

KLASIFIKASI BIDANG ILMU PADA PUBLIKASI TERINDEKS SCOPUS MENGGUNAKAN METODE *K-NEAREST NEIGHBOR*

Aina Nurul Maulida, Andi Riansyah, Imam Much Ibnu Subroto
Fakultas Teknologi Industri, Universitas Islam Sultan Agung Semarang

Correspondence Author : ainanurulmaulida@std.unissula.ac.id, andi@unissula.ac.id,
imam@unissula.ac.id

Abstract

Daftar publikasi nasional yang terakreditasi disediakan oleh SINTA, situs web atau portal ilmiah online yang dikelola oleh Kemendikbud. SINTA juga dapat dilihat sebagai repository atau pusat data jurnal nasional yang terakreditasi. Saat ini permasalahan yang dihadapi oleh SINTA ialah klasifikasi bidang ilmu pada publikasi terindeks Scopus yang belum sesuai dengan 5 bidang ilmu. Berlandaskan permasalahan tersebut maka tujuan dilaksanakannya penelitian ini yaitu untuk mengklasifikasi bidang ilmu setiap judul publikasi menjadi 5 bidang ilmu. Untuk mendapatkan hasil klasifikasi maka penelitian ini memakai metode *K-Nearest Neighbor*. Pengklasifikasian artikel publikasi dapat dilakukan melalui beberapa proses, diantaranya seperti Data Cleaning, Case Folding, Tokenizing, Stopword Removal, serta Stemming. Pada penelitian ini penulis menggunakan data sebanyak 750, dimana masing-masing bidang terdapat 150 data, kemudian data tersebut dibagi menjadi data training dan data testing yaitu masing-masing 675 dan 75 dengan 6 kali percobaan menggunakan beberapa parameter *K* yaitu 5,15,20,25,45,50. Dari hasil percobaan tersebut menghasilkan nilai tertinggi dengan *K*=5 yaitu akurasi sebesar 0.5384%, presisi sebesar 0.3768%, dan recall sebesar 0.2977%. Hasil tersebut belum sesuai dengan hasil yang diharapkan penulis dikarenakan pada saat pelabelan data yang diambil tidak tepat.

Keyword: Klasifikasi, KNN, Scopus.

1. PENDAHULUAN

Scopus ialah pangkalan data sitasi dan abstrak yang berisikan lebih dari 81 juta judul artikel yang berasal dari sekitar 43.000 jurnal internasional yang terdapat pada web SINTA. *Scopus* memberi gambaran paling luas dari beberapa hasil riset yang dihasilkan dunia mencakup atas bidang *Art & Humanities, Engineering & Technology, Life Sciences & Medicine, Natural Sciences, Sosial Sciences & Management*. Jenis sumber yang dicakup pada *Scopus* yakni publikasi berseri (serial publikasi) yang mempunyai nomor seri terstandar (ISSN) misalnya jurnal, buku berseri, publikasi hasil konferensi berseri (seri konferensi) ataupun publikasi non-berseri yang mempunyai nomor buku terstandar (ISBN). *Scopus* juga mengindeks kutipan dari sejumlah buku ilmiah, artikel penelitian, jurnal *peer-review*, editorial serta menyediakan maupun menyajikan informasi mengenai pemeringkatan jurnal yang dirilis setiap tahunnya. [1]

Saat ini permasalahan yang dihadapi oleh SINTA adalah pencarian kategorisasi bidang ilmu setiap artikel publikasi belum sesuai dengan 5 bidang ilmu utama secara efektif dan otomatis, untuk itu perlu dikembangkan suatu metode untuk mengklasifikasi semua artikel yang ada menggunakan pendekatan *data mining*. Dalam metode klasifikasi *data mining* ada sejumlah algoritma yang bisa dipakai satu diantaranya yakni metode *K-Nearest Neighbor* (KNN). Metode tersebut dipakai guna melangsungkan klasifikasi objek berlandaskan data pembelajaran melalui cara melihat kedekatan jarak satu data terhadap data lainnya, kelebihan dari metode KNN diantaranya algoritma tersebut tangguh terhadap *training* data yang efektif jika data latihnya besar. Berlandaskan latar belakang tersebut maka dalam penelitian ini menawarkan sistem dengan memakai metode KNN sebagai solusi untuk mengklasifikasikan publikasi terindeks *Scopus* menjadi 5 bidang ilmu yakni *Art & Humanities, Engineering & Technology, Life Science & Medicine, Natural Sciences, Sosial Sciences & Management*. Sistem penelitian ini diharapkan bisa memberikan bantuan kepada sejumlah pembaca dalam pencarian suatu artikel dari 5 bidang keilmuan dengan mudah dan tepat.

