

Terakreditasi SINTA Peringkat 4

Surat Keputusan Dirjen Penguatan Riset dan Pengembangan Ristek Dikti No. 28/E/KPT/2019
masa berlaku mulai Vol.3 No. 1 tahun 2018 s.d Vol. 7 No. 1 tahun 2022

Terbit online pada laman web jurnal:
<http://publishing-widyagama.ac.id/ejournal-v2/index.php/jointecs>



Vol. 6 No. 1 (2021) 35 - 42

JOINTECS

(Journal of Information Technology and Computer Science)

e-ISSN:2541-6448

p-ISSN:2541-3619

Sistem Pendukung Keputusan Rekomendasi Topik Skripsi Menggunakan *Naïve Bayes Classifier*

Farid¹, Ultach Enri², Yuyun Umidah³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang

¹farid.16086@student.unsika.ac.id, ²ultach@staff.unsika.ac.id, ³yuyun.umidah@staff.unsika.ac.id

Abstract

Every student is required to carry out an obligation, one of which is in the form of research. As a tangible form of the final process towards undergraduate each student is required to make scientific articles in the form of books which are named thesis. So far, the process of determining student thesis topics is done manually, both thesis supervisors who provide input or ideas are obtained from various research papers. And the process of determining the thesis topic without using a computerized system. Therefore, researchers made this research in order to assist students in determining the thesis topic according to student competencies. This research method uses data mining and software development methods by applying the Naïve Bayes Classifier algorithm to a website-based system. The result of this research is a decision support system that can provide thesis topic recommendations based on the value data of the elective courses. The best accuracy model value implemented in this system is 69.27%. The accuracy value is not good because the amount of data is not balanced in each category of the thesis topic.

Keywords: decision support system; thesis topic recommendations; naïve bayes classifier.

Abstrak

Setiap mahasiswa dituntut untuk melakukan kewajiban, salah satunya berupa penelitian. Sebagai wujud nyata proses akhir menuju sarjana setiap mahasiswa diharuskan membuat artikel ilmiah dalam bentuk buku yang diberi nama skripsi. Selama ini proses menentukan topik skripsi mahasiswa dilakukan secara manual, baik pembimbing skripsi yang memberi masukan atau ide diperoleh dari berbagai makalah penelitian. Dan proses penentuan topik skripsi tanpa menggunakan sistem terkomputerisasi. Maka dari itu peneliti membuat penelitian ini agar dapat membantu Mahasiswa dalam menentukan topik skripsi yang sesuai dengan kompetensi Mahasiswa. Metode penelitian ini menggunakan metode pengembangan *data mining* dan perangkat lunak dengan menerapkan algoritma *Naïve Bayes Classifier* ke sistem berbasis *website*. Hasil dari penelitian ini adalah sistem pendukung keputusan yang dapat memberikan rekomendasi topik skripsi berdasarkan data nilai mata kuliah pilihan. Nilai *accuracy* model terbaik yang diimplementasikan pada sistem ini adalah sebesar 69,27%. Nilai akurasi kurang baik karena jumlah data yang tidak seimbang pada setiap kategori topik skripsi.

Kata kunci: sistem pendukung keputusan; rekomendasi topik skripsi; *naïve bayes classifier*.

© 2021 Jurnal JOINTECS

1. Pendahuluan

Tuntutan akademik pada tiap jenjang pendidikan di Indonesia berbeda-beda[1]. Berdasarkan Tri Dharma Perguruan Tinggi, setiap Mahasiswa dituntut untuk melakukan kewajiban salah satunya berupa penelitian. Sebagai wujud nyata proses akhir menuju sarjana adalah setiap Mahasiswa diharuskan membuat artikel ilmiah dalam bentuk buku yang diberi nama skripsi[2]. Selama ini proses menentukan topik skripsi mahasiswa dilakukan secara manual. Yakni baik pembimbing skripsi yang memberi masukan atau ide diperoleh dari berbagai makalah penelitian[3]. Berdasarkan hasil observasi penulis secara langsung pada Fakultas Ilmu Komputer (FASILKOM) Universitas Singaperbangsa Karawang (UNSIKA), proses penentuan topik skripsi yang telah berjalan masih tradisional dan manual tanpa menggunakan sistem yang terkomputerisasi.

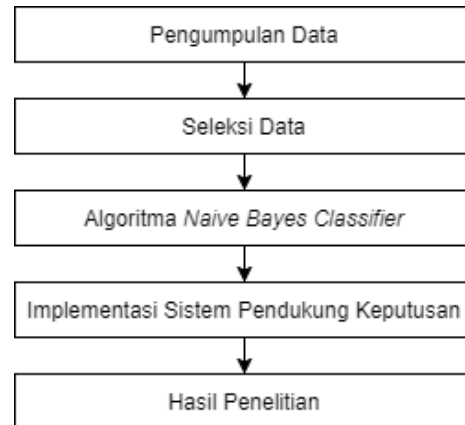
Menentukan judul/topik skripsi termasuk dalam tingkat kesulitan kategori sedang sebagai penyebab terlambatnya seorang Mahasiswa untuk lulus tepat waktu[4]. Selain itu Mahasiswa juga seringkali mengalami kendala dalam menentukan topik skripsi yang akan diambil karena minimnya informasi topik-topik skripsi Mahasiswa terdahulu[5]. Banyaknya mata kuliah yang telah ditempuh, justru membuat Mahasiswa semakin sulit menentukan pilihan pada satu bidang yang akan difokuskan sebagai tugas akhir[6]. Berbagai hal tersebut menemukan bahwa cukup banyak faktor yang menyebabkan Mahasiswa sulit menemukan ide topik skripsi. Berdasarkan hal tersebut, diperlukannya suatu sistem yang dapat melakukan klasifikasi sebagai alat bantu dalam menentukan ide topik skripsi yang sesuai dengan kompetensi Mahasiswa.

