

Klasifikasi Bidang Ilmu Pada Publikasi Terindeks GARUDA Kemdikbud Menggunakan Metode K-Nearest Neighbor (KNN)

Farikhatus Sa'adah, Imam Much Ibnu Subroto, Andi Riansyah

Fakultas Teknologi Industri, Universitas Islam Sultan Agung Semarang

rikaa@std.unissula.ac.id, imam@nissula.ac.id, andi@unissula.ac.id

Abstract

GARUDA merupakan salah satu portal yang memuat referensi ilmiah nasional Indonesia dan memberikan akses terhadap karya ilmiah yang dihasilkan oleh akademisi dan peneliti Indonesia. Klasifikasi merupakan proses penemuan model (fungsi) yang menggambarkan dan membedakan kelas data. Sebelumnya belum ada penelitian yang menggunakan 5 bidang ilmu utama yaitu arts & Humanities, Engineering & Technology, Life Sciences & Medicine, Natural Sciences, Social Sciences & Management, untuk mengklasifikasikan judul artikel pada aplikasi terindeks Garuda. Dalam penelitian ini akan mengklasifikasikan judul artikel yang ada pada aplikasi terindeks Garuda kedalam 5 bidang ilmu yang akan menggunakan metode algoritma K-Nearest Neighbor (K-NN), dan akan dilakukan dengan cara mengumpulkan data pada aplikasi terindeks Garuda, lalu dilakukan tahap preprocessing agar data memiliki kualitas yang baik. Setelah itu mencari data dengan jarak terdekat dengan tetangga terdekat pada data latih dengan data yang akan diuji, lalu evaluasi yang akan digunakan adalah accuracy, precision, recall, dan F-score. Hasil yang diharapkan dari penelitian ini adalah metode algoritma K-Nearest Neighbor (K-NN) dapat menghasilkan tingkat akurasi yang tinggi dalam mengklasifikasikan judul artikel pada aplikasi terindeks Garuda, kedalam 5 bidang ilmu utama yaitu arts & Humanities, Engineering & Technology, Life Sciences & Medicine, Natural Sciences, Social Sciences & Management secara cepat dan otomatis. Hasil dari penelitian ini mendapatkan nilai akurasi tertinggi 0,499 atau 49,9% dengan menggunakan parameter nilai K = 285.

Kata Kunci: Garuda, Klasifikasi, K-NN

1. PENDAHULUAN

GARUDA merupakan salah satu portal referensi ilmiah nasional Indonesia dan memberikan akses terhadap karya ilmiah yang dihasilkan oleh akademisi dan peneliti Indonesia. Garuda Kemdikbud (Kementerian Pendidikan dan Kebudayaan) Dikti pertama diluncurkan tahun 2018 oleh kemenristek (Kementerian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia), sistem ini sebelumnya bernama *Indonesia Publication Index (IPI)* yang mulai beroperasi pada tahun 2018. Saat ini jumlah terindex Garuda sejumlah 1859556 artikel dari 14086 jurnl nasional. Portal Garuda ini digunakan oleh Kementerian Dikbud (Kementerian Pendidikan dan Kebudayaan), untuk memonitor perkembangan publikasi secara nasional dan untuk menilai kualitas publikasi melalui akreditasi jurnal nasional. KNN adalah metode yang digunakan untuk mengklasifikasikan objek yang ada pada data pembelajaran (*neighbor*) yang jaraknya paling dekat dengan objek tersebut. Jarak *neighbor* biasanya dihitung dengan menggunakan jarak *Eclidien*. Memerlukannya sistem klasifikasi yang akan digunakan untuk mencari informasi. Metode KNN dapat dibagi menjadi 2 fase yaitu training dan testing [1]. Proses metode KNN dikerjakan dengan mencari nilai K yang cocok untuk data latih pada jumlah data dari tetangga terdekat pada objek data uji. Berikut adalah langkah proses menggunakan metode KNN:

- Menentukan parameter k (jumlah tetangga paling dekat)
- Menghitung jarak antara data yang akan dievaluasi dengan semua data training
- Mengurutkan jarak yang muncul
- Menentukan jarak terdekat sampai urutan k
- Memasangkan kelas yang bersesuaian
- Mencari jumlah dari data tetangga yang paling dekat dan menentukan kelas tersebut sebagai kelas data yang akan dievaluasi.

