

Penerapan Text Mining pada Sistem Rekomendasi Pembimbing Skripsi Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier

Muhammad Ikhsan¹, Rakhmat Kurniawan R²

mhd.ikhsan@uinsu.ac.id, rakhmat.kr@uinsu.ac.id

Universitas Islam Negeri Sumatera Utara Medan

Informasi Artikel	Abstrak
Diterima: 4 Nov 2023 Direview: 9 Nov 2023 Disetujui: 30 Des 2023	Penunjukan dosen pembimbing skripsi oleh program studi akan disesuaikan dengan rumpun keilmuan dosen dengan topik dari judul proposal skripsi mahasiswa. Selain rumpun keilmuan, program studi juga harus berusaha mendistribusikan dosen pembimbing secara proporsional, sehingga potensi terjadinya kelebihan beban kerja pada dosen pembimbing dapat dihindari. Hal ini berlaku untuk dosen pembimbing I dan dosen pembimbing II. Selanjutnya dirasa perlu melakukan penelitian untuk mengatasi masalah tersebut. Dengan harapan memudahkan pengelola program studi dalam menentukan Pembimbing Skripsi yang sesuai antara topik dengan keilmuan dosen. Selanjutnya mempercepat proses penentuan judul skripsi mahasiswa dan tentunya juga akan mempercepat proses penyelesaian studi mahasiswa strata 1. Maka dirancanglah sebuah sistem yang mampu merekomendasi dosen pembimbing yang sesuai dengan topik skripsi yang diajukan melalui teknologi text mining menggunakan algoritma NBC. Algoritma NBC mampu untuk mendeteksi topik yang ada pada suatu koleksi dokumen beserta besarnya kemunculan topik tersebut.
Kata Kunci	
Skripsi, Judul, Pembimbing, Text Mining, NBC	
Keywords	Abstract
Thesis, Title, Supervisor, Text Mining, NBC	<i>The appointment of a thesis supervisor by the study program will be adjusted to the lecturer's scientific field with the topic of the title of the student's thesis proposal. Apart from scientific groups, study programs must also try to distribute supervisors proportionally, so that the potential for work overload on supervisors can be avoided. This applies to supervisor I and supervisor II. Furthermore, it is deemed necessary to carry out research to overcome this problem. With the hope of making it easier for study program managers to determine a Thesis Supervisor that matches the topic and the lecturer's knowledge. Furthermore, it speeds up the process of determining the title of a student's thesis and of course will also speed up the process of completing undergraduate students' studies. So a system is designed that is able to recommend supervisors that are appropriate to the proposed thesis topic through text mining technology using the NBC algorithm. The NBC algorithm is able to detect topics in a document collection along with the number of occurrences of these topics.</i>

A. Pendahuluan

Skripsi merupakan karya tulis ilmiah yang disusun oleh mahasiswa pada tingkat Strata I yang dilaksanakan melalui *field research* ataupun *library research*. Penelitian ini dilakukan untuk memenuhi salah syarat untuk memperoleh Gelar Sarjana. Setiap mahasiswa yang telah memenuhi syarat, diwajibkan untuk mengajukan judul skripsi. Dalam menyelesaikan skripsi, mahasiswa akan didampingi oleh dosen pembimbing. Penunjukan dosen pembimbing skripsi oleh program studi akan disesuaikan dengan rumpun keilmuan dosen dengan topik dari judul proposal skripsi mahasiswa. Selain rumpun keilmuan, program studi juga harus berusaha mendistribusikan dosen pembimbing secara proporsional, sehingga potensi terjadinya kelebihan beban kerja pada dosen pembimbing dapat dihindari. Hal ini berlaku untuk dosen pembimbing I dan dosen pembimbing II.

Proses penunjukan dosen pembimbing skripsi dilakukan secara manual, dimana setiap dosen akan dicatat judul skripsi dan mahasiswa yang didampinginya. Pencatatan dilakukan secara manual dengan menggunakan Microsoft Excel dan kemudian direkapitulasi untuk mengetahui berapa jumlah mahasiswa yang sedang didampinginya. Setelah pencatatan dilakukan, maka program studi akan mengeluarkan nota tugas pembimbing skripsi. Meskipun pencatatan sudah dilakukan, namun metode pencatatan seperti ini berpotensi menimbulkan kekeliruan berupa ketidaksesuaian antara rumpun keilmuan dosen dengan topik skripsi mahasiswa, sehingga dapat menimbulkan potensi menurunnya kualitas dari skripsi mahasiswa. Pendistribusian dosen pembimbing juga berpotensi tidak proporsional, sehingga memberikan beban kerja yang berlebih. Dengan berlebihnya beban kerja juga akan mempengaruhi performansi dosen pembimbing dalam membimbing skripsi mahasiswa, sehingga dikhawatirkan akan memperlambat masa penyelesaian skripsi.

