

**PENERAPAN *NAIVE BAYES* UNTUK PENENTUAN DOSEN
PEMBIMBING TUGAS AKHIR PADA PRODI TEKNIK INFORMATIKA
STMIK PALANGKARAYA**

TUGAS AKHIR

Disusun Untuk Memenuhi Syarat Kelulusan Program Strata I pada
Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) Palangkaraya



OLEH

DIESY FRIDAYASTUTI
NIM C1455201025
PROGRAM STUDI TEKNIK INFORMATIKA

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA**

2021

**PENERAPAN *NAIVE BAYES* UNTUK PENENTUAN DOSEN
PEMBIMBING TUGAS AKHIR PADA PRODI TEKNIK INFORMATIKA
STMIK PALANGKARAYA**

TUGAS AKHIR

Disusun Untuk Memenuhi Syarat Kelulusan Program Strata I pada
Sekolah Tinggi Manajemen Informatika dan Komputer
(STMIK) Palangkaraya

OLEH

DIESY FRIDAYASTUTI
NIM C1455201025
PROGRAM STUDI TEKNIK INFORMATIKA

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
2021**

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama Mahasiswa : **DIESY FRIDAYASTUTI**

NIM : **C1455201025**

menyatakan bahwa Tugas Akhir dengan judul :

**PENERAPAN *NAIVE BAYES* UNTUK PENENTUAN DOSEN
PEMBIMBING TUGAS AKHIR PADA PRODI TEKNIK INFORMATIKA
STMIK PALANGKARAYA**

Adalah hasil karya saya dan bukan merupakan duplikasi Sebagian atau seluruhnya dari karya orang lain, kecuali bagian yang sumber informasi dicantumkan.

Pernyataan ini dibuat dengan sebenar-benarnya secara sadar dan bertanggungjawab dan saya bersedia menerima sanksi pembatalan Tugas Akhir apabila terbukti melakukan duplikasi terhadap Tugas Akhir atau karya ilmiah lain yang sudah ada.

Palangka Raya, 03 Agustus 2021

Yang membuat pernyataan,

Materai

DIESY FRIDAYASTUTI

MOTTO DAN PERSEMBAHAN

“Lebih baik mencoba lalu menyerah dari pada tidak pernah mencoba lalu menyerah begitu saja dan merasakan penyesalan kemudian.”

(Diesy Fridayastuti)

PERSEMBAHAN :

- Untuk kedua orang tua :
Muhammad Muchdi dan Elsy Meidiarti dan juga adik-adik yang selalu mendukung saya selama ini ini.
- Untuk Bapak Ibu dosen STMIK
Palangkaraya yang telah mengajarkan dan memberikan ilmu pada saya.
- Untuk teman – teman seperjuangan saya :
Haris Rilophambudi dan Mariah Riska yang sama-sama mengejar gelar kita.
- Untuk orang *special* :
Arief Wicaksono yang selama ini selalu mengingatkan dan mendukung saya selama mengerjakan tugas akhir.

INTISARI

Diesy Fridayastuti, C1455201025, 2021. *Penerapan Naive Bayes Untuk Penentuan Tugas Akhir Pada Prodi Teknik Informatika STMIK Palangkaraya.* Pembimbing I Hotmian Sitohang, M.Kom., Pembimbing II Catharina Elmayantie, M.Pd.

Tugas akhir dikerjakan oleh mahasiswa tingkat akhir agar dapat lulus dan mendapatkan gelarnya. Dalam pengerjaan tugas akhir para mahasiswa harus mempunyai dosen pembimbing yang akan membantu pengerjaan tugas akhir tersebut. Namun, saat pemilihan dosen pembimbing mahasiswa juga harus memperhatikan poin – poin yang menjadi pertimbangan dalam menentukan dosen pembimbing. Penelitian ini bertujuan untuk membantu mahasiswa dalam menentukan dosen pembimbing tugas akhir mereka.

Penelitian ini menggunakan metode *Naïve Bayes* untuk menghitung nilai probabilitas dari data yang akan digunakan. Selain itu, penulis juga menggunakan *Rapidminer* sebagai software yang membantu mempermudah penelitian. Penulisan penelitian juga menggunakan metode pengumpulan data dengan melakukan pengamatan atau observasi, metode wawancara, metode dokumentasi, dan juga metode literatur.

Dilihat dari hasil penelitian yang dilakukan penulis, nilai akurasi bergantung pada ketepatan mahasiswa dalam memilih pembimbing tugas akhir. Semakin tepat mahasiswa memilih dosen yang memiliki keahlian sesuai dengan topik tugas akhir, maka semakin tinggi nilai akurasinya. Pada penelitian ini, penulis mendapatkan nilai akurasi yaitu 80% hingga 90%. Diharapkan penelitian ini dapat membantu mahasiswa dalam menentukan dosen pembimbing tugas akhir.

Kata kunci : Dosen Pembimbing, *Naive Bayes*, *Rapidminer*, Tugas Akhir.

ABSTRACT

Diesy Fridayastuti, C1455201025, 2021. *Penerapan Naive Bayes Untuk Penentuan Tugas Akhir Pada Prodi Teknik Informatika STMIK Palangkaraya.* Advisor I Hotmian Sitohang, M.Kom., Advisor II Catharina Elmayantie, M.Pd.

The final project is done by final year students in order to graduate and get their degree. In completing the final project, students must have a supervisor who will assist in the completion of the final project. However, when selecting a supervisor, students must also pay attention to the points that are considered in determining the supervisor. This study aims to assist students in determining their final project supervisor.

This study uses the *Naive Bayes* method to calculate the probability value of the data to be used. In addition, the author also uses Rapidminer as software that helps facilitate research. Research writing also uses data collection methods by making observations or observations, interview methods, documentation methods, and also literature methods.

Judging from the results of the research conducted by the author, the accuracy value depends on the accuracy of the student in choosing the final project supervisor. The more precise students choose lecturers who have expertise in accordance with the topic of the final project, the higher the accuracy value. In this study, the authors get an accuracy value of 80% to 90%. It is hoped that this research can help students in determining the final project supervisor.

Keywords: Final Project, *Naive Bayes*, Rapidminer, Supervisor.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa atas berkat dan rahmatnya sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan judul **“PENERAPAN NAIVE BAYES UNTUK PENENTUAN DOSEN PEMBIMBING TUGAS AKHIR PADA PRODI TEKNIK INFORMATIKA STMIK PALANGKARAYA”**.

Pada kesempatan ini penulis juga ingin menyampaikan banyak terima kasih yang sebesar – besarnya kepada pihak – pihak yang telah memberikan bimbingan dan dorongan dalam penulisan Tugas Akhir ini, yaitu kepada :

1. Suparno, M.Kom., selaku Ketua Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya.
2. Hotmian Sitohang, M.Kom., selaku pembimbing I yang memberikan masukan, saran, dan bantuan kepada penulis selama penulisan Tugas Akhir ini.
3. Catharina Elmayantie. M.Pd., selaku pembimbing II yang juga selalu memberikan bimbingan penulisan untuk penulis dalam penyusunan Tugas Akhir ini.
4. Bapak dan Ibu Dosen STMIK Palangkaraya yang telah memberikan ilmunya sebagai bekal ilmu pengetahuan kepada penulis.
5. Kepada Orang Tua, Saudara, Teman dan Sahabat, serta Orang Tercinta yang telah memberikan dukungan moril dan materi selama ini.

Dengan demikian, penulis juga mengharapkan kritik dan saran yang membangun untuk menyempurnakan Tugas Akhir ini agar lebih baik di masa yang akan datang.

Palangkaraya, 03 Agustus 2021

Penulis

DAFTAR ISI

LEMBAR PERNYATAAN	
PERSETUJUAN.....	Error! Bookmark not defined.
PENGESAHAN	Error! Bookmark not defined.
MOTTO DAN PERSEMBAHAN.....	iii
INTISARI.....	iv
ABSTRACT	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah.....	2
1.3 Batasan Masalah.....	2
1.4 Tujuan dan Manfaat.....	3
1.5 Sistematika Penulisan	3
BAB II LANDASAN TEORI	5
2.1 Tinjauan Pustaka	5
2.2 Kajian Teori.....	7
a. Data.....	7
b. Metode Naive Bayes	8
c. Rapid Miner	9
BAB III METODE PENELITIAN	11
3.1 Lokasi Penelitian	11

3.2	Bahan Penelitian.....	11
3.3	Teknik Pengumpulan Data	11
3.4	Pengolahan Data.....	14
3.5	Tahap Perhitungan Data	19
BAB IV HASIL DAN PEMBAHASAN.....		26
4.1	Pengujian Algoritma.....	26
4.2	Evaluasi dan Validasi Algoritma.....	29
4.3	Analisis Algoritma.....	38
BAB V KESIMPULAN DAN SARAN		39
5.1	Kesimpulan.....	39
5.2	Saran	39

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR TABEL

Tabel 1. Perbandingan Penelitian.....	5
Tabel 2. Data jenis peminatan Prodi Teknik Informatika	15
Tabel 3. Data dosen sesuai dengan kompetensi keahlian	15
Tabel 4. Data dosen Prodi Teknik Informatika	16
Tabel 5. Data Dosen Lintas Prodi	16
Tabel 6. Data Tugas Akhir Mahasiswa	17
Tabel 7. Jumlah data training dan testing.....	18
Tabel 8. Nilai Probabilitas Jumlah Kelas Dosen Pembimbing 1	19
Tabel 9. Nilai Probabilitas Jumlah Kelas Dosen Pembimbing 2	20
Tabel 10. Nilai Probabilitas Dosen Peminatan Jaringan.....	21
Tabel 11. Nilai Probabilitas Dosen Peminatan Sistem Cerdas	22
Tabel 12. Data Testing	23
Tabel 13. Nilai Probabilitas Dosen Pembimbing 1 : Peminatan Jaringan	23
Tabel 14. Nilai Probabilitas Dosen Pembimbing 1 : Peminatan Sistem Cerdas.....	24
Tabel 15. Hasil Perhitungan Nilai Probabilitas Masing-Masing Peminatan.....	24

