

BAB II

LANDASAN TEORI

2.1 Kecerdasan buatan

Kecerdasan buatan atau yang sering disebut dengan AI (*Artificial Intelligence*) merupakan salah satu bagian ilmu komputer yang membuat agar mesin (komputer) dapat melakukan pekerjaannya seperti dan sebaiknya yang dilakukan oleh manusia. Teknologi kecerdasan buatan dipelajari dalam bidang-bidang seperti robotika, penglihatan komputer (*computer visio*), jaringan saraf tiruan (*artificialneural system*), pengolahan bahasa alami (*natural language processing*), pengenalan suara (*speech recognition*), dan sistem pakar (*expert system*). Kecerdasan buatan memiliki arti yang merujuk pada mesin yang mampu berpikir, menimbang tindakan yang akan diambil dan mampu mengambil keputusan seperti yang dilakukan manusia. (Jusniwati, 2013).

Kecerdasan buatan (*Artificial Intelligence*) adalah ide – ide untuk membuat suatu perangkat lunak komputer yang memiliki kecerdasan sehingga perangkat lunak komputer tersebut dapat melakukan suatu pekerjaan yang dilakukan oleh manusia.

Kecerdasan buatan bertujuan untuk membuat komputer menjadi lebih cerdas, bisa bernalar, dan berguna untuk manusia. Kecerdasan buatan dapat juga digunakan untuk membantu meringankan beban kerja (Wijaya & Prasetiyowati, 2012).

Dari pernyataan diatas menunjukkan bahwa AI adalah bagian dari ilmu komputer yang mempelajari bagaimana membuat mesin (komputer) dapat melakukan pekerjaan seperti dan sebaik yang dilakukan oleh manusia bahkan bisa lebih baik daripada yang dilakukan manusia.

2.2 Sistem Pakar

Sistem pakar sangat membantu untuk pengambilan keputusan, dimana sistem pakar ini dapat mengumpulkan dan menyimpan pengetahuan dari seorang atau beberapa orang pakar dalam suatu basis pengetahuan dan menggunakan sistem penalaran yang menyerupai seorang pakar dalam memecahkan masalah (Wijaya & Prasetyowati, 2012).

Dari pernyataan diatas menunjukkan bahwa Sistem Pakar (*Expert System*) adalah suatu sistem komputer yang dirancang dengan menggunakan pengetahuan, fakta dan teknik penalaran untuk menyelesaikan suatu masalah yang biasanya hanya dapat dilakukan oleh seorang pakar.

Ada beberapa keunggulan sistem pakar, diantaranya adalah sebagai berikut :

1. Menghimpun data dalam jumlah yang sangat besar.
2. Menyimpan data tersebut untuk jangka waktu yang panjang dalam suatu bentuk tertentu.
3. Mengerjakan perhitungan secara cepat dan tepat serta tanpa jemu mencari kembali data yang tersimpan dengan kecepatan tinggi.

Sistem pakar mempunyai beberapa kemampuan dalam menyelesaikan masalah, diantaranya :

1. Menjawab berbagai pertanyaan yang menyangkut bidang keahliannya.
2. Bila diperlukan, dapat menyajikan asumsi dan alur penalaran yang digunakan untuk sampai ke jawaban yang dikehendaki.
3. Menambah fakta kaidah dan alur penalaran sah yang baru ke dalam otaknya.

Perbandingan kemampuan antara seorang pakar dengan sebuah sistem pakar seperti :

Tabel 2. 1 Perbandingan Seorang Pakar dengan Sistem Pakar

Factor	Human Expert	Expert System
<i>Time availability</i>	hari kerja	setiap saat
Geografis	lokal/tertentu	dimana saja
Keamanan	tidak tergantikan	dapat digantikan
dapat habis	Ya	Tidak
Perfomansi	variable	Konsisten
Kecepatan	variable	Konsisten
Biaya	tinggi	Terjangkau

2.3 Naive bayes

Naive bayes merupakan pengklasifikasi probabilitas sederhana berdasarkan pada teorema Bayes. Teorema Bayes dikombinasikan dengan “Naive” yang berarti setiap atribut/variabel bersifat bebas (*independent*). Naive bayes dapat dilatih dengan efisien dalam pembelajaran terawasi (*supervised learning*). Keuntungan dari klasifikasi adalah bahwa ia hanya

membutuhkan sejumlah kecil data pelatihan untuk memperkirakan parameter (sarana dan varians dari variabel) yang diperlukan untuk klasifikasi. Karena variabel independen diasumsikan, hanya variasi dari variabel untuk masing-masing kelas harus ditentukan, bukan seluruh matriks kovarians (Setiawan & Ratnasari, 2014).

