Pohon Keputusan

2.5.1 Pengertian Pohon Keputusan

Pohon keputusan (*decision tree*) merupakan salah satu teknik terkenal dalam *data mining* dan merupakan salah satu metode yang populer dalam menentukan keputusan suatu kasus (Widodo, et al., 2013).

Menurut (Widodo, et al., 2013), masalah utama pada pembuatan pohon keputusan adalah variabel manakah yang menjadi akar dari pohon tersebut. Akar di sini adalah pemisah pertama dari pohon keputusan. Dikenal dengan istilah *Bayesian Score* yang menilai suatu variabel, atau dalam Pohon Keputusan terkenal dengan sebutan *Entropy*. *Entropy* dihitung dengan rumusan sebagai berikut:

$$Entropy(p_1, p_2, \dots, p_n) = -p_1 log_2 p_1 - p_2 log_2 p_2 \dots - p_n log_2 p_n \qquad (4)$$

Di mana p_1, p_2, ..., p_n adalah probabilitas tiap kondisi pada atribut yang diprediksi dan:

$$p_1 + p_2 + \dots p_n = 1 \qquad (5)$$

2.5.2 Model Pohon Keputusan

Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. *Decision tree* merupakan metode klasifikasi yang paling popular digunakan. Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun mudah untuk dipahami. Pada *decision tree* terdapat 3 jenis *node*, yaitu:

a. *Root Node*, merupakan *node* paling atas, pada node ini tidak ada input dan bisa tidaknya mencapai output atau mempunyai output lebih dari satu.
b. *Internal Node*, merupakan node percabangan, pada node ini hanya terdapat satu input dan mempunya output minimal dua.
c. *Leaf node* atau *terminal node*¸ merupakan node akhir, pada node ini hanya terdapat satu input dan tidak mempunyai nilai angkut.

2.6 Algoritma C4.5

2.6.1 Pengertian Algoritma C4.5

Algoritma C4.5 merupakan struktur pohon di mana terdapat simpul yang mendeskripsikan atribut-atribut, setiap cabang menggambarkan hasil dari atribut yang diuji dan setiap daun menggambarkan kelas (Widodo, et al., 2013).

Menurut (Han, et al., 2012), algoritma C4.5 menggunakan konsep *information gain* atau *entropy reduction* untuk memilih pembagian yang optimal.

2.6.2 Tahapan Algoritma C4.5

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5 (Widodo, et al., 2013), yaitu:

a. Menyiapkan *data training*. *Data training* biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas-kelas tertentu.
b. Menentukan akar dari pohon. Akar akan diambil dari atribut yang terpilih, dengan cara mengitung nilai *gain* dari masing-masing atribut, nilai *gain* yang paling tinggi

yang menjadi akar pertama. Sebelum mengitung nilai *gain* dari atribut, hitung dulu nilai *entropy*. Untuk menghitung nilai *entropy* digunakan rumus:

$$Entropy(S) = \sum_{i=0}^{n} - pi \cdot \ log2\, pi \qquad (6)$$

Keterangan:

S = himpunan kasus

n = jumlah partisi S

p_i = proporsi Si terhadap S

c. Kemudian hitung nilai *gain* menggunakan rumus:

$$Gain(S, A) = Entropy(S) - \sum_{i=0}^{n} \frac{|Si|}{|S|} * Entropy\ (Si) \qquad (7)$$

Keterangan:

S = himpunan kasus

A = fitur

n = jumlah partisi atribut A

| Si |= proporsi Si terhadap S

| S | = jumlah kasus dalam S

d. Ulangi langkah ke-2 hingga semua *record* terpartisi.

e. Proses partisi pohon keputusan akan berhenti saat:

1) Semua *record* dalam simpul N mendapat kelas yang sama.

2) Tidak ada atribut di dalam *record* yang dipartisi lagi.

3) Tidak ada *record* di dalam cabang yang kosong.

2.7 Naïve Bayes

Menurut (Prasetyo, 2012). *Bayes* merupakan teknik prediksi probalistik sederhana yang berdasar pada penerapan teorema *Bayes* (atau aturan *bayes*) dengan asumsi independensi (ketidak tergantungan) yang kuat (naif). Dengan kata lain, dalam *Naïve Bayes*, model yang di gunakan adalah ”model fitur independen”.

Dalam *Bayes* (terutama *Naïve Bayes*), maksud independensi yang kuat pada fitur adalah bahwa sebuah fitur pada sebuah data tidak berkaitan dengan ada atau tidak adanya fitur lain dalam data yang sama. Contohnya, pada kasus klasifikasi hewan dengna fitur penutup kulit, melahirkan, berat dan menyusui. Disini ada ketergantungan pada fitur menyusui karena hewan yang menyusui biasanya melahirkan, atau hewan bertelur tidak menyusui. Dalam *Bayes*, hal tersebut tidak dipandang sehingga masing-masing fitur seolah tidak memiliki hubungan apa pun.