Dari hasil penelitian yang sudah dilangsungkan oleh Siti Julaiha dan Marteli Bettiza pada tahun 2021 bisa dilakukan penarikan kesimpulan bahwa penerapan telah dihasilkan sistem aplikasi berbasis web untuk mengklasifikasikan calon peserta bidikmisi di kampus Universitas Maritim Raja Ali Haji. Informasi yang

dihasilkan pada sistem yaitu berupa data klasifikasi, data parameter, dan nilai bobot parameter serta data pendaftar bidikmisi, dan hasil klasifikasi. Dimana data yang dipakai yakni sejumlah 331 data training dan 83 data testing dengan nilai $K=5$ menghasilkan nilai akurasi sebesar 83,13%, nilai presisi sebesar 82,35%, dan nilai *recall* sebesar 89,36%. [2]

Berlandaskan hasil riset yang sudah dilangsungkan oleh Kurniawan dan Barokah pada tahun 2020 bisa dilakukan penarikan kesimpulan bahwa terdapat masalah yang dihadapi pada pemberian kartu kredit pada nasabah bank yang sudah terdaftar ialah sukarnya melakukan penentuan terhadap kategori kartu kredit yang sesuai terhadap nasabah bank. Maka dari itu dengan melangsungkan riset ini diharapkan bias memberikan kemudahan kepada pihak bank guna melakukan penentuan terhadap kategori kartu kredit bagi nasabah bank dengan tepat dengan menggunakan metode KNN. Metode KNN dipakai guna melakukan pencarian terhadap beberapa pola dalam data nasabah misalnya penghasilan setiap tahunnya, profesi, jumlah tanggungan, jenis kelamin, serta status. Penelitian dengan judul Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan *K-Nearest Neighbor* menunjukkan hasil rata-rata nilai presisi sebesar 92%, nilai *recall* sebesar 83%, dan nilai akurasi sebesar 93%. [3]

Berdasarkan hasil riset yang sudah dilaksanakan oleh Diki dan Nurina pada tahun 2022 bisa dilakukan penarikan kesimpulan yakni implementasi algoritma KNN pada klasifikasi judul berita hoaks memberikan hasil yang benar-benar baik dengan nilai akurasi sebesar 93,33%, presisi sebesar 100%, *f1-score* sebesar 88,89%, dan *recall* sebesar 80%. Sehingga bias dinyatakan bahwa algoritma KNN cocok guna melakukan pengklasifikasian judul berita hoaks. [4]

Berdasarkan penelitian dari Rahmat Dian Nugraha pada tahun 2020 dilangsungkannya pembuatan maupun pengujian aplikasi Implementasi Metode *K-Nearest Neighbor* (KNN) Untuk Seleksi Calon Karyawan Baru, maka bisa dilakukan penarikan kesimpulan antara lain yakni sistem tersebut bisa dilaksanakan di browser *Google Chrome* serta *Microsoft Edge* secara baik. Berlandaskan hasil perhitungan menggunakan algoritma *K-Nearest Neighbor* dengan nilai parameter $K=7$ menggunakan metode *Euclidean Distance* didapat nilai akurasi 91%, nilai presisi 87%, dan nilai *recall* 100%. [5]

