

PERBANDINGAN METODE CBR, NN-5 DAN NAÏVE BAYES UNTUK KLASIFIKASI INFEKSI NOSOKOMIAL

Taufiq Rizaldi¹, Aji Seto Arifianto²

²Jurusan Teknologi Informasi, Politeknik Negeri Jember

E-mail: mailfor.taufiq@gmail.com

E-mail: ajiset@gmail.com

ABSTRACT

Classification method is a method that is widely used for the completion of the diagnosis of disease, particularly in humans. One type of disease that is dangerous is potentially emerging nosocomial infection in patients postoperatively. Approach classification method based upon the similarity of character data, there is also a reference to the emergence of statistical data. Some methods of classification needs to be tested in order to obtain the appropriate reference method for this case. The purpose of this study was to compare the performance of three methods of classification that Case Based Reasoning method (CBR), Nearest Neighbor-5 (NN-5) and Naïve Bayes for diagnosis of nosocomial infection. Test results showed that the method of CBR and NN-5 has a very good degree of accuracy than Naïve Bayes, Naïve Bayes but has a faster processing time.

Kata Kunci: Klasifikasi, CBR, Infeksi Nosokomial, Naïve Bayes, NN-5

PENDAHULUAN

Infeksi *nosokomial* merupakan infeksi yang disebabkan oleh mikroba *pathogen*, peristiwa ini berpotensi terjadi pada pasien pasca operasi atau dalam proses perawatan di rumah sakit kurang lebih 72 jam [1]. Infeksi ini bisa jadi disebabkan dari faktor *intern* atau *ektern* pasien. Salah satu usaha untuk menekan terjadinya infeksi adalah dengan melakukan diagnosis awal dengan menggunakan tanda-tanda fisik untuk melakukan proses diagnosis. Proses diagnosis dilakukan dengan menerapkan metode-metode klasifikasi berdasarkan data latih. Data latih merupakan kumpulan data pasien yang telah melalui diagnosis dokter spesialis/ahli. Pengambilan keputusan pada proses klasifikasi dengan menghitung nilai kemiripan data uji dan data latih, pendekatannya dapat melalui *similarity* ataupun statistik.

Penerapan metode *Case Based Reasoning* (CBR) untuk melakukan diagnosis terjadi atau tidaknya infeksi menunjukkan hasil yang cukup

memuaskan ditunjukkan dengan tingkat akurasi yang mencapai 96%[2]. Akan tetapi belum didapatkan perbandingan dengan metode lain untuk mengetahui perbandingan tingkat akurasi dan waktu pemrosesan. *Case Based Reasoning* (CBR) adalah sebuah metode pengambilan keputusan untuk menyelesaikan masalah baru dengan berdasarkan solusi kasus-kasus terdahulu yang serupa dan diterapkan dalam membangun sistem komputer cerdas [3].

Naïve Bayes adalah metode yang menggabungkan teori *Bayes* yang memanfaatkan probabilitas dan statistik digabungkan dengan teori *Naïve* yang melakukan asumsi bahwa kondisi antar atribut saking bebas untuk melakukan klasifikasi[4]. *Nearest Neighbor* adalah sebuah metode sederhana yang memanfaatkan nilai *similarity* untuk menentukan dimana suatu kelompok apakah lebih dekat dengan suatu kelompok tertentu[5].

Pada penelitian ini akan dilakukan perbandingan antara tiga metode yaitu metode *Case Based Reasoning*(CBR),

Nearest Neighbor-5 (NN-5) dan *Naïve Bayes* untuk mengetahui tingkat akurasi dan waktu pemrosesan dari setiap metode.