Rumus untuk menghitung jarak pada metode KNN:

$$d(P, Q) = \sqrt{\sum_i^n (p_i - q_i)^2} \quad (1)$$

Dimana :

- d (P, Q) : Jarak *euclidian*
 n : Jumlah data training
 P : Inputan data ke -1 dari data training
 Q : Inputan data ke -1 dari data testing[2]

Data mining adalah suatu proses untuk mengolah data dari berbagai sumber, yang hasilnya adalah sebuah informasi. Secara teknis data mining dapat disebut sebagai suatu proses mencari korelasi atau pola dari ratusan atau ribuan *field* dari sebuah database yang sangat besar [3].

Klasifikasi merupakan tahapan dalam perancangan model (fungsi) yang memisahkan kelas dari data yang berguna untuk memprediksi kelas dari objek yang tidak mengetahui label kelasnya. Adapun tahap-tahap untuk melakukan klasifikasi data yaitu yang pertama adalah learning, yang mana algoritma klasifikasi digunakan untuk menganalisa data training setelah itu dipresentasikan dengan rule klasifikasi. Proses yang kedua yaitu klasifikasi, proses ini yang akan dipakai pada data testing untuk menentukan hasil akurasi dari rule klasifikasi [2]. Dari tahap-tahap klasifikasi data adapun data testing yang digunakan untuk menentukan akurasi dari rule klasifikasi. Adapun proses klasifikasi yang didasarkan pada empat komponen yaitu:

1. *Class*
 Class merupakan variabel dependen yang berupa katagori yang menampilkan label dari objek, seperti contohnya: resiko penularan covid, jenis buah, jenis ikan.
2. *Predictor*
Prediktor merupakan variabel independen yang menampilkan atribut atau karakteristik pada data. Diantara Contohnya adalah: merokok, tabungan, aset, gaji.
3. *Training Dataset*
Training dataset memuat data yang akan dilatih yang digunakan untuk menentukan class yang cocok untuk *predictor*.
4. *Testing Dataset*
Testing Dataset memuat data yang digunakan untuk diuji pada klasifikasi oleh model dan hasil akurasi pada klasifikasi akan dievaluasi [4].

Evaluasi dalam penelitian ini digunakan untuk mengukur kerja metode penelitian agar berjalan sesuai alur, efektivitas metode klasifikasi ini diukur dengan empat parameter statistik yang digunakan untuk mengujinya yaitu, *accuracy*, *precision*, *recall*, dan *F-score*. *Precision* adalah tingkat nilai positif pada hasil klasifikasi yang dirumuskan sebagai berikut:

$$Precision = \frac{tp}{tp+fp} \quad (2)$$

Keterangan:

tp = Jumlah positif benar

fp = Jumlah positif palsu

Sedangkan *recall* adalah persentasi kegagalan yang diidentifikasi oleh model. Rumusnya adalah sebagai berikut:

$$Recall = \frac{tp}{tp+fn} \quad (3)$$

Keterangan:

tp = Jumlah positif benar

fn = Jumlah negatif palsu

Accuracy berarti pengukuran seberapa bagus sebuah model dengan benar mengidentifikasi kelas. *Accuracy* dirumuskan sebagai berikut :

$$Accuracy = \frac{tp+tn}{tp+tn+fp+fn} \quad (4)$$

Keterangan:

tp = Jumlah positif benar

tn = Jumlah negatif benar

fp = Jumlah positif palsu

fn = Jumlah negatif palsu

F-score atau *F-measure* adalah pengukuran tingkat akurasi dengan menggunakan nilai *presisi* dan *recall*, sehingga di dalam pengukuran *F-measure* tidak menyertakan nilai *true negative* .

$$F - measure = \frac{2 \times precision \times recall}{precision + recall} \quad (5)$$

