

JURNAL TELEMATIKA MIKOM



MAGISTER ILMU KOMPUTER
PROGRAM PASCA SARJANA TEKNOLOGI INFORMASI
UNIVERSITAS BUDI LUHUR



TELEMATIKA MKOM

ISSN : 2085-725X (Print) | 2527-9033 (Online)

Mitra Bestari

Reviewer

Universitas Budi Luhur Jakarta

Dr. Indra, S.Kom., M.T.I.

ID Scopus: 57209452300 | ID Sinta: 5978751 | ID Google Scholar: W7dV3C0AAAAJ

Dr. Mardi Hardjianto, S.Kom., M.T.I.

ID Scopus: 57190344469 | ID Sinta: 5977626 | ID Google Scholar: 2JaUYzAAAAAJ

Dr. Mohammad Syafrullah, M.Sc., M.Kom.

ID Scopus: 57201861985 | ID Sinta: 5984532 | ID Google Scholar: Ei49YMQAAAAJ

Dr. Rusdah, S.Kom, M.Kom.

ID Scopus: 57189247170 | ID Sinta: 5979534 | ID Google Scholar: 6UT-LCMAAAAJ

Dr. Gandung Triyono, M.Kom

ID Scopus: 57194158314 | ID Sinta: 6673525 | ID Google Scholar: UIAIPYMAAAAJ

Institut Teknologi PLN, Jakarta, Indonesia

Dr.Dra. Dwina Kuswardani, M.Kom

ID Scopus: 57196455492 | ID SINTA: 5986435 | ID Google Scholar: MtIztwQAAAAJ

Politeknik Negeri Jakarta

Risna Sari

ID Scopus: 57221371746 | ID SINTA: 6659971 | ID Google Scholar: 3vtycCcAAAAJ

Prihatin Oktivasari

ID Scopus: 57202816309 | ID SINTA: 6013616 | ID Google Scholar: 1xZTIgMAAAAJ

Editorial Tim

Editor in Chief

Dr. Rusdah, S.Kom, M.Kom, Universitas Budi Luhur, Jakarta, Indonesia

Scopus ID: 57198358682 | SINTA ID: 5979534 | ID Google Scholar: 6UT-LCMAAAAJ

Assistant Journal In Chief :

Wahyu Pramusinto, S.Kom, M.Kom, Universitas Budi Luhur, Jakarta, Indonesia

Scopus ID: 57215896652 | SINTA ID: 5977936 | Google Scholar : hhH1HnYAAAAJ

Associate (Handing) Editor :

Nidya Kusumawardhany, S.Kom, M.Kom, Universitas Budi Luhur, Jakarta, Indonesia

Scopus ID: 57216853240 | SINTA ID: 6661917 | Google Scholar : mgyZEcgAAAAJ

Devit Setiono, S.Kom, M.Kom, Universitas Budi Luhur, Jakarta, Indonesia

Scopus ID:- | SINTA ID: 6725996 | Google Scholar : qgUDcTYAAAAJ

Vol 18, No 1 (2026)

Jurnal Telematika MKOM Vol. 18 No.1 Maret 2026

Title : Telematika MKOM **Publication Schedule :** Maret, September **ISSN :** 2527-9033 (online) dan 2085-725X (cetak)

Accreditation : - Subject : Network Computer and Security, Data Mining, Sistem pakar (Expert System), Jaringan syaraf tiruan (Artificial Neural Network), Natural Language Processing (NLP), Game Development, Visi Komputer, Temu balik Informasi, Artificial Intelligent, Internet of Things, E-Learning, IT Governance, Serta topik lainnya yang masuk ke dalam rumpun ilmu tersebut **Indexed at :** Google Scholar, Portal Garuda **DOI :** 10.36080 **OAI: - Publisher :** Universitas Budi Luhur

Jurnal Telematika MKOM diterbitkan oleh Program Pascasarjana Magister Ilmu Komputer, Fakultas Teknologi Informasi, Universitas Budi Luhur.

