



Penerapan *Text Mining* Dalam Klasifikasi Judul Skripsi Yang Diusulkan Mahasiswa Menggunakan Metode *Naïve Bayes*

Sukriadi¹, Ismail², A. M Andzar³

Prodi Teknik Informatika, Universitas Lamappapoleonro^{1,3}

Prodi Sistem Informasi, Universitas Lamappapoleonro²

Jl. Kesatria No. 60 Watansoppeng, Soppeng Sulawesi Selatan-Indonesia^{1,2,3}

sukriadi@unipol.ac.id^{*1}, ismail@unipol.ac.id², anchaandi59@gmail.com³

Kata Kunci :

Text Mining,
Klasifikasi Judul
Skripsi, *Naïve*
Bayes.

ABSTRAK

Klasifikasi judul skripsi sesuai konsentrasi mahasiswa secara otomatis dapat dilakukan dengan cara penerapan *text mining* berdasarkan judul skripsi. Tujuan dilakukan klasifikasi judul skripsi berdasarkan konsentrasi mahasiswa untuk memudahkan program studi memonitoring kesesuaian road map penelitian, berdasarkan konsentrasi program studi dengan riset yang dilakukan mahasiswa. Adapun dampak positif atau keuntungan yang didapatkan program studi dalam menerapkan klasifikasi judul skripsi secara otomatis yaitu memudahkan mengembangkan roadmap penelitian mahasiswa ditingkatkan program studi serta selalu meningkatkan kualitas lulusan prodi melalui hasil penelitian mahasiswa. Metode pengembangan sistem menggunakan *system development life cycle* dan metode pengklasifikasian judul menggunakan algoritma *naïve bayes*. Adapun tool yang digunakan untuk memproses data training dan data testing yaitu *rapid miner*. Hasil implementasi Sistem *Text Mining* klasifikasi Judul Skripsi Mahasiswa berjalan dengan baik dan menghasilkan dua bentuk klasifikasi berdasarkan konsentrasi yaitu Y (sesuai) dan T (tidak sesuai). Hasil implementasi mendapatkan 49 judul yang sesuai dengan konsentrasi (Y) dan 26 judul yang tidak sesuai dengan konsentrasi (T) dari total judul sebanyak 75 judul skripsi. Atau sebanyak 65% (Y) dan 35%(T).

Keywords

Text Mining,
Thesis Title
Classification,
Naïve Bayes

ABSTRACT

Classification of thesis titles according to student concentration can automatically be done by applying text mining based on thesis titles. The aim of classifying theses titles is based on student concentration to make it easier for study programs to monitor the suitability of the research road map, based on the concentration of the study program and the research carried out by students. The positive impacts or benefits that study programs get from implementing automatic classification of thesis titles, namely making it easier to develop a student research roadmap at the study program level and always improving the quality of study program graduates through student research results. The system development method uses the system development life cycle and the title classification method uses the naïve Bayes algorithm. The tool used to process training data and testing data is rapid miner. The results of the implementation of the Text Mining System for classifying Student Thesis Titles went well and produced two forms of classification based on concentration, namely Y (suitable) and T (not suitable). The implementation results obtained 49 titles that were in accordance with the concentration (Y) and 26 titles that were not in accordance with the concentration (T) out of a total of 75 thesis titles. Or as much as 65% (Y) and 35% (T)

---Jurnal JISTI @ 2023---



PENDAHULUAN

Jumlah data *text* dilihat dari perkembangan data sudah mencapai jumlah yang sangat besar. Hal ini karena industri teknologi informasi yang berkembang pesat, yang terdiri dari data *text* internal yang terdiri dari media online seperti blog, situs berita online. Situs jejaring sosial merupakan sumber data teks yang sangat potensial untuk dianalisis secara lebih mendalam. Namun, data yang berasal dari teks memiliki karakteristik yang sangat tidak terstruktur dan bertumpuk. Oleh Karena itu dibutuhkan sebuah metode untuk menganalisa dengan mudah agar mendapatkan pengelompokan data yang terstruktur.

Metode penambangan teks merupakan suatu cara mempelajari data teks untuk penyusunan dan pengelompokan teks yang terstruktur. Penambangan teks memiliki keunggulan penting di bidang penambangan data. Dengan menggunakan prosedur penambangan teks, dimungkinkan untuk mengekstraksi data, informasi, dan wawasan yang berpotensi berguna dari data tekstual.

Penambangan teks adalah salah satu proses yang dapat dilakukan dengan *text mining*. Penambangan teks dapat didefinisikan sebagai proses menempatkan dokumen teks ke dalam kelas tertentu. Beberapa algoritma dapat digunakan untuk melakukan proses klasifikasi teks, antara lain *Support Vector Machines (SVM)*, *Naive Bayes*, *K-Nearest Neighbors (KNN)*, *Decision Tree*, *Artificial Neural Networks (ANN)*, dan *K-Means*.

Berbagai bidang telah menerapkan *text mining* untuk mengolah data lebih lanjut untuk mempermudah pengolahan informasi, salah satunya perguruan tinggi. Penerapan *text mining* pada perguruan tinggi dapat dimanfaatkan dalam prediksi kelulusan mahasiswa, klasifikasi mahasiswa berprestasi, pengelompokan peminatan konsentrasi mahasiswa, prediksi penerimaan mahasiswa baru, dan pengelompokan judul skripsi mahasiswa.

Proses pengelompokan judul penelitian mahasiswa (Skripsi) yang terjadi selama ini di program studi Teknik Informatika Universitas Lamappapoleonro masih dilakukan dengan konvensional, pengelompokan judul skripsi mahasiswa hanya dilakukan berdasarkan konsentrasi dan dilakukan dengan cara mahasiswa mengumpulkan judul skripsi pada program studi kemudian ketua program studi memvalidasi judul serta mengelompok judul mahasiswa berdasarkan konsentrasi yang ditulis dalam formulir pengajuan judul skripsi.

Cara ini dipandang masih kurang membantu program studi untuk mengelompokkan judul skripsi mahasiswa berdasarkan konsentrasi mahasiswa antara konsentrasi jaringan atau konsentrasi rekayasa perangkat lunak. Untuk klasifikasi judul skripsi mahasiswa sesuai konsentrasi mahasiswa secara otomatis dapat dilakukan dengan cara penerapan *text mining* berdasarkan judul skripsi. Hal tersebut dilakukan agar diharapkan dapat membantu dalam memberikan gambaran kepada program studi untuk menentukan tema judul skripsi yang relevan dengan bidang konsentrasi mahasiswa tersebut. Ada beberapa bidang pengelompokan judul skripsi yang dapat dilakukan di program studi Teknik Informatika Universitas Lamappapoleonro adalah, bidang web programming, desktop programming, mobile programming, artificial Intelligence, Jaringan Komputer.