Text Mining merupakan suatu proses eksplorasi dan analisis dataset yang besar dalam bentuk teks untuk mendapatkan suatu informasi yang bermanfaat untuk tujuan tertentu. Salah satu algoritma yang dapat digunakan dalam *text mining* adalah *Naïve Bayes Classifier*. Algoritma *Naïve Bayes Classifier* mampu untuk melakukan klasifikasi terhadap suatu data atau teks.

Berdasarkan permasalahan dan uraian penyelesaian masalah diatas, maka akan diajukan proposal penelitian dasar berbasis program studi melalui BOPTN Kemenag tahun 2023 dengan judul: "Penerapan Text Mining Pada Sistem Rekomendasi Pembimbing Skripsi Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier di Program Studi Ilmu Komputer UIN Sumatera Utara Medan".

B. Metode Penelitian

Jenis penelitian yang diadopsi dalam penelitian ini adalah penelitian R&D (Research and Development). Pemilihan metode R&D berdasarkan pada tujuan penelitian untuk menghasilkan produk yang efisien. Metode penelitian Research and Development, memiliki beberapa tahapan sebagaimana dikemukakan oleh (Sugiyono, 2016) yaitu: Potensi dan masalah, Pengumpulan Data, Desain Produk, Validasi Desain, Uji Coba

Produk, Revisi Produk, Uji Pelaksanaan Lapangan, Penyempurnaan Produk, dan Implementasi

Pandangan keliru terhadap metode R&D dapat terjadi ketika peneliti yang menggunakannya menginterpretasikan metode ini sesuai dengan susunan kata yakni penelitian dan pengembangan. Mereka berasumsi ada dua kegiatan yang berlangsung dalam suatu pekerjaan yaitu meneliti dan mengembangkan sesuatu. Dalam konteks R&D sebagai suatu metode penelitian yang utuh. Penelitian dan pengembangan adalah suatu kesatuan istilah yang secara kontekstual tidak dapat dipisahkan antara kata penelitian (research) dan pengembangan (development) baik secara struktur maupun arti/makna. Tentu penamaan tersebut terjadi karena penelitian dan pengembangan mempunyai tujuan yang berbeda dengan jenis metode penelitian yang lain.

1. Pengumpulan Data

Data yang digunakan dalam penelitian ini berupa dokumen proposal skripsi mahasiswa Program Studi Ilmu Komputer UIN Sumatera Utara Medan yang diperoleh langsung dari Program Studi Ilmu Komputer. Data yang digunakan adalah proposal skripsi Program Studi Ilmu Komputer dari tahun 2021 – 2023 dengan jumlah 132 record. Data ini akan terus bertambah seiring berjalannya waktu. Data yang diperoleh dari Program Studi Ilmu Komputer berupa file spreadsheet dalam format gsheets. File tersebut berisi field-field judul, dimana tiap row terdiri dari 1 (satu) judul proposal. Judul proposal pada setiap row ini nantinya akan diolah dengan Menentukan topik dari judul tersebut, kemudian memberikan rekomendasi dosen pembimbing untuk judul proposal tersebut.

C. Hasil dan Pembahasan

a. Representasi Data

Data yang digunakan dalam representasi data ini berupa sampel data dari judul skripsi yang diajukan mahasiswa kepada program studi Ilmu Komputer Universitas Islam Negeri Sumatera Utara dengan jumlah data 10 data judul skripsi mahasiswa yang dibagi menjadi dua bagian yaitu 8 data latihan dan 2 data uji.

b. Hasil Analisis Data

Hasil analisis data digunakan untuk mengetahui apakah data yang digunakan sesuai dengan penelitian. Tahap awal adalah preprocessing teks yang bertujuan untuk merubah teks menjadi data terstruktur. Berikut ini merupakan *sample teks* berupa judul yang digunakan.

Tabel 1. Data Latihan Judul Skripsi Mahasiswa

No.	Nama Mahasiswa	Judul Skripsi	Nama Dosen	Kode Dosen
1	Muhammad Ihsan	Penerapan Metode C4.5 Dalam Data Mining Prediksi Penjualan Produk Pada Cv. Chavasindo Berbasis Web	Yusuf Ramadhan Nasution, M.Kom	D1
2	Yazid Munali	Sistem Pakar Penyakit Hipertensi Dengan Pendekatan Statistikal Probabilistik Naïve Bayes	Armansyah, M.Kom	D2
3	Julkarnain Hidayat	Sistem Pakar Identifikasi Gaya Belajar Menggunakan Metode	Raissa Amanda Putri, S.Kom., M.T.I	D3