Table of Contents

Articles

DETEKSI DINI PENYAKIT KANKER MENGGUNAKAN DECISION TREE

 DOI : 10.36080/telematikamkom.4312 |  Abstract views : 2 times

PDF
1-8

IMPLEMENTASI SISTEM E-COMMERCE BERBASIS CONTENT MANAGEMENT SYSTEM (CMS) WORDPRESS UNTUK MENINGKATKAN PENJUALAN PADA PRODUK UMKM DAPUR BU IDA

 DOI : 10.36080/telematikamkom.4429 |  Abstract views : 2 times

PDF
9-18

IMPLEMENTASI KRIPTOGRAFI AES-128 MODE CBC UNTUK PENGAMANAN FILE BERBASIS WEB PADA VIA COMPUTER

PDF
19-27

KLASIFIKASI BIDANG PENELITIAN ARTIKEL ILMIAH MENGGUNAKAN ALGORITMA NAIVE BAYES

 DOI : 10.36080/telematikamkom.4733 |  Abstract views : 0 times

PDF
28-35

Klasifikasi Bidang Penelitian Artikel Ilmiah Menggunakan Algoritma Naive Bayes

Nurwati^{1*}, Yudi Santoso²

^{1,2} Sistem Informasi, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia
Alamat lengkap semua penulis termasuk nama negara
email koresponden: nurwati@budiluhur.ac.id

(received: 03/02/2026, revised: 13/03/2026, accepted: 29/03/2026)

Abstrak

Peningkatan jumlah publikasi artikel ilmiah di berbagai bidang menyebabkan proses pengelompokan artikel menjadi semakin kompleks. Klasifikasi yang masih dilakukan secara manual membutuhkan waktu yang lama, berpotensi menimbulkan inkonsistensi, dan kurang efektif ketika jumlah artikel terus bertambah. Penelitian ini bertujuan mengembangkan model klasifikasi bidang penelitian artikel ilmiah menggunakan algoritma Naive Bayes untuk membantu proses pengelompokan artikel secara otomatis berdasarkan metadata yang dimiliki. Dataset penelitian diperoleh dari Portal Garuda sejumlah 500 artikel yang terdiri atas artikel pada bidang Rekayasa Perangkat Lunak, Sistem Pendukung Keputusan, dan Sistem Informasi yang dipublikasikan dalam kurun waktu lima tahun terakhir. Tahapan penelitian meliputi pengumpulan data, *text preprocessing* melalui proses *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Dilanjutkan proses klasifikasi menggunakan algoritma Naive Bayes. Kinerja model dievaluasi menggunakan *confusion matrix* dengan Accuracy sebesar 91,00%, Precision sebesar 91,19%, Recall sebesar 91,00%, dan F1-Score sebesar 91,01%. Hasil penelitian menunjukkan bahwa model yang dikembangkan mampu mengklasifikasikan artikel ilmiah ke dalam bidang penelitian yang sesuai dengan tingkat kinerja yang baik. Model ini diharapkan dapat mendukung pengelolaan publikasi ilmiah, mempercepat proses klasifikasi artikel, serta membantu analisis tren penelitian secara lebih efektif.

Kata kunci: Bidang Penelitian, Klasifikasi Artikel Ilmiah, Naive Bayes, Text Mining

Classification of Scientific Research Fields Using the Naive Bayes Algorithm

Abstract

The rising number of scientific article publications across various fields has made the article categorization process increasingly complex. Manual classification is time-consuming, prone to inconsistency, and inefficient as the volume of articles grows. This study aims to develop a classification model for scientific research fields using the Naive Bayes algorithm to automate the categorization process based on article metadata. The dataset, comprising 500 articles published over the last five years, was obtained from the Garuda Portal and covers the fields of Software Engineering, Decision Support Systems, and Information Systems. The research methodology involved data collection and text preprocessing—specifically cleaning, case folding, tokenizing, stopwords removal, and stemming—followed by classification using the Naive Bayes algorithm. Model performance was evaluated using a confusion matrix, yielding an accuracy of 91.00%, precision of 91.19%, recall of 91.00%, and an F1-score of 91.01%. The results demonstrate that the developed model effectively classifies scientific articles into their respective research fields with a high level of performance. This model is expected to support scientific publication management, accelerate the article classification process, and facilitate more effective analysis of research trends.

Keywords: Research Field, Scientific Article Classification, Naive Bayes, Text Mining

1. Pendahuluan

Perkembangan ilmu pengetahuan yang pesat diikuti oleh meningkatnya aktivitas penelitian dalam beberapa tahun terakhir telah mendorong pertumbuhan jumlah publikasi ilmiah secara pesat di berbagai bidang, termasuk Sistem Informasi. Fenomena ini memperkaya sumber pengetahuan yang dapat dimanfaatkan oleh peneliti sebagai landasan dalam mengembangkan penelitian baru. Menurut Ariani dkk, peran artikel ilmiah sangat penting sebagai media komunikasi ilmiah yang memungkinkan peneliti



berbagi gagasan, temuan, dan inovasi kepada komunitas akademik secara global. Di tengah perkembangan teknologi dan revolusi informasi, artikel ilmiah menjadi instrumen utama dalam memajukan ilmu pengetahuan [1]. Seiring meningkatnya jumlah publikasi setiap tahun, pengelolaan dan pengelompokan artikel ilmiah berdasarkan bidang penelitian menjadi semakin kompleks sehingga diperlukan sistem yang mampu melakukan klasifikasi secara otomatis [2].