DAFTAR GAMBAR

Gambar 1. Algoritma <i>Naive Bayes</i>	9
Gambar 2. Tahap Penelitian.....	13
Gambar 3. Penambahan Data ke RapidMiner.....	26
Gambar 4. Seret Data yang Digunakan ke Halaman Kerja RapidMiner	27
Gambar 5. Pemilihan Algoritma yang Dipakai.....	27
Gambar 6. Pemilihan Operator “ <i>Apply Model</i> ”	28
Gambar 7. Pemilihan Operator “ <i>Performance</i> ”	28
Gambar 8. Design Dengan Menggunakan RapidMiner	29
Gambar 9. Hasil Perhitungan	29
Gambar 10. Hasil Nilai Prediksi	30
Gambar 11. Model Confusion Matrix.....	31
Gambar 12. Hasil Data Training Tahap Pertama	32
Gambar 13. Hasil Data Training Tahap Kedua.....	33
Gambar 14. Hasil Data Training Tahap Ketiga	34
Gambar 15. ROC Kurva Tahap Pertama.....	36
Gambar 16. ROC Kurva Tahap Kedua	37
Gambar 17. ROC Kurva Tahap Ketiga	37

DAFTAR LAMPIRAN

Lampiran 1. Surat Tugas Penguji Tugas Akhir

Lampiran 2. Kartu Konsultasi Bimbingan Tugas Akhir

Lampiran 3. Berita Acara Sidang Tugas Akhir

Lampiran 4. Surat Ijin Penelitian

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Tugas akhir adalah syarat mahasiswa untuk mendapatkan gelar pendidikannya. Pemilihan dosen pembimbing dalam pengerjaan tugas akhir terkadang membuat mahasiswa bingung untuk memilih dosen yang sesuai dengan tugas akhir yang diambil. Tugas akhir harus dibimbing oleh dosen yang berkompeten sesuai dengan pembahasan yang diambil oleh mahasiswa. Dosen pembimbing yang dipilih dan berkompeten sesuai dengan pembahasan dapat mempermudah membantu mahasiswa untuk mengerjakan tugas akhir yang dikerjakan. Penentuan dosen pembimbing terkadang dipilih langsung oleh mahasiswa karena alasan tertentu atau juga rekomendasi dari Ketua Prodi dengan mencocokkan pembahasan tugas akhir yang diambil oleh mahasiswa dan keahlian para dosen.

Pada penelitian ini, akan menggunakan metode *Naive Bayes* dalam pengerjaannya. *Naive Bayes* memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. *Naive Bayes* bekerja sangat baik disbanding dengan model lainnya. Hal ini dibuktikan pada jurnal Xhemali, Daniela, Chris J. Hinde, and Roger G. Stone. “*Naive Bayes vs. decision trees vs. neural network in the classifications of training web pages.*”, mengatakan bahwa “*Naive Bayes Classifier* memiliki tingkat akurasi yang lebih baik dibanding model classifier lainnya”.

Berdasarkan hal tersebut, maka penulis memutuskan untuk membuat penelitian dengan judul “Penerapan *Naive Bayes* Untuk Penentuan Dosen Pembimbing Tugas Akhir Pada Prodi Teknik Informatika STMIK Palangkaraya” yang diharapkan dapat membantu mahasiswa lain untuk menentukan dosen pembimbing tugas akhir.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah disampaikan diatas, maka pokok permasalahannya adalah “Bagaimana menerapkan Algoritma *Naive Bayes* untuk penentuan dosen pembimbing tugas akhir pada Prodi Teknik Informatika STMIK Palangkaraya ?”

1.3 Batasan Masalah

Agar permasalahan tidak melebar, maka penulis memberikan batasan-batasan masalah yaitu:

1. Penelitian ini dibuat untuk membantu mahasiswa dalam menentukan dosen pembimbing tugas akhir.
2. Penelitian ini menggunakan algoritma *Naive Bayes*.
3. Penelitian ini menggunakan aplikasi *RapidMiner*.
4. Penelitian ini akan menghasilkan nilai akurasi dari perhitungan data menggunakan algoritma *Naive Bayes* dengan aplikasi *RapidMiner*.

1.4 Tujuan dan Manfaat

a. Tujuan

Untuk menerapkan *Naive Bayes* dalam penentuan dosen pembimbing tugas akhir pada Prodi Teknik Informatika STMIK Palangkaraya.

b. Manfaat

1) Manfaat bagi Penulis

Sebagai penerapan ilmu yang didapatkan selama masa kuliah di STMIK Palangkaraya dan mengerti lebih dalam tentang pengetahuan *Naive Bayes*.

2) Manfaat bagi STMIK Palangkaraya

Penelitian ini diharapkan dapat menjadi salah satu referensi untuk pembuatan tugas akhir mahasiswa selanjutnya di masa mendatang dan juga menambah literatur pustaka pada perpustakaan STMIK Palangkaraya.

3) Manfaat bagi Pengguna

Mahasiswa dapat menggunakan ini untuk menentukan dosen pembimbing yang sesuai dan kompeten sesuai dengan permasalahan yang diambil untuk tugas akhir.

1.5 Sistematika Penulisan

Penelitian ini menggunakan sistematika penulisan yang terdiri dari beberapa bab yang tersusun sebagai berikut:

BAB I : PENDAHULUAN

Pada bab ini berisi tentang latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan yang digunakan untuk menjelaskan permasalahan.

BAB II : LANDASAN TEORI

Bab ini menjelaskan mengenai teori-teori yang mendukung penelitian dan menjadi pemecah masalah dalam penelitian ini.

BAB III: METODE PENELITIAN

Pada bab ini membahas secara rinci mengenai metode yang dilakukan untuk melakukan penelitian. Metode ini terdiri dari beberapa sub bab yang secara garis besar mencakup yaitu lokasi penelitian, pengumpulan data, dan desain sistem.

BAB IV: HASIL DAN PEMBAHASAN

Bab ini membahas tentang pengujian terhadap data yang ada dan juga membahas hasil pengujian data selama melakukan penelitian.

BAB V : KESIMPULAN DAN SARAN

Bab ini membahas tentang kesimpulan dari penelitian yang sudah dilakukan dan juga saran-saran untuk bisa dikembangkan lagi nanti di masa yang akan datang

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Untuk melakukan sebuah penelitian dibutuhkan hasil dari penelitian-penelitian terkait yang sudah dilakukan sebelumnya sebagai referensi. Berikut adalah beberapa hasil penelitian yang berkaitan dengan penelitian yang sedang dilakukan.

Tabel 1. Perbandingan Penelitian

No	Penulis/Tahun	Topik Penelitian	Metode	Hasil	Perbedaan
1	T.Zakia Mulyani, Zulfan Khairil Simbolon, Amirullah / 2019	Implementasi Algoritma <i>Naive Bayes Classifier</i> Dalam Menentukan Topik Tugas Akhir Mahasiswa Berbasis Web	<i>Naive Bayes</i>	Sistem ini memiliki hasil akhir atau output berupa nilai probabilitas tertinggi yang menjadi acuan dalam penentuan class yang direkomendasikan, serta mahasiswa dapat melihat referensi judul jurnal-jurnal yang berkaitan dengan class yang direkomendasikan.	Penerapan ini memiliki hasil yaitu nilai probabilitas dari dosen pembimbing yang sesuai dengan topik tugas akhir yang diambil mahasiswa.
2	Uli Rizki, Abdul Malik Zuhdi, Kusrini / 2019	Sistem Pendukung Keputusan Dengan Metode <i>Naive Bayes</i> Untuk Pemilihan Dosen Pembimbing	<i>Naive Bayes</i>	Sistem ini dapat memberikan informasi dosen pembimbing dengan menampilkan persentase perbandingan lulus tepat waktu atau tidak dengan tingkat akurasi 77,50%.	Penerapan ini hanya akan menghitung nilai probabilitas dosen pembimbing yang dipilih sesuai dengan kompetensi dosen pembimbing.