TINJAUAN PUSTAKA

2.1. Infeksi Nosokomial

Infeksi adalah adanya suatu organisme pada jaringan atau cairan tubuh yang disertai suatu gejala klinis baik lokal maupun sistemik [6]. Secara umum infeksi yang muncul gejalanya setelah 72 jam seorang pasien dirawat atau berada di lingkungan rumah sakit disebut *infeksi nosokomial* [7]. Konsekuensi negatif bagi penderita infeksi *nosokomial*, antara lain lama hari perawatan bertambah, biaya meningkat dan resiko kematian meningkat. Beberapa ciri dari infeksi nosokomial adalah:

a) Panas

Suhu yang meningkat disekitar lokasi terjadinya infeksi, hal ini disebabkan adanya peningkatan aliran darah ke area tersebut

b) Rasa Sakit (*Dolor*)

Pada bagian yang mengalami infeksi, kadar PH dan konsentrasi ion berubah. Hal ini akan menyebabkan keluarnya zat histamine yang menimbulkan rasa nyeri. Pembengkakan daerah infeksi juga sering menimbulkan rasa sakit.

c) Kemerahan (*Rubor*)

Arterioli yang mensuplai darah ke area tersebut semakin melebar, dengan demikian lebih banyak darah yang mengalir ke dalam mikro sirkulasi local sehingga menimbulkan kemerahan pada kulit.

d) Pembengkakan

Cairan dan sel-sel dari sirkulasi darah dikirim ke jaringan interstisial sehingga menyebabkan bengkak.

e) *Functiolaesa*

Malfungsi organ yang disebabkan perubahan secara superficial bagian yang bengkak dan sakit disertai sirkulasi dan

lingkungan kimiawi lokal yang abnormal[6].

2.2. Case Based Reasoning (CBR)

Case Based Reasoning (CBR) adalah sebuah metode pengambilan keputusan dengan menggunakan solusi dari kasus-kasus lama yang pernah terjadi untuk menyelesaikan sebuah kasus baru[5]. Aplikasi dari CBR dapat dikelompokkan menjadi dua tipe yaitu [8]:

a) Tipe Klasifikasi (*classification tasks*)

Bertujuan membagi fitur-fitur tertentu pada kelas-kelas tertentu. Sebuah kasus dicocokkan dengan database untuk menentukan kelas terdekat yang paling sesuai. Tugas klasifikasi antara lain diagnosis, prediksi, penaksiran, proses kontrol, perencanaan.

b) Tipe Sintesis (*synthesis tasks*)

Tipe ini digunakan apabila perlu penempatan atau pemilihan fitur yang sesuai pada tempat yang tepat. Tugas-tugas yang memerlukan sintesa antara lain dalam bidang: perdagangan/industri tetapi sangat sulit untuk dilaksanakan. Hal ini karena pada umumnya lebih mudah mencocokkan sebuah hasil karya manusia terhadap satu kumpulan prototipe hasil karya manusia daripada harus membangun sebuah hasil karya dari sebuah perincian.

2.3. Naïve Bayes

Algoritma *Naïve Bayes* yang dikemukakan oleh ilmuwan Inggris Thomas Bayes adalah sebuah metode yang menggabungkan metode *bayes* yang menggunakan pengalaman di masa lalu untuk memprediksi masa datang dengan metode *naïve* yang mengasumsikan bahwa ada atau tidaknya ciri pada suatu kelas tidak ada hubungannya dengan kelas lain[4]. Metode *Naïve Bayes* menggunakan probabilitas dan statistik untuk memperoleh hasil. Persamaan dari teorema bayes adalah:

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

Dimana:

X : Data dengan class yang belum diketahui

- H : Hipotesis data X merupakan suatu class spesifik
- P(H|X) : Probabilitas hipotesis H berdasar X kondisi X
- P(H) : Probabilitas hipotesis H
- P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X) : Probabilitas X