Tujuan dilakukan klasifikasi judul skripsi berdasarkan konsentrasi mahasiswa untuk memudahkan program studi memonitoring kesesuaian road map penelitian, tema unggulan penelitian program studi dengan riset yang akan dilakukan mahasiswa. Adapun dampak positif atau keuntungan yang didapatkan program studi dalam menerapkan klasifikasi judul skripsi secara otomatis yaitu memudahkan mengembangkan roadmap penelitian mahasiswa ditingkatkan program studi serta Selalu meningkatkan kualitas lulusan prodi melalui hasil penelitian mahasiswa.

KAJIAN PUSTAKA

1. Pengertian Aplikasi.

Aplikasi adalah perangkat lunak yang dapat digunakan untuk mengolah data, masalah atau pekerjaan pada media dan menerapkannya dalam format baru. Fungsi aplikasi secara khusus dan terpadu sesuai kemampuan yang dimilikinya aplikasi merupakan suatu perangkat komputer yang siap pakai bagi user (S. A. Nursakti, 2023).



Seperti yang dipahami, aplikasi adalah program siap pakai yang dirancang untuk menjalankan fungsi layanan aplikasi yang menghadap pengguna dan dirancang untuk menggunakan aplikasi lain yang tersedia untuk penerimanya. Aplikasi masuk akal. Artinya, gunakan salah satu teknik pengolahan data aplikasi untuk mengatasi masalah tersebut. Biasanya berdasarkan perhitungan yang diharapkan atau diharapkan dan pemrosesan data yang diharapkan (Aksa, 2023).

Dari beberapa pengertian aplikasi diatas, dapat disimpulkan aplikasi adalah program siap pakai atau program yang direka untuk melaksanakan suatu fungsi bagi pengguna atau aplikasi yang lain.

2. Pengertian Penerapan

Implementasi adalah tindakan menerapkan teori, metode, dll. ke dalam praktik untuk tujuan atau minat tertentu yang diinginkan oleh kelompok atau kelas yang telah direncanakan dan dikoordinasikan sebelumnya (Ismail, 2020).

Implementasi (Implementation) merupakan perpanjangan dari kegiatan mengkoordinasikan proses interaksi antara tujuan dan sarana untuk mencapainya, membutuhkan jaringan pelaksana, birokrasi yang efektif (Suherman, 2021).

3. Text Mining

Penambangan teks (Text Mining) adalah bidang baru yang berupaya mengekstraksi informasi bermakna dari teks bahasa alami. Bidang ini dapat lebih dikenal sebagai proses analisis teks untuk mengekstraksi informasi yang berguna untuk tujuan tertentu. Dibandingkan dengan jenis data yang disimpan dalam basis data, penambangan teks menggunakan data teks yang tidak terstruktur, tidak memiliki format yang terdefinisi dengan baik, dan sulit diuraikan oleh algoritme.

Text mining adalah proses mendapatkan informasi dari teks. Informasi yang diperoleh berupa pola dan trend pembelajaran dalam pola statistik. Penambangan teks adalah bidang khusus penambangan data.

Tujuan dari *text mining* adalah untuk mengekstrak informasi yang berguna dari sekumpulan dokumen, tetapi tujuan utama dari Penambangan data adalah untuk mendukung proses penemuan pengetahuan dari kumpulan dokumen yang besar.

4. Klasifikasi Data

Klasifikasi adalah tugas mengevaluasi objek data sebagai milik kelas tertentu dari banyak kelas yang ada. Klasifikasi dibagi menjadi dua tugas utama. Yaitu, membuat model sebagai prototipe yang disimpan sebagai penyimpanan, dan menggunakan model tersebut untuk mengklasifikasikan prediksi objek data lain sehingga kita tahu di kelas mana objek data.

5. Pengertian Skripsi

Tugas Akhir (Skripsi) adalah karya akademik di mana seorang mahasiswa sarjana mendiskusikan topik atau bidang tertentu berdasarkan hasil penelitian literatur profesional, penelitian lapangan, atau pengembangan. Saat mengerjakan skripsi, mahasiswa dibimbing oleh minimal dua dosen pembimbing yang ditunjuk oleh universitas masing-masing. Petunjuk ini dimaksudkan untuk memastikan bahwa hasil tugas akhir (Skripsi) mahasiswa berkualitas tinggi, baik dari segi isi maupun pelaksanaannya (Widigda et al., 2018).

Tugas Akhir (Skripsi) adalah karya akademik berdasarkan hasil penelitian lapangan dan literatur yang disusun oleh mahasiswa sesuai dengan mata kuliah yang dianggap sebagai mata kuliah akhir studi formal pada program studi masing-masing.. Sementara proposal skripsi adalah usulan penelitian yang disusun dan disiapkan sedemikian rupa sebelum melakukan penelitian dan penulisan skripsi (Wibawa & F., 2017)



6. Metode Naïve Bayes

Naive bayes merupakan metode pengklasifikasian paling populer digunakan dengan tingkat keakuratan yang baik. Banyak penelitian tentang pengklasifikasian yang telah dilakukan dengan menggunakan algoritma ini. Berbeda dengan metode pengklasifikasian dengan logistic regression ordinal maupun nominal, pada algoritma naive bayes pengklasifikasian tidak membutuhkan adanya pemodelan maupun uji statistik (Amin Abdullah Sidiq & Febrian Wahyu Christanto, 2020).

Naive bayes merupakan metode pengklasifikasian berdasarkan probabilitas sederhana dan dirancang agar dapat dipergunakan dengan asumsi antar variabel penjelas saling bebas (independen). Pada algoritma ini pembelajaran lebih ditekankan pada pengestimasi probabilitas. Keuntungan algoritma naive bayes adalah tingkat nilai error yang didapat lebih rendah ketika dataset berjumlah besar, selain itu akurasi naive bayes dan kecepatannya lebih tinggi pada saat diaplikasikan ke dalam dataset yang jumlahnya lebih besar (Wijaya & Dwiasnati, 2020).

Teorema Bayes dinyatakan secara matematis dalam persamaan berikut (Wijaya & Dwiasnati, 2020):

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Dimana $P(B) \neq 0$

- Pada dasarnya, kita mencoba mencari peluang kejadian A, apabila kejadian B bernilai benar. Kejadian B juga disebut sebagai bukti.
- $P(A)$ adalah apriori dari A (probabilitas sebelumnya, yaitu probabilitas peristiwa sebelum bukti terlihat). Bukti adalah nilai atribut dari instance yang tidak diketahui (peristiwa B).
- $P(A|B)$ adalah probabilitas posteriori dari B, yaitu probabilitas kejadian setelah bukti terlihat.