Prediksi *Bayes* didasarkan pada teorema *Bayes* dengan formula umum sebagai berikut:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} \tag{8}$$

Keterangan:

P(H|E) = Probabilitas akhir bersyarat (*conditional probability*) suatu hipotesis H terjadi jika diberikan bukti (*evidence*) E terjadi.

P(E|H) = Probabilitas sebuah bukti E terjadi akan mempengaruhi hipotesis H.

P(H) = Probabilitas awal hipotesis H terjadi tanpa memandang bukti apapun.

P(E) = Probabilitas awal bukti E terjadi tanpa memandang hipotesis atau bukti yang lain.

Ide dasar dari aturan *Bayes* adalah bahwa hasil dari hipotesis atau peristiwa (H) dapat diperkirakan pada beberapa bukti (E) yang diamati. Ada beberepa hal penting dari aturan *Bayes* yaitu:

1) Sebuah probabilitas awal/prior H atau (PH) adalah probabilitas dari suatu hipotesis sebelum bukti diamati.
2) Sebuah probabilitas akhir H atau (PH) adalah probabilitas dari suatu hipotesis setelah bukti diamati.

Teorema Bayes juga bisa menangani beberapa bukti, misalnya ada E1, E2 dan E3 sehinga akhir untuk hipotesis (H) dapat dihitung dengna cara berikut.

$$P(H|E_1, E_2, E_3) = \frac{P(E_1, E_2, E_3|H)P(H)}{P(E_1, E_2, E_3)} \quad (9)$$

Karena yang digunakan untuk bukti adalah independen, bentuk di atas dapat diubah menjadi.

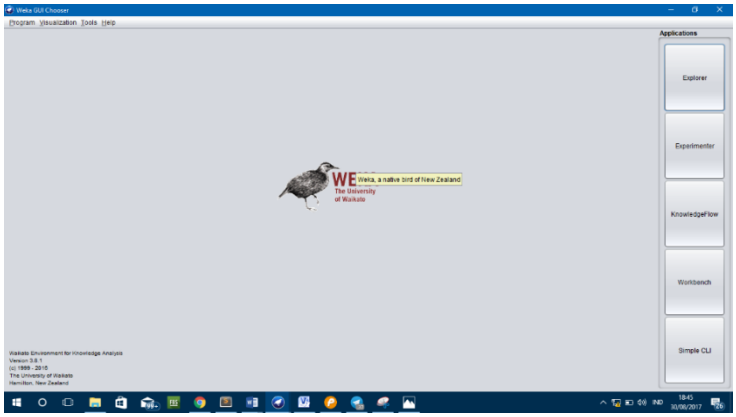
$$P(H|E_1, E_2, E_3) = \frac{P(E_1|H)P(E_2|H)P(E_3|H)P(H)}{P(E_1)P(E_2)P(E_3)} \quad (10)$$

2.8 Weka

The Waikato Environment for Knowledge Analysis (Weka) adalah rangkaian lengkap perpustakaan kelas *Java* yang mengimplementasikan banyak *state-of-the-art* pembelajaran mesin dan algoritma *Data Mining*. Weka tersedia secara bebas di *World*

Wide Web dan menyertai teks baru pada dokumen *Data Mining* dan sepenuhnya menjelaskan semua algoritma yang dikandungnya. Aplikasi yang ditulis menggunakan *library class* pada Weka yang dapat dijalankan pada komputer manapun dengan kemampuan *Browsing Web*, ini memungkinkan pengguna untuk menerapkan teknik pembelajaran mesin untuk data mereka sendiri terlepas dari platform komputer. (Witten, et al., 2011). Empat tombol di atas dapat digunakan untuk menjalanankan Aplikasi yang berfungsi untuk:

a. *Explorer* digunkan untuk menggali lebih jauh data dengan aplikasi Weka.
b. Experimenter digunakan untuk melakukan percobaan dengan pengujian statistik skema belajar.
c. Knowledge Flow digunakan untuk pengetahuan pendukung.
d. Simple CLI antar muka dengan menggunakan tampilan command-line yang memungkinkan langsung mengeksekusi perintah weka untuk Sistem Operasi yang tidak menyediakan secara langsung.