No.	Nama Mahasiswa	Judul Skripsi	Nama Dosen	Kode Dosen
		Case Base Reasoning Dan Dempster Shafer.		
4	Putri Khairani Ritonga	Analisis Perbandingan Algoritma Clustering K-Means Untuk Seleksi Siswa Baru	Muhammad Siddik Hasibuan, M.Kom	D4
5	Alvina Damayanti	Clustering Padi Dengan Algoritma K-Means Penentuan Beras Terlaris Pada Kilang Padi Di Kecamatan Pantai Cermin	Raissa Amanda Putri, S.Kom., M.T.I	D3
6	Mhd. Agung Perdana	Implementasi Logika Fuzzy Tsukamoto Sebagai Kendali Sistem Ruangan Inkubasi DOC Ayam Kampung	Muhammad Ikhsan, S.T., M.Kom	D5
7	Devita Anggraini	Initial Centroid Pada Algoritma K-Means Dan K-Medoids	Muhammad Siddik Hasibuan, M.Kom	D4
8	Rama Suhari	Sistem Pemilihan Calon Anggota DPRD Menggunakan Fuzzy Tahani	Sriani, M.Kom	D6

Tabel 2. Data Uji Judul Skripsi Mahasiswa

No.	Nama Mahasiswa	Judul Skripsi	Kode Dosen
1	Selly Handayani	Penerapan Data Mining Untuk Pengelompokan Data Plat Kendaraan Bermotor Dengan Algoritma Clustering	?
2	Putra Salim	Perancangan Prototype Sistem Implementasi Logika Fuzzy Sugeno Pada Deteksi Longsor Menggunakan Sensor Accelerometer dan Kelembaban Berbasis Mikrokontroler	?

1) Text Preprocessing

Adapun tahapan yang akan dilakukan pada data latih dan data uji yaitu melakukan text preprocessing seperti tahapan berikut:

- **Case Folding dan Stop Character Removal**

Tahapan ini bertujuan untuk penyeragaman *case* menjadi huruf kecil dan penghilangan karakter tanda baca pada penulisan judul skripsi. Hal ini dikarenakan bahasa yang digunakan dalam penerapan *topic modeling* merupakan *case sensitive* yang menganggap judul skripsi yang ditulis dengan huruf kapital dan huruf kecil berbeda walaupun sebenarnya kedua judul tersebut sama dan hal ini dapat menyebabkan terjadinya data kembar.

Tabel 3. Hasil Case Folding dan Stop Character Removal Data Latih

No.	Nama Mahasiswa	Judul Skripsi	Kode Dosen
1	Muhammad Ihsan	penerapan metode c4 5 dalam data mining prediksi penjualan produk pada cv chavasindo berbasis web	D1
2	Yazid Munali	sistem pakar penyakit hipertensi dengan pendekatan statistik probabilitas naïve bayes	D2
3	Julkarnain Hidayat	sistem pakar identifikasi gaya belajar menggunakan metode case base reasoning dan dempster shafer	D3
4	Putri Khairani Ritonga	analisis perbandingan algoritma clustering k means untuk seleksi siswa baru	D4
5	Alvina Damayanti	clustering padi dengan algoritma k means penentuan beras terlaris pada kilang padi di kecamatan pantai cermin	D3

No.	Nama Mahasiswa	Judul Skripsi	Kode Dosen
6	Mhd. Agung Perdana	implementasi logika fuzzy tsukamoto sebagai kendali sistem ruangan inkubasi doc ayam kampung	D5
7	Devita Anggraini	initial centroid pada algoritma k means dan k medoids	D4
8	Rama Suhari	sistem pemilihan calon anggota dprd menggunakan fuzzy tahani	D6

Tabel 4. Hasil Case Folding dan Stop Character Removal Data Uji

No.	Nama Mahasiswa	Judul Skripsi	Kode Dosen
1	Selly Handayani	penerapan data mining untuk pengelompokan data plat kendaraan bermotor dengan algoritma clustering	?
2	Putra Salim	perancangan prototype sistem implementasi logika fuzzy sugeno pada deteksi longsor menggunakan sensor accelerometer dan kelembaban berbasis mikrokontroler	?

