

Sistem Pakar Untuk Mendiagnosa Gangguan Somatisasi Menggunakan Metode *K-Nearest Neighbors* (KNN)

Rima Ruktiari Ismail^{1*}, Hamid Wijaya², Juarni Siregar³, Nurhasan Nugroho⁴

Fakultas Teknologi Informasi, Universitas Sembilanbelas November Kolaka^{1,2}
Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Nusa Mandiri³
Program Studi Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Bina Bangsa⁴
rimaruktiarii@gmail.com^{1*}, hamidwijaya35@gmail.com², juarni.jsr@nusamandiri.ac.id³,
nurhasan.nugroho@binabangsa.ac.id⁴

*) Corresponding Author

(received: 13-09-24, revised: 20-09-24, accepted: 28-10-24)

Abstract

This study aims to develop an expert system capable of diagnosing somatic symptom disorder using the K-Nearest Neighbors (KNN) method. Somatic symptom disorder is a psychological condition that is difficult to diagnose due to its physical symptoms, which actually stem from psychological issues. The ambiguity of these symptoms often leads to unnecessary and costly medical examinations, adding to the burden on both patients and the healthcare system. KNN was chosen for its ability to classify by comparing test data with training data based on proximity, using Euclidean Distance. Euclidean Distance is used to measure the shortest distance between two points in feature space, calculated by taking the square root of the sum of the squared differences between the feature values of the two points. The results show that the developed expert system achieved a high accuracy rate of 92.5%, indicating that the KNN method with Euclidean Distance is effective in diagnosing somatic symptom disorder. Factors such as optimal K value selection and data normalization play a crucial role in the system's success. A significant contribution of this research is the demonstration that KNN can be effectively implemented in expert systems to support medical professionals in diagnosing somatic symptom disorder with high accuracy and reliability.

Keywords: Somatic Symptom Disorder, Expert System, K-Nearest Neighbor, KNN, Euclidean Distance

Abstrak

Penelitian ini bertujuan untuk mengembangkan sistem pakar yang mampu mendiagnosa gangguan somatisasi menggunakan metode *K-Nearest Neighbors* (KNN). Gangguan somatisasi merupakan kondisi psikologis yang sulit didiagnosis karena gejalanya yang bersifat fisik namun berasal dari masalah psikologis. Ketidakjelasan gejala ini sering kali mengarah pada pemeriksaan medis yang tidak diperlukan dan mahal, menambah beban bagi pasien dan sistem kesehatan. KNN dipilih karena kemampuannya untuk melakukan klasifikasi dengan membandingkan data uji dengan data pelatihan berdasarkan kedekatan menggunakan *Euclidean Distance*. *Euclidean Distance* digunakan untuk mengukur jarak terpendek antara dua titik dalam ruang fitur, yang dihitung dengan mengakar kuadrat dari jumlah perbedaan kuadrat antara nilai-nilai fitur dari dua titik tersebut. Hasil penelitian menunjukkan bahwa sistem pakar yang dikembangkan memiliki akurasi yang tinggi, yaitu mencapai 92,5%, yang mengindikasikan bahwa metode KNN dengan *Euclidean Distance* efektif dalam mendiagnosa gangguan somatisasi. Faktor-faktor seperti pemilihan nilai K yang optimal dan normalisasi data berperan penting dalam keberhasilan sistem ini. Kontribusi signifikan dari penelitian ini adalah pembuktian bahwa KNN dapat diimplementasikan secara efektif dalam sistem pakar untuk mendukung tenaga medis dalam melakukan diagnosis gangguan somatisasi dengan akurasi yang tinggi dan keandalan yang baik.

Kata Kunci: Gangguan Somatisasi, Sistem Pakar, *K-Nearest Neighbors*, KNN, *Euclidean Distance*

I. Pendahuluan

Gangguan somatisasi merupakan salah satu jenis gangguan mental di mana individu mengalami berbagai gejala fisik yang tidak dapat dijelaskan secara medis, seringkali berkaitan dengan stres atau kondisi psikologis.

Gejala-gejala ini sering kali menyerupai penyakit fisik, seperti nyeri, kelelahan, atau gangguan pencernaan, tanpa adanya penyebab medis yang jelas [1]. Pasien gangguan somatisasi sering kali menjalani berbagai pemeriksaan medis yang invasif dan mahal tanpa menemukan penyebab yang jelas, yang pada akhirnya dapat menyebabkan stres tambahan bagi pasien serta beban finansial bagi sistem kesehatan [2]. Menurut data dari World Health Organization (WHO), kurang lebih 10-15% dari populasi umum mengalami gejala somatisasi yang tidak dapat dijelaskan secara medis [3]. Di Indonesia, prevalensi gangguan somatisasi diperkirakan mencapai 5-7% dari populasi, dengan banyak kasus yang tidak terdiagnosis atau salah didiagnosis [4]. Kondisi ini menciptakan tantangan signifikan dalam diagnosis dan pengelolaan kesehatan, mengingat ketidakpastian diagnosis sering kali mengarah pada penanganan medis yang tidak efektif, serta meningkatkan beban pada sistem kesehatan. Dampak dari ketidakjelasan diagnosis ini tidak hanya merugikan pasien secara individu, tetapi juga menimbulkan beban ekonomi yang signifikan karena biaya medis yang tidak perlu. Dalam konteks akademik, masalah ini menarik perhatian karena kompleksitas diagnosis yang melibatkan interaksi antara aspek psikologis dan fisik, menuntut pendekatan diagnostik yang lebih canggih. Secara praktis, peningkatan alat bantu diagnostik yang lebih akurat dan efisien sangat dibutuhkan untuk mengurangi kesalahan diagnosis dan meningkatkan efektivitas perawatan.

Permasalahan dalam mendiagnosa gangguan somatisasi sering kali disebabkan oleh kompleksitas gejala yang tidak spesifik dan adanya kesamaan gejala dengan penyakit fisik lainnya. Hal ini membuat tenaga medis sering kali mengalami kesulitan dalam menentukan diagnosis yang tepat, yang pada gilirannya dapat menyebabkan penanganan yang tidak efektif. Selain itu, kurangnya alat bantu diagnostik yang spesifik untuk gangguan somatisasi juga menjadi salah satu hambatan dalam proses diagnosis. Dalam konteks ini, penggunaan teknologi informasi dan sistem pakar menjadi sangat relevan. Meniru keahlian manusia dalam memecahkan masalah spesifik, sistem pakar merupakan program komputer yang dirancang untuk memberikan solusi layaknya seorang ahli [5]. Untuk mengembangkan sistem pakar, diperlukan metode khusus yang digunakan sebagai mesin inferensi, yang berfungsi untuk memproses informasi dari basis pengetahuan dan menerapkannya pada situasi spesifik pengguna [6].