Gambar 2 Tampilan Weka

BAB III

METODE PENELITIAN

3.1 Metode Penelitian

Penelitian ini adalah penelitian eksperimen dengan langkah-langkah atau metode penelitian sebagai berikut:

a. Pengumpulan Data

Adapun data yang digunakan dalam penelitian ini adalah data primer dan data sekunder.

1) Data Primer adalah data yang diperoleh secara langsung dari sumber, yaitu diperoleh secara langsung dari STMIK Palangkaraya yaitu dari data akademik mahasiswa angkatan 2015 dan 2016.
2) Data Sekunder yaitu data yang diperoleh secara tidak langsung, misalnya dari dokumentasi, literatur, buku, jurnal dan informasi lainnya yang ada hubungannya dengan masalah yang diteliti.

b. Pengolahan Awal Data

Data yang digunakan adalah data akademik mahasiswa, data terdiri dari NIM, tempat tanggal lahir, asal sekolah, IPK semester 1 sampai semester 6 dan tanggal kelulusan. Data diambil dari data mahasiswa tahun 2015 dan 2016. Dari data yang sudah ada di pecah lagi menjadi beberapa atribut yaitu NIM, Jenis Kelamin, IPK semester 1, IPK semester 2, IPK semester 3, IPK semester 4, IPK semester 5, IPK semester 6 dan kelulusan.

c. *Cleaning Data*

Proses *cleaning data* atau pembersihan dilakukan terhadap data yang redudan atau ganda, inkonsisten, *missing value* atau *missing data* dan *outlier data*. Hal tersebut perlu dilakukan agar tidak mempengaruhi performa proses klasifikasi yang dilakukan. Pembersihan data yang dilakukan juga menghapus fitur yang tidak diperlukan dalam proses klasifikasi nantinya.

d. *Training Data* dan Pengujian Metode dengan *software* Weka

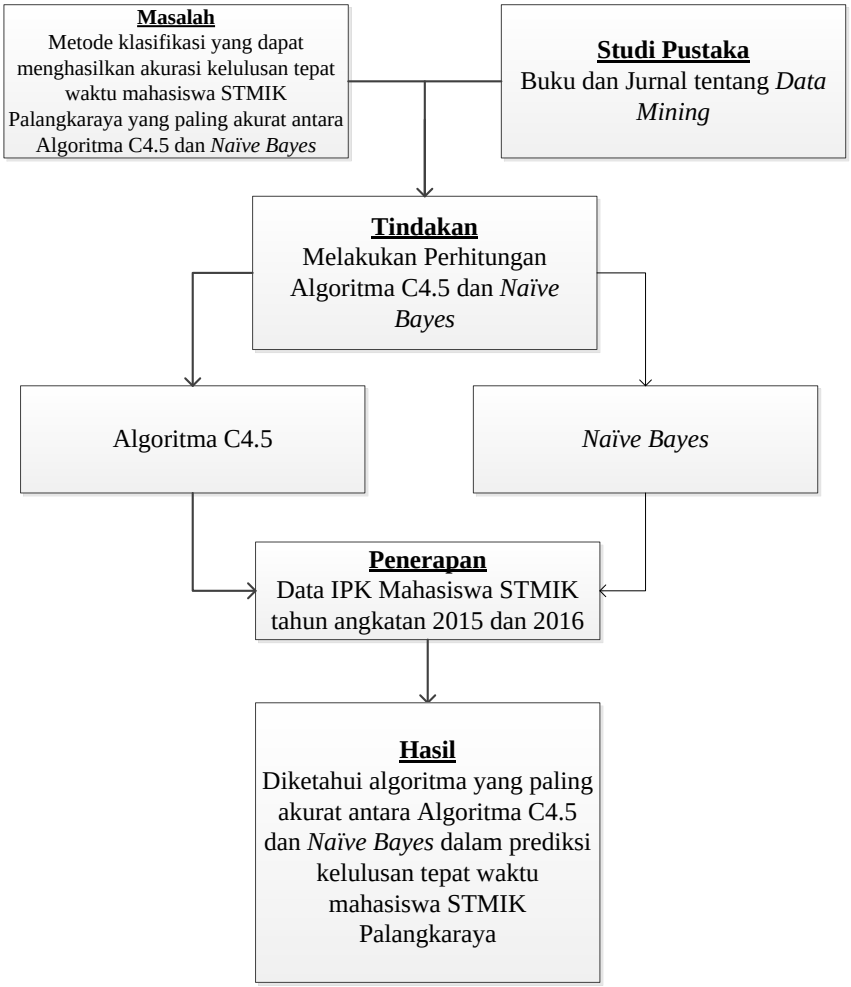
Pemrosesan *training data* diambil dari sebagian *dataset* mahasiswa. Pengujian dilakukan menggunakan *software* Weka, yang mana memasukkan data menggunakan file berformat *excel*. Besarnya proporsi data yang dilakukan pengujian *training data* adalah 60%, sedangkan sisanya digunakan untuk uji coba metode. Pengujian *training data* sehingga akan menghasilkan beberapa aturan dan akan membentuk sebuah pohon keputusan. Pada penelitian ini komparasi algoritma yang akan digunakan ada dua yaitu algoritma klasifikasi C4.5 dan *Naïve Bayes*.

