



Penerapan Metode *Naïve Bayes* Untuk Rekomendasi Topik Tugas Akhir Berdasarkan Daftar Hasil Studi Mahasiswa di Perguruan Tinggi

Anggra Triawan^{1*}, Della Lintang Melinda²

¹Sistem Informasi/STIKOM Binaniga

Email: rajib@stikombinaniaga.ac.id

²Sistem Informasi/STIKOM Binaniga

Email: dellalintangmelinda@gmail.com

ABSTRAK

STIKOM Binaniaga merupakan sekolah tinggi swasta yang menjalankan pendidikan dibidang komputer. Pada akhir studi setiap mahasiswa wajib untuk membuat sebuah penelitian atau tugas akhir. Salah satu tahapan yang harus dilalui dalam proses penyusunan tugas akhir adalah menentukan topik apa yang akan dipilih untuk dasar pembahasan dalam sebuah penelitian yang akan di buat. Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari data set yang diberikan. Input dari aplikasi ini adalah nama mahasiswa semester akhir kelas A, B dan C yang sedang mengerjakan tugas akhir. Lalu aplikasi akan mengambil dan mengolah nilai. Nilai yang diproses adalah nilai yang terkait dengan topik yang disuguhkan dan akan diklasifikasikan kedalam topik yang telah disediakan. Hasil yang diperoleh adalah presentase kelayakan ahli materi sebesar 100% dan di interpretasikan sangat layak sedangkan untuk hasil presentase kelayakan pengguna sebesar 79%, maka aplikasi dapat dikategorikan kedalam interpretasi layak.

Keywords: *Naïve Bayes; data mining; klasifikasi; presentase kelayakan.*

A. PENDAHULUAN

1. Latar Belakang

Pada akhir studi atau pembelajaran setiap mahasiswa yang berasal dari perguruan tinggi negeri maupun perguruan tinggi swasta, baik yang mengambil jenjang Strata maupun Diploma diwajibkan membuat sebuah tugas akhir yang berupa penelitian. Biasanya penelitian yang dilakukan berhubungan dengan program studi yang diambilnya dalam perkuliahan. Dalam pengerjaan tugas akhir ini melalui beberapa proses dan tahapan yang harus dilalui sebelum menghasilkan sebuah penelitian yang dapat dipertanggung jawabkan. Di setiap proses yang dilalui mahasiswa pasti mengalami beberapa kesulitan, salah satu kesulitan yang dihadapi adalah pada saat pemilihan topik tugas akhir yang akan di pakai.

Menurut (Aminudin) pengertian topik adalah “sebuah ide yang mendasari sebuah pembahasan dan sebuah cerita. Dan topik itu akan menjadi sebuah tolak ukur akan pembahasan dan cerita yang ditulis oleh seorang penulis”. Begitu juga yang dilakukan di STIKOM Binaniaga Bogor sebelum menentukan judul, mahasiswa harus menentukan topik apa yang hendak dipilih sebagai acuan atau tolak ukur dalam pengerjaan tugas akhir yang sesuai dengan kemampuan dan masalah yang akan diangkat. Topik yang tepat akan memudahkan mahasiswa dalam mengerjakan tugas akhir dan dengan pemilihan topik yang tepat diharapkan dapat membantu mahasiswa mendapatkan hasil dari tujuan tugas akhir yang sesuai dengan harapan.

Pemilihan topik yang tidak tepat akan menyulitkan mahasiswa dalam proses pengerjaan tugas akhir. Oleh karena itu dibutuhkan sebuah acuan atau tolak ukur yang dapat membantu merekomendasikan mahasiswa dalam menentukan topik untuk tugas akhir. Rekomendasi merupakan suatu pendapat yang telah dipertimbangkan mengenai suatu situasi tertentu dan mencerminkan pengetahuan penilaian (Titis Suwartiningsih, Bagus Setya Rintyarna, Deni Arifianto). Dalam hal ini daftar hasil studi mahasiswa (DHS) bisa dipakai sebagai acuan atau tolak ukur mahasiswa untuk memilih topik tugas akhir apa yang tepat sesuai dengan minat dan bakat yang dimilikinya. Dan dalam penelitian ini juga akan dibuat sebuah sistem informasi yang dapat membantu memberikan rekomendasi untuk menentukan topik tugas akhir berdasarkan daftar hasil studi mahasiswa.

Dalam penelitian ini akan dikembangkan sebuah aplikasi yang dapat merekomendasikan topik untuk tugas akhir mahasiswa dengan menggunakan daftar hasil studi mahasiswa sebagai dasar atau acuan yang jelas dengan menggunakan metode *Naive Bayes*.

2. Permasalahan

Sebagaimana telah ditetapkan oleh bidang studi, ada 6 daftar pilihan topik yang telah ditetapkan untuk program studi sistem informasi kelas A dan 10 daftar pilihan topik yang ditetapkan untuk program studi sistem informasi kelas B dan C yang bisa diambil oleh mahasiswa. Setelah memberi tahu daftar topik apa saja yang dapat dipilih oleh mahasiswa sesuai dengan gambar 1.1 dan 1.2 diatas, mahasiswa diharuskan memilih salah satu diantara topik tersebut tanpa ada acuan atau tolak ukur. Hal ini dirasa kurang efektif dan akan menyulitkan mahasiswa, dengan proses yang sedang berjalan saat ini kebanyakan mahasiswa kebingungan dalam penentuannya dan membutuhkan suatu alat bantu atau sistem yang membantu mahasiswa memberikan rekomendasi untuk pemilihan tugas akhir berdasarkan acuan yang dijadikan pedoman yang telah ditetapkan untuk mahasiswa. Tentunya dasar yang dipakai untuk memilih salah satu dari pilihan itu bisa dilihat dari kemampuan mahasiswa terutama pada indikasi atau indikator keberhasilan studi. Apabila salah satu indikator materi perkuliahan yang berindikasi atau cenderung pada satu topik, maka kecenderungan itu mungkin lebih menjadi bahan pertimbangan bagi mahasiswa untuk memilih salah satu dari beberapa daftar pilihan topik tugas akhir.