Pada penelitian sebelumnya yang mengangkat judul *K-Nearest Neighbor* Berbasis Particle Swarm Optimization untuk Analisis Sentimen Terhadap Tokopedia menghasilkan akurasi paling baik yakni dalam skenario 90:10 melalui nilai $K=1$ yakni sejumlah 88,10%. Hal tersebut dikarenakan dengan memakai $K=1$ perbedaan jarak diantara data latih serta data uji tidak begitu jauh apabila dilakukan perbandingan terhadap nilai K yang lain. Kemudian dilangsungkan *feature selection* dengan memakai algoritma *particle swarm optimization* melalui parameter 20 iterasi serta 10 partikel menghasilkan nilai evaluasi paling baik yakni akurasi sejumlah 97,9%, presisi sejumlah 96,17%, *recall* sejumlah 96,62%, serta *f-measure* sejumlah 96,39%. [6]

Pada penelitian terdahulu dengan judul Penerapan Algoritma *K-Nearest Neighbor* (K-NN) Untuk Klasifikasi Penyakit Diabetes Melitus dilangsungkan proses normalisasi dan perhitungan jarak tetangga terdekat memakai metode *Euclidean Distance* dengan nilai $K=9$ memperoleh hasil klasifikasi sebanyak 4 orang positif serta sebanyak 50 orang negatif diabetes mellitus. Uji algoritma KNN dalam melakukan penentuan terhadap klasifikasi penyakit diabetes mellitus memakai *confusion matrix* memperoleh hasil akurasi yang baik sejumlah 93%, presisi sejumlah 100%, *recall* sejumlah 60% serta *f1-score* sejumlah 75%. [7]

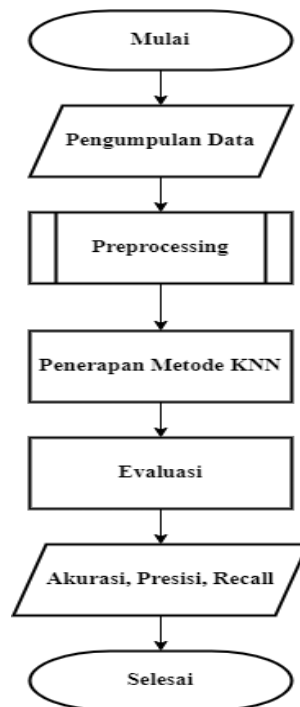
Pada penelitian sebelumnya yang mengangkat judul Implementasi Algoritma *K-Nearest Neighbor* Untuk Menganalisis Sentimen *User Twitter* dengan menggunakan dataset yang sudah diseleksi ada sejumlah 389 data *tweet* yang kontra serta sejumlah 237 data *tweet* yang pro memperlihatkan bahwa sentiment warga terkait kejadian pada Tol Cikampek cenderung kontra. Pemakaian metode KNN dalam riset ini menghasilkan nilai akurasi paling tinggi yakni sejumlah 83,1% ketika $K=9$, nilai presisi paling tinggi yakni sejumlah 66,7% ketika $K=5$, dan nilai *recall* paling tinggi yakni sejumlah 87,5% ketika $K=9$. [8]

Berlandaskan latar belakang yang diterangkan sebelumnya maka rumusan permasalahan dalam penelitian ini yakni bagaimana cara mengklasifikasi bidang ilmu pada publikasi terindeks *Scopus* menggunakan metode KNN. Dan tujuan penelitian ini ialah guna mengklasifikasikan 5 bidang ilmu pada publikasi terindeks *Scopus* menggunakan metode klasifikasi *data mining* dengan algoritma KNN.

2. METODE PENELITIAN

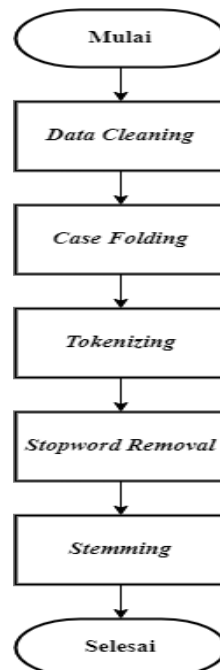
Klasifikasi Bidang Ilmu pada Publikasi Terindeks Scopus Menggunakan Metode K-Nearest Neighbor (Aina Nurul Maulida)

Pada tahapan penelitian, metode yang akan dipakai ialah metode *K-Nearest Neighbor*. Metode ini akan mengklasifikasikan data *Scopus* yang ada pada *database* SINTA ke dalam 5 bidang ilmu. Adapun langkah-langkah pada penelitian ini ialah