- Tokenizing**

Tahapan *tokenizing* merupakan tahapan pemecahan *teks* judul skripsi menjadi elemen kecil berupa kata tunggal. Adapun hasil *tokenizing* sebagai berikut:

Tabel 5. Hasil Tokenizing Data Latih

1	penerapan, metode, c4, 5, dalam, data, mining, prediksi, penjualan, produ, pada, cv, chavasindo, berbasis, web
2	sistem, pakar, penyakit, hipertensi, dengan, pendekatan, statistik, probabilistik, naïve, bayes
3	sistem, pakar, identifikasi, gaya, belajar, menggunakan, metode, case, base, reasoning, dan, Dempster, share
4	analisis, perbandingan, algoritma, clustering, k, means, untuk, seleksi, siswa, baru
5	clustering, padi, dengan, algoritma, k, means, penentuan, beras, terlaris, pada, kilang, padi, di, kecamatan, pantai, cermin
6	implementasi, logika, fuzzy, tsukamoto, sebagai, kendali, sistem, ruangan, inkubasi, doc, ayam, kampung
7	initial, centroid, pada, algoritma, k, means, dan k, medoids
8	sistem, pemilihan, calon, anggota dprd, menggunakan, fuzzy, tahani

Tabel 6. Hasil Tokenizing Data Uji

1	penerapan, data, mining, untuk, pengelompokan, data, plat, kendaraan, bermotor, dengan, algoritma, clustering
2	perancangan, prototype, sistem, implementasi, logika, fuzzy, sugeno, pada, deteksi, longsor, menggunakan, sensor, accelerometer, dan kelembaban, berbasis, mikrokontroler

- Stopword Removal dan Stemming**

Stopword removal adalah sebuah proses penghapusan kata yang tidak berhubungan. Tujuan *stopword* ini untuk meminimalisir munculnya kata yang dianggap tidak penting dalam memudahkan proses klasifikasi. Hasil *Stopword Removal* untuk data latih dan data uji disajikan pada Tabel 7 dan Tabel 8 dibawah ini:

Tabel 7. Stopword Removal dan Stemming Data Latih

1	penerapan, metode, c4, data, mining, prediksi, jual, produk, cv, chavasindo, basis, web
2	sistem, pakar, sakit, hipertensi, dekat, statistik, probabilistik, naïve, bayes
3	sistem, pakar identifikasi, gaya, belajar, guna, metode, case, base, reasoning,

	demster, share
4	analisis, banding, algoritma, clustering, k, means, seleksi, siswa, baru
5	clustering, padi, algoritma, k, means, tentuan, beras, laris, kilang, padi, kecamatan, pantai, cermin
6	implementasi, logika, fuzzy, tsukamoto, kendali, sistem, ruang, inkubasi, doc, ayam, kampung
7	initial, centroid, algoritma, k, means, k, medoids
8	sistem, pilih, calon, anggota dprd, guna, fuzzy, tahani

Tabel 8. Hasil Stopword Removal dan Stemming Data Uji

1	penerapan, data, mining, kelompok, data, plat, kendaraan, bermotor, algoritma, clustering
2	rancang, prototype, sistem, implementasi, logika, fuzzy, sugeno, pada, deteksi, longsor, guna, sensor, accelerometer, lembab, basis, mikrokontroler

2) Klasifikasi Naïve Bayes

Pada tahap klasifikasi naïve bayes proses dari pencarian nilai probabilitas kelas pada data uji yaitu dari data latih. Perhitungan nilai probabilitas setiap kata untuk masing-masing kelas dari data latih dapat dilihat sebagai berikut:

- Melakukan perhitungan probabilitas setiap kategori atau kelas.

$$P(D1 = \text{Yusuf Ramadhan Nasution, M. Kom}) = \frac{12}{79} = 0,15$$

$$P(D2 = \text{Armansyah, M. Kom}) = \frac{9}{79} = 0,11$$

$$P(D3 = \text{Raissa Amanda Putri, S. Kom., M. T. I}) = \frac{24}{79} = 0,30$$

$$P(D4 = \text{Muhammad Siddik Hasibuan, M. Kom}) = \frac{15}{79} = 0,19$$

$$P(D5 = \text{Muhammad Ikhsan, S. T., M. Kom}) = \frac{11}{79} = 0,14$$

$$P(D6 = \text{Sriani, M. Kom}) = \frac{8}{79} = 0,10$$

- Melakukan perhitungan probabilitas setiap kata dari data yang ingin di uji

Data Uji 1: penerapan, data, mining, kelompok, plat, kendaraan, motor, algoritma, clustering

Tabel 9. Hasil Perhitungan Probabilitas Data Uji 1

Kata	D1	D2	D3	D4	D5	D6
Penerapan	0,0219	0,0212	0,0194	0,0212	0,0222	0,0229
Mining	0,0219	0,0212	0,0194	0,0212	0,0222	0,0229
Kelompok	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Plat	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Kendaraan	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Motor	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Algoritma	0,0439	0,0454	0,0388	0,0425	0,0444	0,0459
Clustering	0,0329	0,0340	0,0291	0,0319	0,0333	0,0344

