

Klasifikasi Bidang Ilmu Publikasi Ilmiah Terindeks SINTA Menggunakan Metode Naïve Bayes

Kusuma Nurnasikha, Sam Farisa Chairul Haviana, Imam Much Ibnu Subroto
Teknik Informatika, Fakultas Teknologi Industri, Universitas Islam Sultan Agung Semarang
Email: kusumanurnasikha@std.unissula.ac.id, sam@unissula.ac.id, imam@unissula.ac.id

Correspondence Author: kusumanurnasikha@std.unissula.ac.id

Abstrak

Publikasi ilmiah menjadi tuntutan akademik untuk menunjukkan tanggungjawab sebagai peneliti. Di Indonesia salah satu kegiatan yang berkaitan dengan status peneliti adalah publikasi yang terindeks SINTA (Science and technology Index). Secara garis besar judul-judul dapat diklasifikasikan ke dalam lima bidang ilmu yaitu Arts & Humanities, Engineering & Technology, Life Sciences & Medicine, Natural Sciences, and Social Sciences & Management. Permasalahannya adalah judul penelitian belum terklasifikasikan secara otomatis. Tujuan penelitian ini adalah untuk menerapkan dan menguji kinerja Naive Bayes dalam klasifikasi lima bidang ilmu pada publikasi terindeks Scopus. Naive Bayes merupakan klasifikasi prediksi yang mudah diinterpretasikan. Naive Bayes memudahkan peneliti mengklasifikasikan lima bidang ilmu karena modelnya sederhana dan cukup fleksibel meningkatkan kualitas keputusan dihasilkan. Berdasarkan hasil dari perhitungan menggunakan algoritma Naive Bayes, diperoleh nilai pada akurasi 0,425 recall 0,20 dan presisi 0,05. Namun hasil tersebut masih cukup rendah jika digunakan pada sistem SINTA dikarenakan pada saat pelabelan dataset yang digunakan belum optimal.

Kata Kunci : Klasifikasi, publikasi ilmiah, *Naïve Bayes*, *SINTA*

1. PENDAHULUAN

Publikasi saat ini menjadi tuntutan bagi sumber daya manusia untuk menunjukkan kreativitas dan kemampuan sebagai prodak dan buah karya pikir. Publikasi adalah informasi atau berita lama dan baru, yang diberikan seseorang ke khalayak media berdasarkan dengan undang-undang yang berlaku di Negara. Selain itu, pengertian publikasi menurut kamus besar Bahasa Indonesia menyatakan bahwa media informasi digunakan untuk menyebarkan informasi kepada masyarakat[1]. Jurnal nasional merupakan jurnal yang telah terakreditasi oleh Kementerian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia melalui Science and Technology Index (SINTA) dengan tingkat akreditasi yang tinggi yaitu jurnal SINTA[2]. Sementara itu Jurnal internasional yang telah bereputasi yaitu jurnal yang terindeks Scopus [3]. Platform database artikel yang dipilih adalah Scopus. Artikel dari database Scopus dijadikan dasar mendapatkan artikel dengan kualitas baik. Scopus merupakan database artikel multidisiplin ilmu yang mengumpulkan berbagai artikel dari jurnal-jurnal bereputasi internasional yang sudah diterbitkan oleh Elsevier, Springer, Wiley, Taylor & Francis, Emerald, Nature [4]. Selain itu pengertian Scopus merupakan salah satu pengindeks jurnal terbesar di dunia yang banyak dijadikan rujukan dan referensi dalam membuat suatu penelitian untuk mempublikasikan hasil penelitian pada jurnal. Scopus tidaklah mudah, perlu melewati beberapa tahapan review oleh pakar[5]. Didalam database scopus terdapat beberapa macam-macam 5 bidang ilmu. Didalam database scopus terdapat beberapa macam-macam 5 bidang ilmu. Ada beberapa referensi dalam membagi bidang ilmu publikasi diantaranya adalah QS RANKING yang merupakan organisasi yang melakukan pemeringkatan perguruan tinggi di seluruh dunia berdasarkan 5 bidang ilmu yaitu Arts & Humanities, engineering & Technology, Life Sciences & Medicine, Natural Sciences, dan Social Sciences & Management. Sejauh ini, pencarian jurnal dalam web SINTA belum bisa mengkategorisasikan sesuai dengan 5 bidang keilmuan sehingga kurangnya efektif dalam melakukan pencarian artikel. Untuk itu diperlukan suatu metode klasifikasi semua artikel dengan menggunakan metode naïve bayes. Dalam penggunaannya, Metode Naive Bayes ini memiliki keuntungan yaitu Interpolation, dimana Metode bayes mempunyai pilihan mengenai seberapa besar waktu dan usaha yang dilakukan oleh manusia vs komputer. Yang kedua Bahasa, dimana Metode bayes mempunyai bahasa tersendiri untuk menetapkan hal prior dan posterior. Yang ketiga Intuisi, Melibatkan prior dan integrasi, dua aktivitas yang berguna secara luas[6]. Naive Bayes Classifier merupakan proses pengklasifikasian probabilitas sederhana yang mengacu pada Teori Bayes. Teori tersebut menyatakan bahwa kemungkinan terjadinya suatu peristiwa sama dengan probabilitas intrinsik (dihitung dari data yang tersedia sekarang) dikalikan probabilitas bahwa hal serupa akan terjadi lagi di masa depan

