HOAQ: JURNAL TEKNOLOGI INFORMASI, Volume 17 Nomor 2, Desember 2026, p373-381
p-ISSN: 2337-5280, e-ISSN: 26207427
DOI: https://doi.org/10.52972/hoaq.vol17no2

PENGEMBANGAN *WEBSITE* REKOMENDASI JUDUL DAN DRAF ARTIKEL KEGIATAN KOMUNIO PALEMBANG MENGGUNAKAN TF-IDF, NAIVE BAYES DAN COSINE SIMILARITY

Yulianus Febry Tri Nugroho[1], Marselinus Dewadaru Bayu Adeodatus[2], Sepiyana Nopitasari[3], Latius Hermawan[4]

[1–4] Program Studi Informatika, Universitas Katolik Musi Charitas
Jl. Bangau No.60, 9 Ilir, Palembang - Sumatera Selatan, Indonesia
Email: [1]frytn13@gmail.com, [2]marselinusbayu145@gmail.com, [3]sepiyana19@gmail.com, [4]tiuz.hermawan@ukmc.ac.id

ABSTRAK

KOMUNIO Palembang telah memiliki arsip artikel digital, tetapi sistem publikasi yang berjalan belum menyediakan dukungan editorial untuk mengubah deskripsi kegiatan baru menjadi prediksi kategori, artikel rujukan, rekomendasi judul, dan draf awal dalam satu alur. Penelitian ini mengembangkan dan mengevaluasi pipeline editorial berbasis TF-IDF, Multinomial Naive Bayes, dan Cosine Similarity pada 2.991 artikel yang dipetakan ke empat kategori. Data dibagi 80:20 dengan *random state* 42 dan *oversampling* diterapkan hanya pada data latih. Model final memperoleh *accuracy* 83,97%, *macro precision* 81,97%, *macro recall* 82,11%, dan *macro* F1-*score* 81,95%. Performa antarkelas tidak seragam: Liturgi mencapai F1-*score* 99,6%, sedangkan Berita dan Kegiatan Gereja hanya 57,0%. Sebanyak 51 dari 114 artikel Berita dan Kegiatan Gereja (44,7%) diprediksi sebagai Orang Muda dan Anak Remaja, yang menunjukkan tumpang tindih kosakata dan batas label. Dibanding *majority-class baseline, accuracy* meningkat dari 38,06% menjadi 83,97%. Pada pengujian rekomendasi, Cosine *Similarity* tertinggi berada sekitar 22–26%, sehingga skor ditafsirkan sebagai kedekatan leksikal relatif untuk pemeringkatan rujukan, bukan kesamaan semantik mutlak. Kontribusi penelitian terletak pada integrasi klasifikasi dan pencarian rujukan ringan untuk bantuan editorial berbahasa Indonesia, disertai analisis kesalahan kelas dan batas interpretasi rekomendasi.

Kata kunci: KOMUNIO, TF-IDF, Naive Bayes, Cosine Similarity, NLP

ABSTRACT

KOMUNIO Palembang maintains a digital article archive, but the existing publication system does not provide an integrated editorial workflow that transforms a new activity description into category prediction, reference-article retrieval, title recommendation, and an initial draft. This study develops and evaluates an editorial pipeline based on TF-IDF, Multinomial Naive Bayes, and Cosine Similarity using 2,991 articles mapped into four categories. The data were split 80:20 with a random state of 42, and oversampling was applied only to the training data. The final model achieved 83.97% accuracy, 81.97% macro precision, 82.11% macro recall, and 81.95% macro F1-score. Class performance was heterogeneous: Liturgy reached an F1-score of 99.6%, whereas Church News and Activities reached only 57.0%. Fifty-one of 114 Church News and Activities articles (44.7%) were predicted as Youth and Children, indicating lexical overlap and non-exclusive class boundaries. Compared with a majority-class baseline, accuracy increased from 38.06% to 83.97%. The highest Cosine Similarity scores in the recommendation test were approximately 22–26%; therefore, the scores are interpreted as relative lexical proximity for ranking references rather than absolute semantic equivalence. The contribution is an integrated, lightweight Indonesian editorial-assistance workflow combining classification and content retrieval with class-level error analysis and a conservative interpretation of recommendation scores.

Keywords: KOMUNIO, TF-IDF, Naive Bayes, Cosine Similarity, NLP

1. PENDAHULUAN

Pengelolaan media digital berbasis komunitas tidak berhenti pada proses menerbitkan artikel. Admin perlu mengubah dokumentasi kegiatan menjadi tulisan yang konsisten, menentukan kategori, menyusun judul, dan mencari arsip lama yang relevan. Pada KOMUNIO Palembang, sistem publikasi yang tersedia telah mendukung penyimpanan, pengelompokan, dan penyajian artikel yang sudah selesai, tetapi belum menyediakan satu alur yang secara otomatis mengolah deskripsi kegiatan baru menjadi prediksi kategori, artikel rujukan, rekomendasi judul, dan draf awal. Karena itu, pencarian arsip dan keputusan awal penulisan masih bergantung pada admin.

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi

Penelitian terdahulu mengenai *content-based filtering* umumnya berfokus pada pemeringkatan konten bagi pengguna berdasarkan kemiripan teks, seperti rekomendasi berita, video pembelajaran, dan film [1], [2], [3], [4]. Sementara itu, penelitian klasifikasi teks berbahasa Indonesia menunjukkan bahwa TF-IDF dan Multinomial Naive Bayes dapat digunakan sebagai pendekatan ringan untuk pemetaan dokumen ke kategori [5], [6], [7]. Pengolahan teks Indonesia juga memerlukan preprocessing dan evaluasi yang konsisten agar model tidak hanya menghasilkan metrik agregat, tetapi dapat dianalisis per kelas [8], [9], [10], [11], [12].