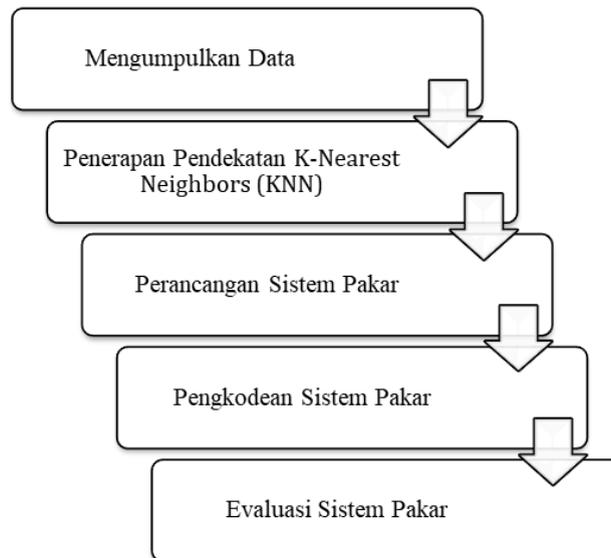
Sebagai salah satu mekanisme penalaran dalam pengembangan sistem pakar, metode K-Nearest Neighbors (KNN) kerap dimanfaatkan sebagai mesin inferensi. Prinsip kerja KNN melibatkan penentuan sejumlah K data terdekat dalam ruang fitur untuk mengklasifikasikan atau memprediksi suatu data baru berdasarkan kedekatan jaraknya. Untuk menentukan kelas dari data baru, KNN melihat kelas dari tetangga-tetangga terdekat tersebut dan mengambil mayoritas sebagai keputusan akhir. Algoritma ini termasuk non-parametrik, yang berarti tidak membuat asumsi apapun tentang distribusi data, sehingga sangat fleksibel dan dapat digunakan pada berbagai jenis data [7]. KNN memiliki kemampuan dalam klasifikasi data dengan cara membandingkan data baru dengan data yang sudah ada berdasarkan kedekatan atau kemiripan [8]. Keunggulan KNN terletak pada kemampuannya untuk menangani data yang tidak terlalu besar dengan tingkat akurasi yang baik, menjadikannya metode yang cocok untuk digunakan dalam sistem pakar di bidang medis [9].

Sejumlah peneliti terdahulu telah menerapkan metode KNN dalam penelitian terkait pembuatan sistem pakar. Terdapat penelitian terkait pengembangan sistem pakar untuk mendeteksi penyakit jantung menggunakan pendekatan KNN [10]. Berdasarkan studi ini, KNN mampu menghasilkan tingkat akurasi 70,40% untuk mendeteksi secara dini penyakit jantung. Penelitian berikutnya mengenai pengembangan sistem pakar untuk mendiagnosa kesehatan jiwa melalui implementasi pendekatan KNN [11]. Pada penelitian ini, KNN dapat mendiagnosa berdasarkan gejala dengan memiliki jarak terdekat dengan akurasi mencapai sebesar 90% pada penyakit ADHD dan 80% pada gangguan perilaku. Ada juga penelitian yang menggunakan KNN untuk diagnosa penyakit osteoarthritis [12]. Hasil dari uji akurasi menunjukkan nilai 91,67%, ini artinya pendekatan KNN mampu melakukan diagnosa dengan baik pada kasus diagnosa penyakit osteoarthritis.

Penelitian ini bertujuan untuk mengembangkan sistem pakar yang dapat melakukan diagnosa gangguan somatisasi dengan menggunakan metode K-Nearest Neighbors (KNN). Selain itu, dalam penerapan KNN digunakan Euclidean Distance untuk menentukan seberapa mirip data uji dengan data pelatihan yang akan didiagnosa. Titik-titik data yang memiliki jarak terdekat dengan data uji dianggap sebagai tetangga terdekat, dan kelas yang paling umum di antara tetangga-tetangga ini digunakan untuk memprediksi kelas dari data uji. Penelitian ini berkontribusi pada pengembangan sistem pakar yang dapat menjadi alat bantu dalam melakukan diagnosis, khususnya dalam diagnosa gangguan psikologis yang manifestasinya bersifat fisik. Untuk meningkatkan aksesibilitas pengguna dalam melakukan diagnosa, sistem pakar yang dirancang diimplementasikan dalam bentuk aplikasi berbasis web. Dengan demikian, penelitian ini tidak hanya memberikan solusi praktis bagi tantangan diagnosis gangguan somatisasi, tetapi juga berkontribusi pada pengembangan teknologi diagnostik yang lebih efisien dan efektif.

II. Metodologi Penelitian

Dalam sebuah penelitian, terutama yang berfokus pada pengembangan sistem berbasis teknologi seperti sistem pakar, tahapan penelitian yang terstruktur sangat penting untuk memastikan bahwa tujuan penelitian dapat tercapai dengan tepat. Setiap tahap dalam penelitian memiliki peran yang krusial, mulai dari pengumpulan data hingga evaluasi hasil [13]. Dengan mengikuti tahapan yang sistematis, peneliti dapat memastikan bahwa proses pengembangan berjalan sesuai rencana, mengidentifikasi dan mengatasi masalah yang mungkin timbul, serta menghasilkan sistem yang efektif dan dapat diandalkan [14]. Tahap-tahap dalam melakukan penelitian yang diterapkan direpresentasikan pada Gambar 1.



Gambar 1. Prosedur Penelitian

Berdasarkan pada Gambar 1, maka secara rinci setiap tahapan yang dilaksanakan dijelaskan sebagai berikut.

Mengumpulkan Data

Sebagai landasan keseluruhan proses riset, pengumpulan data menjadi langkah awal yang krusial dalam penelitian ini. Dalam konteks diagnosis gangguan somatisasi, data yang dikumpulkan berupa gejala, penyakit, dan bobot keyakinan. Pengumpulan data dilakukan melalui wawancara dan mengumpulkan informasi dengan psikiater. Tahap ini penting karena kualitas data yang dikumpulkan akan sangat mempengaruhi akurasi dan efektivitas sistem yang akan dikembangkan.

Penerapan Pendekatan *K-Nearest Neighbors* (KNN)

Setelah data terkumpul, langkah selanjutnya adalah menerapkan pendekatan *K-Nearest Neighbors* (KNN). KNN adalah metode klasifikasi yang sederhana namun efektif, yang bekerja dengan membandingkan data baru dengan data yang sudah ada berdasarkan kedekatan atau kemiripan [15]. Dalam penelitian ini, KNN digunakan untuk mengklasifikasikan gejala yang dialami pasien ke dalam kategori gangguan somatisasi atau bukan, berdasarkan data pasien sebelumnya. Penerapan KNN melibatkan proses seperti normalisasi data, pemilihan parameter K , dan perhitungan jarak antara data baru dengan data yang ada [16]. Tahap ini penting untuk memastikan bahwa sistem pakar memiliki model klasifikasi yang akurat dan dapat diandalkan [17].