No	Penulis/Tahun	Topik Penelitian	Metode	Hasil	Perbedaan
3	Marsanu Asfi, Nopi Fitrianingsih / 2020	Implementasi Algoritma <i>Naive Bayes Classifier</i> Sebagai Sistem Rekomendasi Pembimbing Skripsi	<i>Naive Bayes</i>	Sistem rekomendasi dosen pembimbing dengan penerapan algoritma <i>Naive Bayes</i> dapat meningkatkan efisiensi waktu dan membantu ketua prodi untuk menentukan dosen pembimbing sesuai bidang ilmu skripsi dan kompetensi dosen.	Penerapan ini hanya dapat digunakan dalam memilih dosen pembimbing 1 karena kriteria yang dipilih adalah kompetensi dosen.
4	Fatmi Kasih / 2018	Integrasi Kategori Skripsi Dan Keahlian Dosen Dalam <i>Naive Bayes</i> Untuk Pemilihan Dosen Pembimbing.	<i>Naive Bayes</i>	Aplikasi menghasilkan informasi dan solusi untuk mahasiswa yang dapat digunakan sebagai pertimbangan pemilihan dosen pembimbing berdasarkan topik tugas akhir dan keahlian dosen	Penerapan melakukan perhitungan nilai probabilitas dengan menggunakan aplikasi Rapidminer dan keahlian dosen sebagai kategori perhitungannya.
5	Febri Yohanes, Aldi Wicaksono, Anita, Chaulina Alfianti Oktaviani / 2019	Sistem Penunjang Keputusan Dosen Pembimbing Menggunakan Algoritma <i>Naive Bayes</i> Studi Kasus STIKI Malang.	<i>Naive Bayes</i> dan <i>Cosine Similarity</i>	Aplikasi ini mempermudah kaprodi memilih dosen pembimbing dengan cara mengeluarkan rekomendasi nama dosen dan menampilkan dokumen yang memiliki kemiripan sebagai pertimbangan pemilihan dosen pembimbing.	Penerapan menghitung nilai probabilitas dosen pembimbing 1 dengan menggunakan aplikasi Rapidminer.

2.2 Kajian Teori

a. Data

Data adalah kumpulan informasi atau keterangan-keterangan yang diperoleh dari pengamatan, informasi itu bisa berupa angka, lambang atau sifat. Data yang diperoleh dari berbagai sumber masih menjadi sebuah anggapan atau fakta karena memang belum diolah lebih lanjut. Setelah diolah melalui suatu penelitian atau percobaan maka data dapat berubah menjadi bentuk yang lebih kompleks misal *database*, informasi atau bahkan solusi pada masalah tertentu. Data yang telah diolah akan berubah menjadi sebuah informasi yang dapat digunakan untuk menambah pengetahuan bagi penerimanya. Data memiliki beberapa fungsi diantaranya :

- 1) Untuk membuat keputusan,
- 2) Sebagai dasar suatu perencanaan,
- 3) Sebagai alat pengendali terhadap pelaksanaan atau implementasi suatu aktivitas, dan
- 4) Sebagai dasar evaluasi terhadap suatu kegiatan.

Berdasarkan jenisnya, data dikelompokkan menjadi dua macam yaitu data kualitatif dan data kuantitatif. Data kualitatif merupakan data yang menunjukkan mutu atau kualitas sesuatu yang ada baik proses, keadaan, peristiwa, kejadian dan lainnya yang dinyatakan ke dalam bentuk pertanyaan atau berupa kata-kata. Data ini umumnya diperoleh dari hasil wawancara dan bersifat subjektif. Sedangkan data kuantitatif merupakan data yang berbentuk angka-angka sebagai hasil pengukuran ataupun hasil observasi. Data

kuantitatif didapatkan dari pengukuran langsung dan dari angka-angka yang diperoleh dengan mengubah data kualitatif menjadi data kuantitatif. Data kuantitatif bersifat objektif dan dapat ditafsirkan sama oleh semua orang.

b. Metode Naive Bayes

Naive Bayes merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistic yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari *Naive Bayes* adalah asumsi yang sangat kuat (naif) akan independensi dari masing-masing kondisi atau kejadian.

Menurut (Gumilang, 2014), menjelaskan bahwa *Naive Bayes* untuk setiap kelas keputusan, menghitung probabilitas dengan syarat bahwa kelas keputusan adalah benar, mengingat vektor informasi proyek. Algoritma ini mengasumsikan bahwa atribut obyek adalah independen. Probabilitas yang terlibat dalam memproduksi perkiraan akhir dihitung sebagai jumlah frekuensi dari “master” tabel keputusan.

Naive Bayes bekerja sangat baik dibanding dengan model *classifier* lainnya. Hal ini dibuktikan oleh Xhemali, Hinde Stone dengan jurnalnya “*Naive Bayes vs. Decision Trees vs. Neural Networks in the Classifications of Training Web Pages*” mengatakan bahwa “*Naive Bayes*

memiliki tingkat akurasi yang lebih baik dibanding model *classifier* lainnya”.

Keuntungan pengguna adalah bahwa metode ini hanya membutuhkan jumlah *data training* yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Karena yang diasumsikan sebagai variabel *independent*, maka hanya varian dari suatu variabel dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari matriks kovarian.

Persamaan Teorema Bayes memiliki bentuk umum sebagai berikut :

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

Gambar 1. Algoritma *Naive Bayes*

Keterangan :

X : data dengan class yang belum diketahui

H : hipotesis data X merupakan suatu class spesifik

$P(H|X)$: probabilitas hipotesis H berdasar kondisi X

$P(H)$: probabilitas hipotesis H

$P(X|H)$: probabilitas X berdasar kondisi hipotesis H

$P(X)$: probabilitas dari X

c. Rapid Miner

Untuk mempermudah dalam menjalankan atau melakukan penelitian tersebut, maka penulis menggunakan Rapidminer pada

pengerjaannya. RapidMiner merupakan perangkat lunak yang bersifat terbuka (*open source*). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap *data mining*, *text mining*, dan analisis prediksi. Rapidminer menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner sebelumnya dikenal sebagai *YALE (Yet Another Learning Environment)*, mulai dikembangkan pada tahun 2001 oleh Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer dari Unit Kecerdasan Buatan Universitas Teknik Dortmund. Mulai tahun 2006, perkembangannya didorong oleh Rapid-I, sebuah perusahaan yang didirikan oleh Ingo Mierswa dan Ralf Klinkenberg pada tahun yang sama. Pada tahun 2007, nama perangkat lunak itu berubah dari YALE menjadi RapidMiner. Pada tahun 2013, perusahaan melakukan rebranding dari Rapid-I menjadi RapidMiner (Aprilla, Dennis, Baskoro, Lia , & Simri, 2012).

BAB III

METODE PENELITIAN

3.1 Lokasi Penelitian

Sesuai dengan judul penelitian yang dibuat penulis yaitu Penerapan Naïve Bayes Untuk Penentuan Dosen Pembimbing Tugas Akhir. Lokasi penelitian yang dilakukan oleh penulis adalah di STMIK Palangkaraya dengan alamat Jl. George Obos No.114, Menteng, Kec. Jekan Raya, Palangkaraya, Kalimantan Tengah, Indonesia

3.2 Bahan Penelitian

Dalam penelitian ini data yang menjadi acuan oleh penulis adalah data mahasiswa yang mengajukan tugas akhir pada Prodi Teknik Informatika STMIK Palangkaraya pada tahun 2020 untuk dianalisis.

3.3 Teknik Pengumpulan Data

Dalam melakukan penelitian ini penulis melakukan beberapa tahap pengumpulan data yang bersangkutan dengan tema tugas akhir yang diambil, yaitu :

a. Tahap Pengumpulan Data

1) Metode Pengamatan (*Observation*)

Metode pengamatan atau observasi ini penulis lakukan dengan cara mengamati mahasiswa-mahasiswa yang mengambil tugas akhir lalu memperhatikan apa saja yang menjadi pertimbangan dalam menentukan dosen pembimbing.

2) Metode Wawancara (*Interview*)

Metode ini dilakukan agar mendengar pendapat para mahasiswa yang mengambil tugas akhir dalam menentukan dosen pembimbing. Ini dilakukan kepada beberapa mahasiswa sebagai acuan untuk yang lainnya.

3) Metode Dokumentasi

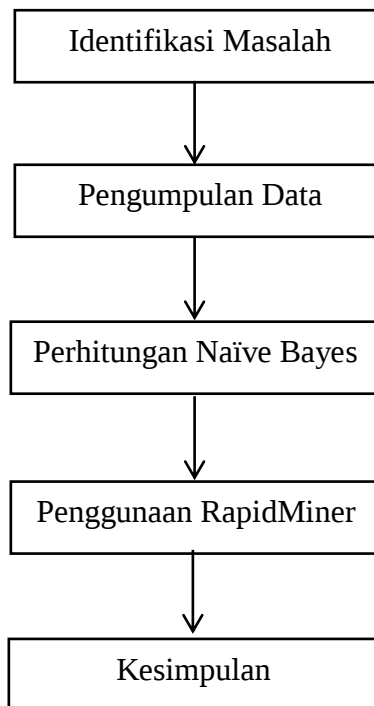
Metode ini memanfaatkan data-data para mahasiswa yang sudah atau telah selesai mengambil tugas akhir yang ada terdapat di ruang Prodi Teknik Informatika.

4) Metode Literatur

Metode ini memanfaatkan data para mahasiswa yang terdahulu yang sudah selesai mengerjakan tugas akhir. Data diambil di Prodi Teknik Informatika dan juga perpustakaan STMIK Palangkaraya.

b. Tahapan Penelitian

Sebuah penelitian memerlukan langkah-langkah yang tersusun secara sistematis agar penelitian tersebut dapat mencapai tujuan yang diinginkan. Maka dari itu tahapan penelitian pada penelitian ini tersusun pada gambar berikut :



Gambar 2. Tahap Penelitian

Penelitian ini dilakukan dengan beberapa langkah sesuai dengan urutan seperti gambar yang ada diatas, yaitu :

1. Identifikasi masalah, tahap ini penulis mengidentifikasi hal yang menjadi masalah dalam pemilihan dosen pembimbing tugas akhir Prodi Teknik Informatika di STMIK Palangkaraya
2. Pengumpulan data, tahap ini penulis membutuhkan data mahasiswa yang mengambil tugas akhir dan pengumpulan data dilakukan dengan cara meminta kepada pihak Prodi Teknik Informatika di STMIK Palangkaraya.