(berdasarkan pengetahuan yang terjadinya di masa lalu). Naïve bayes adalah algoritma pembelajaran probabilitas yang berasal dari teori Keputusan Bayesian [7].

Naive bayes adalah salah satu metode dalam data mining. Data mining adalah proses mengekstraksi dan mengidentifikasi informasi yang berguna dan relevan dari basis data besar menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning [8]. Di dalam data mining terdapat klasifikasi. Klasifikasi adalah proses menemukan sekumpulan model (fitur) yang dapat menggambarkan kelas data atau konsep dan membedakannya sehingga dapat digunakan untuk memprediksi kelas dari objek yang tidak diketahui [9]. Klasifikasi bisa dilakukan dengan metode Naïve Bayes. Dan Publikasi menjadi bukti utama keaslian penelitian yang dilakukan. Publikasi dilakukan tentunya menggunakan media untuk membantu penyebaran informasi sehingga informasi dapat disampaikan secara efektif dan efisien ada beberapa macam publikasi diantaranya: publikasi konvensional, publikasi elektronik, dan publikasi interpersonal [10]. Berdasarkan latar belakang di atas penelitian ini mengusulkan metode naïve bayes yang bertujuan untuk menerapkan dan menguji performa algoritma Naïve 2 Bayes dalam mengklasifikasi 5 bidang ilmu yaitu Arts & Humanities, engineering & Technology, Life Sciences & Medicine, Natural Sciences, and Social Sciences & Management pada publikasi terindeks Scopus.

Terdapat beberapa penelitian terdahulu tentang algoritma Naïve bayes yang pernah dilakukan Devita dan kawan-kawan dalam mengklasifikasikan barang atau benda tertentu, termasuk Perbandingan kinerja Metode Naïve Bayes dan K-Nearest Neighbor untuk mengklasifikasikan artikel Bahasa Indonesia. Berdasarkan penelitian ini diperoleh hasil bahwa metode Naive Bayes mempunyai efisiensi yang lebih baik dengan akurasi sebesar 70%, sedangkan metode K-Nearest Neighbor memiliki tingkat akurasi yang cukup rendah yaitu sebesar 40% [11].