2.4. *Nearest Neighbor*

Nearest Neighbor adalah sebuah metode yang menghitung kedekatan nilai sebuah kasus dengan kasus yang lain berdasarkan pencocokan pada fitur. Pada *Nearest Neighbor-5* pencocokan sebuah kasus dilakukan dengan lima kasus yang memiliki nilai persamaan atau *similarity* tertinggi kemudian diambil rata-ratanya. Persamaan yang digunakan adalah:

$$\sum_{i=1}^n \frac{f(T_i, S_i)}{W_i} \dots\dots\dots(2)$$

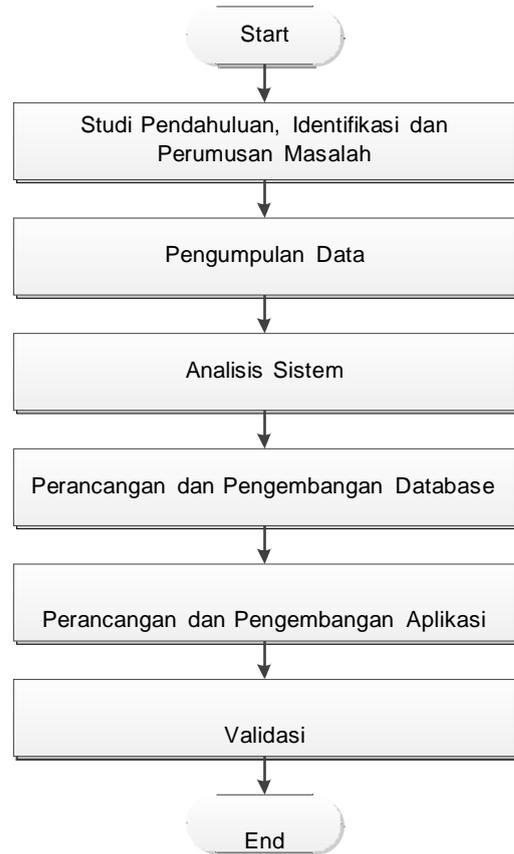
- Dimana:
- T: Kasus Target
 - S: Kasus Sumber
 - i: Atribut individu dari 1 sampai n
 - f: fungsi persamaan untuk atribut I dalam kasus T dan S
 - W: Bobot penting dari atribut i.

PERANCANGAN SISTEM

Pada tahap ini dilakukan pengumpulan data dan informasi terkait infeksi nosokomial serta teori tentang metode CBR, NN-5 dan *Naive Bayes*. Data dan informasi didapatkan dari Rumah Sakit Umum Daerah Ngudi Waluyo Wlingi Blitar serta hasil dari wawancara beberapa pakar dari Fakultas Kedokteran Universitas Brawijaya Malang. Data dan informasi yang telah terkumpul telah divalidasi oleh pakar. Variabel yang nantinya akan digunakan antara lain: luka bersih, nyeri, bengkak, kemerahan, pernanahan, ada infeksi bernilai antara 0 dan 1, suhu tubuh (⁰Celcius), Leukosit (ribu/ μ L) dan Laju endap Darah (mm/jam).

Berdasarkan informasi dan data yang telah didapat maka dilakukan perbandingan antara ketiga metode yaitu

CBR, NN-5 dan *Naive Bayes*. Metode yang digunakan untuk melakukan uji coba perbandingan adalah dengan menguji ulang data sample yang ada pada database apakah hasil keluarannya sesuai dengan data aslinya[9]. Tahapan penelitian ditunjukkan pada alur proses seperti pada Gambar 1.



Gambar 1. Tahapan Penelitian.

Sesuai dengan gambar 1, penelitian diawali dengan studi pendahuluan dengan melakukan survey ke lokasi obyek penelitian membaca literatur, kemudian dilakukan identifikasi dan perumusan masalah. Tahap selanjutnya dilakukan pengumpulan data dengan cara observasi atau pengamatan. Berdasarkan hasil pengumpulan data dilakukan analisis sistem serta perancangan *database*. Tahapan terpenting setelah perancangan adalah implementasi atau pengembangan aplikasi. Terakhir dilakukan validasi terhadap metode dan hasil keluaran aplikasi, untuk mengetahui performa masing-masing metode klasifikasi yang dibandingkan.