3. Perhitungan *Naive Bayes*, penulis menggunakan metode *Naive Bayes* dan perhitungan ini dilakukan secara manual dengan menggunakan rumus *Naive Bayes* sesuai dengan data yang telah didapat.
4. Penggunaan RapidMiner, tahap ini penulis menggunakan aplikasi RapidMiner untuk dapat mengolah data yang didapat dengan operasi *Naive Bayes*.
5. Kesimpulan, yang terakhir pada tahap ini adalah kesimpulan dari hasil yang didapat setelah data diolah baik secara manual maupun menggunakan aplikasi RapidMiner.

3.4 Pengolahan Data

Pengolahan data pada materi ini penulis menggunakan metode teorema bayes sesuai dengan rumus yang sudah dijelaskan diatas. Data yang digunakan antara lain: data peminatan Prodi Teknik Informatika, data dosen sesuai dengan peminatan yang dikuasai para dosen, data dosen Prodi Teknik Informatika, data dosen lintas prodi yang dapat dijadikan dosen pembimbing tugas akhir, dan juga data mahasiswa yang mengambil tugas akhir. Data tersebut di tampilkan pada tabel 2, tabel 3, tabel 4, tabel 5, tabel 6, dan tabel 7.

Data dibawah ini merupakan data dari jenis peminatan matakuliah yang ada pada Prodi Teknik Informatika.

Tabel 2. Data jenis peminatan Prodi Teknik Informatika

No.	Peminatan
1	Jaringan
2	Sistem Cerdas

Selanjutnya, berikut adalah data para dosen pengampu yang sesuai dengan kompetensi keahlian dari masing-masing peminatan yang ada. Dari data ini juga diharapkan agar menjadi bahan pertimbangan mahasiswa yang akan memilih dosen pembimbing sesuai dengan keahlian para dosen dan juga materi tugas akhir yang akan dilakukan.

Tabel 3. Data dosen sesuai dengan kompetensi keahlian

No.	Nama Dosen	
	Jaringan	Sistem Cerdas
1	Elia Zakaria, M.T.	Hotmian Sitohang, M.Kom.
2	Herkules, S.Kom., M.Cs.	Lili Rusdiana, M.Kom.
3	Sam'ani, S.T., M.Kom.	Sam'ani, S.T., M.Kom.
4	Amaya Andri Damaini, S.Kom., M.Kom.	Sulistiyowati, S.Kom., M.Cs.
5	Lili Rusdiana, M.Kom.	Veny Cahya Hardita, M.Kom.
6	Ir. Hj. Siti Maryamah, MM.	Maura Widyaningsih, S.Kom., M.Cs.
7	Heri Setiawan, M.Kom	

Selanjutnya, berikut ini merupakan data nama-nama dosen Prodi Teknik Informatika.

Tabel 4. Data dosen Prodi Teknik Informatika

No.	Nama Dosen
1	Catharina Elmayantie, M.Pd.
2	Elok Faiqotul Himmah, S.Si., M.Cs.
3	Rudini, M.Pd.
4	H.Suratno, S.Kom., M.Kom.
5	Ir. Siti Maryamah, M.M.
6	Herkules, S.Kom., M.Cs.
7	Lili Rusdiana, M.Kom.
8	Elia Zakaria, M.T.
9	Sulistiyowati, S.Kom., M.Cs.
10	Sam'ani, S.T., M.Kom.
11	Hotmian Sitohang, M.Kom.
12	Amaya Andri Damaini, S.Kom., M.Kom.
13	Veny Cahya Hardita, M.Kom.
14	Maura Widyaningsih, S.Kom., M.Cs.
15	Heri Setiawan, M.Kom.

Selain data para dosen yang ada pada Prodi Teknik Informatika, berikut juga data para dosen lintas prodi. Dimaksud dengan dosen lintas prodi, karena para dosen tersebut merupakan dosen dari prodi lain yang ada di STMIK Palangkaraya.

Tabel 5. Data Dosen Lintas Prodi

No.	Nama Dosen
1	Drs. Heri Purwanto, M.Pd.
2	Dewanto Zulkarnain, M.Pd.
3	Norhayati, M.Pd.
4	Drs. Sartana, M.Si.
5	Hafiz Riyadli, M.Kom.

6	Rommi Kaestria, M.Kom.
No.	Nama Dosen
7	Rosmiati, M.Kom.
8	Susi Hendartie, M.Kom.
9	H. Agung Prabowo, S.Kom., M.MSi.
10	Deden Andriawan, M.Kom.
11	Fenroy Yedithia, S.Kom., M.TI.
12	Arliyana, M.Kom.
13	Ferdiyani Haris, M.Kom.
14	Bayu Pratama Nugroho, S.Kom., M.T.
15	Frengklin Matatula, S.Kom., M.MSi.
16	Moch Ichsan, S.T., M.Kom.
17	Sherly Jayanti, S.T., M.Cs.
18	Suparno, M.Kom.
19	Drs. Frands, MM., PhD.

Pada tabel 6 menampilkan data-data mahasiswa yang sedang ataupun telah selesai mengambil tugas akhir. Data tersebutlah yang nantinya akan menjadi data training dalam pengerjaan tugas akhir penulis.

Tabel 6. Data Tugas Akhir Mahasiswa

No.	Nama Mahasiswa	Peminatan	Dosen Pembimbing 1	Dosen Pembimbing 2
1	Glorita	Jaringan	Hotmian Sitohang, M.Kom.	Elok Faiqotul Himmah, S.Si, M.Sc.
2	Jumanto	Jaringan	Heri Setiawan, M.Kom.	Elia Zakharia, MT.
3	Jayanto Ade Hernanda	Jaringan	Maura Widyaningsih, S.Kom., M.Cs.	Rudini, M.Pd.
4	Dwi Arianto Akbar	Jaringan	Sulistyowati, S.Kom., M.Cs.	Lili Rusdiana, M.Kom.
5	Eko Sultoni Khoiril M.	Sistem Cerdas	Sulistyowati, S.Kom., M.Cs.	H. Suratno, S.Kom., M.Si
6	Denny Aditama	Jaringan	Heri Setiawan, M.Kom.	Ir. Hj. Siti Maryamah, MM.
7	Andreas Chavalery	Jaringan	Heri Setiawan, M.Kom.	Hotmian Sitohang, M.Kom.

No.	Nama Mahasiswa	Peminatan	Dosen Pembimbing 1	Dosen Pembimbing 2
8	Puteri Diah Puspita E.	Sistem Cerdas	Hotmian Sitohang, M.Kom.	Lili Rusdiana, M.Kom.
9	Yuf Heri	Sistem Cerdas	Lili Rusdiana, M.Kom.	Herkules, S.Kom., M.Cs.
10	Teti Sumarni	Sistem Cerdas	Lili Rusdiana, M.Kom.	Elok Faiqotul Himmah, S.Si., M.Sc.
11	Sulastri Merni	Sistem Cerdas	Hotmian Sitohang, M.Kom.	Lili Rusdiana, M.Kom.
12	Aminudin	Jaringan	Sulistyowati, S.Kom., M.Cs.	Hotmian Sitohang, M.Kom.
13	Krisdianto	Jaringan	Sulistyowati, S.Kom., M.Cs.	Ir. Hj. Siti Maryamah, MM.
14	Winerdie Rafriandi	Jaringan	Herkules, S.Kom., M.Cs.	Catharina Elmayantie, M.Pd.
15	Dwi Handoko Kabuto	Jaringan	Maura Widyaningsih, S.Kom., M.Cs.	Ir. Hj. Siti Maryamah, MM.
16	Yohana Setia Ningrum	Jaringan	Hotmian Sitohang, M.Kom.	Elok Faiqotul Himmah, S.Si., M.Sc.
17	Nurul Suciani	Sistem Cerdas	Maura Widyaningsih, S.Kom., M.Cs.	Rudini, M.Pd.
18	Try Adi Wibowo	Jaringan	Herkules, S.Kom., M.Cs.	Rudini, M.Pd.
...	...	...	...	...
171	M. Nor Ardiatsyah	Jaringan	Herkules, S.Kom., M.Cs.	Catharina Elmayantie, M.Pd.

Data training digunakan untuk melatih algoritma, sedangkan data testing digunakan untuk mengetahui performa algoritma yang sudah dilatih sebelumnya terhadap data baru. Tabel dibawah ini berisikan jumlah data yang dijadikan data training dan juga data testing yang nantinya akan digunakan pada tugas akhir penulis.

Tabel 7. Jumlah data training dan testing

No.	Jenis Data	Jumlah
1	Data Training	171
2	Data Testing	15

3.5 Tahap Perhitungan Data

Pada tahap ini akan menghitung nilai probabilitas sesuai dengan metode Teorema Bayes. Pada tabel 8 ini merupakan tabel dari nilai probabilitas untuk dosen pembimbing 1.

Tabel 8. Nilai Probabilitas Jumlah Kelas Dosen Pembimbing 1

No.	Probabilitas Dosen	Jumlah Kelas/ Total Kelas	Nilai Probabilitas
1	$P(C = \text{Elia Zakaria})$	30/171	0,175439
2	$P(C = \text{Herkules})$	22/171	0,128655
3	$P(C = \text{Sam'ani})$	0/171	0
4	$P(C = \text{Amaya})$	1/171	0,005848
5	$P(C = \text{Lili Rusdiana})$	29/171	0,169591
6	$P(C = \text{Hotmian Sitohang})$	32/171	0,187135
7	$P(C = \text{Sulistiyowati})$	33/171	0,192982
8	$P(C = \text{Veny Cahya})$	0/171	0
9	Ir. Hj. Siti Maryamah, MM	1/171	0,005848
10	Maura Widyaningsih, S.Kom., M.Cs.	18/171	0,105263
11	Heri Setiawan, M.Kom.	5/171	0,02924

Tabel 9 dibawah ini merupakan tabel dari nilai probabilitas untuk dosen pembimbing 2.