Pada penelitian-penelitian sebelumnya dengan mencari performa metode KNN dengan parameter K. pada penelitian sebelumnya dengan mengklasifikasikan pengguna protokol komunikasi pada trafik jaringan menggunakan metode KNN. Trafik jaringan merupakan jalur lintas komunikasi data dalam jaringan yang telah diidentifikasi dengan informasi dari header paket data. Klasifikasi pada trafik digunakan untuk mendesain rancangan arsitektur jaringan. Data yang akan uji didapat dari *capture* trafik jaringan dengan bantuan aplikasi *wireshark*. Hasil klasifikasi pada penelitian ini menghasilkan akurasi yang sangat tinggi yaitu 99,14% dari perhitungan K = 3 [5].

Penelitian selanjutnya mengklasifikasikan pangajuan kartu kredit menggunakan metode KNN. Pada saat pengajuan kartu kredit biasanya terdapat kategori yang berbeda yang bisa menentukan nasabah termasuk dalam kategori kartu kredit yang mana. Dengan menggunakan metode KNN data pada nasabah bank BNI Syariah Surabaya mendapatkan variabel yang dapat digunakan untuk faktor pendukung yaitu, jenis kelamin, status rumah tangga, status, jumlah anak, pekerjaan, dan penghasilan per tahun. Hasil dari klasifikasi yang menggunakan parameter K = 5 mendapatkan akurasi sebanyak 93%. Dengan hasil akurasi yang begitu besar sistem ini efektif untuk membantu pihak analis kartu kredit dalam mengklasifikasi kategori kartu kredit pada nasabah [6].

Pada penelitian selanjutnya menganalisa ulasan pengguna pada aplikasi bibit dan bareksa dengan metode KNN. Aplikasi bibit dan bareksa merupakan aplikasi investasi online. Investasi online adalah menanamkan modal dalam jangka waktu yang lama agar mendapatkan keuntungan yang dilakukan secara online. Aplikasi bibit dan bareksa dapat diunduh di *google play store*. Jumlah ulasan yang akan dijadikan data sebanyak 998 data untuk aplikasi bareksa yang terdiri dari 484 ulasan positif dan 514 ulasan negatif, sedangkan data untuk aplikasi bibit menggunakan sebanyak 1063 data yang terdiri dari 541 ulasan positif dan 522 untuk ulasan negatif. Setelah data diproses dengan menggunakan model CRISP-DM (*Cross Industry Standard Process for Data Mining*) dan metode KNN serta perbandingan 60:40, mendapatkan hasil tertinggi 91,91% untuk bibit dan 87,15 % untuk bareksa [7].

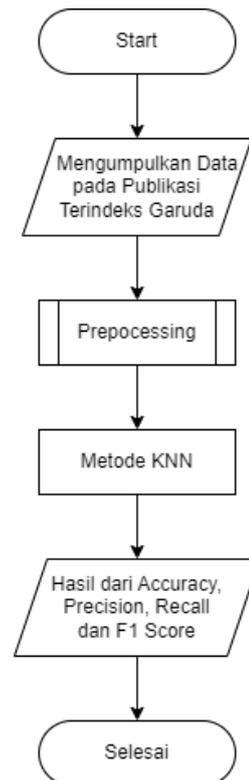
Penelitian selanjutnya mengklasifikasikan jenis tanaman hias daun Philodendron menggunakan metode KNN berdasarkan nilai HSV. Tanaman hias daun biasanya digunakan sebagai hiasan rumah agar rumah terlihat lebih berwarna. Namun dari beragam jenis tanaman hias daun yang ada membuat para pecinta tanaman hias dan kesulitan dalam membedakan jenisnya. Jadi dibutuhkannya sistem untuk mendeteksi jenis tanaman hias daun, dengan memanfaatkan citra dari tanaman hias daun. Cara kerja sistem ini nantinya akan membandingkan citra pada data *training* yang sudah tersimpan pada database dengan citra data *testing*. Citra data *testing* yang telah diterapkan metode KNN berguna untuk menghitung jarak tetangga terdekat antar pixel dari data *training*. Setiap pixel pada citra dilakukan proses konversi *red, green, blue* (RGB) kedalam ekstraksi HSV. Hasil dari data yang digunakan 5 citra data latih dan 10 citra data uji, menghasilkan nilai akurasi 92% [8].