Dari uraian diatas, maka dapat diidentifikasi masalah dalam penelitian ini adalah :

- Tidak akuratnya mahasiswa dalam memilih topik untuk tugas akhir
- Belum dapat diperoleh atau ditentukan tingkat ketepatan metode *Naive Bayes* untuk membantu merekomendasikan topik tugas akhir dalam pendekatan aplikasi

3. Tujuan

Adapun tujuan dari penelitian ini adalah :

- Akuratnya mahasiswa dalam memilih topik
- Mengembangkan *prototype* aplikasi dengan menggunakan metode *Naive Bayes* untuk memilih topik tugas akhir
- Mengukur tingkat ketepatan dalam merekomendasikan pemilihan topik untuk tugas akhir

4. Tinjauan Pustaka

a. Data Mining

Menurut Alfa Saleh (2015, p.208) : “data mining merupakan proses pengekstraksian informasi dari sekumpulan data yang sangat besar melalui penggunaan algoritma dan teknik penarikan dalam bidang statisti, pembelajaran mesin dan sistem manajemen basis data. Data mining adalah proses menganalisa data dari perspektif yang berbeda dan menyimpulkan menjadi informasi-informasi penting yang dapat dipakai untuk meningkatkan keuntungan, memperkecil biaya pengeluaran atau bahkan keduanya”.

Menurut Putri Elfa Mas’udia (2015) salah satu tugas utama dari data mining adalah klasifikasi. Klasifikasi digunakan untuk menempatkan bagian yang tidak diketahui pada data ke dalam kelompok yang sudah diketahui. Klasifikasi menggunakan variabel target dengan nilai nominal. Dalam satu set pelatihan, variabel target sudah diketahui. Dengan

pembelajaran dapat ditemukan hubungan antara fitur dengan variabel target. Ada dua langkah dalam proses klasifikasi :

- Pembelajaran (learning) : pelatihan data dianalisis oleh algoritma klasifikasi
- Klasifikasi : data yang diujikan digunakan untuk mengkalkulasi akurasi dari aturan klasifikasi. Jika, akurasi dianggap dapat diterima, aturan dapat diterapkan pada klasifikasi data tuple yang baru.

b. Naïve bayes Classifier

Menurut Affrida Manalu, Fricles Ariwisanto Sianturi, Mamed Rofendy Manalu (2017, p.17), Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. Naive Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. Dengan kata lain, diberikan nilai output, probabilitas mengamati secara bersama adalah produk dari probabilitas individu. Keuntungan penggunaan Naive Bayes adalah bahwa metode ini hanya membutuhkan jumlah data pelatihan (Training Data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naive Bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

Persamaan Naive Bayes

Menurut Alfa Saleh (2015) prediksi Bayes didasarkan pada teorema Bayes dengan formula umum sebagai berikut :

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

Dimana :

- X : Data dengan *class* yang belum diketahui
- H : Hipotesis data merupakan suatu *class* spesifik
- P(H|X) : Probabilitas hipotesis *H* berdasar kondisi *X* (posteriori probabilitas)
- P(H) : Probabilitas hipotesis *H* (prior probabilitas)
- P(X|H) : Probabilitas *X* berdasarkan kondisi pada hipotesis *H*
- P(X) : Probabilitas *X*

Untuk menjelaskan metode Naive Bayes, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, metode Naive Bayes di atas disesuaikan sebagai berikut:

$$P(C|F1 \dots Fn) = \frac{P(C)P(F1 \dots Fn|C)}{P(F1 \dots Fn)}$$

Di mana Variabel *C* merepresentasikan kelas, sementara variabel *F1 ... Fn* merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas *C* (Posterior) adalah peluang munculnya kelas *C* (sebelum masuknya sampel tersebut, seringkali disebut prior), dikali dengan peluang kemunculan karakteristik-karakteristik sampel pada kelas *C* (disebut juga likelihood), dibagi dengan peluang kemunculan karakteristik-karakteristik sampel secara global (disebut juga evidence). Karena itu, rumus di atas dapat pula ditulis secara sederhana sebagai berikut :

$$\text{Posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}}$$

Nilai Evidence selalu tetap untuk setiap kelas pada satu sampel. Nilai dari posterior tersebut nantinya akan dibandingkan dengan nilai-nilai posterior kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut rumus Bayes tersebut dilakukan dengan menjabarkan ($C|F_1, \dots, F_n$) menggunakan aturan perkalian sebagai berikut:

$$\begin{aligned} P(C|F_1, \dots, F_n) &= P(C)P(F_1, \dots, F_n|C) \\ &= P(C)P(F_1|C)P(F_2, \dots, F_n|C, F_1) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3, \dots, F_n|C, F_1, F_2) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2)P(F_4, \dots, F_n|C, F_1, F_2, F_3) \\ &= P(C)P(F_1|C)P(F_2|C, F_1)P(F_3|C, F_1, F_2) \dots P(F_n|C, F_1, F_2, F_3, \dots, F_{n-1}) \end{aligned}$$

Dapat dilihat bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor - faktor syarat yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisa satu persatu. Akibatnya, perhitungan tersebut menjadi sulit untuk dilakukan. Di sinilah digunakan asumsi independensi yang sangat tinggi (naif), bahwa masing-masing petunjuk ($F_1, F_2, \dots, F_n$) saling bebas (independen) satu sama lain. Dengan asumsi tersebut, maka berlaku suatu kesamaan sebagai berikut:

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i)$$

Untuk $i \neq j$, sehingga

$$P(F_i|C, F_j) = P(F_i|C)$$

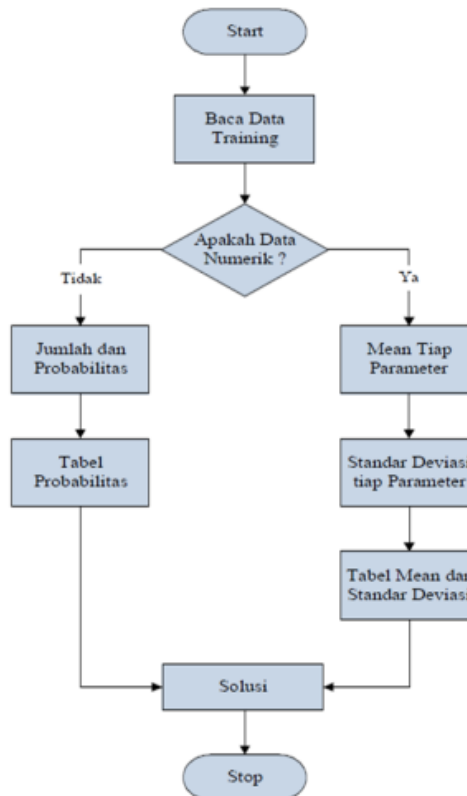
Persamaan di atas merupakan model dari teorema Naive Bayes yang selanjutnya akan digunakan dalam proses klasifikasi. Untuk klasifikasi dengan data kontinyu digunakan rumus Densitas Gauss :

$$P(X_i = x_i | Y = y_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma_{ij}^2}}$$

B. METODE

Menurut Sugiyono (2009) metode penelitian dan pengembangan (R&D) merupakan metode penelitian yang dipakai untuk menghasilkan produk tertentu, dan menguji keektifan produk itu. Agar bisa menghasilkan suatu produk tertentu yang dipakai untuk penelitian yang bersifat analisis kebutuhan (digunakan metode survey atau kualitatif) dan untuk menguji keefektifan produk tersebut supaya bisa berfungsi di masyarakat luas, maka diperlukan penelitian guna menguji keektifan produk tersebut.

Lebih lanjut lagi, Borg and Gall (1989) menyatakan bahwa untuk melakukan penelitian analisis kebutuhan sehingga bisa dihasilkan suatu produk yang bersifat hipotetik, tidak jarang memakai metode penelitian dasar (basic research). Berikutnya guna menguji produk yang masih bersifat hipotetik itu, menggunakan eksperimen atau action research. Sesudah produk teruji, maka bisa diaplikasikan secara luas. Proses pengujian produk dengan menggunakan penelitian eksperimen ini disebut dengan penelitian terapan (applied research). Penelitian R & D itu sendiri, bertujuan untuk menemukan, mengembangkan, dan memvalidasi suatu produk, dengan begitu penelitian R & D bersifat longitudinal.



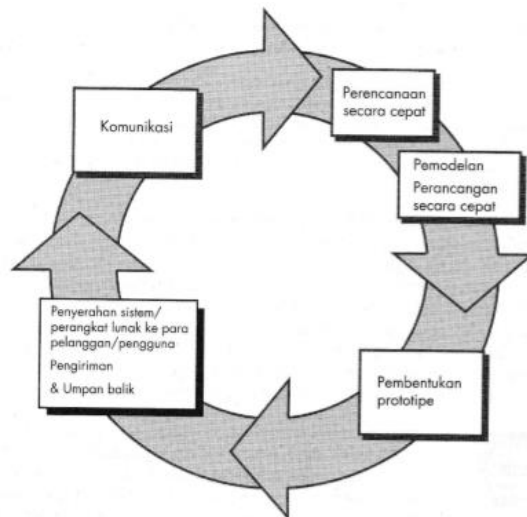
Gambar 1 Alur metode Naive Bayes
 (Sumber :Alfa Saleh,2015,p.211)

1. Model Yang Di Usulkan

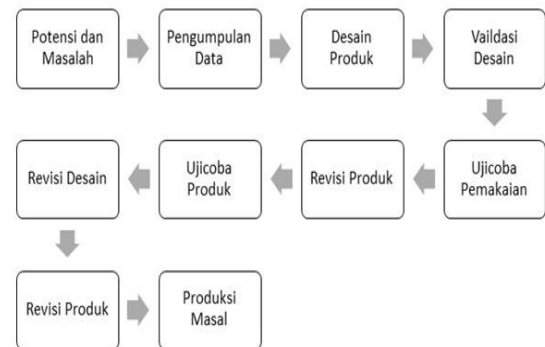
Menurut Ogedebe, dkk (2012, p.55), menyampaikan bahwa prototyping merupakan metode pengembangan perangkat lunak, yang berupa model fisik kerja sistem dan berfungsi sebagai versi awal dari sistem. Dengan metode prototyping ini akan dihasilkan prototype sistem sebagai perantara pengembang dan pengguna agar dapat berinteraksi dalam proses kegiatan pengembangan sistem informasi. Agar proses pembuatan prototype ini berhasil dengan baik adalah dengan mendefinisikan aturan-aturan pada tahap awal, yaitu pengembang dan pengguna harus satu pemahaman bahwa prototype dibangun untuk mendefinisikan kebutuhan awal. Prototype akan dihilangkan atau ditambahkan pada bagiannya sehingga sesuai dengan perencanaan dan analisis yang dilakukan oleh pengembang sampai dengan ujicoba dilakukan secara simultan seiring dengan proses pengembangan.