Dalam penelitian sebelumnya oleh Annur tahun 2018 tentang klasifikasi penduduk miskin dengan menggunakan metode naïve bayes. berdasarkan hasil matriks konfusi yang dilakukan dengan metode split-validation, dengan menggunakan metode klasifikasi Bayesian, yang diuji dataset mencapai akurasi 73% dan tergolong dalam kategori baik. Nilai presisi dan perolehan juga masing-masing 92% dan 86% [12]. Kemudian penelitian sebelumnya oleh Utami dan kawan-kawan pada tahun 2022 tentang Metode weighted naive bayes dengan Laplace smoothing digunakan untuk mengklasifikasikan kelayakan penerima bantuan Program Keluarga Harapan (PKH). Berdasarkan hasil perhitungan klasifikasi, dengan menggunakan metode naïve bayes laplace smoothing, dengan 56 data training dan 24 data uji, hasil uji performa dengan menggunakan confusion matrix menunjukkan akurasi sebesar 95,83%, tingkat kesalahan 4,17%, Sensitivitas 100,00%, dan Spesifitas 94,12% [13]. Selain itu Dalam penelitian sebelumnya pada tahun 2019 oleh Agus Sugianto & Rizky Maulana tentang Klasifikasi Penerima Bantuan Pangan Non Tunai (Studi Kasus Kelurahan Utama) Menggunakan Algoritma Naïve Bayes. Algoritma Naïve Bayes mempunyai tingkat Accuracy sebesar 58,29%. Sedangkan algoritma Decision Tree mempunyai tingkat Accuracy sebesar 73,97%. Dengan menggunakan alpha kurang dari 0,000 untuk hasil uji-t antara algoritma Naive Bayes dan algoritma Decision Tree, dapat disimpulkan bahwa uji-t antara algoritma Naive Bayes dan keputusan algoritma Pohon Keputusan memiliki hasil yang signifikan [14].

Dalam penelitian sebelumnya oleh Yuwono dan kawan-kawan tahun 2021 tentang metode Naïve Bayes digunakan untuk mengklasifikasi penghasilan pedagang kaki lima dan bisnis online yang terkena dampak pandemi COVID-19. Hasil pengujian berdasarkan confusion matrix menggunakan teknik split validasi dengan metode naïve bayes. Akurasi yang dicapai untuk dataset yang telah diambil sebagai objek penelitian adalah 75% atau termasuk dalam kategori yang sesuai. Sedangkan nilai Precision sebesar 100% dan Recall sebesar 66.7% [15].

Dalam penelitian sebelumnya oleh Prakoso dan kawan-kawan tahun 2019 tentang Algoritma Naïve Bayes Dengan Seleksi Fitur Dan Boosting digunakan untuk mengklasifikasi Berita. Hasil evaluasi nilai akurasi, recall, dan presisi diperoleh 73.2%. Namun, dalam metode Naive Bayes Classifier, pengklasifikasi Bayesian Boosting memiliki skor yang sama yaitu 73.2%. Hasil evaluasi model implementasi menunjukkan bahwa penerapan Information Gain feature selection tidak berpengaruh signifikan terhadap peningkatan kinerja pada kondisi label Polynomial [16].

Dalam penelitian sebelumnya oleh Imandasari dan kawan-kawan tahun 2019 tentang Algoritma Naive Bayes Dalam Klasifikasi Lokasi Pembangunan Sumber Air. Berdasarkan hasil perhitungan yang dilakukan dengan menggunakan algoritma Naive Bayes, diperoleh hasil klasifikasi dari 19 alternatif yang digunakan menunjukkan 8 kelas memenuhi persyaratan dan 11 kelas yang gagal, dengan akurasi keseluruhan sebesar 78,95% [17].

Dalam penelitian sebelumnya oleh Rohmat Indra Borman & Mina Wati tahun 2020 tentang Algoritma Naïve Bayes digunakan untuk mengimplementasikan Data Mining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandar Lampung. Dari hasil pengujian WEKA dan Rapidminer memperoleh akurasi 70,33%, recall

70,33%, dan presisi 100%, atau 1, menggunakan 1064 catatan pelatihan dan 300 catatan pengujian. Semakin banyak data uji yang digunakan, semakin besar dampaknya terhadap akurasi [18].