Pada tahapan uji coba sistem dilakukan pengujian dengan menggunakan tiga set data yaitu dengan set data berjumlah 50, kemudian 75 dan 100 yang nantinya akan divalidasi dengan Receiver Operator Characteristic (ROC) Curve atau kurva ROC[10]. Fokus uji coba adalah pada tingkat akurasi dan waktu pemrosesan yang dibutuhkan setiap metode.

HASIL PENELITIAN

Penerapan metode Nearest Neighbor-5 (NN-5) alur prosesnya sama dengan metode CBR [2] hanya saja pada proses pemilihan kasus terdekat jika CBR memilih kasus dengan nilai similarity tertinggi sedangkan NN-5 memilih lima kasus dengan nilai similarity tertinggi lalu menghitung jumlah kasus yang bernilai “Benar” atau “Salah” kemudian memilih kasus dengan jumlah terbanyak, seperti dicontohkan pada Tabel 1. Meskipun pada Tabel 1 kasus dengan nilai similarity tertinggi bernilai “Benar” akan tetapi jumlah kasus bernilai “Salah” lebih banyak dibandingkan yang bernilai “Benar” sehingga nilai akhirnya adalah “Salah”.

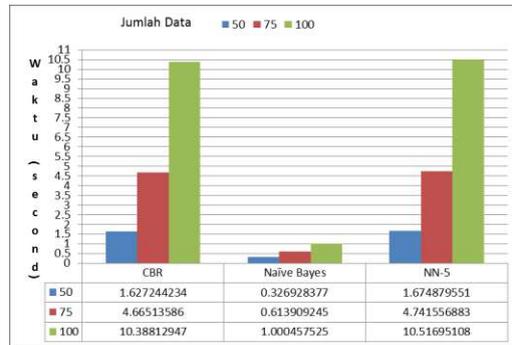
TABEL 1 KASUS TERDEKAT PADA NN-5

Kasus	Nilai Similarity	Nilai
1	0.987654321	Benar
2	0.965432187	Benar
3	0.897652341	Salah
4	0.876543210	Salah
5	0.867543219	Salah

TABEL 2. HASIL PERBANDINGAN

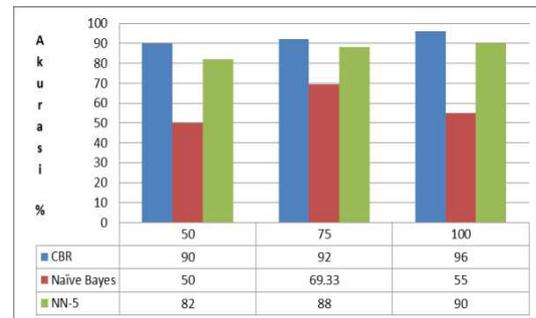
Metode	Data	Waktu(s)	Akurasi
CBR	50	1.6272442	90 %
	75	4.6651358	93.3 %
	100	10.388129	96 %
Naïve Bayes	50	0.3269283	50 %
	75	0.6139092	69.3 %
	100	1.0004575	55 %
NN-5	50	1.6748795	82 %
	75	4.7415568	88 %
	100	10.516951	90 %

Berdasarkan uji coba yang dilakukan maka kemudian dilakukan evaluasi dari hasil keluaran sistem untuk perbandingan tiga metode yaitu CBR, NN-5 dan Naïve Bayes. Hasil uji coba dari ketiga metode tersebut seperti yang ditampilkan pada Tabel 2.



Gambar 2. Perbandingan berdasarkan waktu.