Salah satu metode klasifikasi adalah *Naïve Bayes*. *Naïve bayes* merupakan sebuah algoritma yang memiliki kesetaraan dengan algoritma lain seperti *decision tree* dan *neural network* dalam kemampuan klasifikasi[7]. Metode *Naïve Bayes* merupakan salah satu metode klasifikasi yang efektif dan efisien karena proses pengklasifikasiannya bekerja secara independen pada setiap fitur objek yang akan diklasifikasi[8]. Metode *Naïve Bayes Classifier* dapat melakukan klasifikasi terhadap tugas akhir lebih baik daripada *K-Nearest Neighbor* dimana akurasi KNN 84% dan *naïve bayes* 87%[9]. Hal yang telah dipaparkan tersebutlah yang menjadikan latar belakang penulis untuk melakukan penelitian ini, dengan adanya sistem pendukung keputusan akan sangat membantu Mahasiswa dalam mendapatkan ide untuk penentuan topik skripsi.

2. Metode Penelitian

Penelitian ini dilakukan di Fakultas Ilmu Komputer Universitas Singaperbangsa Karawang. Data yang digunakan pada penelitian ini merupakan data akademik yang didapatkan dari pihak FASILKOM UNSIKA. Penelitian ini melakukan proses *data mining*

untuk mendapatkan model yang terbaik dari data yang ada menggunakan algoritma *naïve bayes classifier*. Dan hasil dari pemodelan yang didapat diimplementasikan menjadi sistem pendukung keputusan rekomendasi topik skripsi. Adapun tahap-tahap penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahap Penelitian

Pada Gambar 1 dapat dijelaskan bahwa pada tahapan pertama dilakukan pengumpulan data. Tahapan kedua dengan melakukan seleksi data. Tahapan ketiga dengan melakukan perhitungan *naïve bayes classifier*. Dalam tahapan keempat dilakukan implementasi model *naïve bayes classifier* menjadi sistem pendukung keputusan. Dan tahap terakhir adalah menjabarkan hasil dari penelitian yang dilakukan.

2.1. Pengumpulan Data

Dalam membuat sistem pendukung keputusan untuk menentukan rekomendasi topik skripsi dilakukan pengumpulan data dengan melakukan permintaan data terdahulu ke pihak FASILKOM UNSIKA berdasarkan prosedur yang berlaku. Data yang digunakan adalah data mahasiswa angkatan 2013 hingga 2015. Data yang didapatkan berjumlah 358, dengan terdapat 15 atribut yakni NPM, nama, topik skripsi dan 12 atribut mata kuliah pilihan. Topik skripsi yang ada terdiri dari 8 kategori yakni *augmented reality*, *data mining*, *IOT*, *kriptografi*, *multimedia*, *networking*, rancang bangun dan sistem pakar.

Tabel 1. Tabel Mata Kuliah

| Kode MK | Mata Kuliah |
|---------|-------------------------------|
| IFK606 | Sistem Pakar |
| IFK607 | Komputer Grafik |
| IFK608 | Sistem Tersebar/Terdistribusi |
| IFK609 | <i>Jaringan Wireless</i> |
| IFK610 | ERP |
| IFK611 | <i>Bussines Intelegent</i> |
| IFK707 | Basis Data Multimedia |
| IFK708 | Kriptografi |
| IFK709 | <i>Data Mining</i> |
| IFK710 | Audit IT |
| IFK711 | Teknologi Selular |
| IFK712 | Digital Forensik |

2.2. Seleksi Data

Pada tahap ini dilakukan pemilihan data untuk digunakan dalam proses *data mining*. Seleksi data dilakukan dengan tujuan mendapatkan hasil *data mining* yang spesifik dan akurat, hanya beberapa data yang diambil dari data mentah dan akan digunakan sebagai atribut untuk selanjutnya diproses[10]. Ada 15 atribut yang tersedia, dan peneliti menggunakan 14 atribut yang ada.

Pada Tabel 1 dapat dilihat berbagai mata kuliah pilihan yang ada di FASILKOM UNSIKA, dimana ke-12 data ini merupakan atribut yang digunakan sebagai atribut proses perhitungan *data mining*. Atribut NPM digunakan sebagai atribut identitas, untuk atribut topik skripsi digunakan sebagai atribut *class target*, dan untuk atribut nama sudah terwakilkan oleh atribut NPM sebagai atribut identitas, maka dari itu atribut nama tidak digunakan dalam proses *data mining* penelitian ini. Tabel 2 dibawah ini adalah sebagian data yang akan digunakan dalam proses *data mining*.

Tabel 2. Tabel *Dataset*

| NPM | IFK606 | ... | IFK712 | Mata Kuliah |
|---------|--------|-----|--------|-------------------|
| 13..012 | NA | ... | B | Augmented Reality |
| 13..234 | B | ... | B | Augmented Reality |
| 13..237 | NA | ... | NA | Augmented Reality |
| ... | ... | ... | ... | ... |
| 15..295 | A | ... | NA | Sistem Pakar |
| 15..038 | NA | ... | NA | Sistem Pakar |
| 15..126 | A | ... | NA | Sistem Pakar |

Tabel 2 berisi *dataset* yang telah dilakukan pemilihan data sesuai dengan kebutuhan proses *data mining*. Total *dataset* berjumlah 358 data. Data yang didapat tersebut selanjutnya diproses untuk digunakan dalam pembuatan model.