Tabel 9. Nilai Probabilitas Jumlah Kelas Dosen Pembimbing 2

No.	Probabilitas Dosen	Jumlah Kelas/ Total Kelas	Nilai Probabilitas
1	P(C= Catharina)	26/171	0,152047
2	P(C= Elok)	19/171	0,111111
3	P(C= Rudini)	20/171	0,116959
4	P(C= Suratno)	9/171	0,052632
5	P(C= Siti Maryamah)	12/171	0,070175
6	P(C= Herkules)	6/171	0,035088
7	P(C= Lili Rusdiana)	10/171	0,05848
8	P(C= Elia Zakaria)	6/171	0,035088
9	P(C= Sulistiyowati)	5/171	0,02924
10	P(C= Sam'ani)	0/171	0
11	P(C= Hotmian Sitohang)	11/171	0,064327
12	P(C= Amaya Andri)	5/171	0,02924
13	P(C= Veny Cahya)	1/171	0,005848
14	P(C= Maura)	3/171	0,017544
15	P(C= Heri Purwanto)	0/171	0
16	P(C= Dewanto)	3/171	0,017544
17	P(C= Norhayati)	0/171	0
18	P(C= Sartana)	1/171	0,005848
19	P(C= Hafiz Riyadli)	1/171	0,005848
20	P(C= Rommi)	4/171	0,023392
21	P(C= Rosmiati)	2/171	0,011696
22	P(C= Susi Hendartie)	2/171	0,011696
23	P(C= Agung Prabowo)	0/171	0,011696
24	P(C= Deden Andriawan)	2/171	0,011696
25	P(C= Fenroy)	2/171	0,011696
26	P(C= Arliyana)	3/171	0,017544
27	P(C= Ferdiani)	2/171	0,011696

No.	Probabilitas Dosen	Jumlah Kelas/ Total Kelas	Nilai Probabilitas
28	P(C= Bayu Pratama)	2/171	0,011696
29	P(C= Frengklin)	5/171	0,02924
30	P(C= Moch Ichsan)	1/171	0,005848
31	P(C= Sherly)	4/171	0,023392
32	P(C= Suparno)	3/171	0,017544
33	P(C= Drs. Frands)	1/171	0,005848

Setelah menghitung nilai probabilitas dari seluruh jumlah data training diatas. Maka selanjutnya dilakukan perhitungan untuk mencari nilai probabilitas dosen berdasarkan peminatan. Pada tabel 10 menampilkan data dari nilai probabilitas dosen peminatan jaringan.

Tabel 10. Nilai Probabilitas Dosen Peminatan Jaringan

No.	Nama Dosen	Jumlah Kasus /Kelas	Nilai Probabilitas
1	Elia Zakaria, M.T.	23/106	0,216981
2	Herkules, S.Kom., M.Cs.	21/106	0,198113
3	Sam'ani, S.T., M.Kom.	0/106	0
4	Amaya Andri Damaini, S.Kom., M.Kom.	1/106	0,009434
5	Lili Rusdiana, M.Kom.	18/105	0,169811
6	Hotmian Sitohang, M.Kom.	16/106	0,150943
7	Sulistiyowati, S.Kom., M.Cs.	15/106	0,141509
8	Veny Cahya Hardita, M.Kom	0/106	0
9	Ir. Hj. Siti Maryamah, MM.	1/106	0,009434
10	Maura Widyaningsih, S.Kom., M.Cs.	6/106	0,056604
11	Heri Setiawan	5/106	0,04717

Selanjutnya pada tabel 11 menampilkan data-data nilai probabilitas dosen dengan peminatan sistem cerdas.

Tabel 11. Nilai Probabilitas Dosen Peminatan Sistem Cerdas

No.	Nama Dosen	Jumlah Kasus/Kelas	Nilai Probabilitas
1	Elia Zakaria, M.T.	7/65	0,107692
2	Herkules, S.Kom., M.Cs.	1/65	0,015385
3	Sam'ani, S.T., M.Kom.	0/65	0
4	Amaya Andri Damaini, S.Kom., M.Kom.	0/65	0
5	Lili Rusdiana, M.Kom.	11/65	0,169231
6	Hotmian Sitohang, M.Kom.	16/65	0,246154
7	Sulistiyowati, S.Kom., M.Cs.	18/65	0,276923
8	Veny Cahya Hardita, M.Kom	0/65	0
9	Ir. Hj. Siti Maryamah, MM.	0/65	0
10	Maura Widyaningsih, S.Kom., M.Cs.	12/65	0,184615
11	Heri Setiawan, M.Kom.	0/65	0

Tabel 12 menampilkan beberapa data testing yang nantinya akan digunakan dalam pada penerapan perhitungan *Naive Bayes*.

Tabel 12. Data Testing

No.	Nama	Peminatan	Dosen Pembimbing
1	Hendro Priyanto	Jaringan	Herkules, S.Kom., M.Cs.
2	Deni Sinaswanto	Sistem Cerdas	Sulistyowati, S.Kom., M.Cs.
3	Daniel Kristiady	Jaringan	Hotmian Sitohang, M.Kom.
4	Nadia Yosy Marlina	Jaringan	Heri Setiawan, M.Kom.
5	Novi Vebrianti	Sistem Cerdas	Lili Rusdiana, M.Kom

Tabel 13 berisikan perhitungan dari data testing diatas dengan data training sesuai dengan nilai probabilitas yang sudah kita hitung sebelumnya sesuai dengan peminatan.

Tabel 13. Nilai Probabilitas Dosen Pembimbing 1 : Peminatan Jaringan

No.	Nama Dosen	Prob Dosen*Prob Peminatan	Nilai Probabilitas
1	Herkules, S.Kom., M.Cs.	$0,128655 \times 0,198113$	0,025488
2	Sulistyowati, S.Kom., M.Cs.	$0,192982 \times 0,141509$	0,027309
3	Hotmian Sitohang, M.Kom.	$0,187135 \times 0,150943$	0,028247
4	Heri Setiawan, M.Kom.	$0,02924 \times 0,04717$	0,001379
5	Lili Rusdiana, M.Kom.	$0,169591 \times 0,169811$	0,028798

Tabel 14 juga sama seperti tabel 13. Namun, pada tabel 14 berisikan nilai probabilitas dengan perhitungan menggunakan nilai probabilitas dari peminatan sistem cerdas.

Tabel 14. Nilai Probabilitas Dosen Pembimbing 1 : Peminatan Sistem Cerdas

No.	Nama Dosen	Prob Dosen*Prob Peminatan	Nilai Probabilitas
1	Herkules, S.Kom., M.Cs.	$0,128655 \times 0,015385$	0,001979
2	Sulistiyowati, S.Kom., M.Cs.	$0,192982 \times 0,276923$	0,053441
3	Hotmian Sitohang, M.Kom.	$0,187135 \times 0,246154$	0,046064
4	Heri Setiawan, M.Kom.	$0,02924 \times 0$	0
5	Lili Rusdiana, M.Kom.	$0,169591 \times 0,169231$	0,0287

Dari perhitungan tabel 13 dan tabel 14 diatas, maka pada tabel 15 ini merupakan hasil dari kedua nilai probabilitas berdasarkan peminatannya.

Tabel 15. Hasil Perhitungan Nilai Probabilitas Masing-Masing Peminatan

No.	Nama Dosen	Hasil Nilai Probabilitas Peminatan	
		Jaringan	Sistem Cerdas
1	Herkules, S.Kom., M.Cs.	0,025488	0,001979
2	Sulistiyowati, S.Kom., M.Cs.	0,027309	0,053441
3	Hotmian Sitohang, M.Kom.	0,028247	0,046064
4	Heri Setiawan, M.Kom.	0,001379	0
5	Lili Rusdiana, M.Kom.	0,028798	0,0287

Perhitungan diatas merupakan perhitungan dosen pembimbing 1 dikarenakan pemilihan dosen pembimbing 2 biasanya membantu mahasiswa dalam penulisan tugas akhir sehingga penulis berpendapat bahwa mahasiswa dapat memilih sendiri dosen pembimbing 2 sesuai dengan kemampuan dosen yang telah diketahui mahasiswa. Alasan lain mengapa saya tidak melakukan perhitungan terhadap dosen pembimbing 2 adalah dikarenakan karena kurang

tahu keahlian dari para dosen, terlebih lagi dosen lintas prodi yang beberapa diantaranya tidak mengajar di Prodi Teknik Informatika.

BAB IV

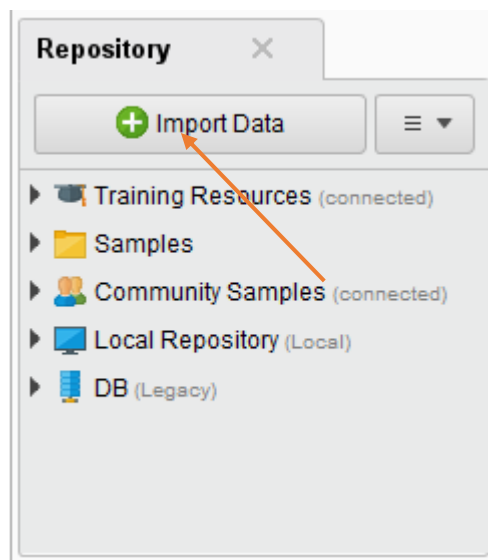
HASIL DAN PEMBAHASAN

4.1 Pengujian Algoritma

Pengujian algoritma yang dilakukan penulis ini menggunakan metode *Naive Bayes* dan juga aplikasi rapidminer. Tahap-tahap yang dilakukan untuk pengujian dengan aplikasi rapidminer ini adalah :

a. Persiapan data.