Data Uji 2: rancan, prototype, sistem, implementasi, logika, fuzzy, sugeno, deteksi, longsor, guna, sensor, accelerometer, lembab, basis, mikrokontroler

Tabel 10. Hasil Perhitungan Probabilitas Data Uji 2

Kata	D1	D2	D3	D4	D5	D6
Rancang	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Prototype	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Sistem	0,0439	0,0454	0,0388	0,0425	0,0444	0,0459
Implementasi	0,0219	0,0212	0,0194	0,0212	0,0222	0,0229
Logika	0,0219	0,0212	0,0194	0,0212	0,0222	0,0229

Fuzzy	0,0329	0,0340	0,0291	0,0319	0,0333	0,0344
sugeno	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Deteksi	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Longsor	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Guna	0,0329	0,0340	0,0291	0,0319	0,0333	0,0344
Sensor	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Accelerometer	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Lembab	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114
Basis	0,0219	0,0212	0,0194	0,0212	0,0222	0,0229
Mikrokontroler	0,0109	0,0113	0,0097	0,0106	0,0111	0,0114

- Melakukan Perkalian Hasil Probabilitas Kelas

Tabel 11. Hasil Perkalian Probabilitas

Data Uji	Probabilitas Dosen Pembimbing					
	D1	D2	D3	D4	D5	D6
1	3,212107E-17	3,238313E-17	2,189465E-17	3,09859E-17	3,438002E-17	3,202558E-17
2	1,625992E-28	2,02857E-28	5,47129E-29	1,32277E-28	1,929146E-28	2,1221191E-28

Dari hasil probabilitas dosen pembimbing judul skripsi pada tabel 4.9 dapat dianalisis nilai probabilitasnya agar dapat menentukan dosen pembimbing judul skripsi dengan cara mencari nilai tertinggi. Adapun nilai probabilitas tertinggi pada data uji 1 terdapat di kode dosen D5 = 3,438002E-17 atas nama dosen Muhammad Ikhsan, S.T., M.Kom., sedangkan untuk data uji 2 terdapat di kode dosen D6 = 2,1221191E-28 atas nama dosen Sriani, M.Kom. seperti pada tabel 4.10 berikut:

Tabel 12. Hasil Prediksi Dosen Pembimbing

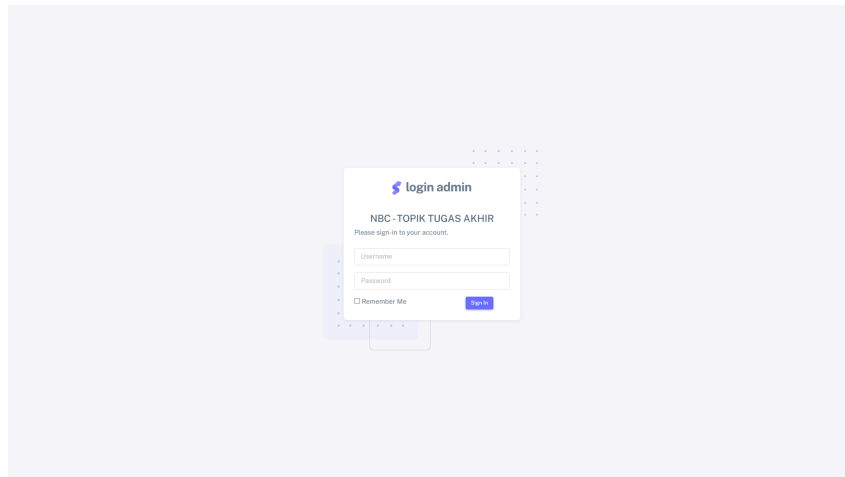
No.	Nama Mahasiswa	Judul Skripsi	Hasil Prediksi	
			Kode Dosen	Nama Dosen
1	Selly Handayani	Penerapan Data Mining Untuk Pengelompokan Data Plat Kendaraan Bermotor Dengan Algoritma Clustering	D5	Muhammad Ikhsan, S.T., M.Kom
2	Putra Salim	Perancangan Prototype Sistem Implementasi Logika Fuzzy Sugeno Pada Deteksi Longsor Menggunakan Sensor Accelerometer dan Kelembaban Berbasis Mikrokontroler	D6	Sriani, M.Kom

c. Penerapan Sistem

Implementasi dari algoritma dan perancangan *user interface* yang telah disajikan pada sub bab sebelumnya, sebagai berikut :