Pada penelitian selanjutnya mengklasifikasi harga *cell phone* menggunakan metode KNN. Pada zaman sekarang ponsel yang memiliki fitur-fitur yang didalamnya semakin canggih teknologinya maka spesifikasinya semakin bertambah bagus dan mahal. Fitur yang ada pada ponsel pada umumnya sekarang memiliki kamera depan belakang yang resolusinya jernih, Ram yang lebih besar, kapasitas memori internal lebih besar, dan kapasitor baterai lebih besar, dimana semakin banyak fitur canggih sudah dimiliki oleh ponsel maka dapat digunakan dalam menetapkan harga dari ponsel. Pada penelitian ini menghitung jarak kemiripan memakai minkoswki, dimana jarak yang ada didalam ruang vektor yang sudah ditentukan dianggap sebagai generalisasi dari kedua jarak. Hasil akurasi yang didapat dari penelitian ini sebesar 82% dengan menggunakan nilai K = 8 [9].

Pada penelitian selanjutnya mengklasifikasikan bumbu dapur indonesia menggunakan metode KNN. Bumbu dapur diindonesia sangat beragam jenis, bentuk, warna dan bau. Kesalahan dalam pemilihan bumbu dapur dapat mempengaruhi rasa pada masakan. Dengan adanya image preprocessing dapat dimanfaatkan untuk mengenali objek citra yang ditangkap menggunakan kamera. Pada penelitian ini hanya digunakan untuk mengklasifikasikan bumbu dapur yang mirip diantaranya, jahe, kencur dan kunyit. Dan hasil penelitian ini mendapatkan akurasi yang didapatkan cukup tinggi sebesar 73% [10].

Berdasarkan penjelasan diatas maka perumusan masalah yang diambil dari penelitian ini adalah bagaimana mengklasifikasikan data publikasi sesuai bidang ilmu masing-masing dari semua artikel yang terindeks Garuda menggunakan metode KNN. Dan tujuan dari penelitian ini adalah untuk mengklasifikasikan bidang ilmu pada publikasi terindeks Garuda kedalam 5 bidang ilmu menggunakan metode KNN.

2. METODE PENELITIAN



Gambar 1 Flowchart metode penelitian

Pada gambar 1 menjelaskan tahapan – tahapan penelitian pada penelitian ini, yang pertama dilakukan adalah mengumpulkan data yang berupa jurnal artikel yang ada pada aplikasi terindeks Garuda, dan setelah itu akan divalidasi oleh tim pengembang aplikasi terindeks Garuda. Setelah data terkumpul maka akan dilakukan tahap – tahap preprocessing yaitu pengkodean judul artikel dari data yang telah dikumpulkan. Kemudian data dikode dan akan dilakukan penerapan metode, dan metode yang digunakan adalah algoritma KNN, dimana data yang ada akan dicari jarak terdekat dengan data uji tetangga terdekat. Setelah itu akan dilakukan evaluasi dengan menggunakan empat parameter statistik yaitu *accuracy*, *precision*, *recall* dan *F1-score*. Yang hasil outputnya akan digunakan oleh sistem pada aplikasi pulikasi terindeks Garuda untuk mengetahui judul jurnal yang diinput oleh *user* termasuk kedalam salah satu dari 5 bidang ilmu. Dalam tahapan penelitian, algoritma atau metode yang akan digunakan adalah KNN, dimana metode ini akan melakukan penentuan keputusan dari data yang ada, dan adapun tahapan yang akan dilaksanakan dalam melakukan penelitian ini:

2.1 Pengumpulan data pada pulikasi terindeks GARUDA

Pengumpulan data dengan menggunakan data pada pulikasi terindeks Garuda. Jumlah data yang digunakan untuk eksperimen sebanyak 500 judul artikel yang terdiri atas 5 bidang ilmu yaitu *Arts & Humanities*, *Engineering & Technology*, *Life Scienses & Madicine*, *Natural Science*, *Social Science & Managemet*, yang masing-masing sebanyak 100 judul artikel.

