

SISTEM PAKAR DIAGNOSA PENYAKIT MENULAR SEKSUAL MENGUNAKAN K-NN

Agung Mustika Rizki

Program Studi Informatika, Fakultas Ilmu Komputer, UPN "Veteran" Jawa Timur

Email: agung.mustika.if@upnjatim.ac.id

Abstrak. Angka penderita Penyakit Menular Seksual (PMS) di Indonesia mulai meningkat, salah satunya di kota Malang. Kota Malang kini mulai berkembang dari segi demografis, sosial, pendidikan hingga wisata. Perkembangan ini juga berbanding lurus dengan resiko penyebaran PMS. Banyak penderita tidak merasakan gejala apapun bahkan tidak terdeteksi hingga melakukan pemeriksaan medis. Permasalahan PMS ini penting untuk diselesaikan mengingat akan berakibat fatal dan sangat berbahaya jika tidak terdeteksi dan ditangani sedini mungkin. Pada penelitian ini, penulis membangun sebuah sistem pakar untuk mendiagnosa penyakit menular seksual dengan menerapkan metode K-Nearest Neighbor (K-NN). Berdasarkan hasil pengujian yang telah dilakukan, akurasi sistem pakar ini sebesar 90%. Metode K-NN cukup optimal digunakan untuk menyelesaikan permasalahan klasifikasi dan permasalahan yang serupa.

Kata Kunci: K-NN, Penyakit Menular Seksual, Sistem Pakar

Angka penderita Penyakit Menular Seksual (PMS) di Indonesia mulai meningkat. Penyakit ini sangat erat kaitannya dengan beberapa infeksi lain seperti HIV (*Human Immunodeficiency Virus*), HPV (*Human Papillomavirus*), HSV (*Herpes Simplex Virus*), *Hepatitis B*, *Sifilis*, *Gonore*, *Klamidia* dan *Trikomoniasis* [1][2]. Salah satu contoh kasusnya ada di kota Malang, pada tahun 2014 sebanyak 466 orang menderita HIV dan 14 orang menderita Sifilis. Menurut laporan kesehatan kota Malang, rata-rata penderita berusia 25 hingga 49 tahun. Ada pula yang asimtomatik, dimana penderita tidak merasakan gejala apapun bahkan tidak terdeteksi hingga melakukan pemeriksaan medis [3].

Kota Malang kini mulai berkembang dari segi demografis, sosial, pendidikan hingga wisata. Perkembangan ini juga berbanding lurus dengan resiko penyebaran PMS. Disamping itu, perawatan penyakit ini membutuhkan biaya yang tidak sedikit. Oleh karena itu, penting bagi kita untuk memeriksa diri sendiri lebih awal agar dapat mengurangi beban pengeluaran perawatan.

Seiring dengan perkembangan di era modern ini, untuk mendeteksi PMS ini dapat memanfaatkan teknologi informasi. Salah satunya yakni membangun sistem pakar yang dapat membantu mendeteksi gejala lebih awal.

Terdapat beberapa penelitian sebelumnya terkait sistem pakar pendeteksi suatu penyakit. Pada tahun 2016, Zainuddin dkk [4] membangun sistem pakar dengan menerapkan *Case Based Reasoning* (CBR) untuk mendiagnosis Penyakit Stroke

menggunakan metode K-Nearest Neighbor (K-NN). Berdasarkan hasil pengujian didapatkan hasil bahwa sistem berbasis CBR ini bersifat fleksibel dalam menambah, menghapus dan memperbaiki data apabila ada data yang kurang sesuai di masa yang akan datang. Penerapan metode K-NN dalam sistem tersebut mampu mendiagnosis dengan tepat sesuai hasil diagnosis pakar sebenarnya dengan akurasi sebesar 93,4%.

Metode CBR yang dikombinasikan dengan K-NN juga diterapkan oleh Fatoni dan Noviantha [5] untuk mendiagnosis penyakit difteri. Metode CBR ini memanfaatkan data terdahulu sebagai pengetahuan dasar pada sistem pakar yang nantinya akan selalu diperbarui. Akurasi penerapan K-NN pada penelitian ini sebesar 95,15% dengan 44 pengujian data kasus baru.