1. Perancangan Halaman Login

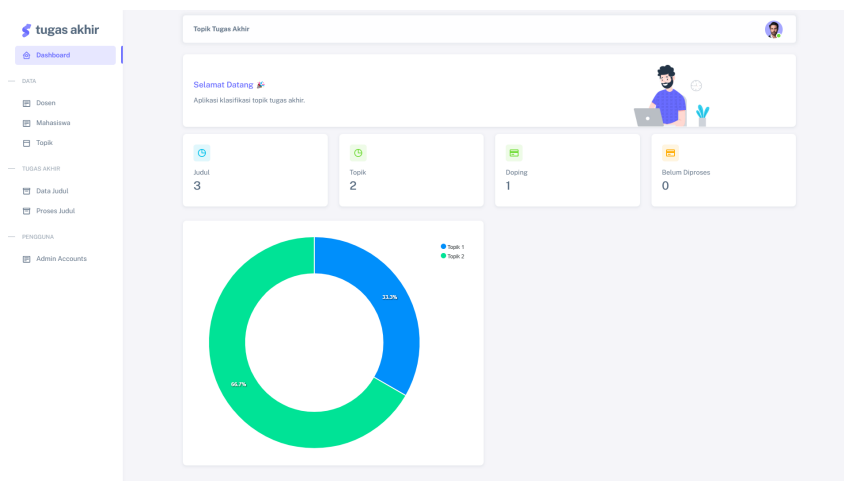
Halaman ini berisi *form login* dengan pilihan level login sebagai administrator dan mahasiswa. Berikut ini adalah desain tampilannya :



Gambar 2. Halaman Login

2. Perancangan Halaman Dashboard (Admin)

Halaman ini berisi informasi terkait jumlah mahasiswa, jumlah proposal serta jumlah kategori yang telah masuk ke dalam *database* sistem.



Gambar 3. Halaman Dashboard (Admin)

3. Halaman *add new dataset* (Admin)

Halaman ini berfungsi untuk menambah *dataset* yang baru, sehingga jumlah kategori yang dapat dipilih lebih banyak variasinya.

Gambar 4. Halaman New Dataset (Admin)

4. Halaman Proses Judul (Admin)

Pada halaman ini admin dapat melihat data judul yang belum diproses. Pada halaman ini juga admin dapat memproses seluruh judul secara sekaligus menggunakan naïve bayes classifier.

ID	NAMA	TOPIK	ABSTRAK	DOKUMEN 1	DOKUMEN 2	AKSI
070175043	Muhammad Hanun	Penerapan Computer Vision Dalam Pengenalan Bahasa Arab Menggunakan Gambar Dan Tulisan Dengan Menggunakan Algoritma Template Matching	---	NDR... Nama: ...	NDR... Nama: ...	Proses
070172057	Nurul Adinda	Implementasi Data Mining Untuk Memrediksi Prestasi Siswa/Siswa Menggunakan Metode Regresi Linier Berganda	---	NDR... Nama: ...	NDR... Nama: ...	Proses
070172022	Imam Utomo	Implementasi Data Mining Menggunakan Metode Clustering Dengan Algoritma K-Means Dalam Proses Pencatatan Pencapaian Kinerja Karyawan Kantor Camat Medan Mamuju	---	NDR... Nama: ...	NDR... Nama: ...	Proses

Gambar 5. Halaman Training Dataset (Admin)

5. Halaman Set Topik

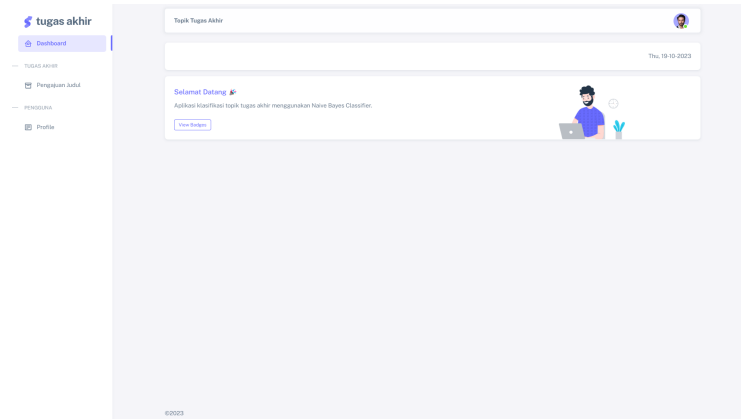
Pada halaman ini admin dapat menentukan topik dari setiap judul yang telah di daftarkan mahasiswa berdasarkan probabilitas dan NBC yang dimiliki oleh setiap judul