2.3. Algoritma *Naïve Bayes Classifier*

Naïve Bayes Classifier merupakan salah satu teknik klasifikasi metode probabilistik dan statistik yang dikemukakan oleh Thomas Bayes yang dikenal dengan *Teorema Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya[11]. Klasifikasi *naïve bayes* berhubungan dengan teori probabilitas sederhana, yang merupakan cabang dari matematika probabilitas dapat digunakan untuk menentukan model dengan data yang tidak pasti dengan tujuan dan hasil yang menarik dengan menggabungkan pengetahuan dari hasil eksperimental dan bukti-bukti pengamatan[12]. Rumus umum *Teorema Bayes* dapat dilihat pada rumus 1 berikut ini[13]:

$$P(c|x) = \frac{P(x|c).P(c)}{P(x)} \tag{1}$$

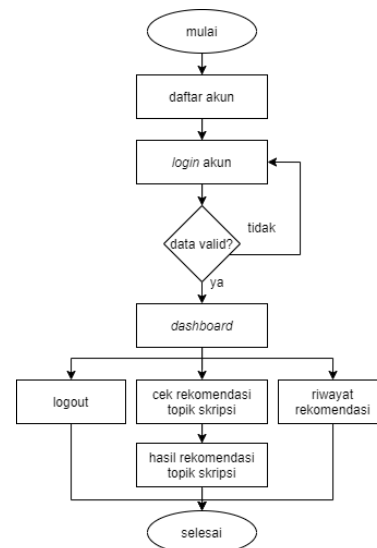
Keterangan rumus 1 ada pada penjelasan ini. x merupakan data dengan *class* yang belum diketahui dan

c yaitu hipotesis dari data x yang merupakan suatu *class* spesifik. Sedangkan $P(c|x)$ merupakan probabilitas hipotesis c yang berdasarkan kondisi x (*posterior probability*) lalu $P(c)$ probabilitas hipotesis c (*prior probability*). Untuk $P(x|c)$ yaitu probabilitas x berdasarkan kondisi hipotesis c dan $P(x)$ yaitu probabilitas dari x [13].

2.4. Implementasi Sistem Pendukung Keputusan

Hasil model *naïve bayes classifier* yang dihasilkan dari proses *data mining* selanjutnya diimplementasikan menjadi sistem pendukung keputusan. Sistem pendukung keputusan adalah bagian dari sistem informasi berbasis komputer termasuk sistem berbasis pengetahuan atau manajemen pengetahuan yang dipakai untuk mendukung pengambilan keputusan dalam suatu organisasi atau perusahaan[14]. Sistem ini tidak dimaksudkan untuk mengotomatisasikan pengambilan keputusan. Tetapi sistem ini memberikan perangkat interaktif yang memungkinkan pengambilan keputusan[11]. Sistem pendukung keputusan dapat dimanfaatkan dalam memberi rekomendasi kepada pengguna dalam menentukan suatu keputusan. Pada penelitian ini sistem dibuat dengan bahasa pemrograman PHP dan *database MySQL*.

Dalam melakukan implementasi menjadi sistem pendukung keputusan peneliti membuat *flowchart* dalam mendesain alur sistem. *Flowchart* dibuat agar mempermudah dalam pembuatan sistem. Pada Gambar 2 dibawah ini dapat dilihat *flowchart* yang telah dibuat.



Gambar 2. *Flowchart* Sistem

Pada Gambar 2 merupakan alur kerja dari sistem yang dibangun. Alur sistem dimulai dengan aksi *daftar akun* kemudian *login* dan sistem melakukan pengecekan, bila data tidak sesuai sistem akan menampilkan tampilan *login* namun bila data sesuai akan dilanjutkan ke tampilan *dashboard*, jika sudah tampil *dashboard* maka pengguna dapat mengakses fitur *cek rekomendasi topik skripsi*, *riwayat rekomendasi* dan *logout*. Dari *flowchart*

yang dibuat kemudian masing-masing dari proses pada *flowchart* diimplementasikan menjadi tampilan sistem.

2.5. Hasil Penelitian

Pada tahap ini bertujuan untuk memvisualisasikan hasil proses *data mining* dan implementasi sistem, dengan visualisasi berupa tabel atau gambar serta menganalisa dan mengevaluasi agar mudah dipahami oleh pembaca. Dalam menemukan hasil model terbaik pada proses *data mining* dilakukan perhitungan nilai akurasi. Akurasi merupakan salah satu indikator dalam menilai sebuah algoritma[15]. Proses penghitungan nilai akurasi dapat dilakukan dengan *confusion matrix*. Sebelum menghitung nilai akurasi dilakukan pembagian *dataset* dengan teknik *percentage split*. *Dataset* akan dibagi menjadi 2, data pelatihan dan data uji[16].

Confusion matrix merupakan tabel matrik yang terdiri dari dua kelas dimana kategori satu kelas yang bernilai positif dan kelas yang lainnya negatif[17]. Pada tahap evaluasi, ukuran atau parameter yang digunakan adalah *confusion matrix*. Tujuan dari tahapan evaluasi adalah untuk melihat efektivitas dari kedua metode pengujian yang dilakukan dalam mengklasifikasikan tugas akhir[9]. Tabel 3 di bawah ini merupakan tabel *confusion matrix*[14].

Tabel 3. Tabel *Confusion Matrix*[14]

| | Yes | No |
|-------|-----------|-----------|
| Yes | <i>tp</i> | <i>fn</i> |
| No | <i>fp</i> | <i>tn</i> |
| Total | <i>P</i> | <i>N</i> |

Tabel 3 adalah perbandingan antara hasil klasifikasi sistem dan hasil klasifikasi sebenarnya. Sehingga nanti akan diketahui nilai yang mendekati antara data aktual dengan data hasil perhitungan sistem. Rumus *accuracy* bisa dilihat pada rumus 2 di bawah ini[14]:

$$accuracy = \frac{tp+tn}{tp+fp+tn+fn} \times 100\% \quad (2)$$

Penjelasan dari Tabel 3 dan rumus 2 adalah *True Positive* (TP) merupakan kasus di mana nilai prediksi (positif) memang benar (*true*) layak. *True Negative* (TN) merupakan kasus di mana hasil prediksi (negatif) tidak layak dan sebenarnya memang (*true*) tidak layak. *False Positive* (FP) merupakan kasus di mana hasil prediksi (positif) layak, ternyata tidak layak. *False Negative* merupakan kasus di mana hasil prediksi (negatif) tidak layak, ternyata sebenarnya (*true*) layak[14].