Dalam penelitian sebelumnya oleh Ericha Apriliyani & Salim pada tahun 2022 tentang menganalisis kinerja pengklasifikasi Naïve Bayes Classifier pada Unbalanced Dataset. Berdasarkan hasil analisis dengan metode Naïve Bayes Classifier, dataset Glass mendapatkan akurasi 46%, presisi 47%, recall 46% dan f-measure 43%, dalam dataset Heart Disease diperoleh akurasi 88%, presisi 88%, recall 88% dan f-measure 88%, dalam dataset Kidney Disease diperoleh akurasi 100%, presisi 100%, recall 100% dan f-measure 100%, dalam dataset Liver Disease diperoleh akurasi 78%, presisi 82%, recall 78% dan f-measure 79%, dalam dataset Diabetes diperoleh akurasi 77%, presisi 76%, recall 77% dan f-measure 76%, dan dalam dataset Breast Cancer diperoleh akurasi 94%, presisi 94%, recall 94% dan f-measure 94%. Hasil penelitian ini menunjukkan bahwa metode Naive Bayes Classifier menghasilkan nilai kinerja yang tidak valid untuk dataset yang tidak seimbang [19]. Serta Dalam penelitian sebelumnya oleh Fani Prasetya & Ferdiansyah tahun 2022 tentang penggunaan Algoritma Naïve Bayes untuk Analisis Data Mining Klasifikasi Berita Hoax COVID 19. Berdasarkan hasil penelitian menggunakan metode naïve bayes dan cross validation mengklasifikasikan berita hoax dengan benar, hasil akurasi sebesar 86.3% dengan 80-90% termasuk dalam kriteria bagus [20].

Tujuan dari penelitian ini adalah merancang sebuah sistem untuk melakukan klasifikasi dengan algoritma Naïve bayes pada publikasi terindeks Scopus. Serta manfaat yang diharapkan dari pembuatan sistem ini yaitu Mempermudah pembaca pada saat melakukan pencarian artikel sesuai dengan bidang ilmu, dan Penerapan metode Naïve Bayes diharapkan mampu mengklasifikasikan data Scopus yang terindeks pada database SINTA sesuai dengan 5 bidang ilmu.

2. METODE PENELITIAN

1. Pengumpulan data scopus

Tujuan dari langkah ini adalah memilih kumpulan data untuk digunakan dalam penelitian ini. Data yang dipilih harus sesuai dengan batasan yang ditetapkan dalam penelitian ini. Pengumpulan data Scopus pada penelitian ini diambil sesuai dengan 5 bidang ilmu yaitu Arts & Humanities, engineering & Technology, Life Sciences & Medicine, Natural Sciences, dan Social Sciences & Management. Dalam penelitian ini, sejumlah 1000 judul publikasi dari lima bidang ilmu berbeda diuji.

2. Preprocessing

Tahap preprocessing dalam penelitian ini yaitu pengkodean data pada judul artikel di SINTA yang terindeks Scopus. Tahapan-tahapan berikut ini dilakukan selama tahapan data preprocessing:

- a. Cleaning merupakan proses memperbaiki atau menghapus karakter yang bukan huruf baca.
- b. Case folding, yaitu proses mengubah menyamaratakan penggunaan huruf kapital pada data menjadi huruf kecil (lowercase).
- c. Tokenizing, yaitu proses pemisahan sebuah kalimat menjadi satu kata dan lazimnya dipisahkan oleh karakter seperti tanda baca spasi, titik (.), dan koma (,).
- d. Stopword, yaitu tahap pemilihan kata yang bernilai dari hasil token dengan menyimpan kata yang bernilai dan membuang kata yang tidak bernilai. Dapat menggunakan algoritma stoplist (membuang kata kurang penting) atau wordlist (menyimpan kata penting) stoplist / stopword adalah “yang”, “dan”, “di”, dan lain-lain.
- e. Stemming Setelah melalui proses stopword removal, langkah selanjutnya yaitu proses stemming.. Stemming merupakan proses pemetaan dan penguraian berbagai bentuk (variasi) kata menjadi bentuk kata dasar.