Proses pengelompokan artikel ilmiah pada umumnya masih mengandalkan proses analisis secara manual terhadap informasi yang terdapat pada judul, abstrak, dan kata kunci, proses ini menjadi kurang efisien karena besarnya volume data dan sifat data yang tidak terstruktur. Pendekatan tersebut membutuhkan waktu yang relatif lama, sangat bergantung pada pemahaman individu, serta berpotensi menimbulkan inkonsistensi dalam penentuan kategori bidang penelitian, terutama ketika jumlah artikel yang harus diklasifikasikan semakin banyak [3]. Kondisi tersebut mendorong perlunya pendekatan otomatis untuk mengekstraksi informasi penting dari kumpulan dokumen dalam jumlah besar menurut R.F.S. Cahyanto dalam N. Aprianti, 2026 [4].

Penelitian ini bertujuan membangun sistem repository berbasis website untuk mendukung pencarian dokumen akademik secara digital, bagaimana algoritma Naive Bayes Classifier diterapkan untuk klasifikasi dokumen berbasis teks, dan bagaimana sistem ini dapat memberikan aksesibilitas serta efisiensi yang lebih baik dibandingkan metode manual [5]. Penelitian sebelumnya menggunakan Naive Bayes yang pernah dilakukan Kusuma Nurnasikha dkk tahun 2023 [2] tentang klasifikasi bidang ilmu publikasi ilmiah terindeks SINTA menggunakan metode Naive Bayes memudahkan peneliti mengklasifikasi lima bidang ilmu karena modelnya sederhana dan cukup fleksibel meningkatkan kualitas keputusan dihasilkan. Hasil dari perhitungan diperoleh nilai akurasi 0,425, nilai *recall* 0,20 dan nilai presisi 0,05. Hasil nilai ini masih cukup rendah jika digunakan pada sistem SINTA karena pelabelan dataset belum optimal.

Penelitian sebelumnya menggunakan Naive Bayes Ainunna'imah dkk tahun 2026 [3] tentang model klasifikasi kesesuaian artikel pada jurnal SINTA berdasarkan metada menggunakan *term frequency-inverse document frequency* dan Naive Bayes Dataset terdiri atas 1.200 metadata artikel mencakup judul, abstrak, dan kata kunci, yang dikumpulkan melalui crawling manual terhadap jurnal-jurnal bidang teknologi pada portal SINTA. Tahapan penelitian meliputi pengumpulan data, prapemrosesan teks (*case folding*, *translasi*, *tokenisasi*, *stopword removal*, dan *stemming*), penggabungan metadata, ekstraksi fitur menggunakan TF-IDF, serta penerapan algoritma Naive Bayes dengan skema 5-fold *cross-validation*. Evaluasi berdasarkan *confusion matrix* menunjukkan bahwa model mencapai *accuracy* 0,7058, *precision* 0,6977, *recall* 0,7133, dan F1-score 0,7065. Hasil ini menegaskan bahwa Naive Bayes mampu memberikan performa klasifikasi yang cukup baik terhadap metadata artikel, serta berpotensi mendukung pengembangan sistem rekomendasi target submission.

Selain itu, penelitian ini bertujuan mengevaluasi performa model menggunakan metrik *accuracy*, *precision*, *recall*, dan F1-score sehingga diperoleh gambaran mengenai efektivitas algoritma dalam mengklasifikasikan artikel ilmiah ke dalam bidang penelitian yang sesuai. Salah satu algoritma klasifikasi yang banyak digunakan pada data teks adalah Naive Bayes. Diharapkan penelitian ini memberikan kontribusi dalam pengembangan penerapan algoritma Naive Bayes untuk klasifikasi artikel ilmiah berdasarkan metadata. Model yang dihasilkan dapat dimanfaatkan sebagai pendukung pengelompokan publikasi secara otomatis serta menjadi referensi bagi pengembangan penelitian dan sistem klasifikasi artikel ilmiah pada masa mendatang.