METODE PENELITIAN

1. Teknik Pengumpulan Data

a. Observasi

Observasi yaitu Dengan melakukan pengamatan langsung pada program studi teknik informatika Universitas Lamappapoleonro. Dari hasil observasi tersebut, kami dapat mengekstraksi data yang akan dijadikan sebagai acuan untuk perancangan dan pembuatan sistem dan material saat membuat laporan penelitian nantinya. Adapun yang akan diobservasi proses pengelompokan judul skripsi dibagian prodi teknik informatika dan berbagai tema judul skripsi di program studi yang ditawarkan.

b. Wawancara

Wawancara (interview) yaitu mengadakan dan melakukan tanya jawab dengan ketua program studi yang terkait dengan objek atau masalah yang sedang diteliti. Adapun yang akan diwawancarai nantinya adalah ketua program studi teknik informatika. Wawancara dilakukan bertujuan untuk mengetahui proses pengelompokan judul skripsi berdasarkan tema penelitian prodi Studi Literatur

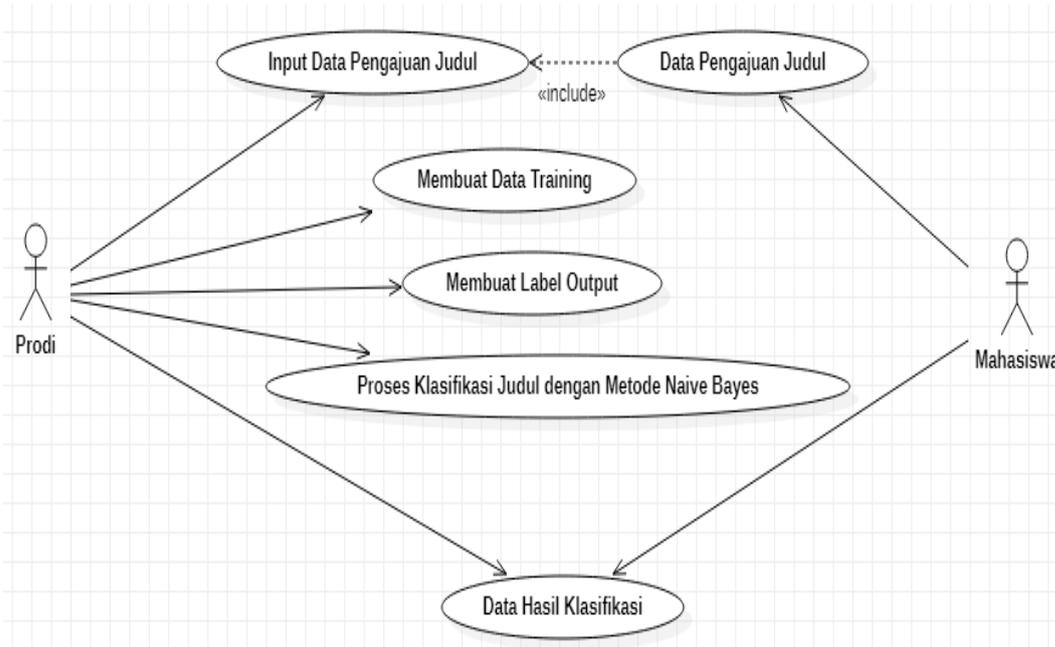
c. Survei literatur

Survei literatur mendapatkan data dari berbagai sumber bacaan yang berkaitan dengan judul penelitian



2. Perancangan Sistem

Upayah-upayah yang diambil dalam merancang sistem ini membuat saran logis dan lainnya untuk pemecahan masalah. Berikut ini adalah blok use case Sistem penerapan text mining untuk klasifikasi judul skripsi mahasiswa. :



Gambar 1 : Use Case Sistem baru

Pada gambar diatas merupakan sistem yang diusulkan, dimana mahasiswa melakukan pengajuan judul kepada prodi. Selanjutnya prodi akan memasukkan data judul skripsi dalam file excel kemudian data judul akan dilakukan proses data training dengan label dan kategori. Setelah itu dilakukan pre-processing dilanjutkan proses klasifikasi judul dengan metode Naive Bayes. Selanjutnya menampilkan hasil klasifikasi judul.

HASIL DAN PEMBAHASAN

1. Implementasi Sistem

1) Pembuatan Dataset

Dataset adalah elemen yang penting dari seluruh proses data analytics. Dataset yang terstruktur dengan baik akan sangat membantu hasil data analysis yang dilakukan, membuat jadi lebih cepat, efisien, dan akurat. Berikut dataset klasifikasi judul skripsi mahasiswa S1 prodi Teknik Informatika Tahun 2022:

Tabel 1: Dataset Judul Skripsi Mahasiswa

No	Nama Mahasiswa	Judul	Konsentrasi
1	Ria Fajriani	Perancangan Robot Pembersih Otomatis Menggunakan Ardiuno	Rpl
2	Nurul Hikmah	Prototipe Pengusir Hama Burung Berbasis Internet Of Things (Iot)	Rpl
3	Nurdin	Rancang Bangun System Pendeteksi Kebakaran Berbasis Iot Menggunakan Ardiuno	Jaringan Komputer

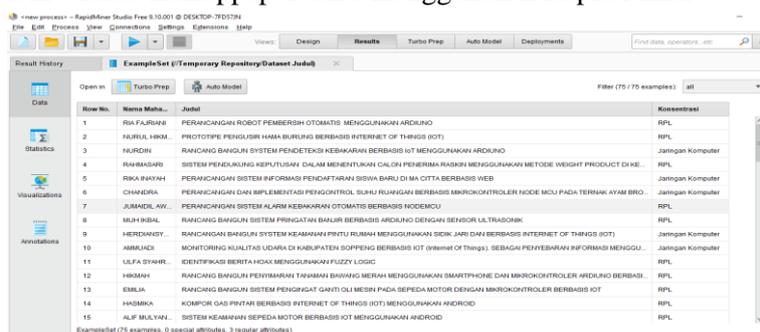


4	Rahmasari	Sistem Pendukung Keputusan Dalam Menentukan Calon Penerima Raskin Menggunakan Metode Weight Product Di Kelurahan Cabenge	Rpl
5	Rika Inayah	Perancangan Sistem Informasi Pendaftaran Siswa Baru Di Ma Citta Berbasis Web	Jaringan Komputer
6	Chandra	Perancangan Dan Implementasi Pengontrol Suhu Ruangan Berbasis Mikrokontroler Node Mcu Pada Ternak Ayam Broiler	Jaringan Komputer
7	Jumaidil Awal	Perancangan Sistem Alarm Kebakaran Otomatis Berbasis Nodemcu	Rpl
8	Muh Iqbal	Rancang Bangun Sistem Pringatan Banjir Berbasis Ardiuno Dengan Sensor Ultrasonik	Rpl
9	Herdiansyah	Rancangan Bangun System Keamanan Pintu Rumah Menggunakan Sidik Jari Dan Berbasis Internet Of Things (Iot)	Jaringan Komputer
10	Ammuadi	Monitoring Kualitas Udara Di Kabupaten Soppeng Berbasis Iot (Internet Of Things). Sebagai Penyebaran Informasi Menggunakan Metode Fuzzy Logic.	Jaringan Komputer
11	Ulfa Syahrani	Identifikasi Berita Hoax Menggunakan Fuzzy Logic	Rpl
12	Hikmah	Rancang Bangun Penyiraman Tanaman Bawang Merah Menggunakan Smartphone Dan Mikrokontroler Ardiuno Berbasis Iot.	Rpl
13	Emilia	Rancang Bangun Sistem Pengingat Ganti Oli Mesin Pada Sepeda Motor Dengan Mikrokontroler Berbasis Iot	Rpl
14	Hasmika	Kompor Gas Pintar Berbasis Internet Of Things (Iot) Menggunakan Android	Rpl
15	Alif Mulyandi	Sistem Keamanan Sepeda Motor Berbasis Iot Menggunakan Android	Rpl