Untuk hasil implementasi sistem dinilai dengan pengujian *black-box*. Pengujian ini dilakukan pada bagian tampilan sistem. Pengujian *black-box* dilakukan dengan menjalankan program dengan maksud menemukan kesalahan serta memeriksa apakah aplikasi dapat berjalan dengan baik sesuai dengan yang direncanakan[18].

3. Hasil dan Pembahasan

Pada penelitian ini proses *data mining* menggunakan bantuan *tools* Rapidminer. Data yang digunakan dalam penelitian ini sebanyak 358 data yang sudah sesuai atribut yang diperlukan untuk proses *data mining*. Dan implementasi model hasil proses *data mining* menjadi sistem pendukung keputusan menggunakan bahasa pemrograman PHP. Berikut pembahasan dan hasil dari penelitian yang dilakukan.

3.1. Hasil Pengujian Model

Pemodelan dilakukan setelah melakukan pembagian data, untuk pembagian data pelatihan dan data tes menggunakan *percentage split* dengan 5 skenario pembagian data yakni sebesar 90% *data training* dan 10% *data testing*, 80% *data training* dan 20% *data testing*, 70% *data training* dan 30% *data testing*, 60% *data training* dan 40% *data testing*, 50% *data training* dan 50% *data testing*. Pembagian yang tepat dapat memberikan kualitas baik terhadap hasil prediksi nantinya. Proses pembagian data dilakukan menggunakan teknik *stratified random sampling*. Teknik *stratified random sampling* dimungkinkan setiap anggota dari populasi mempunyai besar peluang yang sama untuk dipilih dan digunakan sebagai sampel, sehingga pengukuran nantinya dapat dilakukan dengan hanya melibatkan sedikit dari beberapa sampel saja[19].

Terdapat 5 skenario pemodelan yang diproses berdasarkan rumus 1 dimana pemrosesan *data mining* dilakukan dengan bantuan *tools* yaitu RapidMiner. Masing-masing skenario model yang diproses kemudian dibandingkan nilai akurasinya dengan menggunakan teknik *confusion matrix*, hasil akurasi yang didapat digunakan sebagai tolak ukur kualitas suatu model. Dari 5 skenario model yang dibuat, model dengan akurasi terbaik diimplementasikan menjadi sistem pendukung keputusan.

accuracy: 61.11%

| | true Augm... | true Data ... | true IOT | true KriPto... | true Ranc... | true Siste... | true Multi... | true Netw... | class prec... |
|----------------|--------------|---------------|----------|----------------|--------------|---------------|---------------|--------------|---------------|
| pred. Aug... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| pred. Data... | 2 | 10 | 1 | 1 | 1 | 1 | 0 | 0 | 62.50% |
| pred. IOT | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| pred. KriPT... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| pred. Ran... | 0 | 4 | 1 | 0 | 12 | 0 | 0 | 0 | 70.59% |
| pred. Siste... | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| pred. Multi... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| pred. Netw... | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0.00% |
| class recall | 0.00% | 62.50% | 0.00% | 0.00% | 85.71% | 0.00% | 0.00% | 0.00% | |

Gambar 3. Hasil 90%:10%

Gambar 3 merupakan hasil dari model skenario 1 dengan pembagian rasio data latih dan data uji 90%:10%, rasio 90%:10% sama dengan 322 data latih dan 36 data uji. Nilai akurasi yang didapatkan adalah sebesar 61,11%. Pada Gambar 3 dapat dilihat hanya 2 kelas target yang berhasil diklasifikasi namun kelas lainnya tidak dapat diklasifikasi. Skenario 1 ini hanya

berhasil mengklasifikasi dengan benar 2 kelas target, yakni kategori *data mining* dan rancang bangun.

accuracy: 26.39%

| | true Augm... | true Data ... | true IOT | true KriPto... | true Netw... | true Ranc... | true Siste... | true Multi... | class prec... |
|-----------------|--------------|---------------|----------|----------------|--------------|--------------|---------------|---------------|---------------|
| pred. Aug... | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0.00% |
| pred. Data... | 0 | 9 | 3 | 0 | 1 | 14 | 2 | 0 | 31.03% |
| pred. IOT | 0 | 4 | 0 | 1 | 0 | 2 | 0 | 0 | 0.00% |
| pred. KriPto... | 0 | 2 | 0 | 1 | 0 | 1 | 0 | 0 | 25.00% |
| pred. Net... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| pred. Ranc... | 3 | 16 | 1 | 0 | 0 | 9 | 0 | 0 | 31.03% |
| pred. Siste... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| pred. Multi... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| class recall | 0.00% | 27.27% | 0.00% | 50.00% | 0.00% | 33.33% | 0.00% | 0.00% | |

Gambar 4. Hasil 80%:20%

Gambar 4 adalah hasil dari model skenario 2 dengan pembagian rasio data latih dan data uji 80%:20%, rasio 80%:20% sama dengan 286 data latih dan 72 data uji. Nilai akurasi yang didapatkan adalah sebesar 26,39%. Pada Gambar 4 dapat dilihat hanya 3 kelas target yang berhasil diklasifikasi namun kelas lainnya tidak dapat diklasifikasi. Skenario 2 ini hanya berhasil mengklasifikasi dengan benar 3 kelas target, yakni kategori *data mining*, rancang bangun dan kriptografi.