Dalam sistem pakar yang menggunakan metode *K-Nearest Neighbors* (KNN) untuk mendiagnosa gangguan somatisasi, proses perhitungan dilakukan dengan membandingkan data gejala yang baru (data uji) dengan data gejala yang sudah ada dalam basis data (data pelatihan). Metode KNN menentukan klasifikasi dari data uji berdasarkan kedekatan atau kemiripan dengan sejumlah data pelatihan terdekat. Untuk setiap data uji, jarak antara data uji dan setiap data pelatihan dihitung [18]. Jarak yang paling umum digunakan dalam KNN adalah *Euclidean Distance*. *Euclidean Distance* adalah ukuran jarak terpendek antara dua titik dalam ruang fitur, yang dihitung dengan mengakar kuadrat dari jumlah perbedaan kuadrat antara nilai-nilai fitur dari dua titik tersebut [19]. Persamaan *Euclidean Distance* antara dua titik data x dan y dengan n fitur dapat menggunakan persamaan (1).

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

dimana $d(x, y)$ adalah jarak *Euclidean* antara titik data x dan y . x_i dan y_i adalah nilai dari fitur ke- i pada data x dan y . n adalah jumlah fitur (misalnya, berbagai gejala dalam konteks diagnosis gangguan somatisasi).

Perancangan Sistem Pakar

Tahap perancangan melibatkan pembuatan arsitektur sistem yang akan digunakan dan penentuan persyaratan yang dibutuhkan dalam pengembangan perangkat lunak [20]. Pada tahap ini, *use case diagram* digunakan untuk memodelkan interaksi antara pengguna dengan sistem pakar. *Use case diagram* membantu dalam mengidentifikasi kebutuhan fungsional dari sistem, termasuk hubungannya antara pengguna dengan sistem yang dikembangkan [21]. Perancangan ini juga mencakup pemilihan teknologi yang akan digunakan serta bagaimana sistem akan berinteraksi dengan pengguna dan data.

Pengkodean Sistem Pakar

Pengkodean sistem merujuk pada proses penerjemahan desain sistem yang telah dirancang ke dalam kode program yang dapat dieksekusi oleh komputer [22]. Sistem pakar ini dibangun berbasis *website* menggunakan bahasa pemrograman PHP dan *database* MySQL. Pemilihan *platform* berbasis *website* bertujuan untuk memastikan aksesibilitas yang lebih luas, memungkinkan tenaga medis untuk mengakses sistem dari berbagai perangkat dengan mudah. Pengkodean melibatkan pembuatan antarmuka pengguna yang intuitif, pengembangan logika sistem sesuai dengan model KNN yang telah dibangun, serta integrasi dengan *database* untuk penyimpanan data.

Evaluasi Sistem Pakar

Tahap akhir dalam penelitian ini adalah evaluasi sistem pakar, di mana sistem yang telah dikembangkan diuji untuk menilai kinerjanya. Tujuannya adalah mengevaluasi keakuratan model dalam memprediksi hasil menggunakan data yang belum pernah diproses sebelumnya [23]. Evaluasi dilakukan dengan menggunakan uji akurasi untuk mengukur seberapa baik sistem dalam mendiagnosa gangguan somatisasi dibandingkan dengan data yang telah terlabeli. Pengujian akurasi dilakukan melalui perbandingan hasil diagnosis sistem dengan ahli dalam hal ini psikiater. Hasil evaluasi akan menjadi dasar untuk perbaikan sistem sebelum diterapkan secara luas. Untuk mendapatkan nilai akurasi maka didapatkan dengan menggunakan persamaan (2).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

dimana TP (*True Positive*) adalah jumlah kasus yang benar-benar positif dan didiagnosa positif oleh sistem, menunjukkan kesesuaian antara diagnosa sistem dan pakar. TN (*True Negative*) adalah kasus di mana tidak ada penyakit yang terdeteksi oleh sistem, sesuai dengan analisis pakar. FP (*False Positive*) terjadi ketika sistem mendeteksi penyakit yang sebenarnya tidak ada, menghasilkan diagnosa positif yang salah. Sebaliknya, FN (*False Negative*) adalah kasus di mana sistem gagal mendeteksi penyakit yang ada, menghasilkan diagnosa negatif yang salah.

III. Hasil dan Pembahasan

Tujuan penelitian ini adalah merancang suatu sistem pakar dengan metode *K-Nearest Neighbors* (KNN) untuk mendiagnosa gangguan somatisasi. Langkah awal dalam proses ini adalah mengumpulkan data yang relevan untuk membentuk basis pengetahuan. Basis pengetahuan merupakan komponen utama dalam sistem pakar yang berfungsi sebagai penyimpanan informasi, aturan, fakta, dan pengetahuan yang diperlukan untuk memecahkan masalah tertentu. Informasi yang terkandung dalam basis pengetahuan diambil dari para ahli di bidang terkait dalam hal ini adalah seorang psikiater. Data yang dikumpulkan sebagai basis pengetahuan yang diperoleh dari psikiater dilakukan dengan wawancara dan observasi. Berdasarkan hasil dari mengumpulkan data diperoleh data-data gejala, jenis gangguan dan bobot masing-masing gejala yang diberikan oleh pakar. Bobot merupakan tingkat keyakinan yang diberikan oleh seorang pakar dengan rentang nilai 0 sampai dengan 1. Data gejala, jenis gangguan, dan bobotnya disajikan dalam Tabel 1.