Research gap penelitian ini berada pada fungsi dan integrasi metode. Sistem rekomendasi berbasis konten pada penelitian terdahulu lebih banyak menempatkan artikel sebagai objek yang direkomendasikan kepada pembaca, sedangkan penelitian ini memanfaatkan arsip sebagai rujukan kerja admin yang sedang menulis artikel baru. Klasifikasi kategori dan pencarian dokumen serupa juga lazim dievaluasi sebagai tugas terpisah. Penelitian ini menghubungkan keduanya dalam satu pipeline editorial dan mengevaluasi perilaku model pada domain artikel gerejawi yang memiliki kelas saling beririsan, terutama Berita dan Kegiatan Gereja dengan Orang Muda dan Anak Remaja.

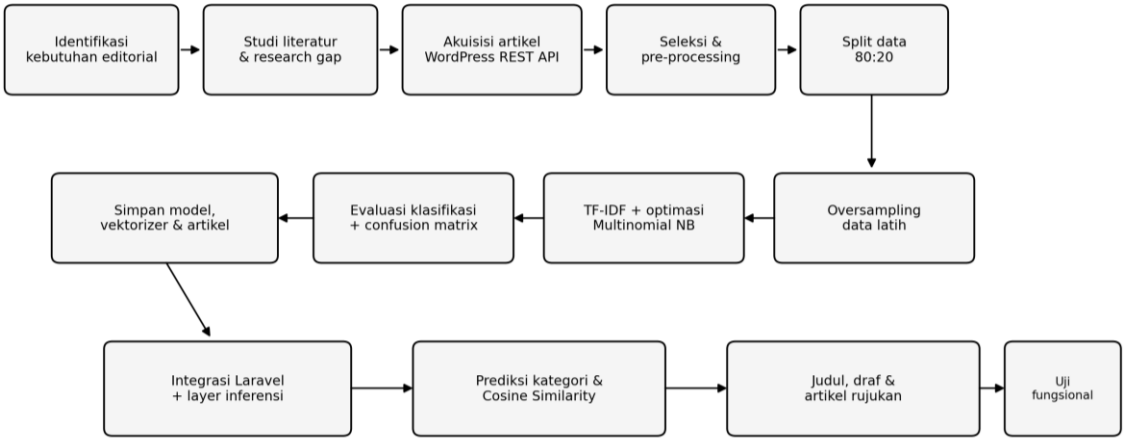
Pemilihan metode disesuaikan dengan ukuran korpus dan kebutuhan implementasi. TF-IDF membentuk representasi sparse yang transparan dan efisien pada ribuan dokumen, serta dapat dipakai untuk klasifikasi dan pengukuran kemiripan [5], [13], [14]. Multinomial Naive Bayes sesuai untuk fitur non-negatif berbasis frekuensi atau bobot kata, mempunyai biaya pelatihan dan inferensi rendah, dan telah digunakan pada klasifikasi teks Indonesia [5], [6], [7]. Cosine Similarity dapat membandingkan arah dua vektor TF-IDF sehingga panjang dokumen tidak menjadi satu-satunya penentu kemiripan [1], [2], [4], [15]. Dalam penelitian ini, *Cosine Similarity* diperlakukan sebagai ukuran kedekatan leksikal untuk pemeringkatan rujukan, bukan sebagai ukuran kesamaan semantik mutlak.

Kontribusi ilmiah penelitian ini bukan berupa algoritma baru. Kontribusinya adalah rancangan dan evaluasi *pipeline* editorial ringan yang menggabungkan representasi TF-IDF, klasifikasi probabilistik, dan pencarian rujukan berbasis konten pada arsip nyata berbahasa Indonesia. Kontribusi empiris meliputi evaluasi per kelas, analisis *confusion matrix* untuk menjelaskan batas kelas, perbandingan dengan *majority-class baseline*, pembahasan pengaruh *oversampling* secara hati-hati, serta interpretasi nilai *Cosine Similarity* yang rendah. Dari sisi implementasi, proses pelatihan dipisahkan dari penggunaan model pada *website* sehingga artefak final dapat digunakan kembali tanpa melakukan *training* pada setiap permintaan.

Berdasarkan *gap* tersebut, penelitian ini bertujuan mengembangkan dan mengevaluasi *website* rekomendasi judul dan draf artikel kegiatan KOMUNIO Palembang menggunakan TF-IDF, Multinomial Naive Bayes, dan *Cosine Similarity*. Evaluasi diarahkan pada kualitas klasifikasi, pola kesalahan antarkelas, fungsi pencarian artikel rujukan, dan keterlaksanaan alur utama *website* sebagai alat bantu awal bagi admin.

2. METODE PENELITIAN

Penelitian dilakukan melalui identifikasi kebutuhan editorial, studi literatur, akuisisi dan seleksi artikel, *pre-processing*, pembagian data, pelatihan model, evaluasi klasifikasi, penyimpanan artefak, integrasi dengan *website*, pengujian fungsional, dan analisis hasil. Gambar 1 memperlihatkan *research pipeline* dengan pemisahan antara proses pelatihan dan proses penggunaan model.