3. Penerapan metode Naïve Bayes

Data set diolah menggunakan metode Naïve Bayes. Pada penelitian ini mengimplementasikan metode naïve bayes. Naïve bayes adalah teknik klasifikasi data mining yang menerapkan teori Bayes. Kelebihannya adalah modelnya sederhana, cepat dan akurasi yang tinggi. Teorema keputusan Bayes adalah pendekatan statistik dasar untuk pengenalan pola. Naïve Bayes didasarkan pada asumsi penyederhanaan bahwa nilai output dan atribut tertentu secara kondisional independen. Artinya, nilai output yang diberikan, mengamati secara bersama adalah produk dari probabilitas masing-masing. Dengan mengganti Persamaan 1 menjadi Persamaan 2 memberikan pendekatan yang digunakan oleh NBC. Teorema Bayes didirikan oleh Thomas Bayes pada tahun 1763. Teorema Bayes digunakan untuk menghitung probabilitas terjadinya suatu

peristiwa yang terjadi berdasarkan pengaruh yang diperoleh dari hasil pengamatan. Perbedaan mendasar antara metode Bayesian dan statistik umumnya adalah bahwa dalam Bayesian parameter dianggap sebagai variabel acak sementara dalam statistik klasik, parameternya tidak diketahui. Nama teorema Bayes berasal dari Pendeta Thomas Bayes, yang menggambarkan hubungan antara probabilitas bersyarat dari dua peristiwa, H dan X.

Dalam teorema bayes Rumus Naïve Bayes Classifier adalah sebagai berikut:

$$P(H|X) = (P(X|H) \frac{P(H)}{P(X)}) \quad (1)$$

X = Data dokumen dengan kelas yang belum diketahui.

H = Hipotesis dari data X merupakan suatu kelas spesifik.

$P(H|X)$ = Probabilitas dari hipotesis H berdasarkan kondisi X (posteriori probability).

$P(H)$ = Probabilitas dari hipotesis H (prior probability).

$P(X|H)$ = Probabilitas X berdasarkan dari kondisi pada hipotesis H.

$P(X)$ = Probabilitas dari X

4. Evaluasi

Tahap ini mencakup hasil eksperimen berdasarkan recall presisi dan akurasi. Tahapan ini digunakan untuk mendapatkan prediksi menggunakan algoritma yang sudah ditentukan. Tahapan ini merupakan langkah penting dalam proses membangun sebuah model. Menggunakan persamaan berikut untuk membuktikan kinerja algoritma yang digunakan:

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \% \quad (1)$$

Hasil akurasi dihitung berdasarkan confusion matrix. Perhitungan pada confusion matrix dilakukan berdasarkan prediksi positif (true positive), prediksi positif yang salah (false positive), prediksi negatif yang benar (True Negatif) dan prediksi negatif yang salah (False Negatif). Semakin tinggi nilai akurasi yang diperoleh maka semakin bagus metode yang dihasilkan.

Precision adalah tingkat ketelitian antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh system.

$$Precision = \frac{TP}{TP+FP} \times 100 \% \quad (2)$$

Recall adalah tingkat keberhasilan sistem dalam mengambil sebuah informasi.

$$Recall = \frac{TP}{TP+FN} \times 100 \% \quad (3)$$

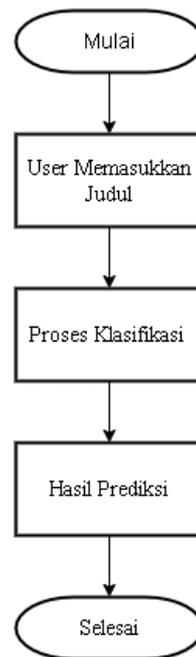
F1 Score adalah perbandingan rata-rata presisi dan recall yang dibitikan.

$$F1 - score = 2x \frac{precision \times recall}{(Precision + Recall)} \quad (4)$$

Setelah hasil klasifikasi diperoleh, Cross Validation (confusion matrix) digunakan untuk mengevaluasi hasil akurasi, presisi dan recall yang dihasilkan oleh model yang diusulkan.

5. Flowchart sistem

Flowchart sistem dalam memprediksi judul publikasi dari lima bidang ilmu untuk menentukan akurasi persentase tertinggi yaitu dengan kesesuaian judul artikel dalam SINTA yang sudah terindeks oleh Scopus dalam 5 bidang ilmu tersebut diperlihatkan seperti gambar 1