Tabel diatas merupakan dataset judul skripsi mahasiswa Prodi Teknik Informatika Universitas Lamappapoleonro tahun 2022. Jumlah dataset yang digunakan pada penelitian klasifikasi judul skripsi sebanyak 75 data. Dataset merupakan data awal yang diproses pada sistem penerapan text mining. Fungsi dataset untuk mendapatkan model awal dalam klasifikasi. Pada dataset diatas akan dilanjutkan ketahap berikutnya yaitu modeling data dengan membuat label output.

2) Proses Data Training

Training dataset adalah himpunan data yang digunakan untuk melatih atau membangun model. *Text Mining* akan merubah parameter pada dirinya untuk menyesuaikan dengan data yang diberikan saat latihan. Berikut data Training Kasifikasi Judul Skripsi Mahasiswa Teknik Informatika Universitas Lamappapoleonro menggunakan Rapid Miner:





Gambar 2: Proses *Data Training*

Gambar diatas merupakan proses data training klasifikasi judul skripsi mahasiswa teknik informatika Universitas Lamappapoleonro. Data training digunakan untuk melatih data pada klasifikasi judul skripsi. Untuk klasifikasi judul skripsi akan dicocokkan judul skripsi dengan konsentrasi mahasiswa apakah judul mahasiswa sesuai dengan konsentrasi. Untuk melakukan klasifikasi maka selanjutnya dilakukan pembentukan label dalam menentukan klasifikasi.

3) Proses Pembentukan Label Text Mining

Label merupakan target output yang akan dicapai pada proses text mining. Berikut ini adalah label text mining klasifikasi judul skripsi mahasiswa:

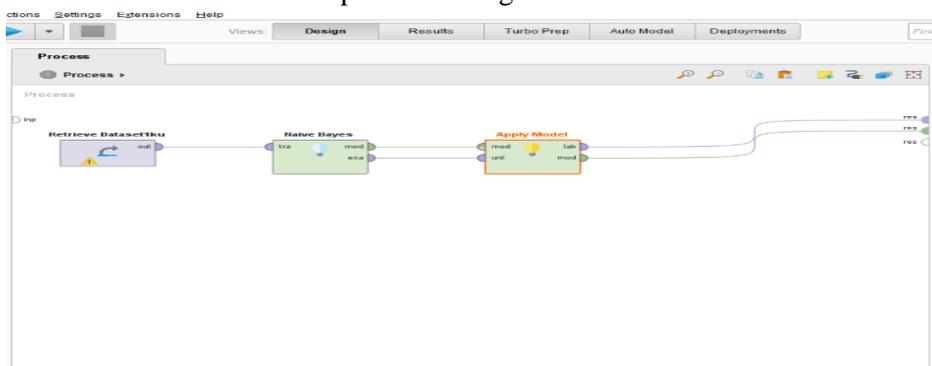
Row No.	Status	Nama Mahasiswa	Judul	Kelompok
1	Y	RIA FAJRIBANI	PERANCANGAN ROBOT PEMBERSIH OTOMATIS MENGGUNAKAN ARDIUNO	RPL
2	T	NURUL HIKMAH	PROTOTYPE PENGUSIR HAMA BURUNG BERBASIS INTERNET OF THINGS (IOT)	RPL
3	Y	NURDIN	RANCANG BANGUN SYSTEM PENDETEKSI KEBAKARAN BERBASIS IOT MENGGUNAKAN ARDIUNO	Jaringan Komputer
4	Y	RAHMASARI	SISTEM PENDUKUNG KEPUTUSAN DALAM MENENTUKAN CALON PENERIMA RASMIN MENGGUNAKAN METODE WEI...	RPL
5	T	RIKA INAYAH	PERANCANGAN SISTEM INFORMASI PEHACFTARAH SISWA BARU DI MA CITTA BERBASIS WEB	Jaringan Komputer
6	Y	CHANDRA	PERANCANGAN DAN IMPLEMENTASI PENGONTROL SUHU RUANGAN BERBASIS MIKROKONTROLER NODE MCU PA...	Jaringan Komputer
7	T	JUMADIL AWAL	PERANCANGAN SISTEM ALARM KEBAKARAN OTOMATIS BERBASIS NODEMCU	RPL
8	T	MUH IKBAL	RANCANG BANGUN SYSTEM PRINGATAN BANJIR BERBASIS ARDIUNO DENGAN SENSOR ULTRASONIK	RPL
9	Y	HERDIANSYAH	RANCANGAN BANGUN SYSTEM KEAMANAN PINTU RUMAH MENGGUNAKAN SIDIK JARI DAN BERBASIS INTERNET OF...	Jaringan Komputer
10	T	AMMADI	MONITORING KUALITAS UDARA DI KABUPATEN SOPPENG BERBASIS IOT (Internet of Things), SEBAGAI PENYEBARAN...	Jaringan Komputer
11	Y	ULFA SYAHRANI	IDENTIFIKASI BERITA HOAX MENGGUNAKAN FUZZY LOGIC	RPL
12	T	HIKMAH	RANCANG BANGUN PENYIARAN TANAMAN BAWANG MERAH MENGGUNAKAN SMARTPHONE DAN MIKROKONTROL...	RPL
13	T	EMILIA	RANCANG BANGUN SISTEM PENGINGAT GANTI OLI MESIN PADA SEPEDA MOTOR DENGAN MIKROKONTROLER BER...	RPL
14	T	HASMIKA	KOMPOR GAS PINTAR BERBASIS INTERNET OF THINGS (IOT) MENGGUNAKAN ANDROID	RPL
15	T	ALIF MULYANDI	SISTEM KEAMANAN SEPEDA MOTOR BERBASIS IOT MENGGUNAKAN ANDROID	RPL

Gambar 3: Proses Pembentukan Label Output

Gambar diatas merupakan proses pembuatan label output. Label merupakan target output yang akan dicapai pada proses text mining. Pada label diatas ditentukan dengan kolom status. Label status akan terisi sesuai dengan hasil training klasifikasi judul dengan konsentrasi atau kelompok. Proses labeling merupakan pencocokan judul skripsi dengan konsentrasi mahasiswa sehingga hasil dari kolom label satatus akan terisi dengan “Y” (ya) atau “T” (tidak),artinya judul skripsi ini sesuai konsentrasi atau tidak.