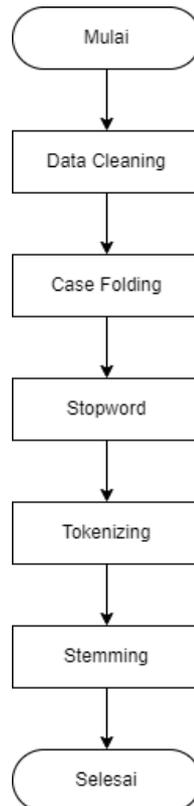
Tabel 1 pengumpulan data sesuai bidang ilmu

| No | Data | Bidang Ilmu |
|----|-------------------------------------|--|
| 1. | <i>Arts & Humanities</i> | Typology Jihatul Ka'bah on Qibla Direction of Mosque in Semarang. |
| 2. | <i>Engineering & Technology</i> | Sistem Konversi Multiguna dan Ensklopedia Berbasis Visual |
| 3. | <i>Life & Madicine</i> | Bisoprolol Therapy and The Risk of Eraclite Dysfunction in Stable Coronary Artery Disease Patients |

- | | | |
|----|--|---|
| 4. | <i>Natural Science</i> | Kajian Eksperimental Konversi Energi Gelombang Dengan Menggunakan Kombinasi Wells Rotor Dan Turbin Darrieus. |
| 5. | <i>Social Science & Management</i> | Dampak Program Garnas Kakao Terhadap Pendapatan Petani Di Desa Manyamba Kecamatan Tammerodo Sendana Kabupaten Majene. |

Pada tabel 1 menjelaskan contoh data yang telah dikumpulkan dari publikasi terindeks Garuda dan telah dilabeli sesuai dengan 5 bidang ilmu.

2.2 Preprocessing



Gambar 2 Flowchart preprocessing

Pada gambar 2 menjelaskan langkah-langkah preprocessing. Preprocessing berisi pengkodean data pada judul artikel yang ada dipublikasi terindeks Garuda. Lalu diklasifikasikan sesuai dengan 5 bidang ilmu yaitu *Arts & Humanities, Engineering & Technology, Life Sciences & Medicine, Natural Science, Social Science & Managemet*. Preprocessing digunakan untuk menghilangkan permasalahan yang dapat mengganggu pada saat pemrosesan data, selain itu preprocessing juga dapat berjalan lebih efektif dan efisien pada saat pemrosesan data. Adapun langkah – langkah preprocessing yaitu Case folding, Data Cleaning, Tokenizing, Stopword, dan Stemming.

2.3 Penarapan Metode KNN

Metode yang diterapkan atau yang digunakan pada penelitian ini menggunakan Algoritma atau metode KNN yang dimana cara kerja metode ini dengan cara, mencari data dengan jarak terdekat dengan tetangga terdekat pada data latih dengan data yang akan diuji.