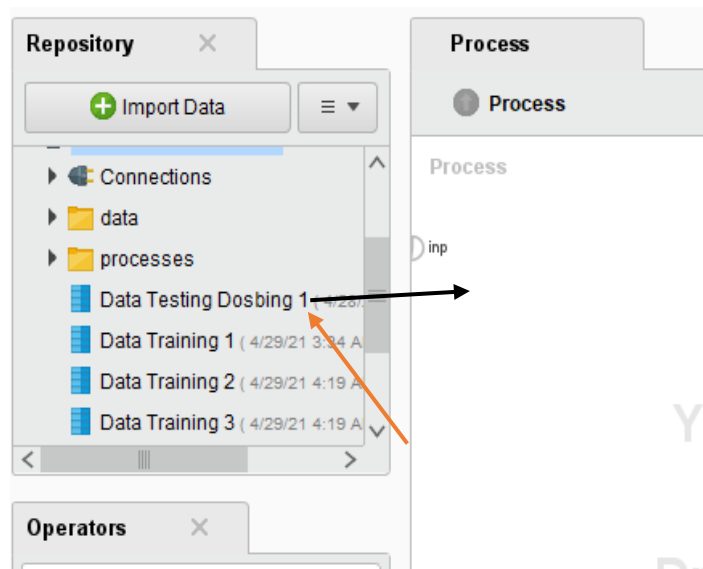
Data yang akan diolah harus dimasukkan terlebih dahulu kedalam aplikasi rapidminer dengan cara klik “*Import Data*” pada tampilan rapidminer.



Gambar 3. Penambahan Data ke RapidMiner

b. Masukkan data yang akan digunakan.

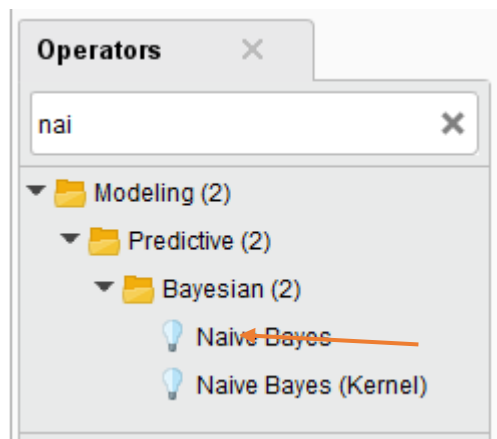
Memasukkan data yang akan digunakan ke halaman kerja rapidminer dengan cara “drag” data yang akan digunakan ke halaman rapidminer.



Gambar 4. Seret Data yang Digunakan ke Halaman Kerja RapidMiner

c. Pemilihan algoritma.

Setelah data sudah masuk ke halaman kerja rapidminer, lalu Langkah selanjutnya memasukkan algoritma yang akan digunakan. Disini penulis memilih algoritma Naive Bayes sesuai dengan topik tugas akhir.

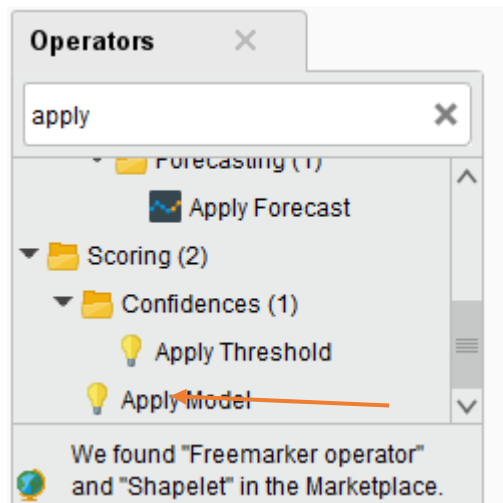


Gambar 5. Pemilihan Algoritma yang Dipakai

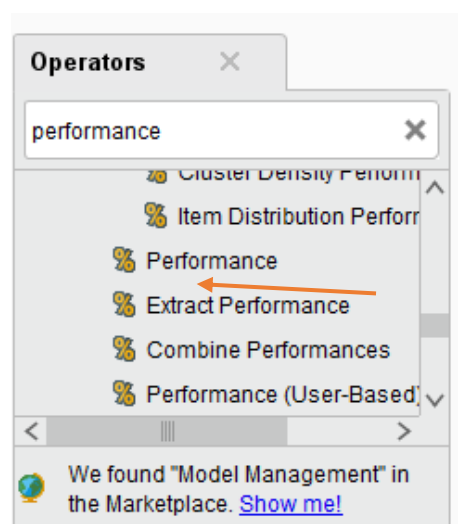
Setelah memilih algoritma, lalu seret ke halaman kerja rapidminer seperti yang sudah dilakukan sebelumnya.

d. Menentukan operator.

Operator diperlukan untuk menghitung data dan juga menampilkan hasil data yang akan dihitung. Operator yang digunakan penulis adalah “Apply Model” dan juga “Performance”. Untuk mencarinya hanya perlu mengetik operator yang akan digunakan dikolom pencarian operator.



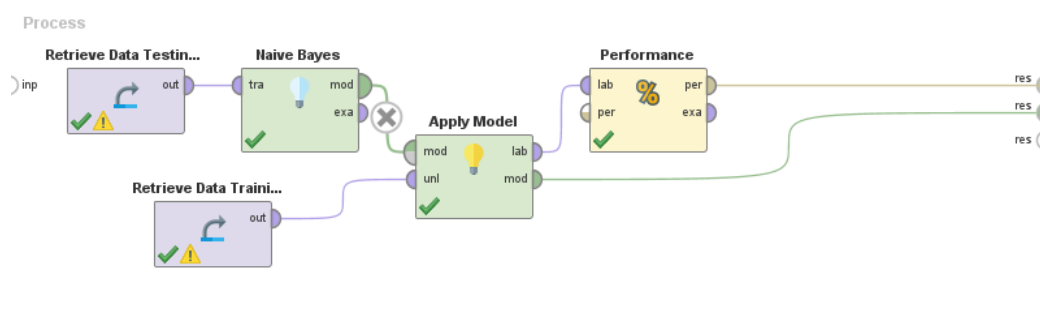
Gambar 6. Pemilihan Operator “*Apply Model*”



Gambar 7. Pemilihan Operator “*Performance*”

e. Hubungkan data dan operator.

Setelah data dan operator yang digunakan telah berada di halaman kerja rapidminer, yang dilakukan selanjutnya adalah menghubungkannya agar bisa mendapatkan hasil perhitungannya.



Gambar 8. Design Dengan Menggunakan RapidMiner

f. Hasil

Setelah data dan operator dihubungkan, maka akan muncul hasil seperti ini.

accuracy: 80.00%

	true Jaringan	true Sistem Cerdas	class precision
pred. Jaringan	2	0	100.00%
pred. Sistem Cerdas	1	2	66.67%
class recall	66.67%	100.00%	

Gambar 9. Hasil Perhitungan

Penggunaan algoritma dan operator tidak selalu sama seperti yang penulis lakukan. Algoritma dan operator yang digunakan menyesuaikan dengan kebutuhan.

4.2 Evaluasi dan Validasi Algoritma

Untuk evaluasi dan validasi perhitungan data menggunakan algoritma dan juga aplikasi rapidminer, maka ada beberapa hal yang akan dijelaskan. Hasil perhitungan data menggunakan rapidminer adalah sebagai berikut :

Row No.	Peminatan	prediction(P...	confidence(...	confidence(...	No.	Nama Maha...	Nama Dosen
1	Jaringan	Jaringan	0.907	0.093	1	Hendro Priya...	Herkules, S.K...
2	Sistem Cerdas	Sistem Cerdas	0.002	0.998	2	Deni Sinasw...	Sulistiyowati, ...
3	Jaringan	Sistem Cerdas	0.321	0.679	3	Daniel Kristia...	Hotmian Sito...
4	Jaringan	Jaringan	0.998	0.002	4	Nadia Yosy M...	Heri Setiawa...
5	Sistem Cerdas	Sistem Cerdas	0.438	0.562	5	Novi Vebrianti	Lili Rusdiana...
6	Sistem Cerdas	Sistem Cerdas	0.193	0.807	6	Fitria Ningsih	Maura Widya...
7	Jaringan	Jaringan	0.612	0.388	7	Herwandi	Elia Zakharia,...
8	Jaringan	Jaringan	0.910	0.090	8	Rivandri	Herkules, S.K...
9	Sistem Cerdas	Sistem Cerdas	0.327	0.673	9	Yunus Rama...	Hotmian Sito...
10	Sistem Cerdas	Sistem Cerdas	0.289	0.711	10	Anchika Utami	Sulistiyowati, ...
11	Jaringan	Jaringan	0.998	0.002	11	Zikri Imam N...	Heri Setiawa...
12	Jaringan	Jaringan	0.911	0.089	12	Muhammad I...	Herkules, S.K...
13	Jaringan	Sistem Cerdas	0.448	0.552	13	Aura Idrati	Lili Rusdiana...
14	Sistem Cerdas	Sistem Cerdas	0.199	0.801	14	Alberamory T...	Maura Widya...
15	Jaringan	Jaringan	0.621	0.379	15	Jaka Saputra	Elia Zakharia,...