Gambar 1. *Research Pipeline* Penelitian

Pengumpulan Data

Dataset diperoleh dari artikel berstatus terbit pada *website* KOMUNIO melalui WordPress REST API. Artikel pada korpus mencakup publikasi dari Maret 2022 hingga Juni 2026. *Snapshot* yang menjadi dasar dokumen eksperimen dibentuk pada 10 Juni 2026; tanggal tersebut digunakan sebagai tanggal akses API dalam pelaporan penelitian karena bertepatan dengan pembentukan dokumen eksperimen dan batas artikel terbaru yang terlihat pada

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi

snapshot. *Field* yang tersedia mencakup id artikel, id WordPress, tanggal terbit, judul, isi, ringkasan, tautan, id kategori, nama kategori, label kategori, sumber, teks gabungan, dan metadata pendukung.

Kriteria inklusi meliputi artikel berstatus publish yang dapat diakses melalui API, memiliki judul dan konten teks, dan kategori asalnya dapat dipetakan ke salah satu dari empat kelas penelitian. Artikel dikeluarkan apabila judul atau isi tidak tersedia, kategori tidak dapat dipetakan ke empat kelas, atau teks tidak lolos penyaringan konten kosong/terlalu pendek pada tahap preprocessing. Naskah eksperimen awal tidak mencatat nilai ambang numerik untuk kriteria "terlalu pendek", sehingga artikel ini tidak menambahkan angka yang tidak dapat diverifikasi.

Tabel 1. Struktur Dataset Artikel KOMUNIO

No	Field	Keterangan
1	*title*	Judul artikel yang telah terbit.
2	*content*	Isi artikel dari *field* konten WordPress setelah pembersihan HTML.
3	*excerpt*	Ringkasan atau potongan artikel dari metadata WordPress.
4	*date*	Tanggal publikasi artikel.
5	*category_name*	Kategori asli yang berasal dari WordPress.
6	*category_label*	Label hasil pemetaan ke empat kelas untuk target klasifikasi.
7	*link*	Tautan sumber artikel pada website KOMUNIO.
8	*text*	Teks gabungan yang disiapkan untuk kebutuhan pemodelan.

***Pre-processing* dan pembentukan label**

Tahap *pre-processing* menghapus elemen HTML, tag gambar dan kutipan yang tidak digunakan sebagai teks, menormalkan spasi serta karakter, menyaring teks kosong atau terlalu pendek, lalu memetakan kategori WordPress menjadi empat kelas utama: Liturgi, Renungan, Orang Muda dan Anak Remaja, serta Berita dan Kegiatan Gereja. Tahap pembersihan seperti normalisasi, tokenisasi, stopword removal, dan stemming merupakan praktik yang banyak digunakan pada pengolahan teks Indonesia [9], [10], [11].

Pada naskah eksperimen awal, *field text* dideskripsikan sebagai gabungan judul, isi, ringkasan, dan informasi kategori, sedangkan *category_label* digunakan sebagai target. Karena informasi kategori dapat berkorelasi langsung dengan target, konfigurasi tersebut dicatat sebagai potensi target *leakage* dan menjadi keterbatasan replikasi. Metrik pada artikel ini dilaporkan sesuai hasil eksperimen yang tersedia, tanpa mengubah *pipeline* secara retrospektif. Replikasi berikutnya harus memastikan *field* kategori hanya berperan sebagai metadata/target dan tidak masuk ke vektor TF-IDF.

Dataset final berjumlah 2.991 artikel. Liturgi menjadi kelas terbesar dengan 1.138 artikel, sedangkan Renungan menjadi kelas terkecil dengan 408 artikel. Rasio kelas terbesar terhadap kelas terkecil sekitar 2,79:1. Distribusi lengkap ditunjukkan pada Tabel 2.

Tabel 2. Distribusi Kategori Final

No	Kategori	Jumlah	Proporsi
1	Liturgi	1.138	38,05%
2	Orang Muda dan Anak Remaja	874	29,22%
3	Berita dan Kegiatan Gereja	571	19,09%
4	Renungan	408	13,64%

Ekstraksi fitur TF-IDF

Teks artikel direpresentasikan menjadi vektor numerik menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Bobot term meningkat ketika term sering muncul dalam satu dokumen tetapi tidak terlalu umum pada keseluruhan korpus. Representasi ini dipilih karena sederhana, dapat dijelaskan, dan cocok untuk ruang fitur teks yang *sparse* [5], [13], [14]. Rumus yang digunakan dituliskan sebagai berikut.

$$TF\text{-}IDF(t,d) = TF(t,d) \times \log(N / DF(t))$$

Pada rumus tersebut, t adalah term, d adalah dokumen, N adalah jumlah dokumen dalam korpus, dan DF(t) adalah jumlah dokumen yang memuat term t. Vektor TF-IDF digunakan sebagai masukan Multinomial Naive Bayes dan sebagai ruang representasi untuk perhitungan *Cosine Similarity*.

Klasifikasi Multinomial Naive Bayes dan Alpha

Multinomial Naive Bayes digunakan untuk memprediksi kategori artikel dari fitur TF-IDF. Parameter alpha menerapkan additive smoothing agar probabilitas term yang jarang atau belum muncul pada suatu kelas tidak menjadi nol. Secara umum, probabilitas term t pada kelas c dapat dituliskan sebagai berikut.

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi

$$P(t|c) = (N_{tc} + \alpha) / (N_c + \alpha|V|)$$

N_tc menyatakan bobot atau jumlah kemunculan term t pada kelas c, N_c menyatakan total bobot term pada kelas c, |V| adalah ukuran *vocabulary*, dan α adalah parameter smoothing. Artefak model final yang tersedia bernama MultinomialNB_alpha_0_10 sehingga nilai alpha yang dapat diverifikasi adalah 0,10. Naskah awal menyebut model sebagai optimized Naive Bayes, tetapi log kandidat alpha dan skema validasi selama pemilihan parameter tidak tersimpan pada dokumen penelitian. Karena itu, artikel ini tidak mencantumkan grid alpha atau jumlah *fold* yang tidak dapat diverifikasi. Nilai 0,10 dilaporkan sebagai konfigurasi final yang benar-benar tersimpan dan dievaluasi.