e. Evaluasi dan validasi penelitian

Pada penelitian ini akan dilakukan evaluasi dengan cara menghitung jumlah prediksi lulus tepat waktu dan prediksi lulus tidak tepat waktu dengan pada algoritma C4.5 dan *Naïve Bayes*. Langkah selanjutnya dilakukan proses analisis data, akan dihitung tingkat akurasi menggunakan rumus *Confusion Matrix*.

3.2 Kerangka Pemikiran

Penulis perlu membuat gambaran singkat sebagai alur penyusunan laporan ini dengan kerangka pemikiran pada Gambar 3.



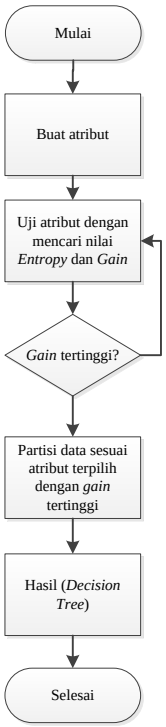
Gambar 3 Kerangka Pemikiran Penelitian

3.1.1 Algoritma

Pada penelitian ini menggunakan 2 algoritma yaitu Algoritma C4.5 dan *Naïve Bayes* untuk mengklasifikasikan kelulusan tepat waktu mahasiswa STMIK Palangkaraya berdasarkan IPK semester 1-6 dan jenis kelamin, kinerja kedua algoritma tersebut diukur berdasarkan *accuracy*, *recall* dan *precision*.

a. Algoritma C4.5

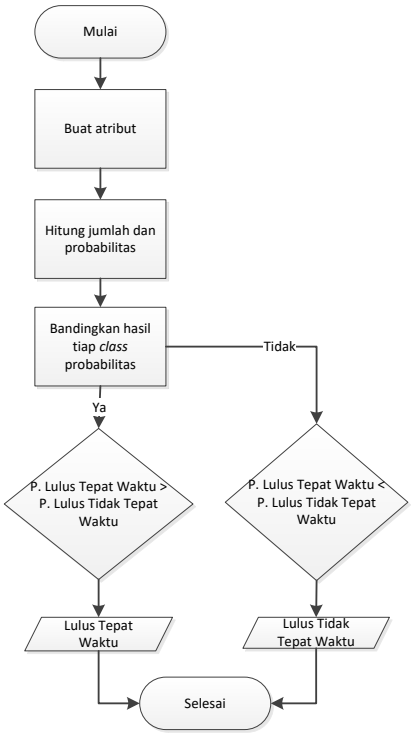
Secara umum, *Decision Tree* (pohon keputusan) merupakan suatu bentuk model dari suatu masalah yang tersusun dari rangkaian keputusan yang menunjuk pada pemecahan masalah yang dihasilkan. Gambar 3 merupakan *flowchart* dari Algoritma C4.5 sampai dengan menghasilkan *Decision Tree*.



Gambar 4 *Flowchart* Algoritma C4.5

b. *Naïve Bayes*

Naïve Bayes merupakan metode *supervised document classification* yang membutuhkan data *training* sebelum melakukan proses klasifikasi. Dalam prosesnya, *Naive Bayes* mengansumsikan ada tidaknya sebuah fitur pada suatu kelas tidak berhubungan dengan ada tidaknya fitur lain di kelas lain. Dalam proses pelatihan, variabel telah ditentukan kategorinya. Kemudian akan diproses untuk membentuk pengetahuan berupa nilai probabilitas pada setiap *term* (kata). Proses ini akan menghasilkan sebuah *term* (kata) pada setiap varibel yang mengkarakteristikkan suatu kategori tertentu. Gambar 5 merupakan *flowchart* dari algoritma *Naïve Bayes*.