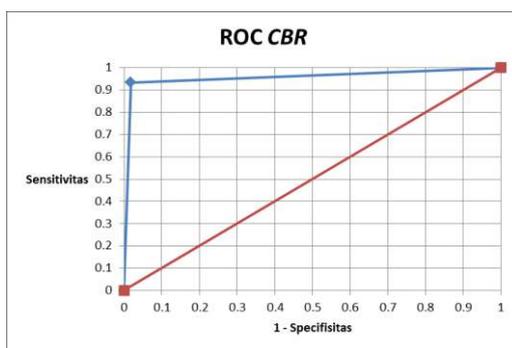
Berdasarkan Gambar 2 diketahui bahwa untuk aspek waktu pemrosesan Naïve Bayes memiliki waktu pemrosesan tercepat dibandingkan dengan metode CBR dan NN-5. Hal tersebut dikarenakan naïve bayes tidak melakukan perhitungan pada setiap data tetapi menggunakan peluang munculnya jumlah kasus pada setiap variabel, sedangkan untuk CBR dan NN-5 membutuhkan proses berulang-ulang untuk menghitung nilai similarity sehingga waktu yang dibutuhkan lebih lama. Sedangkan detail dari perbandingan berdasarkan akurasi seperti pada Gambar 3.



Gambar 3. Perbandingan berdasarkan akurasi.

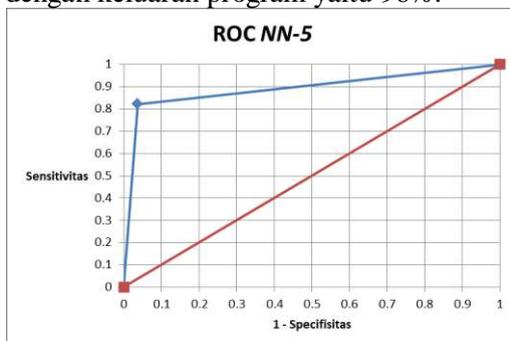
Untuk metode CBR dan NN-5 memiliki tingkat akurasi yang hampir sama, walaupun tingkat akurasi CBR lebih tinggi dibanding NN-5 Sedangkan

metode *naïve bayes* tingkat akurasinya masih di bawah NN-5 dan CBR, hal tersebut dikarenakan pada *naïve bayes* setiap variabel dianggap penting sehingga mengesampingkan bobot yang terdapat pada setiap variabel. Sedangkan walaupun memiliki alur proses yang hampir sama dengan CBR, NN-5 mengambil lima kasus dengan nilai similarity tertinggi sehingga jika diantara lima kasus tersebut kasus dengan nilai similarity tertingginya bernilai true akan tetapi empat kasus lainnya bernilai false maka keluaran akan bernilai false.



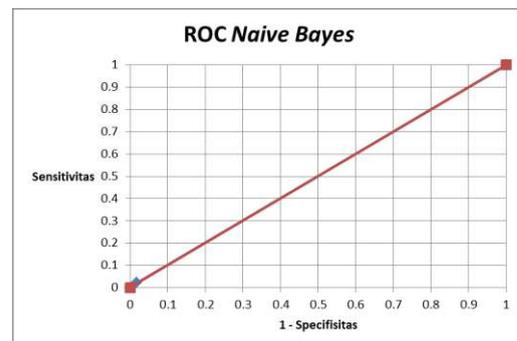
Gambar 4. Kurva ROC CBR.

Dari hasil percobaan tersebut kemudian dilakukan proses validasi menggunakan *Receiver Operator Characteristic (ROC) Curve* atau kurva ROC. Hasil dari kurva ROC untuk setiap metode dengan data set 100 adalah seperti yang ditampilkan pada Gambar 4, Gambar 5 dan Gambar 6. Berdasarkan pada Gambar 4 maka pengujian akurasi dengan metode CBR sudah baik karena bentuk kurva ROC menjauhi garis lintang dan mendekati 1, sedangkan untuk akurasi dihitung dari luas area kurva ROC sesuai dengan keluaran program yaitu 96%.



Gambar 5. Kurva ROC NN-5.