Dataset dibagi menjadi 80% data latih (2.392 artikel) dan 20% data uji (599 artikel) dengan *random state* 42. Oversampling dilakukan setelah pembagian data dan hanya pada data latih untuk menambah representasi kelas minoritas; data uji tetap menggunakan distribusi asli. Naskah eksperimen yang tersedia tidak menyimpan nama operator, parameter, atau jumlah akhir sampel setelah *oversampling*. Oleh sebab itu, metode dilaporkan pada tingkat prosedural sebagai oversampling pada data latih dan tidak diberi nama operator spesifik yang tidak tercatat pada eksperimen. Pengaruhnya dibahas pada bagian hasil tanpa klaim kausal yang melampaui data eksperimen.

Tabel 3. Konfigurasi Pemodelan

Parameter	Nilai
Metode	TF-IDF + Multinomial Naive Bayes + *Cosine Similarity*
Jumlah *dataset*	2.991 artikel
Jumlah kategori	4 kategori
Data latih	2.392 artikel (80%)
Data uji	599 artikel (20%)
Random state	42
Alpha model final	0,10
Oversampling	Diterapkan hanya pada data latih setelah split; data uji tidak diubah

Pencarian Artikel Rujukan dengan *Cosine Similarity*

Setelah kategori diprediksi, deskripsi kegiatan ditransformasikan ke ruang TF-IDF dan dibandingkan dengan vektor artikel pada korpus. *Cosine Similarity* mengukur sudut antara dua vektor dan digunakan untuk mengurutkan artikel dari yang paling dekat secara leksikal [1], [2], [4], [15]. Rumusnya adalah sebagai berikut.

$$cos(A,B) = (A \cdot B) / (||A|| \; ||B||)$$

A adalah vektor input kegiatan, B adalah vektor artikel, A·B adalah dot product, sedangkan ||A|| dan ||B|| adalah norma masing-masing vektor. Skor tidak diperlakukan sebagai probabilitas relevansi. Karena penelitian belum memiliki label relevansi manual untuk setiap hasil top-k, *Cosine Similarity* digunakan sebagai fungsi pemeringkatan kandidat rujukan.

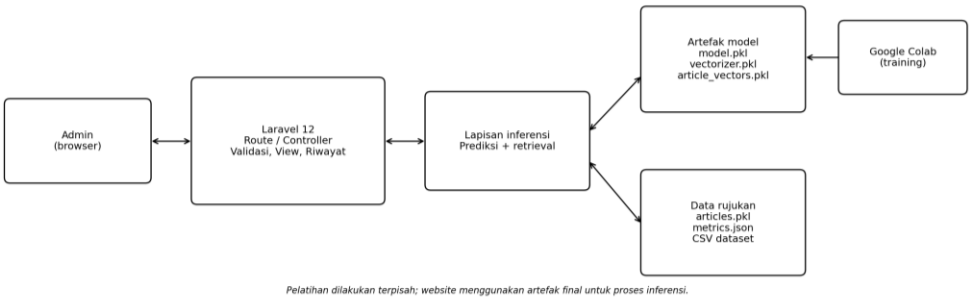
Pembuatan Rekomendasi Judul dan Draf Artikel

Rekomendasi judul dibentuk dari nama kegiatan, lokasi, kategori hasil prediksi, dan kata kunci deskripsi. Draf artikel disusun menggunakan template berita kegiatan yang memuat penyelenggara, waktu/lokasi, peserta, tujuan, rangkaian kegiatan, dan poin penting. Keluaran ini merupakan draf awal yang dapat diedit oleh admin dan tidak dimaksudkan untuk menggantikan keputusan editorial.

Arsitektur Integrasi *Website* dan Model

Website menggunakan Laravel 12 sebagai aplikasi utama, sedangkan pelatihan model dilakukan terpisah di Google Colab. Artefak final disimpan pada storage/app/ml, yaitu model.pkl, vectorizer.pkl, article_vectors.pkl, articles.pkl, metrics.json, dan dataset_artikel_komunio_final.csv. Pemisahan ini membuat website menggunakan hasil training yang sudah tersedia dan tidak melakukan pelatihan ulang pada setiap permintaan. Secara logis, Laravel menangani input, validasi, tampilan, dan riwayat, sedangkan lapisan inferensi menggunakan artefak model untuk prediksi kategori dan pencarian artikel rujukan. Dokumen implementasi yang tersedia tidak mencatat apakah komunikasi Laravel ke lapisan inferensi dilakukan melalui subprocess, HTTP API, atau mekanisme transport lain. Karena itu, Gambar 2 menyajikan arsitektur pada tingkat fungsional tanpa mengasumsikan protokol *transport* tertentu.

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi



Gambar 2. Arsitektur Fungsional Integrasi *Website* dan Artefak Model

Prosedur Pengujian

Evaluasi dilakukan pada dua lapisan. Lapisan pertama mengukur klasifikasi pada 599 data uji menggunakan *accuracy, precision, recall*, F1-*score*, dan *confusion matrix*. Lapisan kedua menggunakan pengujian *black-box* terhadap fungsi utama *website*. Pengujian berfokus pada keluaran yang dapat diamati dari antarmuka: pemuatan status model, pemrosesan satu input kegiatan pada generator, penampilan artikel rujukan, dan penyimpanan draf. Pendekatan ini menyediakan lebih dari satu kasus fungsional tanpa membuat skenario yang tidak tercatat pada implementasi.