Gambar 5 *Flowchart Naïve Bayes*

3.1.2 Simulasi Perhitungan Manual

Simulasi perhitungan manual dengan menggunakan Algoritma C4.5 dan *Naïve Bayes* dapat menjadi acuan pembuktian terhadap beberapa sampel data mengenai tingkat akurasi kebenarannya. Pada proses perhitungan manual terdapat beberapa kebutuhan yang harus terpenuhi seperti data *training* dan kelas kategorial. Berikut akan disajikan simulasi perhitungan manual pada perbandingan kedua metode ini:

a. Data *Training*

Tabel 4 Sampel Data *Training*

Mahasiswa	IPK Semester						Kelulusan
	1	2	3	4	5	6	
Mahasiswa 1	3.11	2.93	3.31	3.45	3.40	3.35	Tepat Waktu
Mahasiswa 2	2.93	3.21	3.11	3.12	3.13	3.05	Tepat Waktu
Mahasiswa 3	2.90	2.61	3.19	3.25	3.27	3.24	Tidak Tepat Waktu
Mahasiswa 4	2.93	2.93	3.21	3.34	3.37	3.31	Tepat Waktu
Mahasiswa 5	2.79	2.77	2.95	3.18	3.24	3.24	Tepat Waktu
Mahasiswa 6	2.48	2.67	2.62	2.74	2.93	2.96	Tidak Tepat Waktu
Mahasiswa 7	2.50	2.72	2.66	2.74	2.75	2.73	Tidak Tepat Waktu
Mahasiswa 8	2.71	2.81	2.69	2.70	2.75	2.62	Tepat Waktu
Mahasiswa 9	3.33	3.44	3.39	3.43	3.45	3.29	Tepat Waktu
Mahasiswa 10	2.43	2.93	3.03	3.10	3.14	3.04	Tidak Tepat Waktu
Mahasiswa 11	3.57	3.67	3.53	3.61	3.59	3.60	Tepat Waktu
Mahasiswa 12	2.95	3.19	3.15	3.26	3.27	3.17	Tidak Tepat Waktu

b. Kelas Kategorial

Tabel 5 Kelas Kategorial

IPK	Predikat
2,00 – 2,75	Memuaskan
2,76 – 3,50	Sangat Memuaskan
3,51 – 4,00	Dengan Pujian (*Cumlaude*)

c. Perhitungan Manual

1) Algoritma C4.5

Dengan menggunakan rumus (6), kita dapat menghitung nilai *Entropy*. Menghitung *Entropy* Total dilakukan dengan cara menghitung jumlah Kelulusan "Tepat Waktu" dan "Tidak Tepat Waktu" dari seluruh kasus yang ada.

$$- Entropy(\text{Total}) = \left(-\frac{7}{12} * \log_2\left(\frac{7}{12}\right)\right) + \left(-\frac{5}{12} * \log_2\left(\frac{5}{12}\right)\right) = 0{,}979868757$$

Entropy Total adalah menghitung nilai total kelulusan Tepat Waktu (7) dan Tidak Tepat Waktu (5), sedangkan 12 adalah jumlah keseluruhan data *training*. Kemudian hitung masing-masing *Entropy* dari semua nilai variabel yang ada. Peneliti menggunakan atribut IPK Semester 1 dan IPK Semester 2 sebagai contoh untuk simulasi perhitungan nilai *Information Gain*. Dan dengan menggunakan rumus (7) kita dapat menghitung nilai *Information Gain* tiap-tiap Atribut.

$$- Gain(\text{Total}, \text{IPK1}) = 0{,}979868757 - \left(\left(\frac{4}{12} * 0{,}811278124\right) + \left(\frac{7}{12} * 0{,}863120569\right) + \left(\frac{1}{12} * 0\right)\right) = 0{,}205955717$$

$$- Gain(\text{Total}, \text{IPK2}) = 0{,}979868757 - \left(\left(\frac{3}{12} * 0\right) + \left(\frac{8}{12} * 0{,}811278124\right) + \left(\frac{1}{12} * 0\right)\right) = 0{,}439016674$$

Setelah semua nilai *Entropy* dan *Information Gain* dihitung, kemudian hasil dari perhitungan tersebut dimasukkan ke dalam Tabel 6. Dari perhitungan nilai *entropy* dan *information gain* diatas, dapat diketahui bahwa nilai *information gain* terbesar adalah IPK Semester 2 dengan nilai 0,439016674 dan yang terkecil adalah IPK Semester 1, dengan nilai 0,205955717. Kemudian, ambil nilai *Information Gain* terbesar yaitu atribut

IPK Semester 2 dan jadikan simpul akar. Hilangkan atribut yang dipilih sebelumnya dan ulangi perhitungan nilai *Entropy*, *Information Gain*, dengan memilih *Information Gain* terbesar dan dijadikan simpul internal pohon. Ulangi perhitungan tersebut hingga semua atribut pohon memiliki kelas.