Tabel 1. Gejala, Jenis Gangguan, dan Bobot

Kode Gejala	Gejala	Jenis Gangguan				Nilai Bobot
		Gangguan Gejala Somatik (P1)	Gangguan Kecemasan Penyakit (P2)	Gangguan Konversi (P3)	Gangguan Nyeri (P4)	
G1	Nyeri yang tidak memiliki penjelasan medis yang jelas.	✓		✓	✓	0,6
G2	Rasa lelah yang tidak hilang meskipun telah beristirahat cukup.	✓				0,8
G3	Merasakan mual, diare, atau sembelit tanpa penyebab fisik yang jelas.	✓				0,6
G4	Merasa kesulitan bernafas tanpa adanya gangguan pernapasan yang terdeteksi.	✓				0,6
G5	Merasa pusing atau kehilangan keseimbangan tanpa sebab yang pasti.	✓	✓			0,8
G6	Kesulitan tidur atau kualitas tidur yang buruk.	✓		✓		0,5
G7	Kelelahan yang terjadi sebagai akibat dari mengatasi nyeri yang terus-menerus.	✓	✓		✓	0,7
G8	Kekhawatiran yang intens dan terus-menerus bahwa seseorang menderita penyakit serius, meskipun pemeriksaan medis menunjukkan hasil negatif.	✓	✓	✓		0,6
G9	Memperhatikan secara berlebihan tanda-tanda fisik yang dianggap sebagai tanda penyakit.	✓	✓	✓		0,8
G10	Menghindari situasi yang dianggap bisa menyebabkan penyakit atau, sebaliknya, sering mencari perawatan medis dan menjalani pemeriksaan.		✓			0,7
G11	Keyakinan bahwa gejala ringan atau normal adalah tanda dari penyakit serius.		✓	✓		0,6
G12	Kekhawatiran tentang penyakit mengganggu pekerjaan, hubungan, atau aktivitas sehari-hari.		✓			0,8
G13	Seperti kelumpuhan, kesulitan berjalan, atau kebutaan, tanpa ada dasar neurologis yang jelas.			✓		0,7
G14	Kejang atau gerakan tiba-tiba yang menyerupai serangan epilepsi, meskipun tanpa adanya kelainan neurologis yang dapat diidentifikasi.			✓		0,6
G15	Kehilangan kemampuan untuk berbicara, meskipun tidak ada cedera otak yang dapat dideteksi.		✓	✓		0,7
G16	Mati rasa atau hilangnya sensitivitas pada bagian tubuh tertentu tanpa alasan medis yang jelas.			✓		0,8
G17	Kesulitan dalam menjaga keseimbangan atau melakukan gerakan yang biasanya dapat dilakukan.			✓		0,6
G18	Nyeri yang berlangsung lama dan tidak proporsional dengan temuan medis.				✓	0,7
G19	Nyeri yang terasa di beberapa bagian tubuh tanpa pola yang jelas atau penyebab yang dapat diidentifikasi.	✓			✓	0,6
G20	Rasa sakit yang menghalangi fungsi normal seseorang dalam menjalankan rutinitas harian, termasuk mobilitas dan pekerjaan.				✓	0,9
G21	Nyeri yang diperburuk oleh faktor emosional seperti stres, kecemasan, atau depresi.		✓		✓	0,8
G22	Penggunaan obat penghilang rasa sakit yang berlebihan atau kecanduan sebagai upaya untuk mengatasi nyeri.				✓	0,7

Untuk studi kasus dalam penyelesaian diagnosa gangguan somatisasi melalui pendekatan KNN, sebagai contoh digunakan beberapa gejala yang dialami oleh pengguna dan bobot atau keyakinan dari pengguna, yaitu:

Nyeri yang tidak memiliki penjelasan medis yang jelas (G1): 0,5

Rasa lelah yang tidak hilang meskipun telah beristirahat cukup (G2): 0,8

Merasa pusing atau kehilangan keseimbangan tanpa sebab yang pasti (G5): 0,7

Kesulitan tidur atau kualitas tidur yang buruk (G6): 0,6

Kelelahan yang terjadi sebagai akibat dari mengatasi nyeri yang terus-menerus (G7): 0,5

Pada sistem pakar yang menggunakan metode *K-Nearest Neighbors* (KNN) untuk mendiagnosa gangguan somatisasi, proses perhitungan dilakukan dengan membandingkan data gejala baru (data uji) dengan data gejala yang telah ada dalam basis data (data pelatihan). KNN menentukan klasifikasi data uji berdasarkan kedekatan atau kemiripan dengan beberapa data pelatihan terdekat. Untuk setiap data uji, jarak antara data pengujian dan setiap data latih dicari menggunakan *Euclidean Distance* melalui persamaan (1). Berdasarkan Tabel 1, maka proses perhitungannya adalah sebagai berikut:

$$d(P_1) = \sqrt{(0,5 - 0,6)^2 + (0,8 - 0,8)^2 + (0,7 - 0,8)^2 + (0,6 - 0,5)^2 + (0,5 - 0,7)^2} = 0,265$$

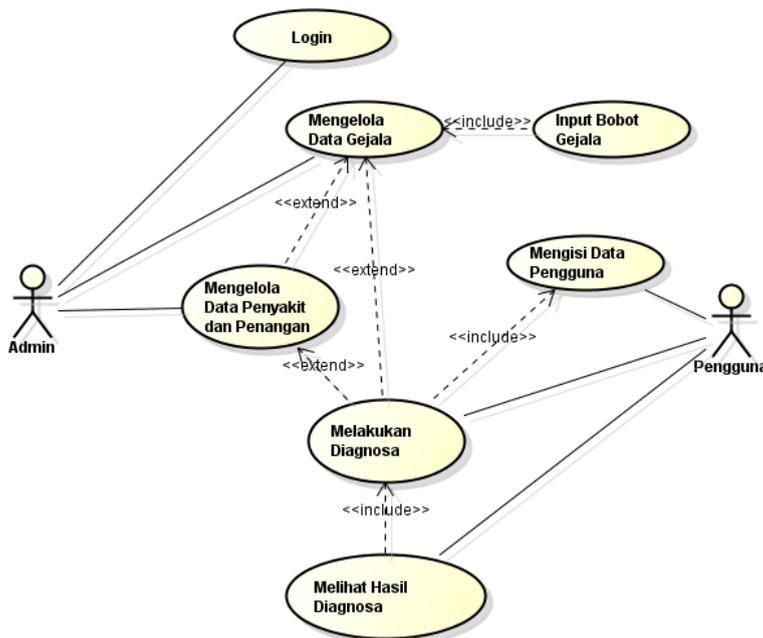
$$d(P_2) = \sqrt{(0,5 - 0)^2 + (0,8 - 0)^2 + (0,7 - 0,8)^2 + (0,6 - 0)^2 + (0,5 - 0,7)^2} = 1,140$$

$$d(P_3) = \sqrt{(0,5 - 0,6)^2 + (0,8 - 0)^2 + (0,7 - 0)^2 + (0,6 - 0,5)^2 + (0,5 - 0)^2} = 1,183$$

$$d(P_4) = \sqrt{(0,5 - 0,6)^2 + (0,8 - 0)^2 + (0,7 - 0)^2 + (0,6 - 0)^2 + (0,5 - 0,7)^2} = 1,241$$

Dalam kasus ini, diperoleh nilai keterdekatan untuk empat jenis gangguan somatisasi antara lain: Gangguan Gejala Somatik (P1) memperoleh nilai 0,265; Gangguan Kecemasan Penyakit (P2) memperoleh nilai 1,140; Gangguan Konversi (P3) memperoleh nilai 1,183; Gangguan Nyeri (P4) memperoleh nilai 1,241. Jarak terdekat dari empat diagnosa tersebut yaitu Gangguan Gejala Somatik (P1) dengan nilai keterdekatan 0,265. Sehingga berdasarkan data uji ini diagnosa yang paling mungkin untuk data uji ini adalah Gangguan Gejala Somatik (P1).