Tabel 4. Skenario Pengujian Fungsional *Website*

No	Kasus uji	Keluaran yang diamati	Hasil
1	Membuka *Dashboard*/Status Model	Jumlah *dataset* 2.991, status model siap, dan *accuracy* sekitar 84% tampil	Berhasil
2	Mengirim input kegiatan pada Generator	Prediksi kategori, rekomendasi judul, dan draf awal tampil	Berhasil
3	Menampilkan artikel rujukan	Daftar artikel diurutkan berdasarkan nilai *similarity* dan memuat tautan sumber	Berhasil
4	Menyimpan draf hasil generator	Draf tersimpan dan jumlah draf pada sistem bertambah/tersedia di riwayat	Berhasil

3. HASIL DAN PEMBAHASAN

Hasil Klasifikasi

Model final Multinomial Naive Bayes dengan alpha 0,10 dievaluasi pada 599 artikel data uji. Model memperoleh *accuracy* 83,97%, *macro precision* 81,97%, macro recall 82,11%, dan *macro* F1-*score* 81,95%. Macro F1-*score* penting pada dataset ini karena memberi bobot yang sama pada setiap kelas, sehingga kelas Liturgi yang berjumlah paling besar tidak mendominasi interpretasi performa.

Tabel 5. Ringkasan Hasil Evaluasi Model

Metrik	Nilai
Accuracy	83,97%
Macro precision	81,97%
Macro recall	82,11%
Macro F1-score	81,95%
Weighted precision	83,60%
Weighted recall	83,97%
Weighted F1-*score*	83,71%

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi

Tabel 6. Hasil Evaluasi Per Kategori

Kategori	*Precision*	*Recall*	F1-*score*	*Support*
Berita dan Kegiatan Gereja	61,0%	53,5%	57,0%	114
Liturgi	99,6%	99,6%	99,6%	228
Orang Muda dan Anak Remaja	72,0%	76,6%	74,2%	175
Renungan	95,3%	98,8%	97,0%	82

Perbandingan dengan *Baseline*

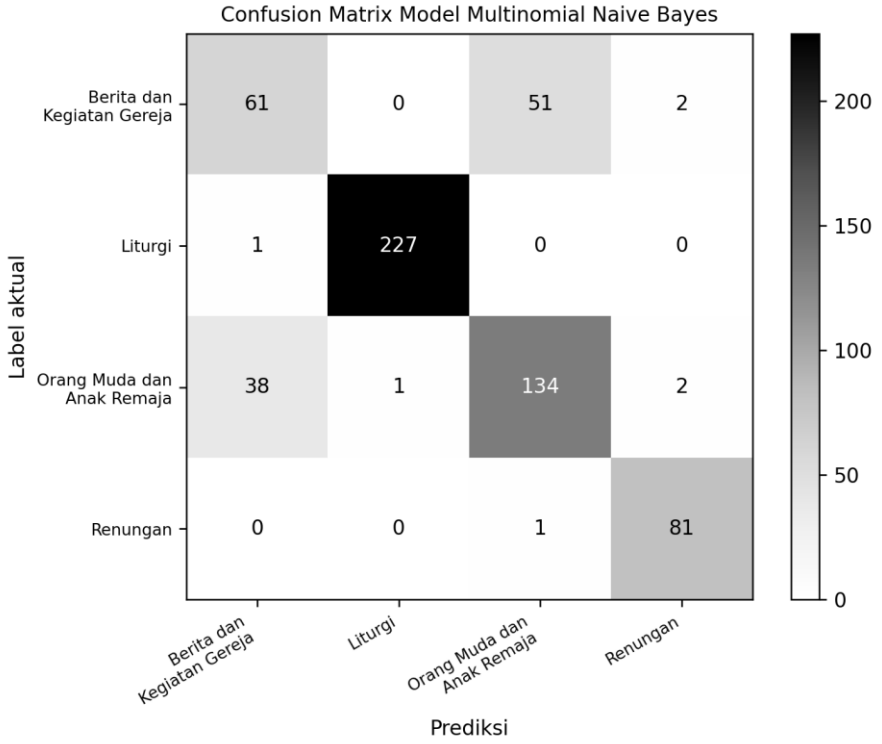
Untuk memastikan bahwa *accuracy* 83,97% tidak hanya mengikuti dominasi kelas terbesar, model dibandingkan dengan *majority-class baseline* yang selalu memprediksi Liturgi. Dari 599 data uji, 228 artikel berlabel Liturgi, sehingga baseline mencapai accuracy 38,06%. Berdasarkan distribusi prediksi tunggal tersebut, macro F1-*score baseline* adalah 13,78%. Model final meningkatkan *accuracy* sebesar 45,91 poin persentase dan macro F1-*score* sebesar 68,17 poin persentase. Perbandingan ini menunjukkan bahwa model mempelajari pola diskriminatif di luar proporsi kelas, meskipun *baseline* mayoritas belum menggantikan kebutuhan perbandingan dengan algoritma lain pada eksperimen lanjutan.