Gambar 1. Tahapan Penelitian

Pada gambar 1 merupakan tahapan yang akan digunakan dalam penelitian ini. Yang mana terdiri dari pengumpulan data *Scopus* pada *database* SINTA kemudian data tersebut akan melalui tahap *preprocessing*. Setelah melalui tahapan *preprocessing* maka akan dilakukan penerapan metode menggunakan algoritma KNN, kemudian data yang sudah melalui penerapan metode akan di evaluasi melalui 3 tahap yaitu, Akurasi, Presisi, dan *Recall*.



Gambar 2. Tahapan Preprocessing

Pada gambar 2 merupakan tahapan pada proses *preprocessing*. Proses pengubah bentuk menjadi data yang terstruktur sesuai dengan kebutuhan untuk proses *data mining*, umumnya akan menjadi nilai numerik.

Sesudah data menjadi terstruktur serta berbentuk nilai numerik maka data bisa dijadikan selaku sumber yang bisa dilakukan pengolahan lebih lanjut. Adapun beberapa tahapan yang akan dilangsungkan dalam penelitian ini, diantaranya yakni :

1. Pengumpulan Data Scopus

Pada tahapan ini dilangsungkan pengumpulan data penelitian karena hanya ketika data yang benar tersedia maka proses penelitian dilanjutkan hingga peneliti memperoleh jawaban dari rumusan permasalahan yang sudah dilakukan penetapan. Data yang dipakai pada penelitian ini diambil dari *Scopus* pada database SINTA dan kemudian data akan diklasifikasi sesuai dengan 5 bidang ilmu yaitu “*Art & Humanities, Engineering & Technology, Life Science & Medicine, Natural Science, Social Science & Management.*”

2. Preprocessing

Didalam proses preprocessing terdapat beberapa tahapan, antara lain :

a. *Data Cleaning*

Data Cleaning merupakan proses penguraian deskripsi yang semula berupa kalimat-kalimat menjadi kata-kata dan menghilangkan delimitter-delimitter seperti tanda titik (.), koma (,), tanda kutip (“”), tanda kurung (), spasi dan karakter angka yang ada pada kata tersebut.

b. *Case Folding*

Case Folding adalah mengubah semua huruf dalam dokumen menjadi huruf kecil. Hanya huruf “a” sampai “z” yang dapat diterima. Karakter selain huruf dihilangkan dan dianggap delimitter.

c. *Tokenizing*

Tokenizing ialah proses melakukan pemecahan terhadap kalimat isi aduan menjadi kata perkata.

d. *Stopword Removal*

Stopword Removal merupakan tahap mengambil kata-kata penting dari hasil token. Bisa menggunakan algoritma *stoplist* (membuang kata kurang penting) atau *wordlist* (menyimpan kata penting). Contoh *stopword* adalah “yang”, “dan”, “di”, dan lain-lain.

e. *Stemming*

Stemming adalah proses pemetaan dan penguraian berbagai bentuk (variasi) dari suatu kata menjadi bentuk kata dasarnya.

3. Penerapan Metode KNN

Pada tahap ini dilakukan klasifikasi menggunakan metode KNN, setelah data di olah melalui tahapan *preprocessing* kemudian data dikelompokkan dalam data *training* serta *testing*. Data *training* dipakai guna melatih algoritma, sementara itu data *testing* dipakai guna menentukan bagaimana kualitas algoritman yang dilatih sebelumnya pada saat menemukan data baru yang belum pernah dilihat sebelumnya.

4. Evaluasi

Pada tahapan ini dilangsungkan oleh peneliti, apakah data yang dikumpulkan telah sesuai terhadap kebutuhan maupun keinginan peneliti ataupun belum. Apabila belum, maka data akan di revisi dengan mengulang beberapa langkah sebelumnya. Namun bila telah sesuai, maka langkah berikutnya akan dilakukan akurasi, presisi, dan *recall*. Pada penelitian ini, hasil akan berupa grafik dengan metode validasi yang dipakai yakni *Confusion Matrix* yang memberi keputusan berlandaskan hasil dari *training* maupun *testing* dan memberi penilaian performasi klasifikasi berlandaskan data secara benar ataupun salah.