Setelah penerapan pendekatan KNN pada sistem pakar selesai dilakukan, langkah berikutnya adalah merancang sistem sebelum diimplementasikan. Pada tahap ini, perancangan sistem pakar dilakukan dengan memanfaatkan *use case diagram* untuk menggambarkan skenario interaksi antara pengguna dan sistem. *Use case diagram* ini berperan dalam memvisualisasikan kebutuhan sistem serta fungsi-fungsi utama yang harus diimplementasikan. Hasil rancangan dalam bentuk *use case diagram* ditampilkan pada Gambar 2.



Gambar 2. Diagram *Use Case* Sistem Pakar Diagnosa Gangguan Somatisasi

Pada Gambar 2, dapat dilihat bahwa terdapat dua aktor yang terlibat dalam sistem antara lain: Admin dan Pengguna. Admin memperoleh fungsi-fungsi seperti *login* kedalam sistem, mengelola data gejala, dan mengelola data penyakit serta penanganannya. Sedangkan untuk pengguna dapat melakukan mengisi data pengguna, melakukan diagnosa, dan melihat hasil diagnosa.

Setelah perancangan selesai, tahap berikutnya adalah pengkodean atau implementasi sistem pakar. Pengembangan sistem dilakukan dalam bentuk aplikasi berbasis web, memanfaatkan PHP sebagai bahasa pemrograman dan MySQL untuk pengelolaan basis datanya. Sistem pakar untuk diagnosa somatisasi yang dikembangkan memiliki dua jenis hak akses, yaitu Admin dan Pengguna. Admin harus melakukan “login” terlebih dahulu untuk mengakses sistem. Setelah berhasil masuk dengan memasukkan “username” dan kata sandi yang benar, Admin dapat memanfaatkan berbagai fitur yang tersedia dalam sistem, seperti pengelolaan gejala, penyakit, dan melihat hasil konsultasi. Pada fitur pengelolaan gejala, Admin memiliki kemampuan untuk memasukkan data gejala serta pembobotannya yang diperoleh dari pakar, melakukan perubahan pada data tersebut, serta menghapus gejala jika sudah tidak relevan. Pada studi kasus ini seluruh gejala diinputkan berdasarkan dari pengumpulan data yang merujuk pada Tabel 1. Tampilan fitur Kelola Gejala dapat dilihat pada Gambar 3.

No	Kode Gejala	Nama Gejala	Bobot	
1	G1	Nyeri yang tidak memiliki penjelasan medis yang jelas.	0.6	✎ ✕
2	G2	Rasa lelah yang tidak hilang meskipun telah beristirahat cukup.	0.8	✎ ✕
3	G3	Merasakan mual, diare, atau sembelit tanpa penyebab fisik yang jelas.	0.6	✎ ✕
4	G4	Merasa kesulitan bernapas tanpa adanya gangguan pemapasan yang terdeteksi.	0.6	✎ ✕
5	G5	Merasa pusing atau kehilangan keseimbangan tanpa sebab yang pasti.	0.8	✎ ✕
6	G6	Kesulitan tidur atau kualitas tidur yang buruk.	0.5	✎ ✕
7	G7	Kelelahan yang terjadi sebagai akibat dari mengatasi nyeri yang terus-menerus.	0.7	✎ ✕
8	G8	Kekhawatiran yang intens dan terus-menerus bahwa seseorang menderita penyakit serius, meskipun pemeriksaan medis menunjukkan hasil negatif.	0.6	✎ ✕

Gambar 3. Fitur Mengelola Data Gejala

Setelah data gejala telah diinput dengan lengkap, Admin dapat melanjutkan ke pengelolaan data penyakit. Pada fitur Kelola Penyakit, Admin memiliki opsi untuk menambahkan data penyakit baru, memperbarui informasi jika diperlukan, atau menghapus data penyakit yang tidak lagi relevan. Untuk menambahkan penyakit baru, Admin cukup menekan tombol “Tambah Data”. Setelah tombol tersebut di klik, sistem akan menampilkan form Tambah Penyakit. Dalam form ini, Admin harus memasukkan informasi seperti “Kode Penyakit”, “Nama Penyakit”, “Keterangan”, dan “Solusi”. Tampilan form untuk penambahan data penyakit yang tervisualisasi pada Gambar 4.

Tambah Penyakit

Kode Penyakit

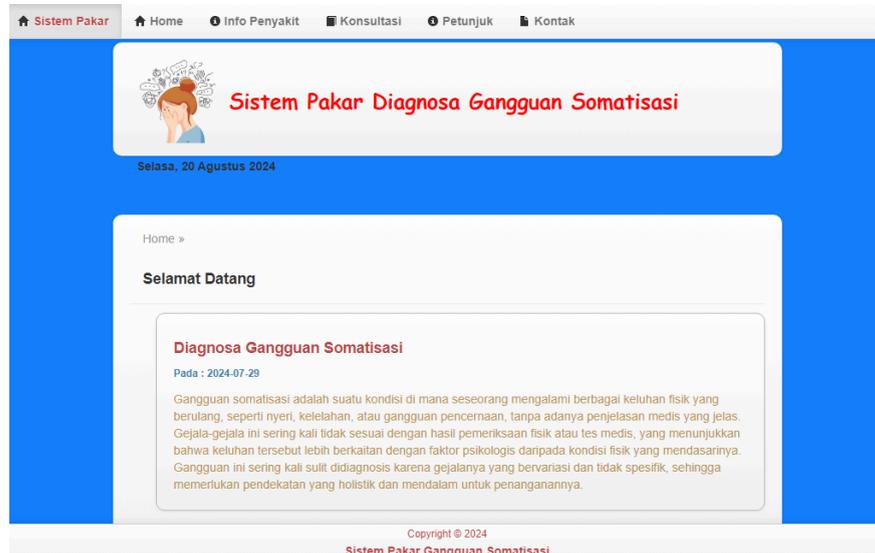
Nama Penyakit

Keterangan

Solusi

Gambar 4. Halaman Untuk Menambah Data Penyakit

Sedangkan untuk hak akses pengguna dapat langsung mengakses menu utama tanpa perlu melakukan login terlebih dahulu. Fitur ini dirancang agar semua orang, termasuk masyarakat umum, dapat dengan mudah melakukan diagnosa. Tampilan menu utama dari sistem yang dikembangkan divisualisasikan pada Gambar 5.