2. Metode Penelitian

2.1 Data Penelitian

Data penelitian yaitu artikel yang termasuk dalam bidang Rekayasa Perangkat Lunak, Sistem Pendukung Keputusan, dan Sistem Informasi. Seluruh data diperoleh dari berbagai macam sumber portal jurnal dengan jumlah 500 artikel yang mewakili ketiga bidang penelitian tersebut selama kurun waktu lima tahun. Dataset ini selanjutnya digunakan sebagai data penelitian untuk membangun dan mengevaluasi model klasifikasi menggunakan algoritma Naive Bayes.

2.2 Tahap Penelitian

Pada gambar 1 bisa dilihat tahapan penelitian yang dilakukan pada penelitian ini



Gambar 1. Tahapan Penelitian

2.2.1 Studi Literatur

Tahapan penelitian di mulai dengan studi literatur. Pada tahap ini dilakukan proses identifikasi permasalahan penelitian berdasarkan informasi yang diperoleh melalui studi literatur, observasi secara langsung, serta berbagai referensi lain yang relevan sebagai dasar dalam merumuskan fokus penelitian [6].

2.2.2 Rumusan Masalah

Pada tahap perumusan masalah, inti permasalahan penelitian diterjemahkan ke dalam sejumlah pertanyaan penelitian yang dirancang secara jelas dan relevan. Pertanyaan tersebut menjadi landasan dalam menentukan tujuan, metode, dan proses penelitian yang akan dilakukan [6].

2.2.3 Pengumpulan data

Pada tahap pengumpulan data, artikel ilmiah dipilih berdasarkan kriteria yang telah ditentukan [2], yaitu berasal dari bidang informatika dan diterbitkan dalam kurun waktu lima tahun terakhir. Metadata artikel yang terdiri atas judul, abstrak, dan kata kunci kemudian dihimpun dan disusun menjadi dataset penelitian [4]. Dataset tersebut selanjutnya digunakan pada tahapan prapemrosesan teks, ekstraksi fitur, serta pengujian model klasifikasi menggunakan algoritma Naive Bayes.

2.2.4 Proses Klasifikasi

Tahap selanjutnya preprocessing. Dalam proses *text mining*, *preprocessing* dilakukan untuk mempersiapkan data teks agar memiliki kualitas yang lebih baik sebelum memasuki tahap analisis. Mengingat data awal masih berupa teks tidak terstruktur, diperlukan proses transformasi sehingga informasi yang terkandung di dalamnya dapat direpresentasikan dalam bentuk numerik dan diolah menggunakan algoritma klasifikasi [7]. Proses *preprocessing* dalam penelitian ini dilakukan melalui beberapa tahapan, yaitu *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Tahapan tersebut diawali dengan pembersihan data dari karakter atau elemen yang tidak diperlukan, kemudian seluruh huruf dikonversi menjadi huruf kecil. Setelah itu, teks dipecah menjadi token-token berupa kata, dilanjutkan dengan penghapusan kata yang tidak memiliki kontribusi signifikan terhadap proses klasifikasi. Tahap terakhir adalah *stemming*, yaitu mengembalikan kata ke bentuk dasarnya agar representasi data menjadi lebih konsisten [8].

Selanjutnya penerapan metode Naive Bayes. Pada tahap klasifikasi, algoritma Naive Bayes diterapkan untuk mengidentifikasi bidang penelitian setiap artikel ilmiah berdasarkan pola distribusi fitur kata yang diperoleh dari metadata dokumen. Proses klasifikasi dilakukan dengan menghitung probabilitas setiap dokumen terhadap seluruh kelas yang tersedia, kemudian dokumen akan ditempatkan pada kategori yang memiliki nilai probabilitas paling tinggi [9]. Algoritma Naive Bayes Classifier memiliki sejumlah keunggulan, di antaranya proses komputasi yang relatif cepat, mekanisme



perhitungan yang sederhana, serta kemampuan menghasilkan tingkat akurasi yang baik pada berbagai kasus klasifikasi teks [10]. Algoritma Naive Bayes adalah metode pembelajaran mesin untuk masalah klasifikasi, terutama digunakan dalam klasifikasi teks dengan data pelatihan berdimensi tinggi. Algoritma ini memungkinkan pembangunan model dengan cepat dan dianggap sebagai salah satu algoritma prediksi yang paling cepat dipelajari. Algoritma ini menggunakan probabilitas objek sebagai dasarnya [11]. Perhitungan pada algoritma Naive Bayes mengacu pada persamaan probabilitas Bayes yang digunakan untuk menentukan kategori suatu dokumen. Persamaan tersebut ditunjukkan pada rumus berikut:

$$P(C|X) = P(C) * P(X|C) / P(X) \dots\dots\dots(1)$$

Dimana $P(C|X)$ menyatakan probabilitas suatu kelas C berdasarkan fitur X , $P(C)$ merupakan probabilitas prior dari kelas C , $P(X|C)$ adalah probabilitas kemunculan fitur X pada kelas C , dan $P(X)$ menyatakan probabilitas dari fitur X [9].