accuracy: 61.68%

| | true Augm... | true Data ... | true IOT | true KriPto... | true Multi... | true Netw... | true Ranc... | true Siste... | class prec... |
|-----------------|--------------|---------------|----------|----------------|---------------|--------------|--------------|---------------|---------------|
| pred. Aug... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| pred. Data... | 3 | 38 | 1 | 1 | 1 | 0 | 9 | 1 | 70.37% |
| pred. IOT | 1 | 3 | 0 | 0 | 0 | 0 | 2 | 1 | 0.00% |
| pred. KriPto... | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 33.33% |
| pred. Multi... | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00% |
| pred. Net... | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0.00% |
| pred. Ranc... | 1 | 7 | 4 | 0 | 0 | 1 | 27 | 1 | 65.85% |
| pred. Siste... | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0.00% |
| class recall | 0.00% | 77.55% | 0.00% | 50.00% | 0.00% | 0.00% | 65.85% | 0.00% | |

Gambar 5. Hasil 70%:30%

Gambar 5 adalah hasil dari model skenario 3 dengan pembagian rasio data latih dan data uji 70%:30%, rasio 70%:30% sama dengan 251 data latih dan 107 data uji. Nilai akurasi yang didapatkan adalah sebesar 61,68%. Pada Gambar 5 dapat dilihat hanya 3 kelas target yang berhasil diklasifikasi namun kelas lainnya tidak dapat diklasifikasi. Skenario 3 ini hanya berhasil mengklasifikasi dengan benar 3 kelas target, yakni kategori *data mining*, rancang bangun dan kriptografi.

accuracy: 58.74%

| | true Augm... | true Data ... | true IOT | true KriPto... | true Multi... | true Netw... | true Ranc... | true Siste... | class prec... |
|-----------------|--------------|---------------|----------|----------------|---------------|--------------|--------------|---------------|---------------|
| pred. Aug... | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0.00% |
| pred. Data... | 4 | 48 | 2 | 1 | 0 | 0 | 14 | 2 | 67.61% |
| pred. IOT | 1 | 5 | 1 | 0 | 0 | 0 | 1 | 1 | 11.11% |
| pred. KriPto... | 0 | 2 | 0 | 1 | 0 | 0 | 1 | 0 | 25.00% |
| pred. Multi... | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0.00% |
| pred. Net... | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0.00% |
| pred. Ranc... | 2 | 10 | 4 | 1 | 0 | 1 | 34 | 1 | 64.15% |
| pred. Siste... | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0.00% |
| class recall | 0.00% | 73.85% | 14.29% | 33.33% | 0.00% | 0.00% | 61.82% | 0.00% | |

Gambar 6. Hasil 60%:40%

Gambar 6 adalah hasil dari model skenario 4 dengan pembagian rasio data latih dan data uji 60%:40%, rasio 60%:40% sama dengan 215 data latih dan 143 data uji. Nilai akurasi yang didapatkan adalah sebesar 58,74%. Pada Gambar 6 dapat dilihat hanya 4 kelas target yang

berhasil diklasifikasi dengan benar namun kelas lainnya tidak dapat diklasifikasi. Skenario 4 ini hanya berhasil mengklasifikasi dengan benar 4 kelas target, yakni kategori *data mining*, rancang bangun, kriptografi dan IOT.

accuracy: 69.27%

| | true Augm... | true Data ... | true IOT | true KriPto... | true Multi... | true Netw... | true Ranc... | true Siste... | class prec... |
|-----------------|--------------|---------------|----------|----------------|---------------|--------------|--------------|---------------|---------------|
| pred. Aug... | 4 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 57.14% |
| pred. Data... | 3 | 61 | 3 | 1 | 0 | 0 | 17 | 3 | 69.32% |
| pred. IOT | 1 | 3 | 4 | 0 | 0 | 0 | 2 | 0 | 40.00% |
| pred. KriPto... | 0 | 1 | 0 | 3 | 0 | 0 | 0 | 0 | 75.00% |
| pred. Multi... | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 100.00% |
| pred. Net... | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 100.00% |
| pred. Ranc... | 1 | 13 | 2 | 0 | 0 | 0 | 47 | 0 | 74.60% |
| pred. Siste... | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 2 | 50.00% |
| class recall | 44.44% | 75.31% | 44.44% | 75.00% | 100.00% | 100.00% | 69.12% | 40.00% | |

Gambar 7. Hasil 50%:50%

Gambar 7 adalah hasil dari model skenario 5 dengan pembagian rasio data latih dan data uji 50%:50%, rasio 50%:50% sama dengan 179 data latih dan 179 data uji. Nilai akurasi yang didapatkan adalah sebesar 69,27%. Pada Gambar 7 dapat dilihat bahwa seluruh kelas target terdapat hasil klasifikasi dengan benar.

3.2. Hasil Evaluasi Model

Pada tahap ini dilakukan evaluasi terhadap hasil akurasi dari 5 skenario model yang menggunakan algoritma *naïve bayes classifier*. Evaluasi dilakukan dengan membandingkan nilai akurasi model manakah yang paling tinggi nilai akurasinya. Penilaian yang digunakan terdapat pada Tabel 4 dengan parameter *accuracy*. Dan bila telah menemukan model dengan akurasi terbaik selanjutnya adalah mengimplementasikan model menjadi sistem pendukung keputusan.

Tabel 4. Tabel Hasil Pemodelan

| Skenario | Data training | Data Testing | Accuracy |
|----------|---------------|--------------|----------|
| 1 | (90%)322 | (10%)36 | 61,11% |
| 2 | (80%)286 | (20%)72 | 26,39% |
| 3 | (70%)251 | (30%)107 | 61,68% |
| 4 | (60%)215 | (40%)143 | 58,74% |
| 5 | (50%)179 | (50%)179 | 69,27% |

Pada Tabel 4 diketahui bahwa nilai *accuracy* tertinggi yang didapat adalah penggunaan *dataset* dengan data latih dan data uji masing-masing 50% mendapatkan nilai akurasi sebesar 69,27%. Model ini menjadi satu-satunya model yang pada setiap kelas targetnya terdapat data yang berhasil diklasifikasi.