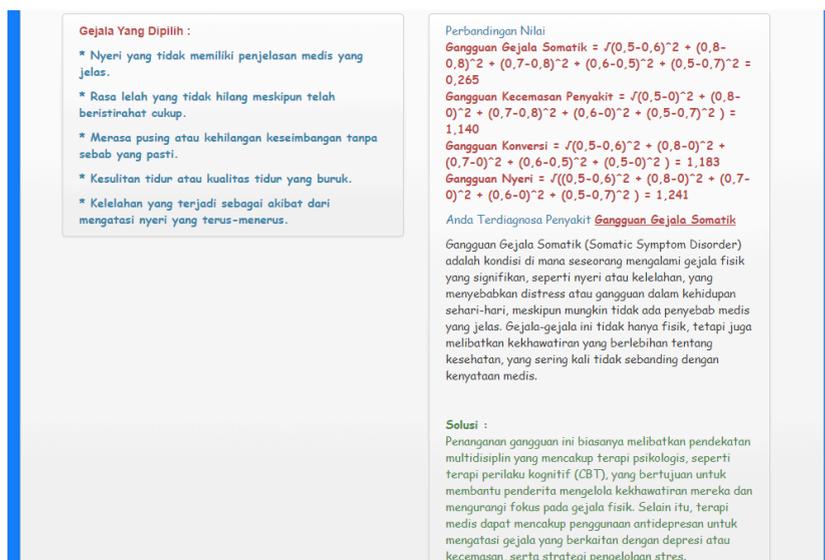
Gambar 5. Menu Utama Sistem Pakar Diagnosa Somatisasi

Gambar 5 menampilkan dashboard atau menu utama bagi pengguna, yang mencakup berbagai fasilitas dalam sistem seperti Info Penyakit, Konsultasi, Petunjuk, dan Kontak. Untuk memulai konsultasi terkait gangguan somatisasi, pengguna dapat memilih opsi Konsultasi. Setelah opsi ini dipilih, pengguna diminta untuk memasukkan “Nama” dan “Usia” untuk data pasien. Setelah data pasien diisi, sistem akan menampilkan daftar gejala-gejala yang terkait dengan gangguan somatisasi. Tampilan form konsultasi pengguna tersaji pada Gambar 6.



Gambar 6. Halaman Konsultasi Pengguna

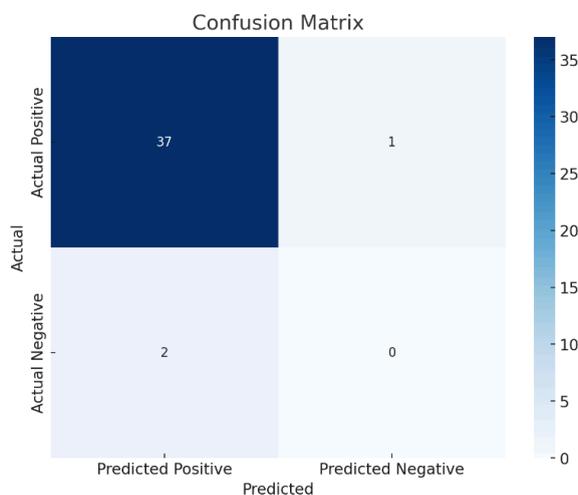
Gambar 6 menampilkan form konsultasi, di mana pengguna diminta untuk mencentang gejala-gejala yang mereka alami. Setelah semua gejala yang relevan dipilih, pengguna dapat mengklik tombol “Periksa Konsultasi”. Sistem kemudian akan memproses data tersebut menggunakan pendekatan KNN dan menampilkan nilai kedekatan menggunakan *Euclidean Distance*. Selain hasil diagnosa, fitur ini juga menyajikan penjelasan mengenai jenis gangguan somatisasi yang didiagnosa dan langkah-langkah penanganan yang disarankan. Tampilan dari fitur hasil diagnosa ini yang terlihat dalam Gambar 7.



Gambar 7. Halaman *Output* Diagnosis Gangguan Somatisasi

Berdasarkan studi kasus yang sama, hasil diagnosa pada Gambar 7 menampilkan nilai keterdekatan yang dihitung oleh KNN menggunakan *Euclidean Distance*. Terlihat bahwa nilai terdekat yang dihasilkan adalah 0,265, yang mengindikasikan Gangguan Gejala Somatik. Dalam metode K-Nearest Neighbors (KNN), nilai KKK merujuk pada jumlah tetangga terdekat yang digunakan untuk menentukan klasifikasi dari data uji. Dalam contoh perhitungan tersebut, meskipun jarak dihitung untuk empat jenis gangguan somatisasi, hanya satu jarak yang digunakan untuk menentukan diagnosis, yaitu jarak terdekat dengan nilai keterdekatan 0,265 (Gangguan Gejala Somatik, P1). Dengan demikian, nilai K pada perhitungan ini adalah 1. Hal ini karena hanya satu tetangga terdekat (dengan jarak terkecil) yang dipilih untuk menentukan klasifikasi atau diagnosis dari data uji. Hasil tersebut didapatkan nilai yang sama dengan perhitungan manual, menunjukkan bahwa implementasi metode KNN telah menghasilkan perhitungan yang akurat dan valid.

Setelah sistem pakar dikembangkan, evaluasi dilakukan untuk memastikan bahwa kinerja sistem telah teruji dan akurat. Pembagian data yaitu 70:30, dimana 70% data sebagai pelatihan dan 30% data sebagai pengujian. Penilaian keakuratan dilakukan melalui perbandingan hasil diagnosa sistem dengan hasil analisis dari seorang ahli, dalam hal ini psikiater. Untuk uji coba, digunakan 40 sampel kasus acak yang mencakup berbagai gejala. Kasus-kasus ini diuji pada sistem pakar dan hasilnya dibandingkan dengan diagnosa pakar. Dari 40 kasus ini, sistem berhasil mendiagnosa dengan tepat sebanyak 37 kasus, sementara terdapat 3 kasus yang tidak tepat, terdiri dari 2 *False Positives* (FP) dan 1 *False Negative* (FN). Berdasarkan hasil ini, *Confusion Matrix* dapat divisualisasikan pada Gambar 8.