Tabel 1. *Confusion Matrix*

Classification	Predicted Class	
	Class = Yes	Class = No
Class = Yes	a (True Positive-TP)	b (False Negative-FN)
Class = No	c (False Positive-FP)	d (True Negative-TN)

Ada sejumlah cara guna melakukan pengukuran terhadap performa, sejumlah cara yang kerap dipakai yakni dengan melakukan perhitungan akurasi, presisi serta *recall*.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

Akurasi dilakukan perhitungan dengan melakukan perbandingan jumlah data yang benar terklasifikasi terhadap jumlah data keseluruhan. Cara perhitungannya bisa diketahui dalam persamaan :

$$Presisi = \frac{TP}{TP + FP}$$

Presisi diartikan selaku ukuran ketepatan. Apabila data diprediksi positif, seberapa sering data prediksi tersebut benar. Nilai presisi bisa diketahui dalam persamaan :

$$Recall = \frac{TP}{TP + FN}$$

Sementara itu *recall* ialah ukuran kelengkapan. Dari jumlah data sebenarnya yang mempunyai nilai positif, sebanyak apakah data yang diprediksi positif. Nilai *recall* bisa diketahui dalam persamaan :

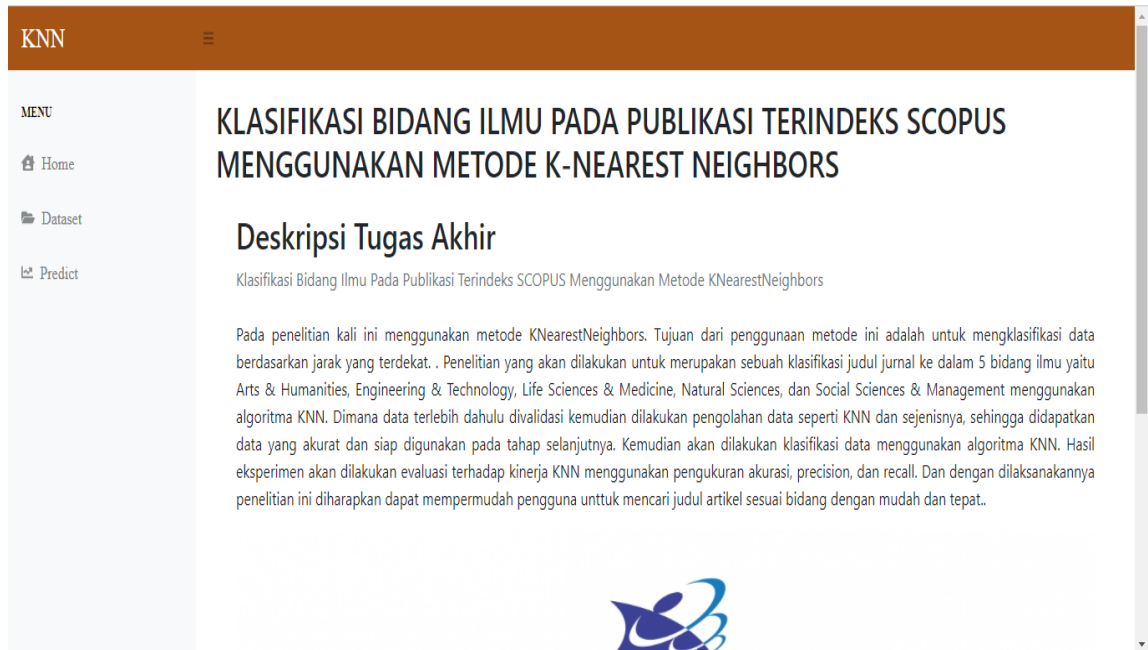
Dengan keterangan :

- TP = nilai *true positive*
- TN = nilai *true negative*
- FP = nilai *false positive*
- FN = nilai *false negative* [9][10]

3. HASIL DAN ANALISA

3.1. Hasil

1. Halaman Home



Gambar 3. Tampilan Home

Menu pertama adalah halaman *Home*. Pada halaman ini menampilkan gambar dan penjelasan singkat tentang deskripsi tugas akhir, pengguna hanya dapat melihat bagaimana peneliti melakukan langkah-langkah sistem serta penjelasan singkat mengenai deskripsi tugas akhir sistem.