Tabel 7. Perbandingan Model Final dengan Majority-Class Baseline

Model	Accuracy	Macro F1-score
Majority-class baseline (selalu Liturgi)	38,06%	13,78%
TF-IDF + Multinomial Naive Bayes (alpha 0,10)	83,97%	81,95%
Selisih	+45,91 poin	+68,17 poin

Analisis Confusion Matrix dan Kesalahan Klasifikasi

Performa tertinggi terjadi pada Liturgi dan Renungan. Liturgi menghasilkan 227 prediksi benar dari 228 data uji, sedangkan Renungan menghasilkan 81 prediksi benar dari 82 data uji. Kedua kelas mempunyai pola bahasa yang lebih khas: artikel Liturgi banyak memuat struktur bacaan, nama kitab, warna liturgi, atau perayaan liturgis, sedangkan Renungan cenderung memiliki struktur reflektif yang konsisten. Kekhasan kosakata tersebut membuat fitur TF-IDF lebih mudah dipisahkan.



Gambar 3. *Confusion Matrix Model* pada 599 Data Uji

Kelas Berita dan Kegiatan Gereja menjadi sumber kesalahan terbesar. Dari 114 artikel aktual pada kelas tersebut, 61 diprediksi benar, sedangkan 51 artikel atau 44,7% diprediksi sebagai Orang Muda dan Anak Remaja.

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi

Kesalahan juga terjadi pada arah sebaliknya: 38 dari 175 artikel Orang Muda dan Anak Remaja atau 21,7% diprediksi sebagai Berita dan Kegiatan Gereja. Pola dua arah ini menunjukkan bahwa masalah tidak hanya berkaitan dengan jumlah data, tetapi juga dengan batas kelas yang secara semantik tidak sepenuhnya eksklusif.

Label Berita dan Kegiatan Gereja mencakup konten yang luas, misalnya pertemuan paroki, pelayanan sosial, perayaan komunitas, pelantikan, atau kegiatan kelompok. Artikel Orang Muda dan Anak Remaja pada dasarnya juga dapat berbentuk berita kegiatan gereja, tetapi pembeda utamanya adalah aktor atau komunitas yang terlibat. Akibatnya, kedua kelas berbagi kata seperti kegiatan, umat, paroki, peserta, pelayanan, perayaan, gereja, dan komunitas. TF-IDF menilai overlap kata, sedangkan Multinomial Naive Bayes menyederhanakan hubungan antarfitur. Informasi kontekstual seperti "siapa pelaku utama kegiatan" tidak selalu tertangkap secara eksplisit, sehingga dua kelas tersebut lebih mudah tertukar dibanding Liturgi dan Renungan.

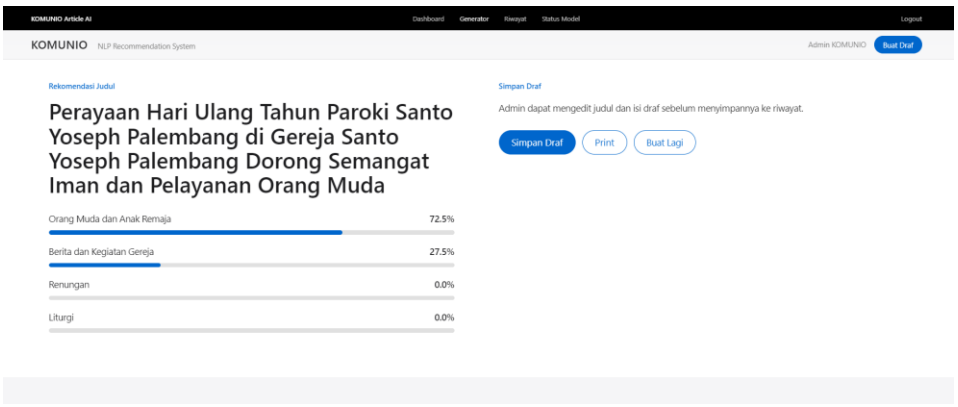
Pengaruh *Oversampling*

Oversampling diterapkan karena distribusi awal tidak seimbang, dengan rasio kelas terbesar terhadap kelas terkecil sekitar 2,79:1. Data uji tidak di-*oversample* sehingga metrik tetap dihitung pada distribusi asli. Hasil menunjukkan bahwa kelas terkecil, Renungan, dapat mencapai *recall* 98,8%, tetapi Berita dan Kegiatan Gereja tetap memiliki recall 53,5%. Temuan ini menunjukkan bahwa penambahan representasi kelas minoritas tidak otomatis menyelesaikan tumpang tindih label. Karena eksperimen identik tanpa *oversampling* tidak tersedia, peningkatan metrik tidak dapat diatribusikan secara kausal kepada *oversampling*. Analisis ablation dengan dan tanpa *oversampling* diperlukan untuk mengukur pengaruhnya secara terpisah.

Hasil Implementasi dan Pengujian Rekomendasi

Pengujian fungsional pada Tabel 4 menunjukkan bahwa alur utama sistem dapat dijalankan dari antarmuka web. *Dashboard* menampilkan status artefak model dan metrik, Generator menerima deskripsi kegiatan dan menghasilkan prediksi kategori serta draf, modul rujukan menampilkan artikel berdasarkan kemiripan, dan hasil generator dapat disimpan sebagai draf. Jumlah gambar antarmuka pada artikel ini dibatasi agar pembahasan berfokus pada temuan model; Gambar 4 digunakan sebagai contoh keluaran yang paling representatif.

Kasus uji utama menggunakan input "Perayaan Hari Ulang Tahun Paroki Santo Yoseph Palembang". Sistem memprediksi Orang Muda dan Anak Remaja sebesar 72,5% dan Berita dan Kegiatan Gereja sebesar 27,5%. Rekomendasi judul yang dihasilkan adalah "Perayaan Hari Ulang Tahun Paroki Santo Yoseph Palembang di Gereja Santo Yoseph Palembang Dorong Semangat Iman dan Pelayanan Orang Muda". Hasil ini memperlihatkan batas model karena kegiatan paroki umum dapat memperoleh probabilitas tinggi pada kelas orang muda ketika deskripsinya memuat fitur leksikal yang diasosiasikan dengan kelompok tersebut. Pola ini konsisten dengan confusion matrix dan menegaskan perlunya koreksi kategori oleh admin ketika konteks kegiatan bersifat lintas kelompok.