Gambar 8. Confusion Matrix Hasil Evaluasi

Gambar 8 memberikan gambaran visual tentang jumlah kasus yang didiagnosa dengan benar sebagai positif (*True Positive*), jumlah kasus yang salah didiagnosa sebagai positif (*False Positive*), serta kasus yang salah didiagnosa sebagai negatif (*False Negative*). Berdasarkan *Confusion Matrix* di atas, akurasi dapat dihitung dengan persamaan (2), sehingga hasilnya adalah sebagai berikut:

$$Accuracy = \frac{37+0}{37+0+2+1} \times 100\% = 92,5\%$$

Dari hasil perhitungan akurasi, diperoleh angka sebesar 92,5%, yang menunjukkan bahwa sistem memiliki tingkat ketidaktepatan diagnosis sebesar 7,5%. Untuk memperjelas hasil ini, visualisasi pengujian akurasi tersebut disajikan melalui grafik pada Gambar 9.



Gambar 8. Hasil Uji Tingkat Keakuratan Sistem

Gambar 9 menunjukkan hasil pengujian akurasi yang telah dilakukan, di mana hasil tersebut dikategorikan berdasarkan kriteria berikut: "Baik" untuk nilai antara 76% hingga 100%, "Cukup" untuk nilai antara 56% hingga 75%, "Kurang Baik" untuk nilai antara 40% hingga 55%, dan "Tidak Baik" untuk nilai di bawah 40% [24]. Dengan tingkat akurasi sebesar 92,5%, sistem ini dikategorikan sebagai "Baik". Hasil ini mencerminkan bahwa sistem mampu memberikan diagnosis yang tepat dalam mayoritas kasus, dengan tingkat kesalahan yang relatif rendah. Ini menandakan bahwa pendekatan yang digunakan oleh sistem, dalam hal ini metode *K-Nearest Neighbors* (KNN) dengan *Euclidean Distance*, sangat efektif untuk menganalisis gejala-gejala yang di input oleh pengguna dan mengklasifikasikannya ke dalam kategori gangguan somatisasi yang sesuai.

Akurasi tinggi yang dicapai oleh KNN dengan *Euclidean Distance* disebabkan oleh beberapa faktor utama. Pertama, *Euclidean Distance* efektif dalam menghitung jarak antara data gejala, memungkinkan identifikasi cepat terhadap pola yang mirip. Kedua, normalisasi data gejala membantu dalam memastikan perhitungan jarak yang lebih akurat, mengurangi bias, dan meningkatkan sensitivitas model terhadap variasi gejala. Ketiga, pemilihan nilai K yang optimal membantu mencegah *overfitting* dan *underfitting*, sehingga prediksi yang dihasilkan lebih stabil dan akurat. Kombinasi faktor-faktor ini berkontribusi pada keandalan model dalam mendiagnosis gangguan somatisasi. Secara keseluruhan, akurasi 92,5% yang dicapai oleh KNN melalui *Euclidean Distance* mengindikasikan bahwa model ini sangat mampu menangani kompleksitas data gejala dan memberikan diagnosis yang dapat diandalkan. Salah satu kelebihan utama KNN adalah kesederhanaannya dan kemampuannya untuk bekerja dengan baik tanpa asumsi distribusi data, yang membuatnya sangat fleksibel. Namun, metode ini juga memiliki kelemahan, seperti sensitivitas terhadap data yang tidak seimbang dan kerentanan terhadap *outlier*, yang dapat mempengaruhi akurasi hasil. Selain itu, KNN memiliki biaya komputasi yang tinggi karena harus menghitung jarak untuk setiap data uji terhadap seluruh data pelatihan, yang dapat menjadi kurang efisien pada dataset yang sangat besar. Sebagai perbandingan, metode lain seperti *Support Vector Machine* (SVM) atau *Decision Tree* memiliki kelebihan masing-masing. SVM, misalnya, dapat lebih efektif dalam kasus dengan margin yang jelas antara kelas, sementara *Decision Tree* dapat memberikan interpretasi yang lebih mudah dipahami. Namun, SVM memerlukan tuning parameter yang lebih kompleks, dan *Decision Tree* cenderung rentan terhadap *overfitting* jika tidak dipangkas dengan baik. Secara keseluruhan, KNN dengan *Euclidean Distance* menunjukkan akurasi yang sangat baik dalam studi kasus sistem pakar diagnosa gangguan somatisasi.

IV. Kesimpulan

Penelitian ini berhasil mengembangkan sistem pakar untuk mendiagnosa gangguan somatisasi dengan menerapkan algoritma *K-Nearest Neighbors* (KNN). Sistem pakar yang dikembangkan menunjukkan tingkat akurasi yang tinggi, yaitu 92,5%, yang menunjukkan bahwa model ini mampu memberikan diagnosis yang akurat dan dapat diandalkan dalam sebagian besar kasus. Temuan utama dari penelitian ini adalah efektivitas KNN dalam mengklasifikasikan gejala-gejala somatisasi dengan menggunakan *Euclidean Distance* sebagai ukuran jarak, yang terbukti sangat efektif dalam menentukan kedekatan atau kemiripan data uji dengan data pelatihan. Penggunaan *Euclidean Distance* yang tepat, normalisasi data gejala, serta pemilihan nilai K yang

optimal menjadi faktor kunci dalam keberhasilan sistem ini. Temuan ini secara langsung menjawab tujuan penelitian yang ditetapkan, yaitu mengembangkan sistem pakar yang mampu mendukung tenaga medis dalam melakukan diagnosis gangguan somatisasi secara efisien dan akurat. Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi penggunaan metode KNN dengan pendekatan jarak lainnya, seperti *Manhattan Distance* atau *Minkowski Distance*, untuk membandingkan kinerjanya dengan *Euclidean Distance* dalam mendiagnosa gangguan somatisasi. Selain itu, peneliti dapat memperluas cakupan dengan mengeksplorasi algoritma lain seperti *Naive Bayes*, *Dempster-Shafer Theory*, atau *Certainty Factor* untuk menentukan apakah ada peningkatan dalam akurasi dan keandalan diagnosis.