2.2.5 Model Uji

Tahap model uji berisi implementasi Algoritma Nive Bayes dengan Python. Persiapan awal dimulai dengan menginstall python, Adapun *library* yang digunakan seperti pandas, numpy, matplotlib.pyplot, graphviz, sklearn.model_selection, sklearn.tree, dan sklearn.metrics untuk membantu dalam pengolahan data, pemisahan dataset, pembuatan model Algoritma C4.5, serta evaluasi hasil klasifikasi [12].

3. Hasil dan Pembahasan

3.1 Hasil Preprocessing tahap 1

Dataset terdiri dari jumlah data 500 artikel, jumlah kolom 3 dan kolom yang tersedia ada kolom No, kolom Judul, kolom kelas (label). Berikut statistik dataset ditampilkan di tabel 1 di bawah ini:

Tabel 1. Statistik dataset

Label Kelas	Jumlah
rekayasa perangkat lunak	225
sistem informasi	176
sistem penunjang keputusan	99

Dataset yang digunakan dalam penelitian ini merupakan data contoh yang dimanfaatkan sebagai bahan pengujian dan evaluasi kinerja model klasifikasi. Proses selanjutnya melakukan proses *preprocessing* tahap 1 dimulai dengan mengolah teks asli yaitu SISTEM PENUNJANG KEPUTUSAN PROFILE MATCHING UNTUK SELEKSI KARYAWAN CAPTURE IT PHOTOBOOTH untuk kemudian dilakukan proses *cleaning* yaitu menghapus tanda baca, menghapus karakter selain huruf, angka, dan spasi [13]. Lalu mengerjakan *case folding* yaitu mengubah seluruh teks menjadi huruf kecil [13]. Kemudian tahap *tokenizing*, memecah judul artikel menjadi token kata, ditampilkan dengan pemisah |. Berikut tabel 2 di bawah ini.

Tabel 2. Hasil Preprocessing Tahap 1

Tahapan	Hasil
Teks asli	SISTEM PENUNJANG KEPUTUSAN PROFILE MATCHING UNTUK SELEKSI KARYAWAN CAPTURE IT PHOTOBOOTH
<i>Cleaning</i>	SISTEM PENUNJANG KEPUTUSAN PROFILE MATCHING UNTUK SELEKSI KARYAWAN CAPTURE IT PHOTOBOOTH
<i>Case Folding</i>	sistem penunjang keputusan profile matching untuk seleksi karyawan capture it photobooth

Tokenizing

sistem | penunjang | keputusan | profile | mathing
| untuk | seleksi | karyawan | capture | it |
photobooth

3.2 Hasil Preprocessing tahap 2

Proses *preprocessing* tahap 2 melakukan proses *stopword removal* dengan menghapus kata-kata umum Bahasa Indonesia yang tidak memberikan informasi penting dalam proses klasifikasi misalnya "dan", "di", "ke", dan "dari" [14]. Kemudian dilanjutkan dengan proses Stemming [14] yaitu melakukan penyederhanaan kata menjadi bentuk yang lebih dasar. Pada file ini dilakukan stemming sederhana sebagai persiapan sebelum penggunaan *stemmer* penuh (misalnya Sastrawi). Berikut tabel 3 hasil preprocessing tahap 2 di bawah ini.

Tabel 3. Hasil Preprocessing tahap 2

Tahapan	Hasil
Stopword removal	sistem penunjang keputusan profile matching untuk seleksi karyawan capture it photobooth
Stemming	sistem penunjang keputusan profile matching untuk seleksi karyawan capture it photobooth

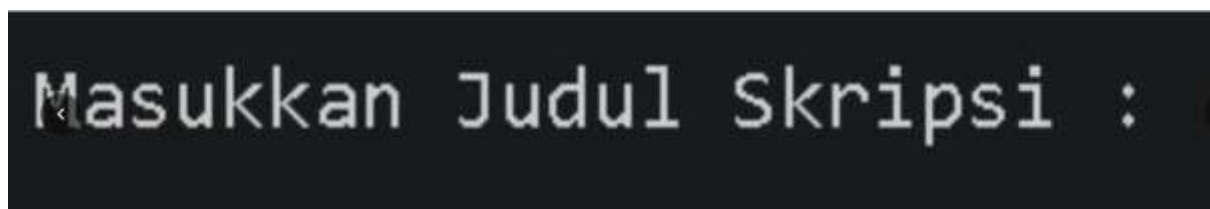
Dari hasil proses preprocessing pada tabel 2 dan tabel 3 dirangkum seperti pada gambar 2 di bawah ini.