4) Pemodelan Algoritma untuk Proses Data Training.

Pemodelan algoritma merupakan tahap proses klasifikasi dengan menggunakan algoritma tertentu. Pada riset ini menggunakan algoritma naïve bayes untuk klasifikasi judul skripsi mahasiswa. Berikut ini adalah hasil pemodelan algoritma:



Gambar 4: Pemodelan Algoritma Klasifikasi menggunakan Naïve Bayes

Gambar diatas merupakan proses pemodelan algoritma untuk klasifikasi judul skripsi. Algoritma yang digunakan adalah naïve bayes. Adapun prosesnya yaitu data training dimasukkan kedalam editor view kemudian dihubungkan pada library algoritma dan modeling naïve bayes kemudian dihubungkan pada titik output.



5) Pembuatan Data Testing

Pembuatan data testing merupakan proses pengujian antara output data training dengan hasil data testing. Data testing dilakukan dengan menggunakan modeling naïve bayes. Berikut hasil data testing:

Row No.	Status	prediction(Status)	confidence(Y)	confidence(T)	Nama Mahasiswa	Judul	Kelompok
1	Y	Y	1,000	0,000	RIA FAJRIANI	PERANCANGAN ROBOT PEMBERSIH OTOMATIS MENGGUNAKAN ARDIUNO	RPL
2	T	T	0,000	1,000	NURUL HIKMAH	PROTOTYPE PENGUSIR HAMA BURUNG BERBASIS INTERNET OF THINGS (IOT)	RPL
3	Y	Y	1,000	0,000	NURDIN	RANCANG BANGUN SYSTEM PENDETEKSI KEBAKARAN BERBASIS IoT MENGGUNAKAN ARDIUNO	Jaringan Komputer
4	Y	Y	1,000	0,000	RAHMASARI	SISTEM PENDUKUNG KEPUTUSAN DALAM MENENTUKAN CALON PENERIMA RASKIN MENGGUNAKAN METODE WEIGHT PRODUCT DI KELURAHAN CABENGE	RPL
5	T	T	0,000	1,000	RIKA INAYAH	PERANCANGAN SISTEM INFORMASI PENDAFTARAN SISWA BARU DI MA CITTA BERBASIS WEB PERANCANGAN DAN IMPLEMENTASI PENGONTROL SUHU RUANGAN BERBASIS MIKROKONTROLER NODE MCU PADA TERNAK AYAM BROILER	Jaringan Komputer
6	Y	Y	1,000	0,000	CHANDRA	PERANCANGAN SISTEM ALARM KEBAKARAN OTOMATIS BERBASIS NODEMCU	Jaringan Komputer
7	T	T	0,000	1,000	JUMAIL AWAL	PERANCANGAN SISTEM ALARM KEBAKARAN OTOMATIS BERBASIS NODEMCU	RPL
8	T	T	0,000	1,000	MUH IKBAL	RANCANG BANGUN SYSTEM PRINGATAN BANJIR BERBASIS	RPL
9	Y	Y	1,000	0,000	HERDIANSYAH	RANCANG BANGUN SYSTEM KEAMANAN PINTU RUMAH MENGGUNAKAN ARDIUNO	Jaringan Komputer
10	T	T	0,000	1,000	AMMUDI	MONITORING KUALITAS UDARA DI KABUPATEN SOPPING BERBASIS ARDIUNO	Jaringan Komputer
11	Y	Y	1,000	0,000	ULFA SYAHRANI	IDENTIFIKASI BERITA HOAX MENGGUNAKAN FUZZY LOGIC	RPL
12	T	T	0,000	1,000	HIKMAH	RANCANG BANGUN PENYIARAN TANAMAN BAWANG MERAH MENGGUNAKAN ARDIUNO	RPL
13	T	T	0,000	1,000	EMILIA	RANCANG BANGUN SYSTEM PENGINGAT GANTI OLI MESIN PADA MOTOR	RPL
14	T	T	0,000	1,000	HASMIKA	KOMPOR GAS PINTAR BERBASIS INTERNET OF THINGS (IoT) MENGGUNAKAN ARDIUNO	RPL

Gambar 5: Data Testing menggunakan Algoritma Naïve Bayes

Gambar diatas merupakan proses pembuatan data testing menggunakan algoritma naïve bayes. Dari hasil yang dilakukan, antara data training dan data testing 100% menghasilkan yang sama dan menghasilkan confidence (Y) dan confidence (T). hasil kalsifikasi dapat dilihat pada tabel 4.2 dibawah ini.

Tabel 2 : Data testing algoritma naïve bayes

N0	Stat us	pre dict	Confi (Y)	Confi (T)	Nama Mahasiswa	Judul	Konsentr asi
1	Y	Y	1,000	0,000	RIA FAJRIANI	PERANCANGAN ROBOT PEMBERSIH OTOMATIS MENGGUNAKAN ARDIUNO	RPL
2	T	T	0,000	1,000	NURUL HIKMAH	PROTOTYPE PENGUSIR HAMA BURUNG BERBASIS INTERNET OF THINGS (IOT)	RPL
3	Y	Y	1,000	0,000	NURDIN	RANCANG BANGUN SYSTEM PENDETEKSI KEBAKARAN BERBASIS IoT MENGGUNAKAN ARDIUNO	Jaringan Komputer
4	Y	Y	1,000	0,000	RAHMASARI	SISTEM PENDUKUNG KEPUTUSAN DALAM MENENTUKAN CALON PENERIMA RASKIN MENGGUNAKAN METODE WEIGHT PRODUCT DI KELURAHAN CABENGE	RPL
5	T	T	0,000	1,000	RIKA INAYAH	PERANCANGAN SISTEM INFORMASI PENDAFTARAN SISWA BARU DI MA CITTA BERBASIS WEB PERANCANGAN DAN IMPLEMENTASI PENGONTROL SUHU RUANGAN BERBASIS MIKROKONTROLER NODE MCU PADA TERNAK AYAM BROILER	Jaringan Komputer
6	Y	Y	1,000	0,000	CHANDRA	PERANCANGAN SISTEM ALARM KEBAKARAN OTOMATIS BERBASIS NODEMCU	Jaringan Komputer
7	Y	Y	1,000	0,000	JUMAIL AWAL	PERANCANGAN SISTEM ALARM KEBAKARAN OTOMATIS BERBASIS NODEMCU	RPL
8	T	T	0,000	1,000	MUH IKBAL	RANCANG BANGUN SYSTEM PRINGATAN BANJIR BERBASIS	RPL