Suatu rancangan cepat berfokus pada representasi semua aspek perangkat lunak yang akan terlihat oleh para pengguna akhir (misalnya rancangan antarmuka pengguna [user interface] atau format tampilan). Rancangan cepat (quick design) akan memulai konstruksi pembuatan prototype. Prototype kemudian akan diserahkan kepada para stakeholder dan kemudian mereka akan melakukan evaluasi-evaluasi tertentu terhadap prototype yang telah dibuat sebelumnya, kemudian akhirnya akan memberikan umpan-balik yang akan digunakan untuk memperhalus spesifikasi kebutuhan iterasi akan terjadi saat prototype diperbaiki untuk memenuhi kebutuhan dari para stakeholder, sementara pada saat yang sama memungkinkan kita untuk lebih memahami kebutuhan apa yang akan dikerjakan pada iterasi selanjutnya (Pressman, 2012, p.51).

Rancang cepat (quick design) akan memulai konstruksi pembuatan prototype, prototype kemudian akan diserahkan kepada para stakeholder dan kemudian akan melakukan evaluasi – evaluasi tertentu terhadap prototype yang telah dibuat sebelumnya, kemudian akhirnya akan memberikan umpan balik yang akan digunakan untuk memperhalus spesifikasi kebutuhan. Iterasi akan terjadi saat prototype diperbaiki untuk memenuhi kebutuhan dari para stakeholder, sementara pada saat yang sama memungkinkan kita untuk lebih memahami kebutuhan apa yang kita kerjakan pada iterasi selanjutnya.



Gambar 2. Prototype Model
 (Sumber : Pressman, 2012, p.50)



Gambar 3. Prosedur Pengembangan

2. Prosedur Pengembangan

Prosedur pengembangan merupakan langkah – langkah dari proses pengembangan yang dilakukan. Prosedur pengembangan dalam penelitian yang akan dilakukan dapat digambarkan pada gambar 3.

3. Sumber Data

Proses pengujian ini bertujuan untuk memperoleh data apasaja yang digunakan untuk mendukung keberhasilan didalam penelitian. Penelitian ini menggunakan dua jenis data yaitu data primer berupa hasil kuisisioner yang digunakan pada tahap penangkapan masalah awal dan data sekunder berupa data nilai mahasiswa STIKOM dari setiap matakuliah yang terkait dengan topik yang disuguhkan. Data ini digunakan untuk menganalisa dan mengetahui kebutuhan dan ketepatan informasi untuk rekomendasi topik tugas akhir berdasarkan daftar hasil studi mahasiswa STIKOM Binaniaga Bogor.

4. Instrumen Penelitian

a. Instrumen Untuk Ahli

Instrument yang digunakan dalam penelitian ini adalah angket atau kuesioner. Sugiyono (2014) menyatakan bahwa “Instrumen penelitian adalah suatu alat pengumpul data yang digunakan untuk mengukur fenomena alam maupun sosial yang diamati”. Instrumen ahli disini adalah ahli teori atau ahli materi. Dalam penelitian ini ahli materi yang dipakai adalah ahli dalam materi metode yang digunakan yaitu Naive Bayes. Instrumen yang digunakan untuk pengukuran uji ahli materi adalah tentang bagaimana alur metode Naive Bayes berjalan dalam aplikasi. Menurut jurnal dari Alfa Saleh 2015 yang berjudul “Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga” langkah-langkah metode Naive bayes adalah seperti yang terlihat pada tabel 1.

Tabel. 1 Instrumen Untuk Ahli

No	Indikator
1	Rules Naive Bayes
2	Jumlah Data dan Probabilitas
3	Tabel Probabilitas
4	Mencari Nilai Likelihood
5	Hasil

Skala Guttman digunakan dalam pengukuran hasil uji ahli materi. Dalam skala Guttman ini menggunakan dua macam jenis pertanyaan pada angket atau kuesioner tersebut, yaitu

jenis pertanyaan tertutup dan jenis pertanyaan terbuka. Jenis pertanyaan tertutup berisi pertanyaan-pertanyaan seputar kesesuaian alur-alur metode Naive Bayes. Sementara jenis pertanyaan terbuka berisi pendapat dan saran dari ahli materi.

Tabel. 2 Skoring skala Guttman

Alternative Jawaban	Skor Alternative Jawaban	
	Positive	Negative
Ya	1	0
Tidak	0	1

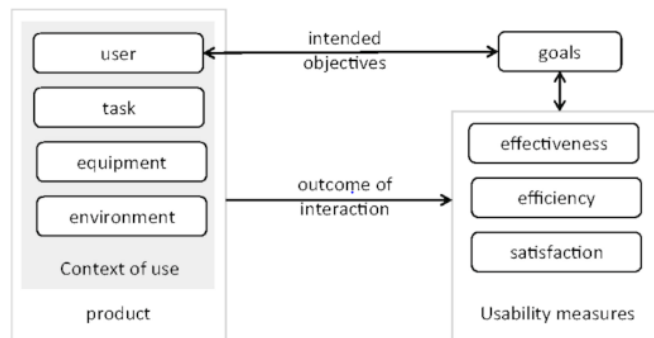
(Sumber: Rizky Djati Munggaran, 2012)

b. Instrumen Untuk Pengguna

International Organization for Standardization (ISO) 9241-11 merupakan standart ISO yang digunakan untuk melakukan pengukuran usability. Pengukuran standart ISO 9241-11 menggunakan beberapa kriteria penilaian yaitu efficiency (efisien), effectiveness (efektivitas), dan satisfaction (kepuasan). Ruang lingkup dalam penggunaan standart ini adalah untuk mengevaluasi dan merancang tampilan sistem atau aplikasi agar sesuai dengan kebutuhan pengguna dan kepuasan pengguna sehingga dapat mencapai sebuah tujuan yang diinginkan. Berikut merupakan pengertian masing-masing tujuan dari usability menurut dokumen ISO 9241-11 2017 :

- 1) Effectiveness (efektivitas) : bagaimana pengguna dapat mencapai tujuan yang diinginkan.
- 2) Efficiency (efisiensi) : sumber daya yang digunakan dalam mencapai tujuan dan hasil yang telah dicapai.
- 3) Satisfaction (kepuasan) : sejauh mana pengguna merasakan kepuasan terhadap suatu produk yang telah digunakan.