Gambar 4. Contoh Hasil Rekomendasi Judul dan Probabilitas Kategori

Tabel 8. Contoh Hasil Pengujian Rekomendasi

Komponen	Hasil
Input kegiatan	Perayaan Hari Ulang Tahun Paroki Santo Yoseph Palembang
Prediksi tertinggi	Orang Muda dan Anak Remaja (72,5%)
Kategori kedua	Berita dan Kegiatan Gereja (27,5%)
Artikel mirip tertinggi	Perayaan Ekaristi dan Pagelaran Seni Tandai Syukur HUT Ke-8 Gedung Serbaguna Santo Louis Palembang (26,07%)

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi

Interpretasi Nilai *Cosine Similarity*

Artikel rujukan pada kasus uji memiliki *Cosine Similarity* tertinggi 26,07%, sedangkan beberapa kandidat lain berada sekitar 22–26%. Nilai tersebut rendah apabila dibaca sebagai tingkat kesamaan absolut, tetapi *Cosine Similarity* pada TF-IDF bukan ukuran "akurasi rekomendasi". Skor merepresentasikan overlap arah vektor berdasarkan term yang sama. Input kegiatan cenderung lebih pendek dan terstruktur, sedangkan artikel arsip lebih panjang serta memuat nama tokoh, lokasi, tanggal, kutipan, dan detail lain yang tidak selalu muncul pada input. Perbedaan kosakata tersebut menurunkan overlap leksikal [11], [15].

Dalam kondisi ini, fungsi yang dapat dipertanggungjawabkan adalah pemeringkatan relatif: artikel dengan skor tertinggi ditampilkan sebagai kandidat rujukan awal untuk membantu admin melihat konteks dan gaya penulisan. Penelitian belum memiliki *ground truth* relevansi dari editor, sehingga tidak dapat menyatakan bahwa skor 26% merupakan ambang artikel relevan. Evaluasi berikutnya perlu meminta penilaian editor terhadap top-k hasil dan menggunakan metrik *information retrieval* seperti Precision@k, Mean Reciprocal Rank, atau nDCG@k.

Keterbatasan Dataset dan Model

Penelitian memiliki beberapa keterbatasan. Dataset hanya berasal dari satu media sehingga gaya bahasa, kebijakan kategori, dan pola editorial KOMUNIO membatasi generalisasi ke media lain. Distribusi kelas tidak seimbang dan label Berita dan Kegiatan Gereja beririsan dengan Orang Muda dan Anak Remaja. TF-IDF bersifat *bag-of-words* sehingga urutan kata dan konteks semantik tidak dimodelkan secara mendalam, sedangkan *Multinomial* Naive Bayes menggunakan asumsi independensi fitur yang menyederhanakan hubungan antarkata. Pengaruh oversampling belum dipisahkan melalui eksperimen ablation, dan evaluasi rekomendasi belum memiliki label relevansi editor.

Dari sisi reproduksibilitas, naskah eksperimen awal tidak menyimpan operator *oversampling, grid/validation log* untuk optimasi alpha, maupun protokol *transport* Laravel ke lapisan inferensi. Selain itu, deskripsi awal field text menyebut informasi kategori sebagai bagian teks gabungan, sehingga terdapat potensi target leakage yang perlu diverifikasi pada notebook atau script asli. Keterbatasan tersebut tidak ditutup dengan asumsi baru pada revisi ini; metrik dilaporkan sebagaimana artefak eksperimen yang tersedia dan klaim dibatasi pada hal yang dapat diverifikasi.

4. SIMPULAN

Penelitian menghasilkan pipeline bantuan editorial yang menggabungkan TF-IDF, Multinomial Naive Bayes, dan Cosine Similarity untuk memanfaatkan 2.991 arsip artikel KOMUNIO Palembang. Pada 599 data uji, model final memperoleh accuracy 83,97% dan macro F1-score 81,95%, lebih tinggi daripada *majority-class baseline* dengan *accuracy* 38,06% dan *macro* F1-*score* 13,78%. Temuan utama terletak pada ketimpangan performa antarkelas. Liturgi dan Renungan dapat dipisahkan dengan sangat baik, sedangkan Berita dan Kegiatan Gereja hanya mencapai F1-*score* 57,0% dan 44,7% datanya salah diprediksi sebagai Orang Muda dan Anak Remaja. Pola tersebut menunjukkan bahwa tumpang tindih definisi label dan kosakata menjadi sumber kesalahan penting pada model.

Nilai *Cosine Similarity* sekitar 22–26% menunjukkan bahwa modul rujukan lebih tepat digunakan untuk pemeringkatan kedekatan leksikal daripada sebagai ukuran kesamaan semantik absolut. Kontribusi penelitian terletak pada penggabungan klasifikasi kategori, pencarian rujukan, rekomendasi judul, dan draf awal dalam alur editorial web yang ringan, disertai analisis kesalahan per kelas, baseline, dan batas interpretasi skor rekomendasi. Pengembangan berikutnya perlu memperjelas skema label, memastikan fitur kategori tidak menyebabkan target leakage, mendokumentasikan seluruh hyperparameter dan mekanisme integrasi, menguji model tanpa oversampling sebagai ablation, membandingkan representasi semantik, serta menilai relevansi top-k menggunakan penilaian editor.