Berdasarkan hasil analisis peneliti pada skenario 1 hingga 4 terdapat perbedaan jumlah data yang cukup jauh antara data latih masing-masing kategori topik skripsi. Seperti pada skenario 1 jumlah data latih kategori rancang bangun adalah 123, sedangkan jumlah data latih kategori *augmented reality* hanya 15. Kedua jumlah data tersebut menunjukkan terdapat jumlah data latih yang sangat jauh jumlahnya hingga perbedaan jumlah data lebih dari 100. Hal inilah yang menjadi

penyebab pada kategori *data mining* terdapat data yang berhasil diklasifikasi sedangkan *augmented reality* tidak berhasil.

Berbeda dengan skenario 5, dimana jumlah data latih kategori rancang bangun jumlahnya 68 dan kategori *augmented reality* jumlahnya 9. Perbedaan jumlah data tidak lagi lebih dari 100 jumlahnya hal ini menunjukkan jumlah perbedaan data yang sudah sangat berkurang. Hal ini lah yang menjadi salah satu faktor mengapa pada skenario 5 setiap kelas target terdapat data yang berhasil diklasifikasi. Maka dari itu model skenario 5 inilah yang diimplementasikan menjadi sistem pendukung keputusan.

3.3. Hasil Implementasi Sistem Pendukung Keputusan

Setelah didapatkan model terbaik dari data yang ada, selanjutnya adalah melakukan implementasi model yang dihasilkan menjadi sistem pendukung keputusan rekomendasi topik skripsi. Dari *flowchart* pada Gambar 2 masing-masing bagian diimplementasikan menjadi tampilan sistem. Implementasi dibuat dalam bentuk sistem berbasis *website*. Berikut adalah hasil dari tampilan sistem.



Gambar 8. Tampilan Daftar

Gambar 8 merupakan tampilan bagian daftar yang berfungsi untuk melakukan pendaftaran akun. Data yang perlu dimasukan adalah *email*, nama lengkap, NPM dan *password*. Akun yang berhasil dibuat digunakan untuk digunakan untuk masuk sistem pada bagian *login*.



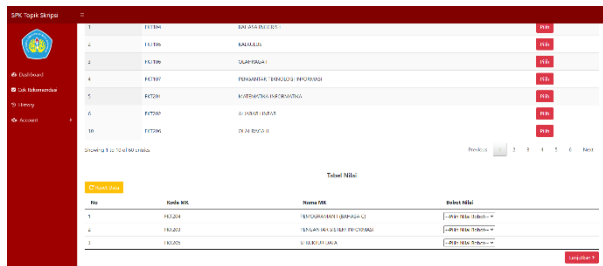
Gambar 9. Tampilan Login

Gambar 9 merupakan tampilan bagian *login* yang berfungsi untuk melakukan validasi akun untuk masuk kedalam sistem dan mengakses fitur yang ada pada sistem. Data yang digunakan untuk sistem *login* adalah NPM dan *password*. Bila data yang divalidasi sesuai maka sistem akan menampilkan halaman *dashboard* dan pengguna dapat mengakses fitur pada sistem.



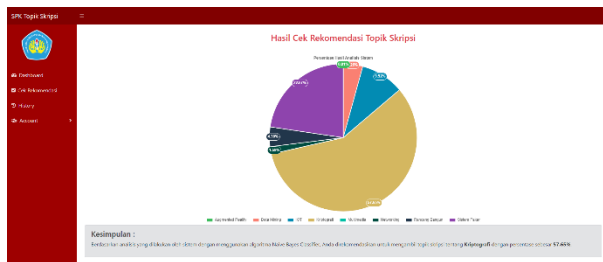
Gambar 10. Tampilan Dashboard

Gambar 10 merupakan tampilan bagian *dashboard*. Tampilan ini merupakan halaman awal saat setelah melakukan *login*. Terdapat 4 menu yang dapat diakses yakni *dashboard*, *cek rekomendasi*, *riwayat* dan *logout*.



Gambar 11. Tampilan Cek Rekomendasi

Gambar 11 merupakan tampilan *form* cek rekomendasi. Tampilan ini berfungsi untuk memasukan *data testing* oleh pengguna, dimana pengguna dapat memilih mata kuliah dan nilai mata kuliah tersebut berdasarkan nilai yang didapat di perkuliahan. Nilai yang dapat dipilih adalah 'A', 'B', 'C', 'D' dan 'E'. Data yang dimasukan kemudian diproses menggunakan algoritma *naive bayes classifier*.



Gambar 12. Tampilan Hasil Rekomendasi

Gambar 12 merupakan tampilan hasil rekomendasi. Tampilan ini tentunya menampilkan hasil rekomendasi untuk pengguna yang datanya telah diproses menggunakan algoritma *naive bayes classifier*. Hasil rekomendasi ditampilkan dalam bentuk *pie chart* agar pengguna dapat dengan mudah memahami persentase dari beberapa kemungkinan rekomendasi yang didapat, dan hasil utama rekomendasi dituliskan dibawah *pie chart* tersebut.

Gambar 13. Tampilan Riwayat Rekomendasi

Gambar 13 merupakan tampilan riwayat hasil rekomendasi. Tampilan ini menampilkan hasil rekomendasi telah didapatkan sebelumnya pada tampilan hasil rekomendasi. Mahasiswa dapat menggunakan fitur ini untuk melihat hasil rekomendasi yang pernah didapatkan. Pada tampilan ini data yang ditampilkan berupa topik rekomendasi, persentase dan tanggal cek rekomendasi.

3.4. Hasil Pengujian Sistem

Pengujian sistem ini menggunakan dua tahap pengujian yaitu *black-box testing*. Pengujian *black-box* berfokus pada fungsi dan kinerja yang terlihat pada tampilan sistem. Pengujian ini menguraikan bagaimana aplikasi bekerja, pengujian ini dilakukan pada seluruh tampilan aplikasi dengan menjalankannya untuk mencari kesalahan yang terjadi pada tampilan aplikasi sehingga aplikasi dapat dilakukan perbaikan. Pada pengujian *black-box* penelitian ini dilakukan pengujian terhadap 5 fitur pada sistem. Yakni fitur pendaftaran, *login*, cek rekomendasi, riwayat dan *logout*. Masing-masing fitur diuji dengan skenario yang berbeda. Berikut adalah hasil dari *black-box testing*.