Gambar 10. Hasil Nilai Prediksi

Gambar diatas merupakan hasil prediksi dari 15 data yang mahasiswa yang memilih para dosen pembimbing 1. Selanjutnya penulis akan menjelaskan hasil dari perhitungan data menggunakan rapidminer yaitu dengan menghitung *Confusion Matrix*.

a. *Confusion Matrix*

Confusion matrix menurut Han dan Kamber (2011) dapat diartikan sebagai suatu alat yang memiliki fungsi untuk melakukan analisis apakah classifier tersebut baik dalam mengenali daftar (urutan) dari kelas yang berbeda. Nilai dari True-Positive dan True-Negative memberikan informasi ketika classifier dalam melakukan klasifikasi data bernilai benar, sedangkan False-Positive dan False-Negative memberikan informasi ketika classifier salah dalam melakukan klasifikasi data.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Gambar 11. Model Confusion Matrix

TP (True Positive) : Jumlah data dengan nilai sebenarnya positif dan nilai prediksi positif.

FP (False Positive) : Jumlah data dengan nilai sebenarnya negatif dan nilai prediksi positif.

FN (False Negative) : Jumlah data dengan nilai sebenarnya positif dan nilai prediksi negatif.

TN (True Negative) : Jumlah data dengan nilai sebenarnya negatif dan nilai prediksi negatif.

Perhitungan yang dilakukan adalah *accuracy*, *precision*, dan *recall*.

Penulis melakukan pengujian secara bertahap pada 15 data training yang ada. Data training di uji dengan 3 tahap.

1. Tahap pertama

Penulis melakukan pengujian pertama dengan menggunakan 5 buah data training dan mendapatkan hasil sebagai berikut.

accuracy: 80.00%

	true Jaringan	true Sistem Cerdas	class precision
pred. Jaringan	2	0	100.00%
pred. Sistem Cerdas	1	2	66.67%
class recall	66.67%	100.00%	

Gambar 12. Hasil Data Training Tahap Pertama

Result Data Training 1

TP (*True Jaringan*) = 2

FP (*False Jaringan*) = 0

TN (*True SC*) = 2

FN (*False SC*) = 1

Acuuracy

$$\begin{aligned}
 & (TP + TN) / (TP + TN + FP + FN) \times 100\% = \\
 & (2 + 2) / (2 + 2 + 0 + 1) \times 100\% = 4/5 \times 100\% \\
 & = 80\%
 \end{aligned}$$

Precision

$$\begin{aligned}
 \text{Jaringan} : TP / (FP + TP) \times 100\% &= 2 / (0 + 2) \times 100\% \\
 &= 2 / 2 \times 100\% \\
 &= 100\%
 \end{aligned}$$

$$\begin{aligned}
 \text{SC} : TN / (FN + TN) \times 100\% &= 2 / (1 + 2) \times 100\% \\
 &= 2 / 3 \times 100\% \\
 &= 66,67\%
 \end{aligned}$$

Recall

$$\begin{aligned}
 \text{Jaringan} : TP / (FN + TP) &= 2 / (1 + 2) \\
 &= 2 / 3
 \end{aligned}$$

$$= 0,6667 (66,67\%)$$

$$\begin{aligned} \text{SC} &: \text{TN} / (\text{FP} + \text{TN}) = 2 / (0+2) \\ &= 2 / 2 \\ &= 1 (100\%) \end{aligned}$$

2. Tahap kedua.

Penulis melakukan pengujian menggunakan 10 data training dari 15 data yang ada dengan hasil seperti dibawah ini.

accuracy: 90.00%

	true Jaringan	true Sistem Cerdas	class precision
pred. Jaringan	4	0	100.00%
pred. Sistem Cerdas	1	5	83.33%
class recall	80.00%	100.00%	

Gambar 13. Hasil Data Training Tahap Kedua

Result Data Training 2

$$\text{TP (True Jaringan)} = 4$$

$$\text{FP (False Jaringan)} = 0$$

$$\text{TN (True SC)} = 5$$

$$\text{FN (False SC)} = 1$$

Acuuracy

$$\begin{aligned} (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \times 100\% &= \\ (4 + 5) / (4 + 5 + 0 + 1) \times 100\% &= 9/10 \times 100\% \\ &= 90\% \end{aligned}$$

Precision

$$\begin{aligned} \text{Jaringan} : \text{TP} / (\text{FP} + \text{TP}) \times 100\% &= 4 / (0 + 4) \times 100\% \\ &= 4 / 4 \times 100\% \end{aligned}$$

$$= 100\%$$

$$\begin{aligned} \text{SC} &: \text{TN} / (\text{FN} + \text{TN}) \times 100\% = 5 / (1 + 5) \times 100\% \\ &= 5 / 6 \times 100\% \\ &= 83,33\% \end{aligned}$$

Recall

$$\begin{aligned} \text{Jaringan} &: \text{TP} / (\text{FN} + \text{TP}) \times 100\% = 4 / (1 + 4) \times 100\% \\ &= 4 / 5 \times 100\% \\ &= 80\% \end{aligned}$$

$$\begin{aligned} \text{SC} &: \text{TN} / (\text{FP} + \text{TN}) \times 100\% = 5 / (0 + 5) \times 100\% \\ &= 5 / 5 \times 100\% \\ &= 100\% \end{aligned}$$

3. Tahap ketiga.

Penulis melakukan pengujian seluruh data training dengan hasil sebagai berikut.

accuracy: 86.67%

	true Jaringan	true Sistem Cerdas	class precision
pred. Jaringan	7	0	100.00%
pred. Sistem Cerdas	2	6	75.00%
class recall	77.78%	100.00%	

Gambar 14. Hasil Data Training Tahap Ketiga

Result Data Training 3

$$\text{TP (True Jaringan)} = 7$$

$$\text{FP (False Jaringan)} = 0$$

$$\text{TN (True SC)} = 6$$

$$\text{FN (False SC)} = 2$$

Accuracy

$$\begin{aligned} (TP + TN) / (TP + TN + FP + FN) \times 100\% &= \\ (7 + 6) / (7 + 6 + 0 + 2) \times 100\% &= 13/15 \times 100\% \\ &= 86,67\% \end{aligned}$$

Precision

$$\begin{aligned} \text{Jaringan} : TP / (FP + TP) \times 100\% &= 7 / (0 + 7) \times 100\% \\ &= 7/7 \times 100\% \\ &= 100\% \end{aligned}$$

$$\begin{aligned} \text{SC} : TN / (FN + TN) \times 100\% &= 6 / (2 + 6) \times 100\% \\ &= 6/8 \times 100\% \\ &= 75\% \end{aligned}$$

Recall

$$\begin{aligned} \text{Jaringan} : TP / (FN + TP) \times 100\% &= 7 / (2 + 7) \times 100\% \\ &= 7 / 9 \times 100\% \\ &= 77,78\% \end{aligned}$$

$$\begin{aligned} \text{SC} : TN / (FP + TN) \times 100\% &= 6 / (0 + 6) \times 100\% \\ &= 6 / 6 \times 100\% \\ &= 100\% \end{aligned}$$

b. ROC Curve

ROC Curve (kurva ROC) menunjukkan visualisasi dari akurasi model dan membandingkan perbedaan antar model klasifikasi. ROC merupakan grafik dua dimensi dimana false positives sebagai garis horizontal sedangkan true positives untuk mengukur perbedaan

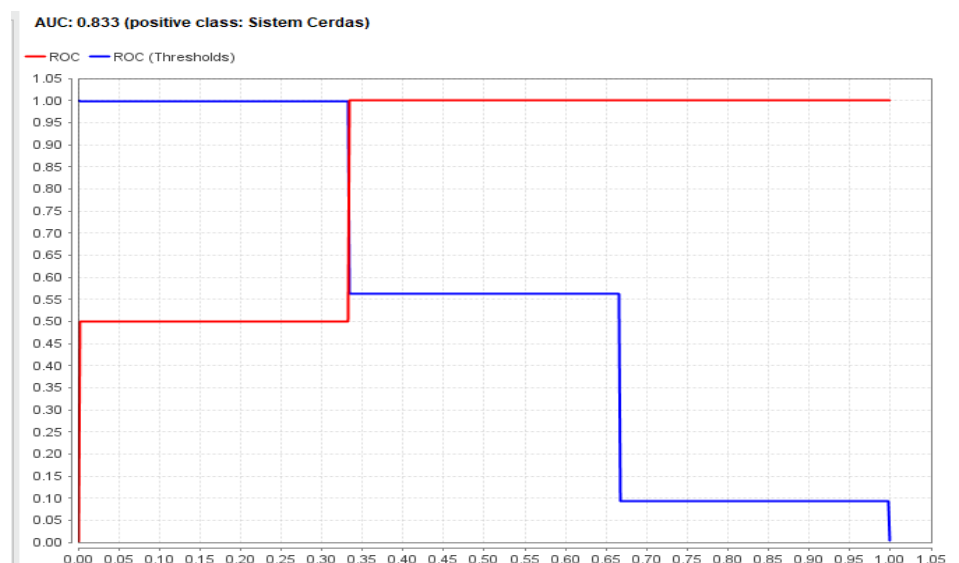
performansi metode yang digunakan. Kurva ROC merupakan teknik untuk memvisualisasi dan menguji kinerja pengklasifikasian berdasarkan performannya (Gorunescu, 2011).