KODE	NAMA	DESKRIPSI	AKSI
T-03	Fuzzy Logic Computing	Fuzzy Logic Computing	Proses
T-04	Artificial Neural Network Computing	Artificial Neural Network Computing	Proses
T-05	Probabilistic Computing	Probabilistic Computing	Proses
T-06	Data Mining	Data Mining	Proses
T-07	Text Mining	Text Mining	Proses
T-08	Data Visualization	Data Visualization	Proses
T-09	Informatics Retrieval	Informatics Retrieval	Proses
T-10	Decision Support System	Decision Support System	Proses
T-11	Robotics	Robotics	Proses
T-12	Intelligent Agents	Intelligent Agents	Proses
T-13	Internet of Things	Internet of Things	Proses
T-14	Image and Signal Processing	Image and Signal Processing	Proses
T-15	Virtual Reality & Augmented Reality	Virtual Reality & Augmented Reality	Proses
T-16	Vision or Pattern Recognition	Vision or Pattern Recognition	Proses
T-17	System Interrelation	System Interrelation	Proses

Gambar 6. Halaman Set Topik

6. Halaman Dashboard (Mahasiswa)

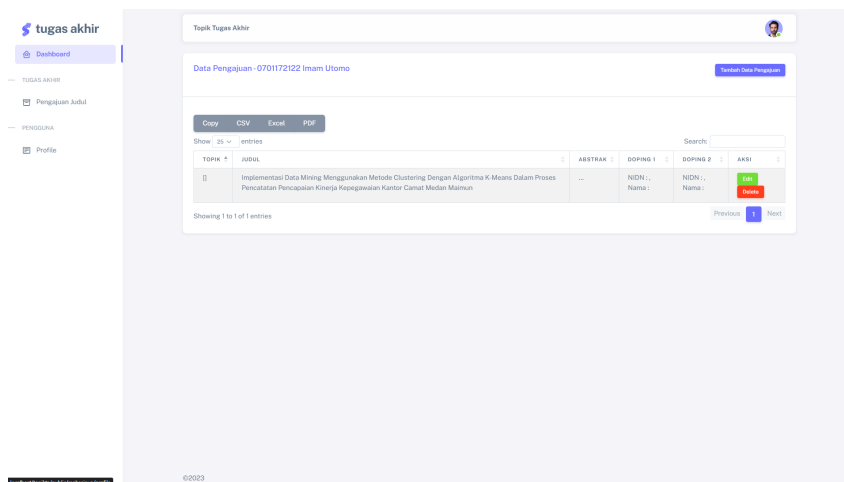
Pada halaman ini mahasiswa dapat melihat data diri dan riwayat proposal yang telah diajukan.



Gambar 7. Halaman Dashboard (Mahasiswa)

7. Halaman Detail Proposal (Mahasiswa)

Pada halaman ini mahasiswa dapat melihat status dari judul skripsi yang telah diajukan secara detail, serta dapat mengetahui status dari judul yang diberikan dan melihat saran yang diberikan admin terkait judul yang diajukan apabila ada.



Gambar 8. Halaman Detail Proposal (Mahasiswa)

D. Simpulan

Berdasarkan hasil penelitian dengan judul “Penerapan Text Mining Pada Sistem Rekomendasi Pembimbing Skripsi Mahasiswa Menggunakan Algoritma Naïve Bayes Classifier di Program Studi Ilmu Komputer UIN Sumatera Utara Medan” didapat kesimpulan bahwa algoritma *Naïve Bayes Classifier* dapat digunakan untuk menentukan topik dari judul skripsi yang diajukan mahasiswa. Sistem mampu memberikan rekomendasi Dosen Pembimbing berdasarkan kesesuaian topik dengan keilmuan dosen. Dengan penggunaan sistem ini, proses penunjukan dosen

pembimbing skripsi dapat di lakukan dalam waktu yang singkat.

E. Ucapan Terimakasih

Terimakasih kami sampaikan kepada LP2M Universitas Islam Negeri Sumatera Utara Medan yang telah menyelenggarakan kegiatan penelitian yang didanai oleh BOPTN sehingga terfasilitasinya kegiatan penelitian ini. Serta ucapan terimakasih kepada seluruh tim penelitian dan publikasi baik dari dosen, alumni dan mahasiswa yang membantu dalam mensukseskan dan menghasilkan penelitian ini.