Tabel 5. *Black-box Testing* Pendaftaran

| No | Skenario Sistem | Hasil yang Diharapkan | Hasil |
|----|--|--|----------------|
| 1 | Mengisi identitas <i>user</i> dengan benar | Berpindah ke halaman <i>Login</i> dan menampilkan Pesan “Berhasil Daftar Akun” | Sesuai Harapan |
| 2 | Mengosongkan nama | Menampilkan Pesan “Harap isi bidang ini” | Sesuai Harapan |
| 3 | Mengosongkan <i>email</i> | Menampilkan Pesan “Harap isi bidang ini” | Sesuai Harapan |
| 4 | Mengosongkan NPM | Menampilkan Pesan “Harap isi bidang ini” | Sesuai Harapan |
| 5 | Mengosongkan <i>password</i> | Menampilkan Pesan “Harap isi bidang ini” | Sesuai Harapan |

Pengujian fitur pendaftaran. Tabel 5 menunjukkan hasil uji *black-box* pada fitur pendaftaran akun. Terdapat 5 skenario pengetesan pada fitur ini. Dan dapat dilihat pada Tabel 5 bahwa seluruh skenario pada tampilan pendaftaran hasil pengujian seluruhnya sesuai harapan.

Tabel 6. *Black-box Testing Login*

| No | Skenario Sistem | Hasil yang Diharapkan | Hasil |
|----|--|--|----------------|
| 1 | Mengisi identitas <i>user</i> dengan benar | Berpindah ke halaman <i>User</i> | Sesuai Harapan |
| 2 | Mengosongkan NPM | Menampilkan Pesan “Harap isi bidang ini” | Sesuai Harapan |
| 3 | Mengosongkan <i>password</i> | Menampilkan Pesan “Harap isi bidang ini” | Sesuai Harapan |

Pengujian fitur *login*. Tabel 6 menunjukkan hasil uji *black-box* pada fitur *login*. Terdapat 3 skenario pengetesan pada fitur ini. Dan dapat dilihat pada Tabel 6 bahwa seluruh skenario pada tampilan *login* hasil pengujian seluruhnya sesuai harapan.

Tabel 7. *Black-box Testing* Cek Rekomendasi

| No | Skenario Sistem | Hasil yang Diharapkan | Hasil |
|----|--|--|----------------|
| 1 | Memilih mata kuliah dan mengisi nilai dengan benar | Berpindah ke halaman hasil rekomendasi | Sesuai Harapan |
| 2 | Tidak memilih mata kuliah | Menampilkan Pesan “Harap pilih minimal 4 mata kuliah” | Sesuai Harapan |
| 3 | Memilih mata kuliah dan tidak memasukkan nilai | Menampilkan Pesan “Harap memasukan nilai mata kuliah” | Sesuai Harapan |
| 4 | Memilih mata kuliah lebih dari 5 | Menampilkan Pesan “Hanya 5 Mata Kuliah yang dapat dipilih! Klik Reset Data untuk memilih ulang!” | Sesuai Harapan |

Pengujian fitur cek rekomendasi. Tabel 7 menunjukkan hasil uji *black-box* pada fitur cek rekomendasi. Terdapat 4 skenario pengetesan pada fitur ini. Dan dapat dilihat pada Tabel 7 bahwa seluruh skenario pada tampilan cek rekomendasi hasil pengujian seluruhnya sesuai harapan.

Tabel 8. *Black-box Testing* Riwayat

| No | Skenario Sistem | Hasil yang Diharapkan | Hasil |
|----|---|-------------------------------------|----------------|
| 1 | Melakukan klik pada menu <i>riwayat</i> | Berpindah ke halaman <i>riwayat</i> | Sesuai Harapan |

Pengujian fitur riwayat. Tabel 8 menunjukkan hasil uji *black-box* pada fitur riwayat. Terdapat 1 skenario pengetesan pada fitur ini. Dan dapat dilihat pada Tabel 8 bahwa seluruh skenario pada tampilan riwayat hasil pengujiannya sesuai harapan.

Tabel 9. *Black-box Testing Logout*

| No | Skenario Sistem | Hasil yang Diharapkan | Hasil |
|----|--|---|----------------|
| 1 | Melakukan klik pada menu <i>logout</i> | Berpindah ke halaman <i>login</i> dan menampilkan Pesan “Berhasil Logout” | Sesuai Harapan |

Pengujian fitur *logout*. Tabel 9 menunjukkan hasil uji *black-box* pada fitur *logout*. Terdapat 1 skenario pengetesan pada fitur ini. Dan dapat dilihat pada Tabel 9 bahwa seluruh skenario pada tampilan *logout* hasil pengujian seluruhnya sesuai harapan.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan dapat diambil kesimpulan bahwa rekomendasi topik skripsi dapat dilakukan dengan menggunakan algoritma *naive bayes classifier*. Model yang dihasilkan dari algoritma *naive bayes classifier* dengan parameter nilai mata kuliah pilihan menghasilkan model dengan akurasi sebesar 69,27%. Karena jumlah data yang tidak seimbang pada setiap kategori topik skripsi menyebabkan nilai akurasi yang kurang baik. Dalam penelitian ini penulis memberikan saran untuk

menambahkan teknik lain agar jumlah data dapat seimbang sehingga nilai akurasi model dapat menjadi lebih baik. Penelitian ini juga bisa sebagai bahan pengembangan untuk penelitian selanjutnya terutama untuk rekomendasi topik skripsi mahasiswa. Data yang digunakan dalam penelitian ini berjumlah 358 data dengan parameter mata kuliah pilihan, sehingga bisa diperbanyak lagi jumlah data dan parameternya untuk pengembangan penelitian selanjutnya.