Tabel 6 Hasil perhitungan *Entropy* dan *Gain*

		Jml Kelulusan	Tepat Waktu	Tidak Tepat Waktu	*Entropy*	*Information Gain*
TOTAL		12	7	5	0,979868757	
IPK SMT 1						0,205955717
	Memuaskan	4	1	3	0,811278124	
	Sangat Memuaskan	7	5	2	0,863120569	
	Cumlaude	1	1	0	0	
IPK SMT 2						0,439016674
	Memuaskan	3	0	3	0	
	Sangat Memuaskan	8	6	2	0,811278124	
	Cumlaude	1	1	0	0	

2) *Naïve Bayes*

Tabel 7 Contoh Data Testing Untuk Perhitungan *Naïve Bayes*

Mahasiswa	IPK Semester 1	IPK Semester 2	Kelulusan
Mahasiswa 1	3.11 (Sangat Memuaskan)	2.93 (Sangat Memuaskan)	????

Tabel 8 Hasil perhitungan *Naïve Bayes*

Step 1				
Probabilitas Lulus Tepat Waktu	PTW	= Lulus Tepat Waktu / Jumlah Data	= 7 / 12	0,583333
Probabilitas Lulus Tidak Tepat Waktu	PTTW	= Lulus Tidak Tepat Waktu / Jumlah Data	= 5 / 12	0,416667

Step 2 : Membandingkan setiap atribut yang dimiliki dataset			
Case 1	IPK Semester 1 = Sangat Memuaskan & Kelulusan = Tepat Waktu	5 / 7	0,7142857142857143
	IPK Semester 1 = Sangat Memuaskan & Kelulusan = Tidak Tepat Waktu	2 / 5	0,4
Case 2	IPK Semester 2 = Sangat Memuaskan & Kelulusan = Tepat Waktu	6 / 7	0,8571428571428571
	IPK Semester 2 = Sangat Memuaskan & Kelulusan = Tidak Tepat Waktu	2 / 5	0,4

Step 3 : Mengalikan seluruh probabilitas Tepat Waktu dan Tidak Tepat Waktu	
Hasil Tepat Waktu	0,6122448976122449
Hasil Tidak Tepat Waktu	0,16
Step 4 : Membandingkan probabilitas Tepat Waktu dan Tidak Tepat Waktu	
Jika Tepat Waktu > Tidak Tepat Waktu	
0,6122448976122449 > 0,16 = **Tepat Waktu**	
Untuk **IPK Semester 1** = Sangat Memuaskan, **IPK Semester 2** = Sangat Memuaskan maka masuk ke kelas Kelulusan = "**Tepat Waktu**"	

DAFTAR PUSTAKA

Anam, C. & Santoso, H. B., 2018. Perbandingan Kinerja Algoritma C4.5 dan Naive Bayes untuk Klasifikasi Penerima Beasiswa. 5 April.Volume 8.

Bahri, S., Madyanti, D. M. & Hidayati, R., 2018. Perbandingan Algoritma Naive Bayes dan C4.5. *Seminar Nasional Aplikasi Teknologi Informasi (SNATi),* 11 Agustus.pp. 24-31.

Fitriani, E., 2020. Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan. *Jurnal Sistem Informasi,* Januari, 9(1), pp. 103-115.

Han, J., Kamber, M. & Pei, J., 2012. *Data mining: concepts and techniques morgan kaufmann.* s.l.:s.n.

N. & A., 2019. Perbandingan Kinerja Algoritma Naive Bayes dan C4.5 Untuk Klasifikasi Harga Pangan. *Jurnal Ilmiah Teknik Elektro,* Mei, 7(1), pp. 20-24.

Prasetyo, E., 2012. *Data Mining, Konsep dan Aplikasi menggunakan MATLAB.* Yogyakarta: Andi.

Widodo, P. P., Trias, H. R. & H., 2013. *Penerapan Data Mining dengan Matlab.* Bandung: Rekayasa Sains.

Wati, R., 2016. Penerapan Algoritma Genetika Untuk Seleksi Fitur Pada Analisis Sentimen Review Jasa Maskapai Penerbangan Menggunakan Naive Bayes. *Jurnal Evolusi,* 4(1), pp. 25-31.

Witten, I., Frank, E. & Hall, M., 2011. *Data Mining: Practical Machine Learning Tools and Techniques.* 3rd ed. s.l.:Elsevier.

Yahya, N. & Jananto, A., 2019. Komparasi Kinerja Algoritma C.45 Dan Naïve Bayes Untuk Prediksi Kegiatan Penerimaan Mahasiswa Baru. *Proceeding SENDI_U,* pp. 221-228.