						ARDIUNO DENGAN SENSOR ULTRASONIK RANCANGAN BANGUN SYSTEM KEAMANAN PINTU RUMAH MENGUNAKAN SIDIK JARI DAN BERBASIS INTERNET OF THINGS (IOT)	Jaringan Komputer
9	T	T	0,000	1,000	HERDIANS YAH		
10	T	T	0,000	1,000	AMMUADI	MONITORING KUALITAS UDARA DI KABUPATEN SOPPENG BERBASIS IOT (Internet Of Things). SEBAGAI PENYEBARAN INFORMASI MENGUNAKAN METODE FUZZY LOGIC.	Jaringan Komputer
11	Y	Y	1,000	0,000	ULFA SYAHRAN I	IDENTIFIKASI BERITA HOAX MENGUNAKAN FUZZY LOGIC	RPL
12	T	T	0,000	1,000	HIKMAH	RANCANG BANGUN PENYIMARAN TANAMAN BAWANG MERAH MENGUNAKAN SMARTPHONE DAN MIKROKONTROLER ARDIUNO BERBASIS IOT.	RPL
13	T	T	0,000	1,000	EMILIA	RANCANG BANGUN SISTEM PENGINGAT GANTI OLI MESIN PADA SEPEDA MOTOR DENGAN MIKROKONTROLER BERBASIS IOT	RPL
14	T	T	0,000	1,000	HASMIKA	KOMPOR GAS PINTAR BERBASIS INTERNET OF THINGS (IOT) MENGUNAKAN ANDROID	RPL
15	Y	Y	1,000	0,000	ALIF MULYAN DI	SISTEM KEAMANAN SEPEDA MOTOR BERBASIS IOT MENGUNAKAN ANDROID	RPL

6) Hasil Klasifikasi Judul

Hasil kasifikasi judul dapat dilihat dalam bentuk grafik maupun tabel. Dari hasil kasifikasi terdapat status sesuai judul dengan konsentrasi dan ada yang tidak sesuai judul dengan konsentrasi. Berikut hasil kalsfikasi judul Skripsi Mahasiswa Teknik Informatika Universitas Lamappapoleonro Tahun 2022:

Attribute	Parameter	Y	T
Nama Mahasiswa	value=RIA FAJRIANI	0.020	0.000
Nama Mahasiswa	value=NURUL HIKMAH	0.000	0.038
Nama Mahasiswa	value=NURDIN	0.020	0.000
Nama Mahasiswa	value=RAHMASARI	0.020	0.000
Nama Mahasiswa	value=RIKA IHAYAH	0.000	0.038
Nama Mahasiswa	value=CHANDRA	0.020	0.000
Nama Mahasiswa	value=JUMAIL AWAL	0.000	0.038
Nama Mahasiswa	value=MUH IKBAL	0.000	0.038
Nama Mahasiswa	value=HERDIANSYAH	0.020	0.000
Nama Mahasiswa	value=AMMUADI	0.000	0.038
Nama Mahasiswa	value=ULFA SYAHRANI	0.020	0.000
Nama Mahasiswa	value=HIKMAH	0.000	0.038
Nama Mahasiswa	value=EMILIA	0.000	0.038
Nama Mahasiswa	value=HASMIKA	0.000	0.038
Nama Mahasiswa	value=ALIF MULYANDI	0.000	0.038

Gambar 6: Hasil Klasifikasi Judul dalam Bentuk Tabel

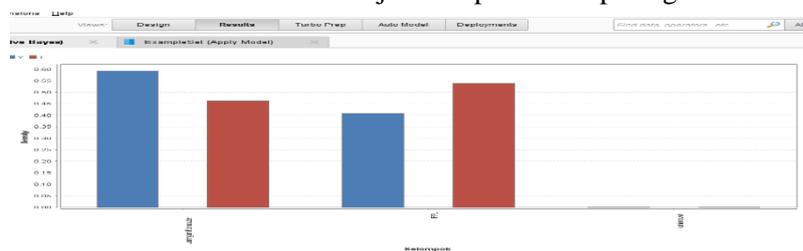
Gambar diatas merupakan hasil klasifikasi judul Skripsi Mahasiswa dalam bentuk tabel yang dapat dilihat pada tabel dibawah ini.



Tabel 3: Dataset Judul Skripsi Mahasiswa

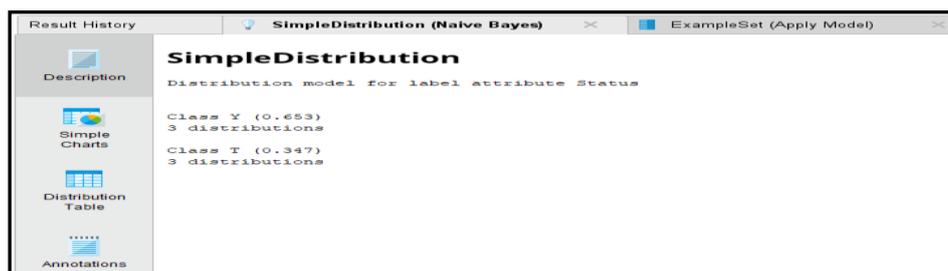
NO	Atitude	Parameter	confidence(Y)	confidence(T)
1	RIA FAJRIANI	value=RIA FAJRIANI	0,020	0,000
2	NURUL HIKMAH	value=NURUL HIKMAH	0,000	0,020
3	NURDIN	value=NURDIN	0,020	0,000
4	RAHMASARI	value=RAHMASARI	0,020	0,000
5	RIKA INAYAH	value=RIKA INAYAH	0,000	0,038
6	CHANDRA	value=CHANDRA	0,020	0,000
7	JUMAIDIL AWAL	value=JUMAIDIL AWAL	0,000	0,038
8	MUH IKBAL	value=MUH IKBAL	0,000	0,038
9	HERDIANSYAH	value=HERDIANSYAH	0,020	0,000
10	AMMUADI	value=AMMUADI	0,020	0,000
11	ULFA SYAHRANI	value=ULFA SYAHRANI	0,000	0,038
12	HIKMAH	value=HIKMAH	0,000	0,038
13	EMILIA	value=EMILIA	0,000	0,038
14	HASMIKA	value=HASMIKA	0,000	0,038
15	ALIF MULYANDI	value=ALIF MULYANDI	0,020	0,000

Hasil klasifikasi terdapat dua status yaitu Y (ya) dan T(tidak). Dari 15 data judul skripsi mahasiswa yang diklasifikasi, mendapatkan hasil sebanyak 49 mahasiswa yang sesuai antara judul dengan konsentrasi dan sebanyak 26 mahasiswa yang tidak sesuai judul dengan konsentrasi. Adapun hasil visualisasi data klasifikasi judul dapat dilihat pada gambar berikut:



Gambar 7: Hasil Klasifikasi Judul dalam Bentuk Grafik

Gambar diatas merupakan hasil klasifikasi judul Skripsi Mahasiswa dalam bentuk tabel. Hasil klasifikasi terdapat dua status yaitu Y (ya) dan T(tidak). Dari 75 data judul skripsi mahasiswa yang diklasifikasi, mendapatkan hasil sebanyak 49 mahasiswa (65%) yang sesuai antara judul dengan konsentrasi dan sebanyak 26 mahasiswa (35%) yang tida sesuai judul dengan konsentrasi. Berikut hasil distribusi klasifikasi judul Skripsi Mahasiswa Prodi Teknik Informatika Universitas Lamappapoleonro:





Gambar 8: Distribusi Hasil Klasifikasi Judul Skripsi

Gambar diatas merupakan hasil distribusi klasifikasi judul skripsi Mahasiswa Prodi Teknik Informatika Universitas Lamappapoleonro Tahun 2022. Hasil yang didapatkan sebanyak 65% judul sesuai dengan Konsentrasi dan 35% judul tidak sesuai dengan konsentrasi.