Sedangkan pada Gambar 5 diketahui bahwa akurasi pengujian dengan metode NN-5 juga sudah baik karena bentuk kurva ROC menjauhi garis lintang dan mendekati 1, untuk akurasi dihitung dari luas area kurva ROC sesuai dengan keluaran program yaitu 90%. Sedangkan pada metode *Naïve Bayes* didapat tingkat akurasi yang tidak terlalu baik hal tersebut dibuktikan dengan hasil kurva ROC yang berbentuk sejajar dengan garis lintang seperti pada Gambar 6.



Gambar 6. Kurva ROC *Naïve Bayes*.

KESIMPULAN

Hasil perbandingan antar metode CBR, *Naïve Bayes* dan *Nearest Neighbor-5 (NN-5)* secara waktu yang dibutuhkan untuk pemrosesan hasilnya *Naïve Bayes* memiliki waktu pemrosesan yang lebih cepat. Hal tersebut dikarenakan *Naïve Bayes* hanya menghitung berdasarkan peluang terjadinya kembali sebuah kasus, sedangkan untuk CBR dan NN-5 membutuhkan proses berulang-ulang untuk menghitung nilai similarity sehingga waktu yang dibutuhkan sedikit lebih lama. Untuk metode CBR dan NN-5 memiliki tingkat akurasi tidak jauh berbeda, walaupun CBR memiliki tingkat akurasi yang lebih baik. Sedangkan metode *Naïve Bayes* tingkat akurasi masih di bawah NN-5 dan CBR.

DAFTAR PUSTAKA

- [1] Ahsan. 2013, *Nursing Care Knowledge Management Based Training Decrease Nosocomial Infection Inciden in Post Sectio Cesarea Patients*. Ners Journal

- Volume 8 No 2 Oktober 2013.
- [2] Rizaldi, Taufiq, M. Aziz M, Erni Y. 2014, *Knowledge Management System* untuk Diagnosis Infeksi Nosokomial Jurnal EECCIS Vol. 8, No. 2, Desember 2014.
- [3] Santoso, P. B. 2012, "*Modul 4-Case Based Reasoning*". Universitas Brawijaya. Malang.
- [4] Basuki, Achmad. 2006, *Metode Bayes*. Surabaya. Politeknik Elektro Negeri Surabaya-Institut Teknologi Sepuluh November.
- [5] Kusriani, Sri Hartati, Retantyo Wardoyo, Agus Harjoko. 2009. *Perbandingan Metode Nearest Neighbor Dan Algoritma C4.5 Untuk Menganalisis Kemungkinan Pengunduran Diri Calon Mahasiswa Di Stmik Amikom Yogyakarta*. Jurnal DASIS Vol. 10 No. 1 Maret 2009.
- [6] Yuddhityarasati. 2007, "*Paradigma Managerialism (Teori Manajemen Publik)*". Penerbit Graha. Jakarta.
- [7] Mandal, B.K., Wilkins, E.G.L., Dunbar, E.M. dan R.T. Mayon-White. 2012, "*Lecture Notes Penyakit Infeksi edisi keenam*". Penerbit Erlanga. Jakarta.
- [8] Santoso, P. B. 2012, "*Modul 4-Case Based Reasoning*". Universitas Brawijaya. Malang.
- [9] O'Leary, D. E. 1993, "*Verification and validation of case-based systems*". *Expert Syst. Applicat.*, vol. 6, pp. 57-66, 1993.
- [10] Sastroasmoro, S., Sofyan Ismail. 1995, "Dasar-Dasar Metode Penelitian Penelitian Klinis". Binarupa Aksara. Jakarta.
- [11] Abu Naser, S., R. Al-Dahdooh, A. Mushtaha, M. El-Naffar. 2010, "*Knowledge Management in ESMDA: Expert System for Medical Diagnostic Assistance*". *ICGST-AIML Journal*, Volume 10, Issue 1, October 2010.