Berikut adalah ROC dari data training sesuai dengan tahap-tahap yang dilakukan penulis.

Untuk keakurasian nilai AUC (Area Under Curve) dalam klasifikasi *data mining* dibagi menjadi lima kelompok (Gorunescu, 2011), yaitu :

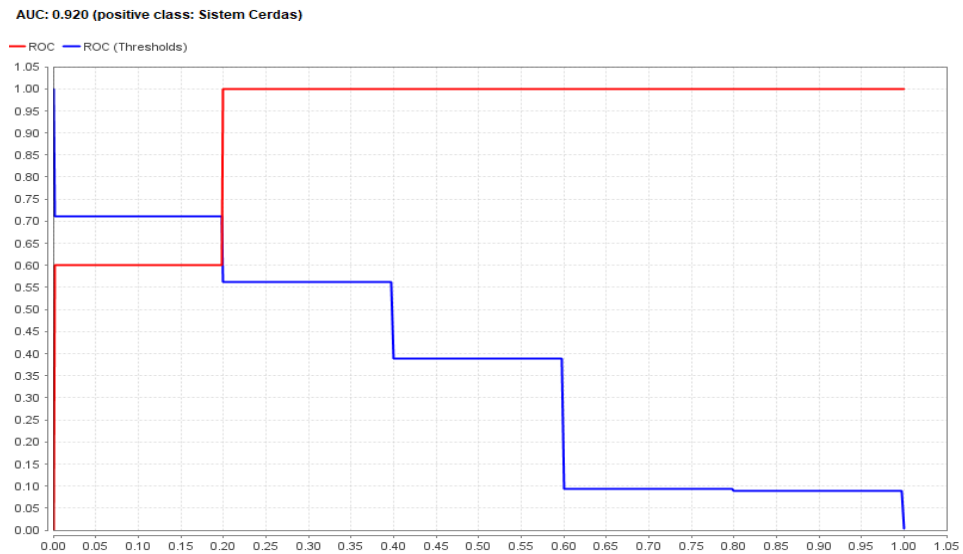
1. 0.90 – 1.00 = klasifikasi sangat baik (*excellent classification*)
2. 0.80 – 0.90 = klasifikasi baik (*good classification*)
3. 0.70 – 0.80 = klasifikasi cukup (*fair classification*)
4. 0.60 – 0.70 = klasifikasi buruk (*poor classification*)
5. 0.50 – 0.60 = klasifikasi salah (*failure*)

1. ROC data training tahap pertama



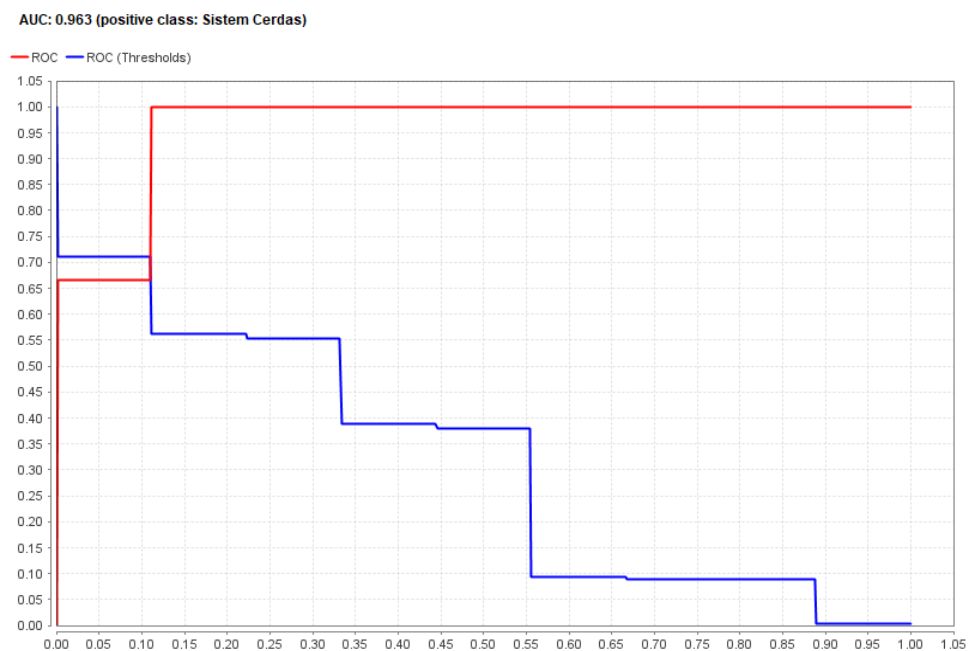
Gambar 15. ROC Kurva Tahap Pertama

2. ROC data training tahap kedua.



Gambar 16. ROC Kurva Tahap Kedua

3. ROC data training tahap ketiga



Gambar 17. ROC Kurva Tahap Ketiga

Dari hasil perbandingan ROC Curve ketiga data diatas kesimpulannya bahwa semakin mahasiswa memilih dosen yang tepat sebagai pembimbing 1 untuk tugas akhir yang dilakukan, maka nilai

AUC akan semakin tinggi. Dari ketiga data ROC di atas, ROC tahap pertama memiliki nilai 0,833 termasuk dalam kelompok *good classification*. Sedangkan untuk ROC tahap kedua dan tahap ketiga memiliki nilai 0,920 dan 0,963 keduanya termasuk dalam kelompok *excellent classification*.

4.3 Analisis Algoritma

Hasil perhitungan *Confusion Matrix* yang dilakukan oleh rapidminer dan perhitungan manual yang dilakukan penulis adalah sama. Dan juga hasil dari pengujian data dengan menggunakan rapidminer adalah bahwa semakin banyak data yang digunakan dan semakin sesuai data testing dan juga data testing akan mempengaruhi hasil dari perhitungan tersebut. Seperti perbedaan pada pada tiap-tiap tahap perhitungan, semakin tepat mahasiswa memilih dosen pembimbing 1 maka nilai akurasi akan semakin bagus. Selain itu, dalam perhitungan hanya bisa menghitung nilai probabilitas untuk dosen pembimbing 1 dikarenakan penentuan berdasarkan keahlian peminatan para dosen sehingga dosen yang tidak memiliki keahlian peminatan tidak dapat dilakukan perhitungannya.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penerapan Naive Bayes dalam membantu pemilihan dosen pembimbing tugas akhir pada Prodi Teknik Informatika STMIK Palangkaraya dapat dilakukan dan menggunakan aplikasi rapidminer dengan hasil :

1. Perhitungan algoritma hanya dapat digunakan untuk dosen pembimbing 1 dikarenakan penulis membatasi pemilihan dosen pembimbing menggunakan keahlian peminatan yang dimiliki dosen.
2. Semakin tepat mahasiswa dalam memilih dosen pembimbing maka nilai akurasi akan semakin tinggi. Seperti hasil yang didapat oleh penulis kali ini nilai akurasinya adalah 80% hingga 90%.

5.2 Saran

Dengan hasil yang didapatkan penulis kali ini diharapkan agar kedepannya dapat dikembangkan lagi. Agar kedepannya tidak hanya dapat memperhitungkan nilai probabilitas dosen yang dipilih mahasiswa, namun dapat merekomendasikan dosen-dosen yang kompeten peminatan yang dikuasai. Juga agar dapat merekomendasikan dosen pembimbing 2 dengan berdasarkan hal lain yang bisa dipertimbangkan sebagai acuan pemilihannya.

DAFTAR PUSTAKA

- Aldi Wicaksono, F. Y., Anita , & Oktavia, C. A. (2019). Sistem Penunjang Keputusan Penentuan Dosen Pembimbing Menggunakan Algoritma Naive Bayes Studi Kasus STIKI Malang. *J-INTECH : Journal of Information and Technology*, VII.
- Aprilla, Dennis, Baskoro, D. A., L. A., & Simri, I. W. (2012). *Belajar Data Mining dengan RapidMiner*. Jakarta: Gramdia Pustaka Utama .
- Fibrianda, M. F., & Bhawiyuga, A. (2018). Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 3112-3123.
- Gumilang, I. C. (2014). *Prediksi Persediaan Obat Dengan Metode Naive Bayes (Studi Kasus: Apotek Saputra)*. Surakarta: Universitas Muhammadiyah Surakarta.
- Han, J. K. (2012). *Data Mining: Concept and Techniques*. Waltham: Morgan Kaufmann Publishers.
- M. A., & N. F. (2020). Implementasi Algoritma Naive Bayes Classifier sebagai Sistem Rekomendasi Pembimbing Skripsi. *Jurnal Nasional Informatika dan Teknologi Jaringan*, V.
- Maulani, T. Z., Simbolon, Z. K., & Amirullah. (2019). Implementasi Algoritma Naïve Bayes Classifier Dalam Menentukan Topik Tugas Akhir Mahasiswa Berbasis Web. *Jurnal Infomedia : Teknik Informatika, Multimedia, dan Jaringan*, IV.
- P. K. (2018). Integrasi Kategori Skripsi Dan Keahlian Dosen Dalam Naïve Bayes Untuk Pemilihan Dosen Pembimbing. *Nusantara of Engineering*, III.
- U. R., Zuhdi, A. M., & Kusrini. (2019). SISTEM PENDUKUNG KEPUTUSAN DENGAN METODE NAIVE BAYES UNTUK PEMILIHAN DOSEN PEMBIMBING. *Jurnal INFORMA Politeknik Indonusa Surakarta*, V.