4.2 Pengujian Sistem

Pengujian menyajikan anomali yang menarik bagi perekayasa perangkat lunak. Pada proses perangkat lunak, perekayasa pertama-tama berusaha membangun perangkat lunak dari konsep abstrak ke implementasi yang dapat dilihat, baru kemudian dilakukan pengujian. Metode pengujian yang dilakukan yakni pengujian *black-box*.

1) Pengujian Cleaning Data

Tabel 4: Pengujian Cleaning Data

Data masuk	Yang diharapkan	Pengamatan	Kesimpulan
Import Data	Menampilkan data missing pada saat import data training	Data missing kelihatan pada tabel data training	sesuai

Screen Shoot

No integer	NIM integer	Nama Mahasis... polynominal	Judul polynominal	Konsentrasi polynominal
1	1904044	RIA FAJRIANI	PERANCANGAN ROB...	RPL
2	1904034	NURUL HIKMAH	PROTOTYPE PENGU...	RPL
3	?	?	?	?
4	1904012	NURDIN	RANCANG BANGUN ...	Jaringan Komputer
5	?	?	?	?
6	1904022	RAHMASARI	SISTEM PENDUKUN...	RPL
7	?	?	?	?
8	1904046	RIKA INAYAH	PERANCANGAN SIST...	Jaringan Komputer
9	1904027	CHANDRA	PERANCANGAN DAN...	Jaringan Komputer
10	?	?	?	?
11	1904033	JUMAILIL AWAL	PERANCANGAN SIST...	RPL
12	1904025	MUH IKBAL	RANCANG BANGUN ...	RPL

2) Pengujian Konsistensi Data

Tabel 4. 3 Pengujian Konsistensi Data Training

Data masuk	Yang diharapkan	Pengamatan	Kesimpulan
Import Data	Tidak Ada data missing pada saat import data training	Sukses, data konsisten	sesuai

Screen Shoot

Nama Mahasiswa polynominal	Judul polynominal	Kelompok polynominal
1 RIA FAJRIANI	PERANCANGAN ROBOT PEMBERSEH ...	RPL
2 NURUL HIKMAH	PROTOTYPE PENGUSIR HAMA BURUN...	RPL
3 NURDIN	RANCANG BANGUN SYSTEM PENDET...	Jaringan Komputer
4 RAHMASARI	SISTEM PENDUKUNG KEPUTUSAN D...	RPL
5 RIKA INAYAH	PERANCANGAN SISTEM INFORMASI P...	Jaringan Komputer
6 CHANDRA	PERANCANGAN DAN IMPLEMENTASI ...	Jaringan Komputer
7 JUMAILIL AWAL	PERANCANGAN SISTEM ALARM KEBAB...	RPL
8 MUH IKBAL	RANCANG BANGUN SISTEM PRINGAT...	RPL
9 HERDIANSYAH	RANCANGAN BANGUN SYSTEM KEAM...	Jaringan Komputer
10 AMMUADI	MONITORING KUALITAS UDARA DI KA...	Jaringan Komputer
11 ULFA SYAHRANI	IDENTIFIKASI BERITA HOAX MENGGU...	RPL
12 HIKMAH	RANCANG BANGUN PENYIMARAN TAN...	RPL



3) Pengujian Hasil Klasifikasi Data Training dengan Data Testing

Tabel 4. 4 Pengujian Data Training dan Data Testing

Data masuk	Yang diharapkan	Pengamatan	Kesimpulan
Proses Klasifikasi	Tingkat Kesamaan dengan Training dan Testing	Data Sukses, 100% Sama dengan Data	Sesuai

Screen Shoot

Row No.	Status	prediction(Status)	confidence(Y)	confidence(T)	Nama Maha...	Judul	Kelompok
1	Y	Y	1.000	0.000	RIA FAJRIANI	PERANCANGAN ROBOT PEMBERSIH OTOMATIS MENGGUNAKAN ...	RPL
2	T	T	0.000	1.000	NURUL HIKM...	PROTOTYPE PENGUSIR HAMA BURUNG BERBASIS INTERNET OF ...	RPL
3	Y	Y	1.000	0.000	NURDIN	RANCANG BANGUN SYSTEM PENDETEKSI KEBAKARAN BERBASIS ...	Jaringan Komputer
4	Y	Y	1.000	0.000	RAHMASARI	SISTEM PENDUKUNG KEPUTUSAN DALAM MENENTUKAN CALO...	RPL
5	T	T	0.000	1.000	RIKA INAYAH	PERANCANGAN SISTEM INFORMASI PENDAFTARAN SISWA BARU ...	Jaringan Komputer
6	Y	Y	1.000	0.000	CHANDRA	PERANCANGAN DAN IMPLEMENTASI PENGONTROL SUHU RUAN...	Jaringan Komputer
7	T	T	0.000	1.000	JUMAILDIL AW...	PERANCANGAN SISTEM ALARM KEBAKARAN OTOMATIS BERBASIS...	RPL
8	T	T	0.000	1.000	MUH IKBAL	RANCANG BANGUN SISTEM PRINGATAN BANJIR BERBASIS ARDI...	RPL
9	Y	Y	1.000	0.000	HERDIANSY...	RANCANGAN BANGUN SYSTEM KEAMANAN PINTU RUMAH MENG...	Jaringan Komputer
10	T	T	0.000	1.000	AMMUADI	MONITORING KUALITAS UDARA DI KABUPATEN SOPPENG BERBA...	Jaringan Komputer
11	Y	Y	1.000	0.000	ULFA SYAHR...	IDENTIFIKASI BERITA HOAX MENGGUNAKAN FUZZY LOGIC	RPL
12	T	T	0.000	1.000	HIKMAH	RANCANG BANGUN PENYIMARAN TANAMAN BAWANG MERAH ME...	RPL
13	T	T	0.000	1.000	EMILIA	RANCANG BANGUN SISTEM PENGINGAT GANTI OLI MESIN PADA ...	RPL
14	T	T	0.000	1.000	HASMIKA	KOMPOR GAS PINTAR BERBASIS INTERNET OF THINGS (IOT) ME...	RPL
15	T	T	0.000	1.000	ALIF MULYAN...	SISTEM KEAMANAN SEPEDA MOTOR BERBASIS IOT MENGGUNAK...	RPL