Berikut merupakan kerangka kerja usability menurut ISO 9241-11



Gambar. 4 Konsep Usability ISO 9241-11

(Sumber: Draft International ISO DIS 9241-11 Standart)

5. Teknik Analisa Data

Data dari angket dalam penelitian ini merupakan data kuantitatif yang akan dianalisis secara deskriptif persentase. Metode analisis data yang digunakan persentase kelayakan. Adapun rumus yang digunakan adalah sebagai berikut:

$$\text{Persentase kelayakan (\%)} = \frac{\text{Skor yang diobservasi}}{\text{Skor yang diharapkan}} \times 100\%$$

Hasil presentase digunakan untuk memberikan jawaban atas kelayakan dari aspek – aspek yang diteliti. Menurut Arikunto (2009, p.44), pembagian kategori kelayakan ada lima. Skala ini memperhatikan rentang dari bilangan presentase. Nilai maksimal yang diharapkan adalah 100% dan minimum 0%. Pembagian rentang kategori kelayakan menurut Arikunto (2009, p.44), dapat dilihat pada Tabel 3.

Tabel 3 Kategori Kelayakan Menurut Arikunto

Presentase Pencapaian	Interpretasi
< 21%	Sangat Tidak Layak
21% - 40%	Tidak Layak

41% - 60%	Cukup Layak
61% - 80%	Layak
81% - 100%	Sangat Layak

(Sumber : Arikunto, 2009, p.44)

6. Uji Hasil

Confusion matrix adalah metode yang digunakan untuk melakukan perhitungan akurasi pada konsep data mining. Presisi atau confidence adalah proporsi kasus yang diprediksi positif yang juga positif benar pada data yang sebenarnya. Recall atau sensitivity adalah proporsi kasus positif yang sebenarnya yang diprediksi positif secara benar.

Tabel. 4 Model Confusion Matrix

Aktual	Classified as	
	+	-
+	True Positives	False Negatives
-	False Positives	True Negatives

(sumber :Triowali Rosandy,2016)

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

C. HASIL DAN PEMBAHASAN

1. HASIL

a. Hasil Analisa Proses

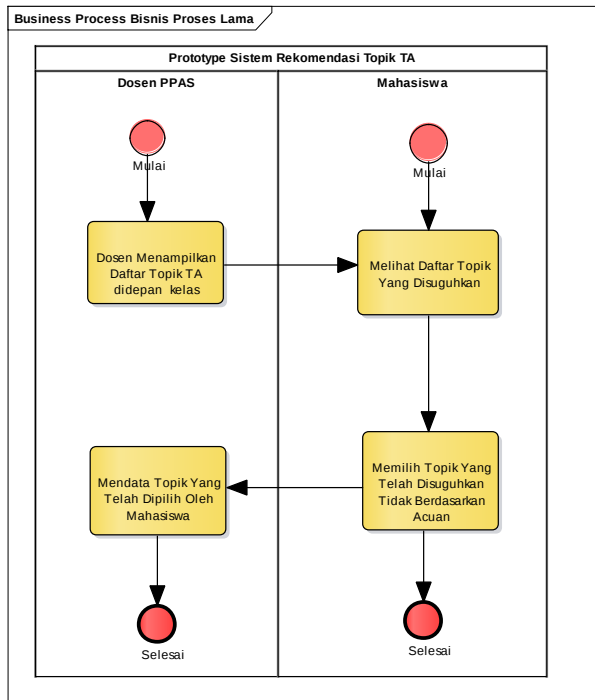
Penentuan topik tugas akhir yang selama ini dilakukan di STIKOM Binaniaga Bogor dijelaskan pada gambar 5 dalam diagram proses bisnis lama yang didapat berdasarkan hasil analisis kelemahan sistem lama yaitu mahasiswa dalam menentukan topik tugas akhir tidak berdasarkan acuan sebagai tolak ukurnya.

b. Hasil Analisa Metode

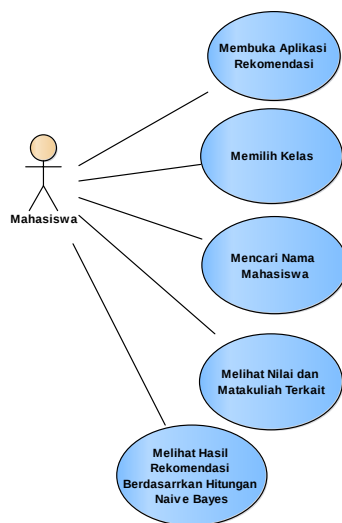
Untuk mengatasi masalah penentuan proses rekomendasi topik tugas akhir yang belum mempunyai kriteria akan dibuat sistem informasi penentuan proses rekomendasi topik tugas akhir yang didalamnya menerapkan metode *Naïve Bayes*. Dapat dilihat pada gambar 1 langkah *Naïve Bayes*.