2. Halaman Dataset

KNN

MENU

- Home
- Dataset
- Predict

Dataset yang Digunakan

KNearestNeighbors

Sampel Data yang digunakan dalam penelitian ini menggunakan Data dari Database SINTA.

Data Scopus

[Unduh Dataset](#)

No.	Judul	Bidang Ilmu
1	Higher-order approximations for frequency domain time series regression	Art_and_Humanities
2	Artificial decision-making and artificial ethics: A management concern	Art_and_Humanities
3	Verbal overshadowing effects on Raven s advanced progressive matrices: Evidence for multidimensional	Art_and_Humanities
4	The politics of space: changing discourses on Chinese burial grounds in post-war Singapore	Art_and_Humanities
5	The Durbin-Watson ratio under infinite-variance errors	Art_and_Humanities
6	"The ""anthropologicality"" of indigenous anthropology"	Art_and_Humanities
7	Data base navigation: an office environment for the professional	Art_and_Humanities
8	Trending time series and macroeconomic activity: Some present and future challenges	Art_and_Humanities

Gambar 4. Tampilan Dataset

Menu kedua adalah halaman *Dataset*. Pada halaman ini menampilkan fitur unduh dan menampilkan tabel berisi judul data *Scopus* dari masing-masing 5 bidang ilmu. Pengguna dapat melihat hasil data pada tabel dan mengunduh *dataset* tersebut pada tombol yang telah disediakan.

3. Halaman Predict

KNN

MENU

- Home
- Dataset
- Predict

Masukan Judul Data Scopus

masukkan judul

Predict

Gambar 5. Tampilan Predict

Menu terakhir adalah halaman *predict*. Halaman ini memiliki fitur menampilkan hasil dari data *predict* yaitu dengan cara pengguna memasukkan judul yang diinginkan kemudian klik tombol *predict* untuk melihat hasil dari data tersebut.

3.2. Analisa

Pada tahapan analisa, tujuan dilakukannya tahapan tersebut yaitu guna bisa mengetahui seberapa tingkat akurasi sistem yang sudah dibuat, tahap uji ini akan menggunakan data *testing*, yang mana data dibagi menjadi data *training* dan data *testing*, dimana pada data tersebut menggunakan perbandingan 9:1 yaitu *training* 90% dan *testing* 10% dari total data 750 judul artikel jurnal.

Tabel 2. Perbandingan Data *Training* dan Data *Testing*

Klasifikasi Bidang Ilmu pada Publikasi Terindeks Scopus Menggunakan Metode K-Nearest Neighbor (Aina Nurul Maulida)

Nama Data	Jumlah	Jumlah Angka
Data <i>Training</i>	90%	675
Data <i>Testing</i>	10%	75

Yang mana nantinya dalam perhitungan akurasi, akan menghasilkan sebuah confusion matrix, yang mana confusion matrix ini akan digunakan dalam mencari nilai akurasi, presisi, dan *recall* dengan berbagai varian jumlah K yaitu 5,15,20,25,45,50 yang nantinya akan mencari nilai K terbaik dalam melakukan klasifikasi judul publikasi.

Tabel 3. Hasil Pengukuran Akurasi, Presisi, *Recall*

Jumlah K	Akurasi	Presisi	Recall
5	0.5384	0.3768	0.2977
15	0.4355	0.1476	0.2133
20	0.4543	0.2409	0.2133
25	0.4464	0.2401	0.2044
45	0.4525	0.1817	0.2177
50	0.4267	0.1389	0.2133

Berdasarkan dari pengujian yang sudah dilaksanakan, bisa dilakukan penarikan kesimpulan bahwa pada pengujian akurasi berdasarkan nilai K yang telah ditentukan sebelumnya, sistem ini memiliki nilai akurasi yang kurang sesuai dengan yang diharapkan, karena nilai akurasi tertinggi pada K=5 yaitu dengan nilai 0.53, nilai presisi yaitu 0.376, dan nilai *recall* yaitu 0.297. Sedangkan nilai terendah, pada nilai K=50 dengan akurasi yaitu 0.42, nilai presisi 0.13, dan nilai *recall* 0.21.