No	Judul	Kelas	Cleaning	Case Folding	Tokenizing	Stopword Removal	Stemming
0	IMPLEMENTASI AES-256 UNTUK MENGAMANKAN DOKUMEN...	NaN	IMPLEMENTASI AES 256 UNTUK MENGAMANKAN DOKUMEN...	implementasi aes 256 untuk mengamankan dokumen...	implementasi aes 256 untuk mengamankan...	implementasi aes 256 untuk mengamankan...	implementasi aes 256 untuk mengamankan...
1	PERBANDINGAN ALGORITMA KRIPTOGRAFI AES-128 DAN...	NaN	PERBANDINGAN ALGORITMA KRIPTOGRAFI AES 128 DAN...	perbandingan algoritma kriptografi aes 128 dan...	perbandingan algoritma kriptografi aes ...	perbandingan algoritma kriptografi aes...	perbandingan algoritma kriptografi aes...
2	ANALISIS MARKET BASKET DENGAN ALGORITMA APRIORI...	NaN	ANALISIS MARKET BASKET DENGAN ALGORITMA APRIORI...	analisis market basket dengan algoritma aprior...	analisis market basket dengan algoritma...	analisis market basket dengan algoritma...	analisis market basket dengan algoritma...
3	ANALISIS SENTIMEN 100 HARI KERJA PRESIDEN PRAB...	NaN	ANALISIS SENTIMEN 100 HARI KERJA PRESIDEN PRAB...	analisis sentimen 100 hari kerja presiden prab...	analisis sentimen 100 hari kerja pre...	analisis sentimen 100 hari kerja pre...	analisis sentimen 100 hari kerja pre...
4	SISTEM PENUNJANG KEPUTUSAN PROFILE MATCHING UN...	NaN	SISTEM PENUNJANG KEPUTUSAN PROFILE MATCHING UN...	sistem penunjang keputusan profile matching un...	sistem penunjang keputusan profile mat...	sistem penunjang keputusan profile mat...	sistem penunjang keputusan profile mat...

Gambar 2. Hasil preprocessing semua tahap

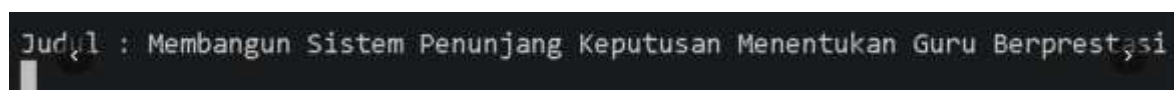
3.3 Proses Klasifikasi

Tahapan proses klasifikasi merupakan implementasi metode Naive Bayes ditunjukkan pada gambar 3. Gambar 3 di bawah ini memperlihatkan rancangan antarmuka yang menyediakan kolom masukan untuk judul artikel. Melalui kolom tersebut, pengguna dapat memasukkan judul artikel yang akan digunakan sebagai data masukan pada proses klasifikasi. Judul yang akan dimasukkan ke dalam kolom input adalah Membangun Sistem Penunjang Keputusan Menentukan Guru Berprestasi



Gambar 3. Rancangan antarmuka

Gambar 3 memperlihatkan rancangan layar kolom masukan diinput judul artikel.



Gambar 4. Rancangan layar diinput judul artikel

Gambar 4 menampilkan hasil prediksi. Gambar 5 menampilkan hasil prediksi dari inputan judul yang diberikan.

```

bangun sistem tunjang putus tentu guru prestasi
Kategori : SPK

5 Rekomendasi Jurnal

1. Sistem Pendukung Keputusan Beasiswa Menggunakan SAW
   Similarity : 0.3398

2. Rancang Bangun Penilaian Karyawan menggunakan Sistem Penunjang Keputusan
   Similarity : 0.2869
  
```

Gambar 5. Hasil prediksi

Gambar 5 hasil prediksi menampilkan judul yang sudah mengalami proses preprocessing dan menampilkan kategori dari inputan judul yaitu bangun sistem tunjang putus tentu guru prestasi. Hasil pemrosesan sistem berupa lima rekomendasi jurnal yang dipilih berdasarkan nilai kesamaan tertinggi antara judul artikel yang diinput dengan judul artikel yang terdapat pada dataset.