Daftar Pustaka

- [1] M. F. Akbar and F. D. Anggraeni, "Teknologi Dalam Pendidikan : Literasi Digital dan Self-Directed Learning pada Mahasiswa Skripsi," *Indig. J. Ilm. Psikol.*, vol. 2, no. 1, pp. 28–38, 2017, doi: 10.23917/indigenous.v1i1.4458.
- [2] A. Abdullah and Sucipto, "Prototipe Sistem Cerdas Dalam Menentukan Topik Skripsi Menggunakan Fuzzy AHP di Universitas Muhammadiyah Pontianak," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 15, no. 2, pp. 97–102, 2020.
- [3] E. Rasywir, "Implementasi Sistem Penentuan Skripsi Mahasiswa," *J. Ilm. Media Process.*, vol. 12, pp. 880–893, 2017.
- [4] Widarto, "Faktor Penghambat Studi Mahasiswa yang Tidak Lulus Tepat Waktu di Jurusan Pendidikan Teknik Mesin FT UNY," *J. Din. Vokasional Tek. Mesin*, vol. 2, pp. 127–138, 2017.
- [5] N. L. P. Merawati and S. Hartati, "Sistem Rekomendasi Topik Skripsi Menggunakan Metode Case Based Reasoning," *J. Ilm. Teknol. Infomasi Terap.*, vol. 4, no. 3, pp. 174–183, 2018.
- [6] T. Z. Maulani and Z. K. Simbolon, "Implementasi Algoritma Naïve Bayes Classifier Dalam Menentukan Topik Tugas Akhir Mahasiswa Berbasis Web," *Infomedia*, vol. 4, pp. 33–41, 2019.
- [7] I. S. Bakti and Ivandari, "Model Prediksi Penyakit Diabetes Menggunakan Bayesian Classification Dan Information Gain Untuk Seleksi Fitur Dan Adaptive Boosting Untuk Pembobotan Data," *J. IC-Tech*, vol. 23, no. 3, pp. 28–37, 2019.
- [8] K. Aji, "Sistem Pakar Tes Kepribadian Menggunakan Metode Naive Bayes," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 4, no. 2, p. 75, 2019, doi: 10.31328/jointecs.v4i2.1010.
- [9] Yusra, D. Olivita, and Y. Vitriani, "Perbandingan Klasifikasi Tugas Akhir Mahasiswa Jurusan Teknik Informatika Menggunakan Metode Naïve Bayes Classifier dan K-Nearest Neighbor," *J. Sains, Teknol. dan Ind.*, vol. 14, no. 1, pp. 79–85, 2016.
- [10] A. S. Yaumi, Z. Zulfiqkar, and A. Nugroho, "Klasterisasi Karakter Konsumen Terhadap Kecenderungan Pemilihan Produk Menggunakan K-Means," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 5, no. 3, pp. 195–202, 2020, doi: 10.31328/jointecs.v5i3.1523.
- [11] D. Audilla, D. Hidayatullah, S. Informasi, U. Nasional, and E. S. Informasi, "Sistem Pendukung Keputusan Penilaian Efektivitas Sistem Informasi Menggunakan Metode Naïve Bayes Classifier," *J. STRING (Satuan Tulisan Ris. dan Inov. Teknol.)*, vol. 4, no. 2, pp. 150–158, 2019.
- [12] E. Priyanti, "Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Bakteri Gram-Negatif," *J. Tek. Komput.*, vol. 3, no. 2, pp. 68–76, 2017.
- [13] D. Laia, E. Buulolo, and M. J. F. Sirait, "Implementasi Data Mining Untuk Memprediksi Pemesanan Driver Go-Jek Online Dengan Menggunakan Metode Naive Bayes (Studi Kasus: Pt. Go-Jek Indonesia)," *J. KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 434–439, 2018, doi: 10.30865/komik.v2i1.972.
- [14] B. Satria and L. Tambunan, "Sistem Pendukung Keputusan Penerima Bantuan Rumah Layak Huni Menggunakan FMADM dan SAW," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 5, no. 3, p. 167, 2020, doi: 10.31328/jointecs.v5i3.1361.
- [15] R. K. Dinata, S. Safwandi, N. Hasdyna, and N. Azizah, "Analisis K-Means Clustering pada Data Sepeda Motor," *INFORMAL Informatics J.*, vol. 5, no. 1, pp. 10–17, 2020, doi: 10.19184/isj.v5i1.17071.
- [16] D. Sartika and D. Indra, "Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian," *J. Tek. Inform. Dan Sist. Inf.*, vol. 1, no. 2, pp. 151–161, 2017.
- [17] R. A. Arnomo, W. L. Y. Saptomo, and P. Harsadi, "Implementasi Algoritma K-Nearest Neighbor Untuk Identifikasi Kualitas Air (Studi Kasus : Pdam Kota Surakarta)," *J. Teknol. Inf. dan Komun.*, vol. 6, no. 1, pp. 1–5, 2018.
- [18] M. Muhammad, S. Novi, and P. Narti, "Implementasi Metode Simple Additive Weighting(SAW) pada Sistem Pendukung Keputusan dalam Menentukan Prioritas Perbaikan Jalan," *Jur. Tek. Inform.*, vol. 5, no. 4, pp. 157–162, 2017.
- [19] S. F. Ulya, Y. Sukestiyarno, and P. Hendikawati, "Analisis Prediksi Quick Count Dengan Metode Stratified Random Sampling Dan Estimasi Confidence Interval Menggunakan Metode Maksimum Likelihood," *Unnes J. Math.*, vol. 7, no. 1, pp. 108–119, 2018, doi: 10.15294/ujm.v7i1.27385.