Gambar 1 *Flowchart* Sistem

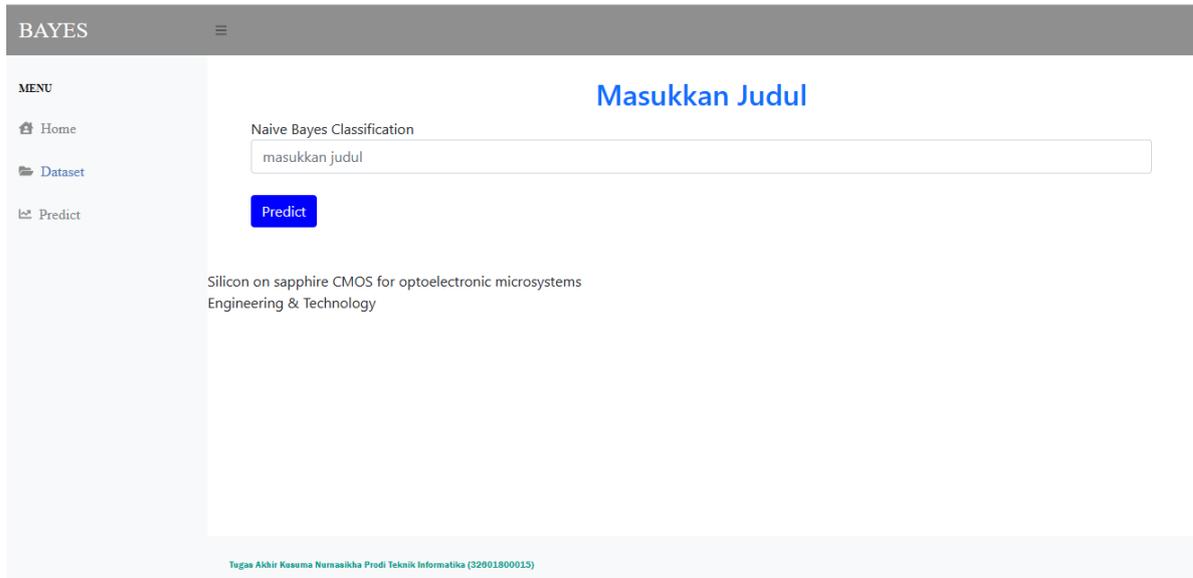
Flowchart sistem dalam memprediksi judul publikasi dari lima bidang ilmu untuk menentukan akurasi persentase tertinggi yaitu dengan kesesuaian judul artikel dalam SINTA yang sudah terindeks oleh Scopus dalam 5 bidang ilmu tersebut yaitu :

1. User Memasukkan Judul.
2. Setelah user memasukkan judul, tekan tombol prediksi untuk menyelesaikan Proses Klasifikasi, dimana sistem mengklasifikasikan judul dengan menggunakan metode Naïve Bayes
3. Menampilkan hasil prediksi judul publikasi dari lima bidang ilmu tersebut, maka alur proses sistem ini telah selesai.

3. HASIL DAN ANALISA

3.1 Hasil Prediksi

Pada gambar dibawah 2 halaman ini menampilkan kolom input seperti contoh user akan memasukkan judul artikel ke dalam kolom tersebut, judul yang akan dimasukkan ke dalam kolom yaitu *Silicon on sapphire CMOS for optoelectronic microsystems*, kemudian di dalam halaman tersebut juga terdapat predict button yang mana tombol tersebut berfungsi untuk memproses prediksi yang akan dilakukan oleh sistem tersebut. Dimana contoh judul yang dimasukkan oleh user menampilkan hasil dari pencarian data prediksi berupa judul artikel yaitu *Engineering & Technology*, dengan tampilan tersebut memudahkan *user* untuk mengetahui judul artikel tersebut masuk kedalam salah satu bidang ilmu.