3.4 Model Uji/Evaluasi

Evaluasi dari penelitian ini menggunakan pembagian data 80% data latih dan 20% data uji dengan TF-IDF dan multinomial Naive Bayes ditunjukkan pada tabel 4 di bawah ini.

Tabel 4. Hasil evaluasi model

Metrik	Nilai
Accuracy	91,00%
Precision	91,19%
Recall	91,00%
F1-score	91,01%

Hasil klasifikasi ditampilkan pada tabel 5 di bawah ini.

Tabel 5. Hasil klasifikasi Naive Bayes

Kelas	Precision	Recall	F1-Score	Support
Rekayasa Perangkat Lunak	0.88	0.93	0.90	45
Sistem Informasi	0.91	0.86	0.88	35
Sistem Penunjang Keputusan	1.00	0.95	0.97	20

Berdasarkan hasil pengujian menggunakan algoritma Naive Bayes dengan pembobotan TF-IDF terhadap dataset artikel ilmiah, diperoleh nilai *accuracy* sebesar **91,00%**, menunjukkan bahwa model mampu mengklasifikasikan artikel ilmiah ke dalam bidang penelitian yang sesuai dengan tingkat ketepatan yang tinggi. Selain itu, diperoleh nilai *precision* sebesar **91,19%**, *recall* sebesar **91,00%**, dan *F1-score* sebesar **91,01%**. Nilai-nilai tersebut menunjukkan bahwa model memiliki keseimbangan yang baik antara kemampuan memberikan prediksi yang tepat (*precision*) dan kemampuan mengenali seluruh data pada kelas yang sebenarnya (*recall*). Dengan demikian, algoritma Naive Bayes dapat digunakan sebagai metode yang efektif untuk melakukan klasifikasi bidang penelitian artikel ilmiah pada dataset yang digunakan dalam penelitian ini.

4. Kesimpulan

Penelitian ini berhasil membangun sistem klasifikasi bidang penelitian artikel ilmiah menggunakan algoritma Naive Bayes dengan pembobotan fitur menggunakan metode TF-IDF. Sistem mampu mengolah data judul artikel ilmiah dan mengelompokkannya ke dalam kategori bidang



penelitian yang telah ditentukan. Hasil pengujian menggunakan dataset sebanyak 500 artikel ilmiah yang terdiri dari tiga kategori, yaitu Rekayasa Perangkat Lunak, Sistem Informasi, dan Sistem Penunjang Keputusan, model menghasilkan nilai *accuracy* sebesar 91,00%, Precision sebesar 91,19%, Recall sebesar 91,00%, dan F1-Score sebesar 91,01%. Hasil tersebut menunjukkan bahwa algoritma Naive Bayes memiliki performa yang baik dalam melakukan klasifikasi bidang penelitian artikel ilmiah dan penggunaan pembobotan TF-IDF mampu meningkatkan representasi kata pada judul artikel sehingga membantu algoritma Naive Bayes dalam membedakan karakteristik setiap kategori penelitian. Hal ini dibuktikan dengan nilai evaluasi yang relatif tinggi dan seimbang pada seluruh metrik pengujian.