c. Hasil Analisa Kebutuhan Sistem

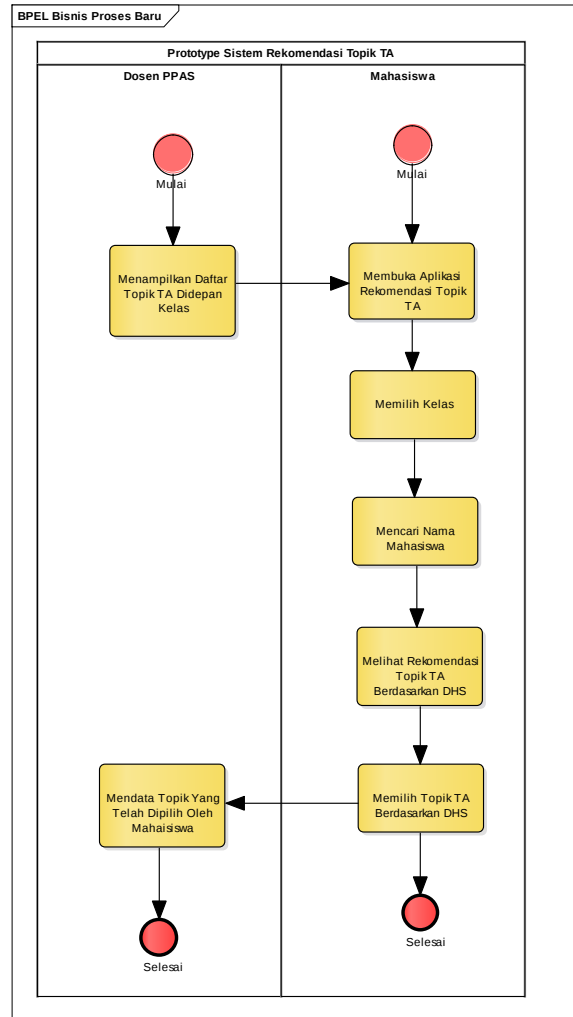
Pemodelan objek pada sistem yang dikembangkan ini dijelaskan dalam bentuk diagram use case berdasarkan pada proses rekomendasi topik tugas akhir pada sistem yang akan dikembangkan untuk memodelkan serta mengorganisasi pada sistem sehingga mendapatkan keluaran sistem sesuai dengan yang diharapkan dan dibutuhkan. Gambar 7 merupakan diagram use case pada sistem yang akan dikembangkan.



Gambar 5. Proses Bisnis Lama



Gambar 7. Diagram Usecase



Gambar 6. Proses Bisnis Baru

2. PEMBAHASAN

a. Uji Hasil *Naïve Bayes*

Perhitungan yang digunakan untuk memperoleh probabilitas dari setiap topik berdasarkan nilai dari setiap matakuliah yang telah dibuatkan rulesnya. Berikut ini adalah penjelasan lebih terperinci mengenai tiap-tiap langkah dalam probabilitas setiap topik dari setiap nilai matkul dengan menggunakan Naive Bayes untuk mengklasifikasikan datanya.

Tabel. 5 Kemunculan Setiap Nilai Untuk Atribut Data and Info Management

DIM	Jumlah Kejadian Dipilih						Probabilitas					
	T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6
A	4	1	1	1	1	1	0.667	0.167	0.167	0.167	0.167	0.167
B	1	1	1	1	1	1	0.167	0.167	0.167	0.167	0.167	0.167

C	1	2	2	2	2	2	0.167	0.333	0.333	0.333	0.333	0.333
D	0	2	2	2	2	2	0.000	0.333	0.333	0.333	0.333	0.333
E	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
Jumlah	6	6	6	6	6	6	1	1	1	1	1	1

Tabel. 6 Kemunculan Setiap Nilai Untuk Knowledge Management

KM	Jumlah Kejadian Dipilih						Probabilitas					
	T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6
A	1	4	1	1	1	1	0.167	0.667	0.167	0.167	0.167	0.167
B	1	1	1	1	1	1	0.167	0.167	0.167	0.167	0.167	0.167
C	2	1	2	2	2	2	0.333	0.167	0.333	0.333	0.333	0.333
D	2	0	2	2	2	2	0.333	0.000	0.333	0.333	0.333	0.333
E	0	0	0	0	0	0	0.000	0.000	0.000	0.000	0.000	0.000
Jumlah	6	6	6	6	6	6	1	1	1	1	1	1

Tabel. 7 Kemunculan Setiap Nilai Untuk Keputusan Topik A

Keputusan	Jumlah Kejadian Dipilih						Probabilitas					
	T1	T2	T3	T4	T5	T6	T1	T2	T3	T4	T5	T6
Jumlah	6	6	6	6	6	6	0.167	0.167	0.167	0.167	0.167	0.167

- a) menghitung likelihood, nilai likelihood didapatkan dari hasil perkalian seluruh probabilitas pada tabel untuk setiap matakuliah dengan topik yang sama lalu dikalikan lagi dengan hasil bagi dari jumlah kemunculan untuk topik tersebut yang dibagi dengan jumlah seluruh data aturan. Apabila mahasiswa bernama Della Lintang Melinda dari reguler A memiliki nilai matakuliah Data and Information Management A, Knowledge Management B, Enterprise Arsitektur B, Information Project Management A, Data Mining and Data Warehousing A, dan Application Development B, maka dihitung likelihoodnya sebagai berikut :

Likelihood Data Management :

$$\left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{6}{36}\right) = 0.0000143$$

Likelihood Knowledge Acquisition :

$$\left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{6}{36}\right) = 0.0000036$$

Likelihood System Development Process :

$$\left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{6}{36}\right) = 0.0000036$$

Likelihood System Project Estimation:

$$\left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{4}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{6}{36}\right) = 0.0000143$$

Likelihood Business Intelligent Support System:

$$\left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{4}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{6}{36}\right) = 0.0000143$$

Likelihood Tools and Techniques in System Development :

$$\left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{1}{6}\right) \times \left(\frac{6}{36}\right) = 0.0000036$$

Setelah didapatkan nilai likelihood dari masing-masing topik, selanjutnya dicari nilai probabilitasnya:

Probabilitas Data and Info Management :

$$0.0000143$$

$$\frac{0.0000143 + 0.0000036 + 0.0000036 + 0.0000143 + 0.0000143 + 0.0000036}{0.0000036} = 0.2667$$

Probabilitas Knowledge Acquisition :

$$0.0000036$$

$$\frac{0.0000143 + 0.0000036 + 0.0000036 + 0.0000143 + 0.0000143 + 0.0000036}{0.0000036} = 0.0667$$

Probabilitas System Development Process :

$$\frac{0.0000036}{0.0000143 + 0.0000036 + 0.0000036 + 0.0000143 + 0.0000143 + 0.0000036} = 0.0667$$

Probabilitas System Project Estimation:
 $\frac{0.0000143}{0.0000143 + 0.0000036 + 0.0000036 + 0.0000143 + 0.0000143 + 0.0000036}$

$$= 0.2667$$

Probabilitas Business Intelligent Support System:
 $\frac{0.0000143}{0.0000143 + 0.0000036 + 0.0000036 + 0.0000143 + 0.0000143 + 0.0000036}$

$$= 0.2667$$

Probabilitas Tools and Techniques in System Development :
 $\frac{0.0000036}{0.0000143 + 0.0000036 + 0.0000036 + 0.0000143 + 0.0000143 + 0.0000036}$

$$= 0.0667$$

Tabel 8 Tabel Hasil Probabilistik Kelas A

NO	TOPIK	PROBABILITAS
1	Data and Info Management	0.2667
2	System Project Estimation	0.2667
3	Business Intelligent Support System	0.2667
4	Knowledge Acquisition	0.0667
5	System Development Process	0.0667
6	Tools and Techniques in System Development	0.0667

b. Hasil Persentase Kelayakan

Hasil persentase kelayakan merupakan jawaban atas kelayakan dari aspek – aspek yang diteliti. Hasil yang diperoleh melalui instrumen uji tingkat kelayakan pengguna dan akurasi komparasi.

1) Hasil Kelayakan Pengguna

Berikut hasil persentase kelayakan yang diperoleh, dapat dilihat pada tabel 9.

Tabel 9. Hasil Persentase Kelayakan

No	Aspek Penilaian	Skor Observasi	Skor Yang Diharapkan	Kelayakan
1	<i>Efficiency</i>	180	225	80%
2	<i>Effectiveness</i>	218	300	73%
3	<i>Satisfaction</i>	233	300	78%
Overall		650	825	79%

Berdasarkan tabel 9, persentase kelayakan terbesar yaitu pada aspek penilaian *Efficiency* sebesar 80%. Namun secara keseluruhan kelayakan dari aspek penilaian *Efficiency*, *Effectiveness* *Satisfaction* adalah sebesar 79%. maka dapat dikategorikan ke dalam interpretasi yang “Sangat Layak” seperti yang ditunjukkan pada tabel 3.

2) Hasil Akurasi Komparasi

Uji akurasi dilakukan untuk mengukur seberapa akurat hasil dari aplikasi yang akan dibuat. Dalam penelitian ini untuk mengukur tingkat akurasi menggunakan rumus confusion matrix. Komparasi dilakukan dengan membandingkan output dari aplikasi yang dibuat dengan output yang dikeluarkan oleh alat data mining yaitu rapidminer seperti yang tertera pada tabel 10 dibawah ini.

Tabel. 10 Hasil Confusion Matrix

Aplikasi rekomendasi	Rapid Miner									
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
T1	0	0	0	0	0	0	0	0	0	0

T2	0	1	0	0	0	0	0	0	0	0
T3	0	0	1	0	0	0	0	0	0	0
T4	0	0	0	0	0	0	0	0	0	0
T5	0	0	0	0	0	0	0	0	0	0
T6	0	0	1	0	0	1	0	0	0	0
T7	0	0	0	0	0	0	0	0	0	0
T8	0	0	1	0	0	0	0	2	0	0
T9	0	0	0	0	0	0	0	0	0	0
T10	0	0	1	0	0	0	0	1	0	1

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$\text{Akurasi} = \frac{6}{10} \times 100\%$$

$$\text{Akurasi} = 60\%$$

Berdasarkan perhitungan akurasi, maka didapatkan hasil akurasi sebesar 60%.

D. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, kesimpulan yang bisa diuraikan antara lain :

1. Penggunaan variabel matakuliah dan atributnya nilai mutu menghasilkan sebuah rekomendasi topik tugas akhir yang akurat bagi mahasiswa, dibuktikan dengan hasil klasifikasi penentuan topik tugas akhir
2. Proses penentuan topik menjadi lebih efektif karena adanya perangkat aplikasi rekomendasi topik tugas akhir, sehingga mahasiswa mendapatkan rekomendasi topik tugas akhir dengan acuan yang tepat tanpa harus mengikuti pilihan teman.
3. Dengan menerapkan metode Naive Bayes dibuat sebuah aplikasi rekomendasi topik tugas akhir berdasarkan daftar hasil studi. Aplikasi membantu memberi rekomendasi kepada mahasiswa dalam menentukan topik untuk tugas akhir. Telah dilakukan uji kelayakan terhadap aplikasi yang dibuat dengan nilai persentase kelayakan secara keseluruhan dari tiga indikator yaitu Efficiency, Effectiveness Satisfaction adalah sebesar 79% yang bermakna aplikasi masuk dalam kategori layak digunakan, hal tersebut didasarkan pada tabel kategori kelayakan skala likert.
4. Komparasi antara output dari aplikasi yang telah dibuat dengan output dari alat data mining yang sudah ada, sehingga dihasilkan tingkat akurasi dari komparasi tersebut sebesar 60%.