Silmina dan Hardiani [6] merancang sebuah sistem pakar untuk mendeteksi penyakit *Pneumonia* pada balita dengan menggunakan K-NN. Prinsip kerja dari K-NN ini adalah mencari jarak terdekat antara data yang akan dievaluasi dengan data tetangga sekitarnya. Hasil pengujian didapatkan dengan menghitung jarak dan nilai *similarity* dari data *testing* kasus baru dengan data *training* kasus lama yang telah disimpan sebelumnya. Akurasi sistem ini dalam mendeteksi penyakit *Pneumonia* adalah sebesar 81%.

Metode K-NN juga diterapkan oleh Ramadhani dan Niswatin [7] dalam membangun sistem diagnosa penyakit diabetes. Berdasarkan hasil pengujian dalam penelitian tersebut didapatkan kesimpulan bahwa metode

K-NN cukup optimal dalam mendiagnosis penyakit diabetes. Akurasi yang dihasilkan adalah sebesar 85,71%.

I. Metodologi

Berdasarkan penelitian yang telah dilakukan sebelumnya, penulis menarik kesimpulan bahwa metode K-NN cocok diterapkan ke dalam sebuah sistem pakar. Permasalahan PMS ini penting untuk diselesaikan mengingat akan berakibat fatal dan sangat berbahaya jika tidak terdeteksi dan ditangani sedini mungkin.

Penyakit Menular Seksual (PMS)

PMS ini merupakan suatu penyakit yang ditularkan dari satu orang ke orang lain melalui kontak seksual. Akibat dari penyakit ini salah satunya adalah gangguan saluran reproduksi. Jika PMS tidak segera ditangani maka akan berakibat lebih fatal dan menyebabkan penderita sakit berkepanjangan. Beberapa gejala yang dialami oleh penderita PMS antara lain timbul rasa sakit saat buang air kecil, adanya benjolan, terjadi pembengkakan hingga keluar cairan atau nanah dari kemaluan. Terdapat beberapa jenis PMS antara lain [8]: *Gonore, Sifilis, Herpes Genitalis, Chlamydia, Trikomoniasis Vaginalis, Kutil Kelamin, Chankroid, Granuloma Inguinale, Limfogranuloma Venereum, Servicitis, Kandidiasis Vulvovaginalis, Vaginosis Bacterial, Moluscum Kontagiosum, Proctitis, Konjungtivitis Neonatorum* dan *Radang Panggul*.

Sistem Pakar

Sistem pakar adalah suatu paket perangkat lunak komputer dimana dapat memecahkan permasalahan tertentu berdasarkan data pengetahuan [9]. Istilah sistem pakar ini merupakan penggunaan pengetahuan seorang pakar yang ditransformasikan menjadi data pada suatu sistem di komputer. Sistem pakar ini tidak menggantikan peran pakar tersebut, tetapi dapat menjadi asisten dalam mengolah data pengetahuan. Disamping itu, sistem pakar ini juga dapat bermanfaat bagi orang yang bukan pakar.

Case Based Reasoning (CBR)

CBR merupakan salah satu penyelesaian permasalahan dengan memanggil kembali pengalaman yang telah dialami atau pernah terjadi sebelumnya. Berdasarkan pengalaman

tersebut, kita dapat menarik kesimpulan berupa informasi dan data yang akan berguna pada masa mendatang. Sehingga solusi-solusi yang telah dilakukan sebelumnya dapat menjadi petunjuk dalam menyelesaikan permasalahan serupa.

Terdapat empat tahapan dalam penerapan CBR ini, antara lain [5]:

1. *Retrive*: mendapatkan permasalahan serupa dengan yang dialami sebelumnya
2. *Reuse*: menggunakan kembali solusi yang telah digunakan sebelumnya
3. *Revise*: merubah solusi sebelumnya berdasarkan data serta informasi kondisi saat ini
4. *Retain*: menggunakan solusi baru untuk permasalahan saat ini

K-Nearest Neighbor (K-NN)

K-NN merupakan suatu metode untuk mengklasifikasi masukan data *testing* yang memiliki jarak terdekat berdasarkan data *training* [4]. K-NN ini termasuk ke dalam metode klasifikasi non parametrik. Secara komputasi lebih sederhana dibandingkan metode lain.