ExampleSet (75 examples, 4 special attributes, 3 regular attributes)

SIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan tentang penerapan text mining kalsifikasi judul skripsi Mahasiswa Prodi Teknik Informatika Universitas Lamappapoleonro, maka dapat ditarik kesimpulan sebagai berikut:

- Hasil analisis data kasifikasi judul skripsi mahasiswa dilakukan dengan cara pendataan menggunakan excel kemudian dianalisa satu persatu judul untuk pencocokan judul dengan konsentrasi mahasiswa. Hasil pengamatan menyimpulkan memakan waktu yang lama dan kurang akurat.
- Hasil perancangan sistem dengan menggunakan *Text Mining* dapat memudahkan proses klasifikasi usulan judul skripsi mahasiswa. Perancangan sistem dapat membentuk kelompok data yang konsisten sehingga memudahkan untuk diolah lebih lanjut dalam klasifikasi menggunakan algoritma naïve bayes.
- Hasil implementasi Sistem Text Mining klasifikasi usulan Judul Skripsi Mahasiswa berjalan dengan baik dan menghasilkan dua bentuk klasifikasi berdasarkan konsentrasi yaitu Y (sesuai) dan T (tidak sesuai). Hasil implementasi mendapatkan 49 judul yang sesuai dengan konsentrasi (Y) dan 26 judul yang tidak sesuai dengan konsentrasi (T) dari total judul sebanyak 75 judul skripsi. Atau sebanyak 65% (Y) dan 35%(T).

DAFTAR PUSTAKA

Aksa, A. N. (2023). *Perancangan Aplikasi Administrasi Instalasi Gawat Darurat Pada Rumah Sakit Latemmamala Kabupaten Soppeng*. Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI), 6(1), 43–51.

Amin Abdullah Sidiq, & Febrian Wahyu Christanto. (2020). *Algoritma Naive Bayes Untuk Penentuan Pkh (Program Keluarga Harapan) Berbasis Sistem Pendukung Kepu-Tusan (Studi Kasus: Kelurahan Karanganyar Gunung Se-Marang)*. Jurnal Riptek, 14(1), 65–71.

Hermiati, R., Asnawati, & Kanedi, I. (2021). *Pembuatan E-Commerce Pada Raja Komputer Menggunakan Bahasa*. Jurnal Media Infotama, 17(1), 54–66.



- Ismail, Mukhlis, A. (2023). *Sistem Pendukung Keputusan Penentuan Jurusan Menggunakan Metode Multi Factor Evaluation Process (MFEP) di SMAN 5 Soppeng*. Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI), 6(1), 9–19.
- Ismail, S. H. (2020). *Penerapan Metode Topsis Pada Sistem Penerimaan Pegawai Pemerintah Non Pegawai Negeri Di Kantor Pertanahan Kabupaten Soppeng*. Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika “JISTI,” 3(1), 1–12.
- Nursakti, N., & Parwati, H. (2022). *Rancang Bangun Aplikasi Android Untuk Pelayanan Pernikahan Pada Kantor Urusan Agama Kecamatan Lamuru*. Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI), 5(2), 67–76.
- Nursakti, S. A. (2023). *Perancangan Aplikasi Online Shop pada Toko Nuzhly Shop Menggunakan Metode Agile*. Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI), 6(1), 26–33.
- Permatasari, R., Arifin, M., & Padilah, R. (2020). *Studi Deskriptif Dampak Psikologis Mahasiswa Program Studi Bimbingan dan Konseling Universitas PGRI Banyuwangi Dalam Penyusunan Skripsi di Masa Pandemi Covid-19*. Jurnal Bina Ilmu Cendekia, 2(1), 128–141.
- Putra, M. P. R., & Wardani, K. R. N. (2020). *Penerapan Text Mining Dalam Menganalisis Kepribadian Pengguna Media Sosial*. JUTIM (Jurnal Teknik Informatika Musirawas), 5(1), 63–71.
- Ramadhani, S., Azzahra, D., & Z, T. (2022). *Comparison of K-Means and K-Medoids Algorithms in Text Mining based on Davies Bouldin Index Testing for Classification of Student’s Thesis*. Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi, 13(1), 24–33.
- Romli, I., & Putra, B. M. (2021). *Evaluasi Penilaian Kinerja Dalam Klasifikasi Data Mining Dengan Metode Naive Bayes*. Jurnal Teknik Industri, 1(01), 36–45.
- Saharuddin, S., & Wisnu Prihatmono, M. (2022). *Pengenalan Dan Pelatihan Dasar Bahasa Pemrograman Python Pada Siswa/I Sma Negeri 3 Makassar*. SELAPARANG: Jurnal Pengabdian Masyarakat Berkemajuan, 6(4), 22-33.
- Suherman. (2021). *Penerapan Algoritma Simple Multi Attribute Rating Technique Pada Sistem Pendukung Penerimaan Siswa*. Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI), 4(1), 58–67.
- Tahir, M. A. (2022). *Perancangan Sistem Pengarsipan Data Pegawai Kantor Camat Donri-Donri Kabupaten Soppeng Menggunakan Electronical Filling System*. Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI), 5(1), 1–8.
- Utomo, D. P., & Mesran, M. (2020). *Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung*. Jurnal Media Informatika Budidarma, 4(2), 437.
- W, B., Riski, I., Dwi, K., Nooraeni, R., Siahaan, T., & Dhea, Y. (2019). *Analisis Text Mining dari Cuitan Twitter Mengenai Infrastruktur di Indonesia dengan Metode Klasifikasi Naïve Bayes*. Eigen Mathematics Journal, 2(2), 92–101.
- Wibawa, J. C., & F., M. R. (2017). *Pengembangan Sistem Informasi Penjadwalan dan Manajemen Keuangan Kegiatan Seminar dan Sidang Skripsi/Tugas Akhir (Studi Kasus Program Studi Sistem Informasi UNIKOM)*. Jurnal Teknik Informatika Dan Sistem Informasi, 3(1), 150–168.
- Widigda, I. R., Setyaningrum, W., Pascasarjana, J., Matematika, P., Yogyakarta, U. N., & Skripsi, M. (2018). *Kecemasan Mahasiswa Pendidikan Matematika Universitas Riau Kepulauan Dalam Menghadapi Skripsi*. Jurnal Pendidikan Matematika Dan Sains, 6(2), 190–199.
- Wijaya, H. D., & Dwiasnati, S. (2020). *Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat*. Jurnal Informatika, 7(1), 1–7.