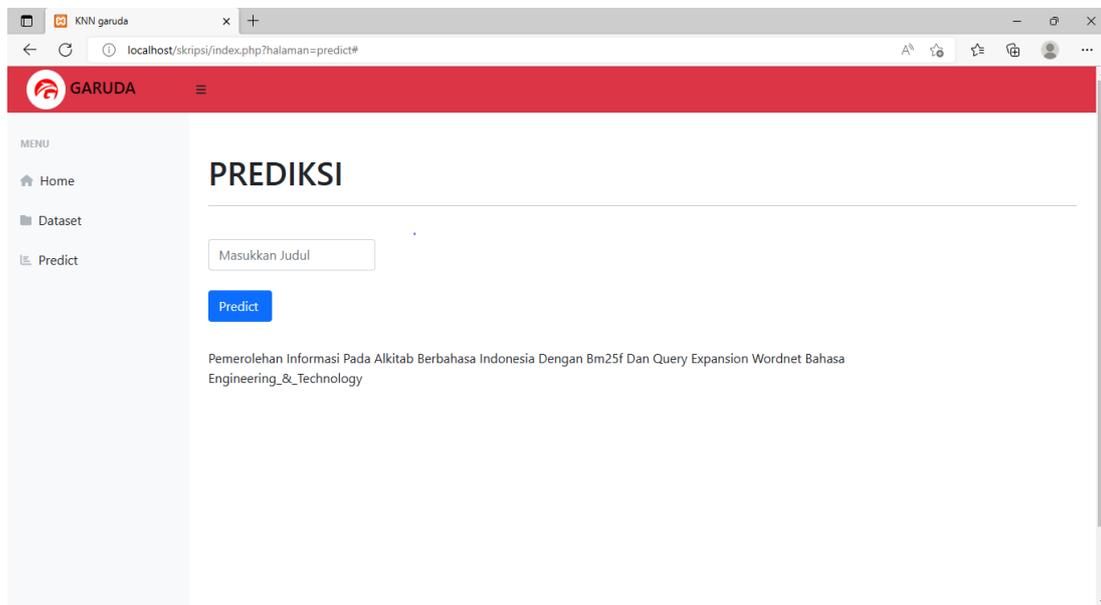
2.4 Evaluasi

Dalam evaluasi menggunakan empat parameter statistik untuk mengujinya yaitu *accuracy, precision, recall, dan F-score*. Dalam *accuracy* mengukur hasil kedekatan dengan nilai sebenarnya atau nilai target. Precision akan membandingkan antara jumlah positif benar dengan banyaknya data

yang telah diprediksi benar. Dan recall akan membandingkan antara data yang positif benar dengan banyaknya data yang sebenarnya positif. Jika F-score mempunyai skor yang baik maka mengidentifikasi bahwa model klasifikasi mempunyai precision dan recall yang baik.

3. HASIL DAN ANALISA (10 PT)

3.1 Hasil



Gambar 3 Halaman Hasil Prediksi

Pada gambar 3 menampilkan halaman prediksi halaman ini akan menampilkan hasil dari proses judul artikel yang telah klasifikasikan kedalam 5 bidang ilmu. Hasil yang ditampilkan adalah nama judul artikel yang telah diinputkan oleh *user* dan dibawahnya hasil dari akurasi bidang ilmu paling tinggi yang akan ditampilkan oleh sistem.

3.2 Analisa

Untuk mengklasifikasikan aplikasi terindeks Garuda dalam penelitian ini menggunakan 500 data yang terdiri dari 20% data testing yaitu 100 data dan 80% data testing yaitu 400 data. Yang mana nantinya dalam perhitungan akurasi, akan menghasilkan sebuah *confusion matrix*, dimana *confusion matrix* ini akan digunakan dalam mencari nilai *accuracy*, *precision*, *recall* dan *f-score* yang nantinya dengan berbagai jumlah K diantaranya yaitu, 41, 85, 87, 193, 197, 285, 293 yang nantinya akan dicari nilai K paling terbaik dalam melakukan klasifikasi aplikasi terindeks Garuda. Pada tabel dibawah adalah hasil dari perhitungan *confusion matrix* dari berbagai nilai K.

Tabel 2 Hasil pengukuran Accuracy, Recall, Precision, F1 Score dari berbagai nilai K

| Jumlah K | Accuracy | Recall | Precision | F1 Score |
|----------|----------|--------|-----------|----------|
| 41 | 0.440 | 0.2 | 0.04 | 0.06 |
| 85 | 0.454 | 0.210 | 0.240 | 0.086 |
| 87 | 0.468 | 0.220 | 0.240 | 0.104 |
| 193 | 0.484 | 0.24 | 0.132 | 0.127 |
| 197 | 0.471 | 0.229 | 0.118 | 0.114 |
| 285 | 0.499 | 0.26 | 0.121 | 0.150 |
| 293 | 0.491 | 0.25 | 0.122 | 0.135 |
| 385 | 0.402 | 0.18 | 0.059 | 0.075 |