K-NN menghitung kedekatan antara kasus baru dengan kasus lama berdasarkan bobot yang cocok dari sejumlah fitur yang ada. Adapun rumus untuk menghitung bobot similaritas seperti ditunjukkan persamaan berikut [8].

$$\text{similaritas}(T, S) = \frac{\sum_{i=1}^n f(T_i, S_i) * w_i}{w_i} \quad (1)$$

Dimana,

- T : kasus baru
 S : kasus lama
 n : atribut dari setiap kasus
 i : individu dengan atribut 1 sampai n
 f : kemiripan atribut antara T dan S
 w : bobot dari atribut i

$$w_i = \frac{1}{d(x', x_i)^2} \quad (2)$$

Kedekatan biasanya terletak pada nilai antara 0 hingga 1. Nilai 0 memiliki arti bahwa kedua kasus tersebut berbeda, sedangkan nilai 1 memiliki arti kedua kasus tersebut sama mutlak.

Langkah-langkah penyelesaian dengan metode K-NN adalah sebagai berikut:

1. Tentukan parameter K (jumlah tetangga terdekat)
2. Hitung kuadrat *Euclidean Distance* pada setiap objek dengan data contoh
3. Urutkan berdasarkan nilai *Euclidean Distance* terkecil. Rumus perhitungan *Euclidean Distance* sebagai berikut:

$$D(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2} \quad (3)$$

4. Mengumpulkan kategori Y (hasil klasifikasi K-NN)
5. Memprediksi nilai berdasarkan kategori mayoritas

II. Hasil dan Pembahasan

Pada penelitian ini, terdapat sebanyak 139 data yang kemudian dibagi menjadi dua yakni untuk kebutuhan *training* dan *testing*. Mekanisme *training* menggunakan sebanyak 109 data sedangkan untuk *testing* menggunakan sebanyak 30 data. Selanjutnya hasil dari sistem tersebut akan dibandingkan dengan data sebenarnya dari basis pengetahuan pakar untuk mengetahui akurasi dari penerapan metode K-NN dengan rumus perhitungan sebagai berikut:

$$\text{Akurasi} = \frac{\sum \text{data yang sesuai}}{\sum \text{data testing}} \times 100\% \quad (4)$$

Hasil pengujian akurasi hasil sistem untuk 30 data *testing* seperti ditunjukkan pada Tabel 1.

Tabel 1 Hasil Pengujian Sistem Pakar

No	Basis Pengetahuan Pakar	Hasil Sistem Pakar
1	<i>Urethritis/Gonore</i>	<i>Urethritis/Gonore</i>
2	<i>Moluskum</i>	<i>Moluskum</i>
3	<i>Kontagiosum</i>	<i>Kontagiosum</i>
4	<i>Proctitis</i>	<i>Proctitis</i>
5	<i>Konjungtivitis neonatorum</i>	<i>Konjungtivitis Neonatorum</i>
6	<i>Vaginosis Bacterial</i>	<i>Vaginosis Bacterial</i>
7	<i>Limfogramulama Venereum</i>	<i>Limfogramulama Venereum</i>
8	<i>Granuloma Inguinale</i>	<i>Granuloma Inguinale</i>
9	<i>Sifilis</i>	<i>Sifilis</i>
10	<i>Kandidiasis</i>	<i>Kandidiasis</i>
11	<i>Trikomoniasis Vaginalis</i>	<i>Trikomoniasis Vaginalis</i>
12	<i>Trikomoniasis Vaginalis</i>	<i>Trikomoniasis Vaginalis</i>
13	<i>Servicitis</i>	<i>Servicitis</i>
14	<i>Sifilis</i>	<i>Sifilis</i>
15	<i>Urethritis/Gonore</i>	<i>Urethritis/Gonore</i>
16	<i>Konjungtivitis Neonatorum</i>	<i>Konjungtivitis Neonatorum</i>
17	<i>Radang Panggul</i>	<i>Radang Panggul</i>