Gambar 2 Hasil Prediksi

3.2 Evaluasi performa

Akurasi yang dihasilkan dihitung berdasarkan confusion matrix. Perhitungan pada confusion matrix dihitung sesuai dengan prediksi positif yang benar (*True Positive*), prediksi positif yang salah (*False Positive*), prediksi negatif yang benar (*True Negatif*) dan prediksi negatif yang salah (*False Negatif*). Semakin tinggi nilai akurasi yang didapat maka semakin baik pula metode yang dihasilkan

Dari data tersebut menghasilkan confusion matrix, confusion matrix ini digunakan untuk menentukan akurasi, presisi, recall & skor F1. Dari hasil training yang diujikan, menghasilkan confusion matrix dengan nilai akurasi, presisi, recall dan F1 sebagai berikut:

Tabel 1 *confusion matrix*

Accuracy	Precision	Recall	F1 score
0.42575	0.0558	0.20133	0.0742

Tabel 1 menunjukkan hasil pengujian dengan data sampel 1000 judul artikel Naive Bayesian menghasilkan nilai akurasi sebesar 0,425 dengan presisi, recall, dimana F1 masing-masing memiliki nilai 0,055, 0,201 dan 0,074 yang menunjukkan bahwa metode Naive Bayesian belum optimal untuk klasifikasi judul artikel jurnal publikasi scopus dimana metode tersebut hanya memberikan akurasi sebesar 0,425, sedangkan hasil tersebut masih tergolong rendah. Nilai akurasi yang rendah juga dapat disebabkan oleh beberapa faktor seperti datanya, dimana data tersebut digunakan peneliti masih sangat kotor dan perlu banyak pembersihan, pembersihan yang lebih dalam akan merubah format nama data tersebut yang berupa judul, sehingga itu tidak bisa dilakukan oleh peneliti yang mana akan mengubah substansi dari arti judul tersebut, kemudian yang kedua banyak ditemukan kemiripan judul antara satu bidang ilmu kesamaan.

Tabel 2 *confusion matrix*

	Positive	Negatif
True	302	1208
False	1198	1198

```
"true positives": 302,
"true negatives": 1208,
"false positives": 1198,
"false negatives": 1198,
```

Gambar 3 Output hasil (*True Positive, True Negative, False Positive, False Negative*)

Berdasarkan pengujian yang dilakukan, dapat disimpulkan bahwa pengujian menggunakan total 5000 data, dimana rincian pembagian data training mencapai 90%=4500 dan data testing mencapai 10%=500. Hasilnya adalah akurasi keseluruhan sebesar 0,425 (42%).

4. KESIMPULAN

Dari penelitian ini dapat disimpulkan bahwa dengan mengklasifikasi bidang ilmu publikasi terindeks Scopus menggunakan metode Naïve bayes menghasilkan output yaitu kesesuaian judul artikel dalam SINTA yang sudah terindeks oleh Scopus dalam 5 bidang ilmu tersebut. Hasil dari pengujian sistem yang cukup dengan akurasi, recall, presisi. Diperoleh nilai akurasi sebesar 0,425, recall 0,20 dan presisi 0,05, Namun hasil tersebut masih cukup rendah jika digunakan pada sistem SINTA dikarenakan pada saat pelabelan dataset yang digunakan belum optimal. Dari percobaan diatas dataset yang diharapkan tidak sesuai, tetapi sistem bekerja dengan baik sebagaimana fungsinya mampu mengklasifikasi.