DAFTAR PUSTAKA

[1] I. A. W. Nandita, I. M. A. D. Suarjaya, and I. P. A. Bayupati, "News Recommendation System Using Content-Based Filtering through RSS Customization Service," Journal of Applied Informatics and Computing, vol. 9, no. 4, pp. 1858–1867, 2025, doi: https://doi.org/10.30871/jaic.v9i4.9807.

[2] G. Yunanda, D. Nurjanah, and S. Meliana, "Recommendation System from Microsoft News Data using TF-IDF and Cosine Similarity Methods," Building of Informatics, Technology and Science, vol. 4, no. 1, pp. 277–284, 2022, doi: https://doi.org/10.47065/bits.v4i1.1670.

[3] F. Ramadhan and A. Musdholifah, "Online Learning Video Recommendation System Based on Course and Sylabus Using Content-Based Filtering," Indonesian Journal of Computing and Cybernetics Systems, vol. 15, no. 3, pp. 265–274, 2021, doi: https://doi.org/10.22146/ijccs.65623.

[4] A. H. J. P. Hiro, J. Permana, and A. T. Wibowo, "Movie Recommendation System Based on Synopsis Using Content-Based Filtering with TF-IDF and Cosine Similarity," International Journal on Information and Communication Technology, vol. 9, no. 2, pp. 1–14, 2023, doi: https://doi.org/10.21108/ijoict.v9i2.747.

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi

[5] A. Gunawan and H. Bunyamin, "Perbandingan Algoritma Machine Learning Umum Berbasis TF-IDF untuk Klasifikasi Artikel Bahasa Indonesia," Journal of Smart Technology and Engineering, vol. 1, no. 1, 2025, doi: https://doi.org/10.28932/jste.v1i1.13161.

[6] M. Luthfi, F. Martanto, and W. Istiono, "Sentiment Analysis of M-Paspor App Reviews Using Multinomial Naive Bayes," Journal of Logistics, Informatics and Service Science, vol. 11, no. 10, pp. 311–326, 2024, doi: https://doi.org/10.33168/JLISS.2024.1017.

[7] A. Sabrani, I. G. W. Wedashwara W., and F. Bimantoro, "Multinomial Naïve Bayes untuk Klasifikasi Artikel Online tentang Gempa di Indonesia," Jurnal Teknologi Informasi, Komputer, dan Aplikasinya, vol. 2, no. 1, pp. 89–100, 2020, doi: https://doi.org/10.29303/jtika.v2i1.87.

[8] M. F. Kono, I. N. Fajri, and Y. Pristyanto, "Public Sentiment Analysis on Corruption Issues in Indonesia Using IndoBERT Fine-Tuning, Logistic Regression, and Linear SVM," Journal of Applied Informatics and Computing, vol. 9, no. 5, pp. 2616–2628, 2025, doi: https://doi.org/10.30871/jaic.v9i5.10537.

[9] M. A. Rosid, A. S. Fitrani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, "Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi," IOP Conference Series: Materials Science and Engineering, vol. 874, Art. no. 012017, 2020, doi: https://doi.org/10.1088/1757-899X/874/1/012017.

[10] K. Nasution, K. Saddami, R. Roslidar, A. Akhyar, F. Fathurrahman, and N. Aulia, "Comparative Study of BiLSTM and GRU for Sentiment Analysis on Indonesian E-Commerce Product Reviews Using Deep Sequential Modeling," Jurnal Teknik Informatika, vol. 6, no. 4, pp. 1881–1896, 2025, doi: https://doi.org/10.52436/1.jutif.2025.6.4.4878.

[11] A. Setiawan, Z. Abidin, and M. Imamudin, "Impact of Preprocessing on Indonesian Extractive Summarization Using LexRank, TextRank, DivRank, and Cosine Similarity," G-Tech: Jurnal Teknologi Terapan, vol. 9, no. 4, pp. 2311–2321, 2025, doi: https://doi.org/10.70609/g-tech.v9i4.8306.

[12] A. Rozaq, R. Fitri, M. S. Akbar, M. H. Sumari, N. Khafiyya, I. Wardiah, and F. Sholihin, "Legal Literacy in Indonesia: Leveraging Semantic-Based AI and NLP for Enhanced Civil Law Access," E3S Web of Conferences, vol. 622, Art. no. 03002, 2025, doi: https://doi.org/10.1051/e3sconf/202562203002.

[13] D. E. Cahyani and I. Patasik, "Performance comparison of TF-IDF and Word2Vec Models for Emotion Text Classification," Bulletin of Electrical Engineering and Informatics, vol. 10, no. 5, pp. 2780–2788, 2021, doi: https://doi.org/10.11591/eei.v10i5.3157.

[14] J. J. Sihombing, A. Arnita, S. I. Al Idrus, and D. Y. Niska, "Implementation of Text Summarization on Indonesian Scientific Articles Using TextRank Algorithm with TF-IDF Web-Based," Journal of Soft Computing Exploration, vol. 5, no. 3, pp. 310–319, 2024, doi: https://doi.org/10.52465/joscex.v5i3.475.

[15] A. Sanjaya, A. B. Setiawan, U. Mahdiyah, I. N. Farida, and A. R. Prasetyo, "Pengukuran Kemiripan Makna Menggunakan Cosine Similarity dan Basis Data Sinonim Kata," Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 10, no. 4, pp. 747–752, 2023, doi: https://doi.org/10.25126/jtiik.2024106864.

Copyright (c) HOAQ (High Education of Organization Archive Quality) : Jurnal Teknologi Informasi