LAMPIRAN

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangkaraya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

SURAT TUGAS

No.8/STMIK-3.C.2/AU/I/2021

Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya menugaskan nama- nama tersebut di bawah ini :

1. Nama : Lili Rusdiana, M.Kom.
   NIK : 198707282011007
   Sebagai Pembimbing I Dalam Pembuatan Program
2. Nama : H. Suratno, S.Kom., M.Si.
   NIK : 196912281997101
   Sebagai Pembimbing II Dalam Penulisan Tugas Akhir

Untuk membimbing Tugas Akhir mahasiswa :

Nama	: Aulia Rahmayanti
NIM	: C1555201057
Program Studi	: TEKNIK INFORMATIKA (55201)
Tanggal Daftar	: 18 Januari 2021
Judul Tugas Akhir	: Perbandingan Metode Algoritma C4.5 dan Naive Bayes Untuk Memprediksi Kelulusan Tepat Waktu Mahasiswa STMIK Palangkaraya

Demikian surat ini dibuat agar dapat dipergunakan sebagaimana mestinya dan dilaksanakan dengan penuh tanggung jawab.

Palangka Raya, 25 Januari 2021
Ketua Program Studi,

Hotmian Sitohang, M.Kom.
NIK. 198503282008002

Tembusan :
1. Pembimbing I dan II
2. Mahasiswa yang bersangkutan
3. Arsip

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangkaraya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

KARTU KEGIATAN KONSULTASI
TUGAS AKHIR

Nama Mahasiswa : Aulia Rahmayanti
NIM : C1555201057
Tanggal Persetujuan Judul : 25 Januari 2021
Judul Tugas Akhir : Perbandingan Metode Algoritma C4.5 dan Naive Bayes Memprediksi Kelulusan Tepat Waktu Mahasiswa STMIK

No.	Tanggal Konsultasi		Uraian	Tanda Tangan
	Terima	Kembali		
	27/2 2021		Perbaiki penulisan	
	8/2		lanjutkan penulisan, sesuaikan pedoman	
	8/2 2021		Lanjutkan revisi yang diminta lengkapi berkas	
	2/3 2021		ACC	
	20/2 2021	21/2	lengkapi bab 3 sesuai topik	
	22/2		Acc seminar	

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangkaraya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

SURAT TUGAS
PENGUJI SEMINAR PROPOSAL TUGAS AKHIR
No.49/STMIK-3.C.2/AK/III/2021

Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya menugaskan kepada nama-nama berikut :

1. Nama : Sulistyowati, S.Kom., M.Cs.
   NIK : 198212162007002
   Sebagai Ketua
2. Nama : Lili Rusdiana, M.Kom.
   NIK : 198707282011007
   Sebagai Sekretaris
3. Nama : H. Suratno, S.Kom., M.Si.
   NIK : 196912281997101
   Sebagai Anggota

Tim Penguji Seminar Proposal Tugas Akhir :

Nama	:	Aulia Rahmayanti
NIM	:	C1555201057
Hari/Tanggal	:	Sabtu, 13 Maret 2021
Waktu	:	08.00 WIB
Judul Proposal	:	Perbandingan Metode Algoritma C4.5 dan Naive Bayes Untuk Memprediksi Kelulusan Tepat Waktu Mahasiswa STMIK Palangkaraya

Demikian surat ini dibuat agar dapat dipergunakan sebagaimana mestinya dan dilaksanakan dengan penuh tanggung jawab.

Palangka Raya, 10 Maret 2021
Ketua Program Studi Teknik Informatika

Lili Rusdiana, M.Kom
NIK. 198707282011007

Tembusan :
1. Dosen Penguji
2. Mahasiswa yang Bersangkutan
3. Arsip Prodi

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangkaraya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

BERITA ACARA
SEMINAR PROPOSAL TUGAS AKHIR

Periode (Bulan) : ...Maret.....Tahun ..2021

1. Hari/Tanggal Seminar : Selasa ..... / 13.....
2. Waktu (Jam) : 8.00..... .WIB sampai dengan ....WIB
3. Nama Mahasiswa : Aulia P.A.Lumayanti.....
4. Nomor Induk Mahasiswa : C1555 201027.....
5. Program Studi : Teknik Informatika
6. Tahun Angkatan : 2015.....
7. Judul Tugas Akhir : PERBANDINGAN METODE ALGORITMA C4.5 DAN NAIVE BAYES UNTUK KELULUSAN TEPAT WAKTU MAHASISWA STMIK PALANGKARAYA.....
8. Dosen Penguji :

	Nama	Nilai	Tanda Tangan
1.	Sulistyowati	=	( )
2.	Uli R	=	( )
3.	SUPATMO	=	( )