F. Referensi

- [1] Balya. (2019). *Analisis Sentimen Pengguna Youtube di Indonesia pada Review Smartphone Menggunakan Naïve Bayes*.
<http://repositori.usu.ac.id/handle/123456789/23217>
- [2] Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84. <https://doi.org/10.1145/2133806.2133826>
- [3] Buntoro, G. A. (2017). Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter. *Integer Journal*, 2(1), 32–41. <https://t.co/jrvaMsgBdH>
- [4] Gilang Kencana, C., & Sibaroni, Y. (n.d.). *Klasifikasi Sentiment Analysis pada Review Buku Novel Berbahasa Inggris dengan Menggunakan Metode Support Vector Machine (SVM)*.
- [5] Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. *JEPIN (Jurnal Edukasi Dan Penelitian Informatika)*, 4(2), 113–118.
- [6] Hakim, D. M. (2019). *Optical Music Recognition Pada Citra Notasi Musik Menggunakan Convolutional Neural Network*. UNIKOM.
- [7] Indrayuni, E. (2018). KOMPARASI ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE UNTUK ANALISA SENTIMEN REVIEW FILM. 14(2), 175.
- [8] Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2014). Foundations in Machine learning. In *SpringerBriefs in Computer Science* (Vol. 0, Issue 9783319056050).
- [9] Muljono, Artanti, D. P., Syukur, A., Prihandono, A., & Setiadi, D. R. I. M. (2018). Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naive Bayes. *Konferensi Nasional Sistem Informasi*, 165–170.
- [10] Nugroho, D. D. A., & Alamsyah, A. (2018). Analisis Konten Pelanggan Airbnb Pada Network Sosial Media Twitter. *EProceedings of Management*, 1622–1628.
- [11] Nurzahputra, A., & Muslim, A. (2016). Analisis Sentimen pada Opini Mahasiswa Menggunakan Natural Language Processing. *Seminar Nasional Ilmu Komputer*.
- [12] Priyono, F., Kanti, S., Dzulfihar, I. I., Amirulloh, I., Sri, E. P., Rosiyadi, D., Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri Jl Salemba Raya No, P., & Pusat, J. (2016). Analisis Sentimen Media Sosial Opini Ujian Nasional Berbasis Komputer menggunakan Metoda Naive Bayes. In *Journal of Electrical and Electronics Engineering* (Vol. 1, Issue 2). <http://www.antarakalten>
- [13] Putra, K. B., & Kusumawardani, R. P. (2017). Analisis Topik Informasi Publik Media Sosial di Surabaya Menggunakan Pemodelan Latent Dirichlet Allocation (LDA). *Jurnal Teknik ITS*, 6(2).
<https://doi.org/10.12962/j23373539.v6i2.23205>

-
- [14] Putu, I., Wirayasa, M., Made, I., Wirawan, A., Pradnyana, A., Kunci, K., Stemming, ., Bali, B., & Bastal, A. (n.d.). *ALGORITMA BASTAL: ADAPTASI ALGORITMA NAZIEF & ADRIANI UNTUK STEMMING TEKS BAHASA BALI* (Vol. 8).
- [15] Rahim, R., Zufria, I., Kurniasih, N., Simargolang, M. Y., Hasibuan, A., Sutiksno, D. U., Nanuru, R. F., Anamofa, J. N., Ahmar, A. S., & Achmad Daengs, G. S. (2018). C4.5 classification data mining for inventory control. *International Journal of Engineering and Technology(UAE)*, 7, 68–72. <https://doi.org/10.14419/ijet.v7i2.3.12618>
- [16] Russell, S., & Norvig, P. (2021). Artificial Intelligence A Modern Approach (4th Edition). In *Pearson Series*.
- [17] S, V., & R, J. (2016). Text Mining: open Source Tokenization Tools – An Analysis. *Advanced Computational Intelligence: An International Journal (ACII)*, 3(1), 37–47. <https://doi.org/10.5121/acii.2016.3104>
- [18] Sevsa, B. A., & Wahyudi, M. D. R. (2019). Analisis Sentimen pada Indeks Kinerja Dosen Fakultas SAINTEK UIN Sunan Kalijaga Menggunakan Naive Bayes Classifier. *Jurnal Buana Informatika*, 10(2), 112–123.
- [19] Sugiyono. (2016). Metode Penelitian Bisnis. Alfabeta: Bandung. *Jurnal Analisis*, 6(2).
- [20] Sunardi, Fadlil, A., & Suprianto. (2018). Analisis Sentimen Menggunakan Metode Naive Bayes Classifier Pada Angket Mahasiswa. *SAINTEKBU: Jurnal Sains Dan Teknologi*, 10(2541–1942), 1–9.
- [21] Suntoro, J. (2019). *Data Minig: Algoritma dan Implementasi dengan Pemrograman PHP*. PT Elex Media Komputindo.
- [22] Triase, & Samsudin. (2020). IMPLEMENTASI DATA MINING DALAM MENGLASIFIKASIKAN UKT (UANG KULIAH TUNGGAL) PADA UIN SUMATERA UTARA MEDAN. *Jurnal Teknologi Informasi*, 4(2).
- [23] Ulfah Siregar, Z., Ruli, R., Siregar, A., & Arianto, R. (2019). *KLASIFIKASI SENTIMENT ANALYSIS PADA KOMENTAR PESERTA DIKLAT MENGGUNAKAN METODE K-NEAREST NEIGHBOR*. 8(1).