18	<i>Pembengkakan Skrotum</i>	<i>Klamidia</i>
19	<i>Vaginosis Bacterial</i>	<i>Servicitis</i>
20	<i>Servicitis</i>	<i>Servicitis</i>
21	<i>Klamidia</i>	<i>Klamidia</i>
22	<i>Herpes Genitalis</i>	<i>Herpes Genitalis</i>
23	<i>Granuloma Inguinale</i>	<i>Granuloma Inguinale</i>
24	<i>Kutil Kelamin</i>	<i>Kutil Kelamin</i>
25	<i>Vaginosis Bacterial</i>	<i>Trikomoniasis Vaginalis</i>
26	<i>Chancroid</i>	<i>Chancroid</i>
27	<i>Herpes Genitalis</i>	<i>Herpes Genitalis</i>
28	<i>Proctitis</i>	<i>Proctitis</i>
29	<i>Klamidia</i>	<i>Klamidia</i>
30	<i>Klamidia</i>	<i>Klamidia</i>

Berdasarkan tabel hasil pengujian, ditemukan sebanyak tiga data *testing* yang tidak sesuai dengan basis pengetahuan dari pakar, sehingga akurasi dari sistem pakar diagnosa penyakit menular seksual ini sebagai berikut:

$$\text{Akurasi} = \frac{27}{30} \times 100\% = 90\%$$

Ketidaksesuaian tiga data tersebut disebabkan oleh ada lebih dari satu gejala yang mirip satu penyakit dengan penyakit lainnya.

III. Kesimpulan

Berdasarkan hasil pengujian tersebut, maka didapatkan kesimpulan bahwa penerapan metode K-NN pada sistem pakar cukup optimal untuk menyelesaikan permasalahan klasifikasi seperti ini dengan akurasi sebesar 90%.

Untuk penelitian selanjutnya, penulis perlu meningkatkan akurasi tersebut dengan mengoptimalkan parameter pada metode K-NN itu sendiri. Selain itu, juga diperlukan jumlah data *training* dan data *testing* yang tepat agar hasil keluaran dari sistem pakar lebih akurat.

IV. Daftar Pustaka

- [1] T. Y. Aditama. (2011). *Pedoman Nasional Penanganan Infeksi Menular Seksual*. Jakarta: Kementerian Kesehatan Republik Indonesia,.
- [2] World Health Organization. (2016). Media Centre-Sexually Transmitted Infections(STIs). <http://www.who.int/mediacentre/factsheets/fs110/en/>.
- [3] Dinas Kesehatan Kota Malang. (2015).

Profil Kesehatan Kota Malang Tahun 2014.

- [4] M. Zainuddin, K. Hidjah, and I. W. Tunjung. (2016). Penerapan Case Based Reasoning (CBR) Untuk Mendiagnosis Penyakit Stroke Menggunakan Algoritma K-Nearest Neighbor. *Citeseee*, pp. 21–26.
- [5] C. S. Fatoni and F. D. Noviandha. (2018). Case Based Reasoning Diagnosis Penyakit Difteri dengan Algoritma K-Nearest Neighbor. *Creat. Inf. Technol. J.*, vol. 4, no. 3, p. 220. doi: 10.24076/citec.2017v4i3.112.
- [6] E. P. Silmina and T. Hardiani. (2018). Perancangan Sistem Pakar Penyakit Pneumonia pada Balita Menggunakan Algoritme K-NN (K-Nearest Neighbor). *J. Pseudocode*, vol. 5, no. 2, pp. 56–63. doi: 10.33369/pseudocode.5.2.56-63.
- [7] R. A. Ramadhani and R. K. Niswatin. (2018). Sistem Diagnosa Diabetes Menggunakan Metode K-NN. *J. Sains dan Inform.*, vol. 4, no. 2, pp. 98–104. doi: 10.34128/jsi.v4i2.121.
- [8] G. E. Yuliasuti, A. N. Alfiyatin, A. M. Rizki, A. Hamdianah, H. Taufiq, and W. F. Mahmudy. (2018). Performance analysis of data mining methods for sexually transmitted disease classification. *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, pp. 3933–3939. doi: 10.11591/ijece.v8i5.pp3933-3939.
- [9] N. Aini, R. Ramadiani, and H. R. Hatta. (2017). Sistem Pakar Pendiagnosa Penyakit Tuberkulosis. *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 12, no. 1, p. 56. doi: 10.30872/jim.v12i1.224.