DAFTAR PUSTAKA

- [1] M. A. Mustopa, M. A. Famuzia, R. Budiawan, dan U. Telkom, "Sema : Sistem Pendataan Dan Publikasi Elektronik Kabupaten," vol. 7, no. 5, hal. 1914–1922, 2021.
- [2] Amaliyah, "Program Kemitraan Masyarakat: Peningkatan Kemampuan Penulisan Karya Ilmiah dan Teknik Publikasi di Jurnal Internasional," *Interv. Komunitas*, vol. 1, no. 1, hal. 48–56, 2019.
- [3] A. Firmansyah, R. A. Qadri, dan A. Arham, "Pelatihan melalui Web Seminar terkait Publikasi Artikel untuk Menembus Jurnal Sinta 2 dan Scopus," *Abdimas J. Pengabd. Masy. Univ. Merdeka Malang*, vol. 5, no. 2, hal. 131–138, 2020, doi: 10.26905/abdimas.v5i2.4244.
- [4] R. Herdianto, N. Windyaningrum, B. Masruroh, dan M. A. Setiawan, "Filsafat pendidikan dan perkembangannya: kajian bibliometrik berdasarkan database Scopus [Philosophy of education and its development: bibliometric studies based on scopus database]," *Belantika Pendidik.*, vol. 4, no. 1, hal. 44–56, 2021.
- [5] C. Ibrahim, "Tren Kepenulisan Publikasi Penelitian," vol. 6003, hal. 13–28, 2019.
- [6] F. Dhany, H. W., & Izhari, "Analisis Algorithms Support Vector Machine Dengan Naive Bayes Kernel Pada Klasifikasi Data. Jurnal Teknik Dan Informatika, 6(2), 30–35.," *Anal. Algorithms Support Vector Mach. Dengan Naive Bayes Kernel Pada Klasifikasi Data. J. Tek. Dan Inform.* 6(2), 30–35., vol. 6, 2019.
- [7] H. Mustofa dan A. A. Mahfudh, "Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes," *Walisono J. Inf. Technol.*, vol. 1, no. 1, hal. 1, 2019, doi: 10.21580/wjit.2019.1.1.3915.
- [8] D. P. Utomo dan M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. 2, hal. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- [9] A. H. Nasrullah, "Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris," *J. Ilm. Ilmu Komput. Fak. Ilmu Komput. Univ. Al Asyariah Mandar*, vol. 7, no. 2, hal. 45–51, 2021.
- [10] E. Kartikasari dan A. Dianti, "Strategi Publikasi Unit Pengelolaan Informasi Majapahit Dalam Menyampaikan Informasi Di Masa Pandemi Covid-19," *eProceedings Manag.*, vol. 8, no. 5, 2021.
- [11] R. N. Devita, H. W. Herwanto, dan A. P. Wibawa, "Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa indonesia," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 4, hal. 427, 2018, doi: 10.25126/jtiik.201854773.
- [12] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2,

- hal. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.
- [13] D. Utami, P. Aisyiyah, dan R. Devi, “Klasifikasi kelayakan penerima bantuan program keluarga harapan (pkh) menggunakan metode weighted naïve bayes dengan laplace smoothing,” vol. 07, hal. 1373–1384, 2022.
- [14] C. Agus Sugianto dan F. Rizky Maulana, “Algoritma Naïve Bayes Untuk Klasifikasi Penerima Bantuan Pangan Non Tunai,” *Techno.COM*, vol. 18, no. 4, hal. 321–331, 2019.
- [15] L. Yuwono, M. egi Fadillah, M. Indrayani, dan ..., “Klasifikasi Pendapatan Pedagang Kaki Lima Dan Pelaku Usaha Online Akibat Dampak Covid-19 Menggunakan Metode Naive Bayes,” *Bull. Appl. ...*, vol. 2, no. 1, hal. 1–6, 2021, [Daring]. Tersedia pada: <http://jim.unindra.ac.id/index.php/baiet/article/view/3947>
- [16] B. S. Prakoso, D. Rosiyadi, H. S. Utama, dan D. Aridarma, “Klasifikasi Berita Menggunakan Algoritma Naive Bayes Classifier Dengan Seleksi Fitur Dan Boosting,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 2, hal. 227–232, 2019, doi: 10.29207/resti.v3i2.1042.
- [17] T. Imandasari, E. Irawan, A. P. Windarto, dan A. Wanto, “Algoritma Naive Bayes Dalam Klasifikasi Lokasi Pembangunan Sumber Air,” *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, hal. 750, 2019, doi: 10.30645/senaris.v1i0.81.
- [18] R. Indra Borman dan M. Wati, “Penerapan Data Mining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naïve Bayes,” *J. Ilm. Fak. Ilmu Komput.*, vol. 09, no. 01, hal. 25–34, 2020.
- [19] E. Apriliyani dan Y. Salim, “Analisis performa metode klasifikasi Naïve Bayes Classifier pada Unbalanced Dataset,” *Indones. J. Data Sci.*, vol. 3, no. 2, hal. 47–54, 2022.
- [20] F. Prasetya dan F. Ferdiansyah, “Analisis Data Mining Klasifikasi Berita Hoax COVID 19 Menggunakan Algoritma Naive Bayes,” *J. Sist. Komput. dan Inform.*, vol. 4, no. 1, hal. 132, 2022, doi: 10.30865/json.v4i1.4852.