Daftar Pustaka

- [1] A. Julia, S. Febriani, and S. Nurjannah, "Publikasi Ilmiah Dan Inovasi Pendidikan," *JUPI J. Ilmu Pendidik. Islam*, vol. 23, no. 04, 2025, doi: <https://doi.org/10.36835/jipi.v23i4.249>.
- [2] K. Nurnasikha, S. F. C. Haviana, and I. M. I. Subroto, "Klasifikasi Bidang Ilmu Publikasi Ilmiah Terindeks SINTA Menggunakan Metode Naïve Bayes," *J. Transistor Elektro dan Inform. (TRANSISTOR EI)*, vol. 5, no. 3, pp. 147–154, 2023, [Online]. Available: <https://jurnal.unissula.ac.id/index.php/EI/article/view/31126/9730>
- [3] A. 'imah, I. Riadi, and H. Yuliansyah, "Model Klasifikasi Kesesuaian Artikel Pada Jurnal SINTA Berdasarkan Metadata Menggunakan Term Frequency-Inverse Document Frequency dan Naive Bayes," *semantik*, vol. 11, no. 1, pp. 1–12, 2026, [Online]. Available: <https://semantik.uho.ac.id/index.php/journal/article/view/270/58>
- [4] N. W. Aprianti, H. Yuliansyah, and M. K. Biddinika, "Pengelompokan Abstrak Jurnal Ilmiah Menggunakan Term Frequency – Inverse Document Frequency dan K-Means," *J. Komtika (Komputasi dan Inform.)*, vol. 10, no. 1, pp. 17–31, 2026, doi: <https://doi.org/10.31603/komtika.v10i1.15516>.
- [5] M. A. Putra, M. Azmi, and Yuhefizar, "Optimalisasi Pencarian Dokumen Akademik Menggunakan Naïve Bayes Classifier : Studi Kasus pada Repository Digital," vol. 1, no. 1, pp. 1–7, 2025, [Online]. Available: <https://jurnal-bitwise.org/index.php/bitwise/article/view/8/10>
- [6] A. Rohmah and D. Anubhakti, "KLASTERISASI KASUS KEKERASAN TERHADAP PEREMPUAN DI PROVINSI JAWA TENGAH MENGGUNAKAN ALGORITMA K-MEANS," *TELEMATIKA*, vol. 17, no. 2, pp. 75–85, 2025, [Online]. Available: <https://journal.budiluhur.ac.id/telematika/article/view/4314/0>
- [7] Apriani, H. Zakiyudin, and K. Marzuki, "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta," *J. Bumigora Inf. Technol.*, vol. 3, no. 1, pp. 19–27, 2021, doi: 10.30812/bite.v3i1.1110.
- [8] D. Fisabilillah, K. Auliasari, and Y. A. Pranoto, "Sistem Rekomendasi Konversi Program MSIB Dengan Mata Kuliah Prodi Informatika ITN Malang Menggunakan Cosine Similarity," *SKANIKASistem Komput. dan Tek. Inform.*, vol. 8, no. 1, pp. 83–94, 2025, [Online]. Available: <https://jom.fti.budiluhur.ac.id/SKANIKA/article/view/3325/1398>
- [9] I. Rahdiana, J. Jonathan, M. Anif, B. H. Prasetyo, and Subandi, "KLASIFIKASI OPINI MASYARAKAT TENTANG PROGRAM MAKAN BERGIZI GRATIS DI TWITTER DENGAN ALGORITMA NAIVE BAYES," *Telemat. MKOM*, vol. 17, no. 2, pp. 128–134, 2026, [Online]. Available: <https://journal.budiluhur.ac.id/telematika/article/viewFile/4348/pdf>
- [10] D. F. Zhafira, B. Rahayudi, and Indriati, "ANALISIS SENTIMEN KEBIJAKAN KAMPUS MERDEKA MENGGUNAKAN NAIVE BAYES DAN PEMBOBOTAN TF-IDF BERDASARKAN KOMENTAR PADA YOUTUBE," *J. Sist. Informasi, Teknol. Informasi, dan Edukasi Sist. Inf.*, vol. 2, no. 1, pp. 55–63, 2021, [Online]. Available: <https://just-si.ub.ac.id/index.php/just-si/en/article/view/24/17>
- [11] M. Nurizki, W. Apriandari, and Asriyanik, "Algoritma Naïve Bayes untuk Rekomendasi Seleksi Peserta Paskibraka Berbasis Website," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1486–1493, 2023, doi: 10.47065/josh.v4i4.3574.
- [12] M. B. Alfayyadh, S. Assegaff, and Fachruddin, "Perbandingan Algoritma C4.5 Dan Naïve Bayes Dalam Machine Learning Untuk Klasifikasi Performa Pelajar," *J. Manaj. Teknol. dan Sist. Inf.*,



- vol. 5, no. 2, pp. 1095–1104, 2025, [Online]. Available: https://ejournal.unama.ac.id/index.php/jms/article/download/2328/1665/12194?__cf_chl_tk=P_hnYQMld1U5SoP.Z41PjwRTInHNAZp9ztTG.RdhKbX8-1785478243-1.0.1.1-fXslfM2xCfNm_L5uzbOA.fFAmV0X73QRjGwbvBtrepE
- [13] R. A. Riyanto, Y. Cahyana, and Rahmat, “PENERAPAN ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN CRYPTOCURRENCY PADA MEDIA SOSIAL TWITTER,” *Sci. Student J. Information, Technol. Sci.*, vol. V, no. 1, pp. 104–111, 2024.
- [14] S. Setyabudi and E. Aryanny, “Analisis Sentimen Penilaian Pengguna Marketplace Lazada Dengan Metode Naive Bayes Dan Support Vector Machine,” *INOVTEK POLBENG-SERI Inform.*, vol. 10, no. 1, pp. 422–433, 2025.