Pada tabel 2 menampilkan hasil dari pengukuran akurasi dari berbagai nilai K. Berdasarkan dari pengujian yang telah dilaksanakan, dapat diambil kesimpulan bahwa dalam akurasi pada penelitian ini berdasarkan K yang telah diuji sebelumnya, dari nilai K yang telah diuji sistem ini mendapatkan nilai akurasi yang kurang dari yang diharapkan. Diantara nilai K yang telah diuji mendapatkan nilai skor *accuracy* paling tinggi dibandingkan nilai K lainnya adalah K = 285 dengan nilai *accuracy* 0.499 atau 49,9%, sedangkan pada K lainnya *accuracy* yang terkecil yaitu 0.402 sampai 0.491. Dan untuk skor dari *recall* dan *f1 score* dari nilai K = 11 juga mendapat nilai tertinggi yaitu 0.26 dan 0.150, akan tetapi pada *precision* mendapat skor 0,122 dimana tidak tertinggi diantara *precision* dari nilai K lainnya

4. KESIMPULAN

Dapat disimpulkan bahwa sistem sudah dapat mengklasifikasi kedalam 5 bidang ilmu utama, akan tetapi hasil dari pengujian data dari aplikasi terindeks Garuda menggunakan metode KNN memperoleh hasil akurasi yang belum sesuai yaitu, 49,9% dengan menggunakan parameter K = 285, yang seharusnya akurasi yang baik itu lebih dari 70%. Hasil pengujian ini tidak sesuai yang diharapkan dikarenakan data yang digunakan masih ada persamaan kata atau kalimat dari satu bidang ilmu dan bidang ilmu lainnya ataupun disebabkan data yang dipakai tidak sesuai menggunakan metode KNN.

ACKNOWLEDGEMENTS

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

DAFTAR PUSTAKA

- [1] M. M. Baharuddin, H. Azis, and T. Hasanuddin, "Analisis Performa Metode K-Nearest Neighbor Untuk Identifikasi Jenis Kaca," *Ilk. J. Ilm.*, vol. 11, no. 3, pp. 269–274, 2019, doi: 10.33096/ilkom.v11i3.489.269-274.
- [2] 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.
- [3] H. Rofiq, K. C. Pelangi, and Y. Lasena, "Penerapan Data Mining Untuk Menentukan Potensi Hujan Harian Dengan Menggunakan Algoritma Naive Bayes," *J. Manaj. Inform. dan Sist. Inf.*, vol. 3, no. 1, pp. 8–15, 2020, [Online]. Available: <http://mahasiswa.dinus.ac.id/docs/skripsi/jurnal/19417.pdf>.
- [4] F. Agus, H. R. Hatta, and M. Mahyudin, "Pengklasifikasian Dokumen Berbahasa Arab Menggunakan K-Nearest Neighbor," *J. SIFO Mikroskil*, vol. 18, no. 1, pp. 43–56, 2017, doi: 10.55601/jsm.v18i1.413.
- [5] K. K. A. Subrata, I. M. O. Widyantara, and L. Linawati, "Klasifikasi Penggunaan Protokol Komunikasi Pada Trafik Jaringan Menggunakan Algoritma K-Nearest Neighbor," *Maj. Ilm. Teknol. Elektro*, vol. 16, no. 1, p. 67, 2016, doi: 10.24843/mite.1601.10.
- [6] K. Neighbor *et al.*, "Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan," vol. 22, no. 1, pp. 73–82, 2020.
- [7] A. Dwiki *et al.*, "Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit Dan Bareksa Dengan Algoritma KNN," vol. 8, no. 2, pp. 636–646, 2021.
- [8] D. Syahid, J. Jumadi, and D. Nursantika, "Sistem Klasifikasi Jenis Tanaman Hias Daun Philodendron Menggunakan Metode K-Nearest Neighbor (KNN) Berdasarkan Nilai Hue, Saturation, Value (HSV)," *J. Online Inform.*, vol. 1, no. 1, p. 20, 2016, doi: 10.15575/join.v1i1.6.
- [9] S. H. Putra, "Klasifikasi Harga Cell Phone menggunakan Metode K-Nearest Neighbor (KNN)," vol. 4, no. 1, pp. 978–979, 2018.
- [10] S. Y. Riska, L. Farokhah, I. Teknologi, and A. Malang, "Klasifikasi Bumbu Dapur Indonesia Menggunakan Metode K- Nearest Neighbors (K-NN)," vol. 11, pp. 37–42, 2021.