E. DAFTAR PUSTAKA

- [1] Argario, Brilian Healto. Nurul Hidayat. Ratih Kartika Dewi (2018). Implementasi Metode *Naive Bayes* Untuk Diagnosis Penyakit Kambing (Studi Kasus: UPTD Pembibitan Ternak dan Hijauan Makanan Ternak Kec. Singosari Malang). Universitas Brawijaya. Malang
- [2] Djadir, dkk (2017). Sumber Belajar Penunjang 2017 Mata Pelajaran/Paket Keahlian Matematika.
- [3] Ghaniy, Rajib, and Kepin Sihotang. "Penerapan Metode Naïve Bayes Classifier untuk Penentuan Topik Tugas Akhir." *Teknois*, vol. 9, no. 1, 16 May. 2019, pp. 63-72, doi:10.36350/jbs.v9i1.7.
- [4] Kasih, Patmi. Intan Nur Farida. (2017). Sistem Bantu Pemilihan Dosen Pembimbing Tugas Akhir Berdasarkan Kategori Pilihan dan Keahlian Dosen menggunakan Naive Bayes. Sekolah Tinggi Informatika dan Komputer Indonesia. Malang
- [5] Kristian, R Ananda. Ida Wahyuni. (2018). Penentuan Topik Judul Tugas Akhir Mahasiswa di STMIK Asia Malang Menggunakan Fuzzy Inference System Tsukamoto. STMIK ASIA. Malang
- [6] Kusnendi. Modul Konsep Dasar Sistem Informasi. Universitas Pendidikan Indonesia. Bandung

- [7] Manalu, Effrida. Fricles Ariwisanto Sianturu. Mamed Rofendy Manalu. (2017). Penerapan Algoritma *Naive Bayes* Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan dan Jumlah Pemesanan Pada CV Papa dan Mama Pastries. Politeknik Trijaya Krama. Medan
- [8] Mas'udia, Putri Elfa. (2015). Klasifikasi Tugas Akhir Untuk Menentukan Dosen Pembimbing Menggunakan *Naive Bayes Classifier* (NBC). Politeknik Negeri Malang
- [9] Matondang, Zulkifli. (2009). Validitas dan Reliabilitas Suatu Instrumen Penelitian. Universitas Negeri Medan. Medan
- [10] Meilantika, Dian. (2017). Rancang bangun Sistem Informasi Administrasi Menggunakan Metode Throwaway Prototyping Development Pada Sultan-Sport. AMIK AKMI Baturaja. Baturaja
- [11] Muchlis, Imam. (2015). Pengaruh Stres Kerja Terhadap Kinerja Karyawan Di PT.Batik Danar Hadi Surakarta. Surakarta.
- [12] Munggaran, Rizky Djati (2012). Pemanfaatan Open Source Software Pendidikan Oleh Mahasiswa Dalam Rangka Implementasi Undang-Undang No.19 Tahun 2002 Tentang Hak Cipta. Universitas Pendidikan Indonesia. Bandung
- [13] Mulyanto, Aunur R. (2008). *Rekayasa Perangkat Lunak* Jilid 1. Jakarta: Direktorat Pembinaan Sekolah Menengah Kejuruan
- [14] Muslihudun, M. (2017). Sistem Informasi Penjualan Batik Basurek Berbasis Web Pada Basurek Collection Bengkulu. STMIK Pringsewu Lampung. Lampung
- [15] Prabowo, Singgih. Aisyah Endah Palupi. Pengembangan Modul Pembelajaran CNCII Untuk Meningkatkan Efektifitas Belajar Mahasiswa Program Studi D3 Teknik Mesin Universitas Negeri Surabaya. Surabaya
- [16] Prasetyo, Markus Guntur Tri. (2014). Sistem Pendukung Keputusan Pemilihan Tema Tugas Akhir Dengan Metode AHP-SAW (Studi Kasus Program Studi Teknik Informatika – S1 UNDINUS). Universitas Dian Nusantoro. Semarang
- [17] Pressman, RS. (2012). *Rekayasa Perangkat Lunak*. Yogyakarta: Andi.
- [18] Purnomo, Dwi. (2017). Model Prototyping pada pengembangan Sistem Informasi. Universitas Widyagama Malang. Malang
- [19] Rosalinda, Sayyidatul Eka Putri (2018). Evaluasi *Usability* Sistem Informasi Manajemen Kepegawaian UIN Sunan Ampel Surabaya Berbasis ISO9241-11. Surabaya
- [20] Triowali, Rosandy (2016). Perbandingan Metode *Naive Bayes Classifier* Dengan Metode Decision Tree (C.45) Untuk Menganalisa Kelancaran Pembiayaan (Study Kasus : KSPPS / BMT Al-Fadhila)
- [21] Saleh, Alfa. (2015). Implementasi Metode Klasifikasi *Naive Bayes* Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. Universitas Potensi Utama Medan. Medan
- [22] Sugianto, Castaka Agus 2015. Penerapan Teknik Data Mining Untuk Menentukan Hasil Seleksi Masuk SMAN 1 Gigeber Untuk Siswa Baru Menggunakan Decision Tree.
- [23] Suwartiningsih, Titis. Rekomendasi Topik Tugas Akhir Mahasiswa Teknik Informatika Di Universitas Muhammadiyah Jember Menggunakan Metode *Naive Bayesian Classifier*. Universitas Muhammadiyah Jember. Jember
- [24] Syarli. (2016). Metode *Naive Bayes* Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi. Universitas AL-Asyariah Mandar.
- [25] Wasiati, Hera. Dwi Wijayanti. (2014). Sistem Pendukung Keputusan Penentuan Kelayakan Calon Tenaga Kerja Indonesia Menggunakan Metode *Naive Bayes* (Studi Kasus: Di P.T. Karyatama Mitra Sejati Yogyakarta). STIMIK AKAKOM. Yogyakarta