Daftar Pustaka

- [1] T. HazellLast and H. Willacy, "Somatisation and somatoform disorders," *patient.info*, 2022. <https://patient.info/mental-health/somatisation-and-somatoform-disorders>
- [2] A. R. Biromo, "Gangguan Somatisasi - Penyebab, Gejala, dan Penanganannya," *Siloam Hospitals*, 2023. <https://www.siloamhospitals.com/informasi-siloam/artikel/apa-itu-gangguan-somatisasi>
- [3] Widia Aina Rohmah and L. T. Merijanti, "Hubungan Stres Akademik Dengan Kecenderungan Gejala Somatisasi Pada Siswa Sma Di Era Pandemi Covid-19," *J. Penelit. Dan Karya Ilm. Lemb. Penelit. Univ. Trisakti*, vol. 9, no. 1, pp. 38–48, 2024, doi: 10.25105/pdk.v9i1.16487.
- [4] Rokom, "Kemenkes Beberkan Masalah Permasalahan Kesehatan Jiwa di Indonesia," *sehatnegeriku.kemkes.go.id*, 2021. <https://sehatnegeriku.kemkes.go.id/baca/rilis-media/20211007/1338675/kemenkes-beberkan-masalah-per-masalahan-kesehatan-jiwa-di-indonesia/>
- [5] Y. Fernando, R. Napianto, and R. I. Borman, "Implementasi Algoritma Dempster-Shafer Theory Pada Sistem Pakar Diagnosa Penyakit Psikologis Gangguan Kontrol Impuls," *Insearch Inf. Syst. Res. J.*, vol. 2, no. 2, pp. 46–54, 2022.
- [6] R. Napianto, Y. Rahmanto, R. I. Borman, O. Lestari, and N. Nugroho, "Dhempster-Shafer Implementation in Overcoming Uncertainty in the Inference Engine for Diagnosing Oral Cavity Cancer," *CSRID (Computer Sci. Res. Its Dev. Journal)*, vol. 13, no. 1, pp. 45–53, 2021, doi: 10.22303/csrid.13.1.2021.46-54.
- [7] F. H. Gurning, M. Ramadhan, and R. Mahyuni, "Sistem Pakar Deteksi Perilaku Penyimpangan Seksual Menggunakan Algoritma K-Nearest Neighbor," *J. Sist. Inf. TGD*, vol. 3, no. 2, pp. 231–240, 2024.
- [8] H. Sajili and Y. M. Djaksana, "Aplikasi Diagnosa Penyakit Saraf Pada Otak Menggunakan Metode Modified K-Nearest Neighbor Berbasis Android," *OKTAL J. Ilmu Komput. dan Sci.*, vol. 3, no. 2, pp. 438–449, 2024.
- [9] O. Br Sinuhaji, D. Nofriansyah, and I. Mariami, "Sistem Pakar Mendiagnosa Penyakit Demam Tifoid (Typhoid Fever) Menggunakan Metode K-Nearest Neighbor," *J. Sist. Inf. TGD*, vol. 3, no. 2, pp. 283–292, 2024.
- [10] T. A. Munandar and A. Q. Munir, "Implementasi K-Nearest Neighbor Untuk Prototype Sistem Pakar Identifikasi Dini Penyakit Jantung," *J. Teknol. Inf.*, vol. XVII, no. 2, pp. 44–50, 2022, doi: 10.35842/jtir.v17i2.457.
- [11] D. P. Aditama, "Sistem Pakar Untuk Mendiagnosis Kesehatan Jiwa Manusia Menggunakan Metode K-Nearest Neighbor," *Teknologipintar.org*, vol. 2, no. 8, pp. 1–8, 2022.
- [12] E. Qiudandra, R. Akram, and N. Novianda, "Sistem Pakar Diagnosa Penyakit Osteoarthritis Dengan Menggunakan Metode K-Nearest Neighbor," *Methodika J. Ilm. Tek. Inform.*, vol. 2, no. 2, pp. 37–48, 2022.
- [13] R. I. Borman, R. Napianto, N. Nugroho, D. Pasha, Y. Rahmanto, and Y. E. P. Yudoutomo, "Implementation of PCA and KNN Algorithms in the Classification of Indonesian Medicinal Plants," in *International Conference on Computer Science, Information Technology and Electrical Engineering (ICOMITEE)*, 2021, pp. 46–50.
- [14] R. I. Borman and M. Wati, "Penerapan Data Maining Dalam Klasifikasi Data Anggota Kopdit Sejahtera Bandarlampung Dengan Algoritma Naïve Bayes," *J. Ilm. Fak. Ilmu Komput.*, vol. 9, no. 1, pp. 25–34, 2020.
- [15] H. Sulistiani, I. Darwanto, and I. Ahmad, "Penerapan Metode Case Based Reasoning dan K-Nearest Neighbor untuk Diagnosa Penyakit dan Hama pada Tanaman Karet," *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 1, p. 23, 2020, doi: 10.26418/jp.v6i1.37256.
- [16] Y. Hamzah and M. H. Botutihe, "Metode K-Nearest Neighbor untuk Sistem Pakar Diagnosa Penyakit Thypoid," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 6, no. 1, pp. 109–114, 2023, doi: 10.32672/jnkti.v6i1.5823.

- [17] Z. Panjaitan, E. Elfitriani, W. R. Maya, and C. D. Siahaan, "Aplikasi Sistem Pakar Untuk Mendiagnosa Lebih Dini Penyakit Kolera Pada Anak Menggunakan Metode K-Nearest Neighbor (KNN)," *J. Sci. Soc. Res.*, vol. 5, no. 2, p. 220, 2022, doi: 10.54314/jssr.v5i2.878.
- [18] F. Tolana, I. Muzakkir, and A. Riadi, "Metode K-Nearest Neighbor (KNN) Untuk Sistem Pakar Diagnosa Penyakit Pneumonia Pada Balita," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 6, no. 3, pp. 363–373, 2023.
- [19] A. Yudhana, S. Sunardi, and A. J. S. Hartanta, "Algoritma K-NN Dengan Euclidean Distance Untuk Prediksi Hasil Penggajian Kayu Sengon," *Transmisi*, vol. 22, no. 4, pp. 123–129, 2020, doi: 10.14710/transmisi.22.4.123-129.
- [20] I. Ahmad, Y. Rahmanto, D. Pratama, and R. I. Borman, "Development of augmented reality application for introducing tangible cultural heritages at the lampung museum using the multimedia development life cycle," *Ilk. J. Ilm.*, vol. 13, no. 2, pp. 187–194, 2021.
- [21] R. I. Borman, A. T. Priandika, and A. R. Edison, "Implementasi Metode Pengembangan Sistem Extreme Programming (XP) pada Aplikasi Investasi Peternakan," *JUSTIN (Jurnal Sist. dan Teknol. Informasi)*, vol. 8, no. 3, pp. 272–277, 2020.
- [22] Y. Rahmanto, J. Alfian, and R. I. Borman, "Penerapan Algoritma Sequential Search pada Aplikasi Kamus Bahasa Ilmiah Tumbuhan," *J. Buana Inform.*, vol. 12, no. 1, pp. 21–30, 2021.
- [23] H. Mayatopani, R. I. Borman, W. T. Atmojo, and A. Arisantoso, "Classification of Vehicle Types Using Backpropagation Neural Networks with Metric and Eccentricity Parameters," *J. Ris. Inform.*, vol. 4, no. 1, pp. 65–70, 2021, doi: 10.34288/jri.v4i1.293.
- [24] R. I. Borman, Y. Fernando, and Y. E. P. Yudoutomo, "Identification of Vehicle Types Using Learning Vector Quantization Algorithm with Morphological Features," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 339–345, 2022.