9. Hasil Ujian : **LULUS / ~~TIDAK LULUS~~ *)** NILAI = 81,47

   Dengan Perbaikan/ T~~anpa Perbaikan~~ *)

10. Catatan Penting : 1. Lama Perbaikan : 14 hari (Maks. 15 hari)
    2. Jika lebih dari 15 hari s/d 1 (satu) bulan dikenakan sanksi berupa denda sebesar Rp. 300.000,- (Tiga ratus ribu rupiah), dan jika lebih dari 1 (satu) bulan dikenakan denda Rp. 600.000,- (Enam Ratus ribu rupiah) per bulan dari tanggal ujian
    3. Jika lebih dari 3 (tiga) bulan dari tanggal ujian maka hasil ujian dibatalkan dan wajib mengajukan judul dan pembimbing baru. Wajib membayar Denda dan membayar biaya seminar ulang.

Palangka Raya, 13 - 3 - 2021

Mengetahui :
Ketua Program Studi Teknik Informatika,

Lili Rusdiana, M.Kom.
NIK. 1987072820110 07

Ketua Penguji,

Sulistyowati
NIK. 87002

Tembusan :
1. Arsip Prodi Teknik Informatika
2. Mahasiswa yang bersangkutan

Dibawa saat konsultasi perbaikan dengan dosen penguji

**) Coret yang tidak perlu*

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp. 0536-3225515 Fax. 0536-3236933 Palangkaraya
Email : stmikplk@gmail.com – website : www.stmikplk.ac.id

KARTU KEGIATAN SEMINAR PROPOSAL SKRIPSI

Nama Mahasiswa : Aulia Rahmayanti
NIM : C1555201057
Jurusan : ~~Sistem Informasi~~/ Teknik Informatika*)

No.	Hari Tanggal	Judul	Mahasiswa Penyaji	Nama Tim Dosen	Tanda Tangan
1	3 Juni 2016	~~Apl~~ Aplikasi kamus Dayak Kahayan - Indonesia - Ma'ayan Berbasis web	RONNIE	Lili Rusdiana, M.kom Rudini M.Pd Ferdiyani Haris	
2.	Sabtu 2/3-16	Pengenalan Monumen Bersejarah Kota Palangka Raya Berbasis AUGMENTED REALITY	SIXCROWN	Maura Widyaningsih, S.kom, M.CS Sulistyowati, S.kom, M.CS Catharina Elwayantie, M.Pd	
3	Jum'at 15/3-16	Rancang Bangun Sistem pengenalan wajah untuk kehadiran mahasiswa menggunakan metode EIGENFACE berbasis Android	M. Dadang Setiawan	Lili Rusdiana M.kom Sulistyowati. S.kom., M.CS Herkolas, S.kom., M.CS	
4	Jum'at 15/3-16	Rancang Bangun layar informasi masjid untuk daerah palangkaraya dan sekitarnya berbasis - web Mobile	AMMAUDIN	MOchammad Ichsan, ST., M.kom. Sulistyowati, S.kom., M.CS Hamlan sitohang, M.kom	
5	Sabtu, 16/3-16	Algoritma Jaringan Syaraf Tiruan untuk Prakiraan Cuaca menggunakan metode Hopfield	Amalia Rohmah	Lili Rusdiana, M.kom Sulistyowati, S.kom., M.CS. Elok Faiqotul H., S.Si, M.Sc.	

Palangka Raya, 3 Juni 2016
Mahasiswa ybs,

Aulia Rahmayanti

Keterangan :
- *) Coret yang tidak perlu
- Harap kartu jangan sampai hilang, digunakan sebagai syarat seminar
- Minimal 5 (lima) kali mengikuti seminar

DAFTAR HADIR PESERTA
SEMINAR PROPOSAL TUGAS AKHIR

1. Nama Penyaji : Aulia Rahmayanti
2. Hari/ Tanggal : Sabtu, 13 Maret 2021
3. Waktu : 08.00 – 09.00 WIB
4. Judul Proposal : Perbandingan Metode Algoritma C4.5 dan Naive Bayes untuk memprediksi kelulusan Tepat Waktu Mahasiswa STMIK Palangkaraya

No.	Nama Mahasiswa	NIM	Tanda Tangan
1	Maulida Anggriani	C1555201061	
2	Samsul Hadi	C1655201018	
3	Nella Saputri	C1555201007	
4	Meysa Rani	C1555201075	
5	Sabiqur Zurdin		
6	Vandri	C1555201017	
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			

Palangka Raya, 13 Maret 2021

Mengetahui :
Ketua Tim Penguji,

Mahasiswa Penyaji,

Aulia Rahmayanti