4. KESIMPULAN

Berlandaskan penelitian ini bias dilakukan penarikan kesimpulan bahwa Klasifikasi Bidang Ilmu pada Publikasi Terindeks Scopus Menggunakan Metode *K-Nearest Neighbor*, menunjukkan hasil yang cukup dengan akurasi sejumlah 0.53, presisi sebesar 0.37, dan *recall* sebesar 0.29. Namun hasil ini masih cukup rendah bila digunakan untuk database SINTA dikarenakan pada saat pelabelan data yang diambil tidak tepat. Akan tetapi sistem sudah berjalan dengan baik sesuai yang peneliti harapkan yaitu untuk mengklasifikasi bidang ilmu sesuai dengan 5 bidang keilmuan tersebut.

ACKNOWLEDGEMENTS

Tugas akhir ini merupakan bagian dari penelitian dosen Teknik Informatika Universitas Islam Sultan Agung berjudul Pengembangan Sistem Klasifikasi otomatis Publikasi Ilmiah pada database SINTA *Science and Technology Index* yang dibiayai dari Kementerian Pendidikan dan Kebudayaan, Pendidikan Tinggi, Riset dan Teknologi Republik Indonesia tahun 2022.

DAFTAR PUSTAKA

- [1] U. Andayani, "Perkembangan Kajian Pendidikan Islam Pada Publikasi Internasional Terindeks Scopus Tahun 2017-2021," *Al Maktab.*, vol. 20, pp. 11–20, 2021.
- [2] D. A. (Universitas M. R. A. H. Siti Julaiha, Marteli Bettiza, "Penerapan Algoritma K-Nearest Neighbor (KNN) Untuk Klasifikasi Calon Penerima Bidikmisi," *Student Online J.*, vol. 2, pp. 230–235, 2021.
- [3] Y. I. Kurniawan and T. I. Barokah, "Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan K-Nearest Neighbor," *J. Ilm. Matrik*, vol. 22, no. 1, pp. 73–82, 2020, doi: 10.33557/jurnalmatik.v22i1.843.
- [4] M. Diki and B. Nurina, "Penerapan Algoritma K-Nearest Neighbor dalam Klasifikasi Judul Berita Hoax," 2022.
- [5] A. Rahmat Dian Nugraha, K. Auliasari, and Y. Agus Pranoto, "IMPLEMENTASI METODE K-NEAREST NEIGHBOR (KNN) UNTUK SELEKSI CALON KARYAWAN BARU (Studi Kasus : BFI Finance Surabaya)," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 4, no. 2, pp. 14–20, 2020, doi: 10.36040/jati.v4i2.2656.
- [6] D. Pajri, Y. Umaidah, and T. N. Padilah, "K-Nearest Neighbor Berbasis Particle Swarm Optimization untuk Analisis Sentimen Terhadap Tokopedia," *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 242–253, 2020, doi: 10.28932/jutisi.v6i2.2658.
- [7] H. A. D. Fasnuari, H. Yuana, and M. T. Chulkamdi, "PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENYAKIT DIABETES MELITUS STUDI KASUS : WARGA DESA JATITENGAH," vol. 16, no. 2, pp. 133–142, 2022.

-
- [8] Danilo Gomes de Arruda, "IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR UNTUK MENGANALISIS SENTIMEN USER TWITTER," p. 6, 2021.
- [9] R. Sagita, U. Enri, and A. Primajaya, "Klasifikasi Berita Clickbait Menggunakan K-Nearest Neighbor (KNN)," vol. 5, no. 2, pp. 230–238, 2020, doi: 10.33633/joins.v5i2.3705.
- [10] I. A. Nikmatun and I. Waspada, "Implementasi Data Mining untuk Klasifikasi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *J. SIMETRIS*, vol. 10, no. 2, pp. 421–432, 2019.