Thomas Bayes menemukan pendekatan penalaran statistik yang jauh lebih maju dibandingkan dengan pola pikir matematis tradisional pada saat itu. Fokus matematika pada saat itu adalah pada tingkah laku sampel dari populasi yang diketahui. Akan tetapi, Bayes mengemukakan ide untuk menentukan property dari populasi berdasarkan sampel tersebut. Dalam “*An Essay Towards the Solving a Problem in the Doctrines of Chance*”, dia menyajikan tentang “Proposition 9”, yang akhirnya dikenal dengan “Teorema Bayes” (Wahyono, 2012). Selanjutnya, teorema ini menjadi dasar dalam pengambilan keputusan modern. Formula Bayes dinyatakan dalam persamaan :

$$P(H | E) = \frac{P(E|H) \cdot P(H)}{P(E)} \quad (2.1)$$

Dimana :

$P(H|E)$ = Probabilitas posterior bersyarat suatu hipotesis H terjadi jika diberikan evidence/bukti E terjadi (*Posterior Probability*).

$P(E|H)$ = Probabilitas sebuah evidence E terjadi akan mempengaruhi hipotesis H (*Likelihood*).

$P(H)$ = Probabilitas awal (priori) hipotesis H terjadi tanpa memandang evidence apapun (*Prior probability*).

$P(E)$ = Probabilitas awal evidence E terjadi tanpa memandang hipotesis/evidence yang lain (*evidence*).

Naive bayes Classifier atau bisa juga disebut sebagai *Multinomial Naive bayes* merupakan model dari penyederhanaan Bayes. Algoritma *Naive bayes* berasumsi bahwa efek suatu nilai variabel di sebuah kelas yang ditentukan adalah tidak terkait pada nilai-nilai variabel lain. *Naive Bayes* dinyatakan sebagai sebuah hipotesa yang disebut dengan HMAP (*Hypothesis Maximum Appriori Probability*) (Wahyono, 2012). Secara matematis HMAP dirumuskan seperti persamaan :

$$\begin{aligned} H_{MAP} &= \arg \max P(h|e) \\ &= \arg \max \frac{P(e|h) * P(h)}{P(e)} \\ &= \arg \max P(e|h) * P(h) \quad (2.2) \end{aligned}$$

Dalam konteks data mining atau *machine learning*, data e adalah set training, dan h adalah ruang dimana fungsi yang akan ditemukan tersebut terletak. HMAP juga seringkali dituliskan seperti persamaan :

$$H_{MAP} = \arg \max_{h_j \in H} \frac{P(a_1, a_2, a_3, \dots, a_n | h) * P(h_j)}{\dots\dots\dots} \quad (2.3)$$

Dimana:

H_{map} = Nilai *output* terbesar hasil klasifikasi *Naive bayes*.

$P(a_1, a_2, a_3, \dots, a_n | h_j)$ = Peluang atribut-atribut (inputan) jika diketahui keadaan / kategori h_j .

$P(h_j)$ = Peluang jenis kategori penyakit h_j .

HMAP menyatakan hipotesa yang diambil berdasarkan nilai probabilitas berdasarkan kondisi prior yang diketahui. HMAP inilah yang digunakan di dalam *machine learning* sebagai metode untuk mendapatkan hipotesis suatu keputusan.

Contoh kasus :

Dalam contoh akan dijelaskan cara melakukan perhitungan berdasarkan data hasil penelitian (Nugroho & Wardoyo, 2013):

- Jumlah pasien: 50 orang
- Penderita Kista Indung Telur: 8 orang, sehingga probabilitas terkena kista indung telur tanpa memandang gejala apapun, $P(\text{Kista indung telur})$ adalah $8/50$.
- Pasien dengan gejala Keputihan adalah 7 orang, sehingga probabilitas pasien dengan gejala keputihan jika menderita kista indung telur $P(\text{keputihan}|\text{Kista}) = 7/8$.
- Jika diketahui gejala keputihan dapat juga menyebabkan kanker indung telur maka probabilitas pasien dengan gejala keputihan jika menderita kanker indung telur, $P(\text{keputihan}|\text{kanker indung})$ adalah $5/6$.
- Sedang probabilitas pasien yang terkena kanker indung telur tanpa memandang gejala apapun, $P(\text{Kanker indung telur})$ adalah $6/50$.

Dengan menggunakan data diatas dapat dihitung:

- (1) probabilitas Kista indung telur jika diketahui gejala keputihan,

$$P(\text{Kista}|\text{Keputihan}) = \frac{P(\text{Keputihan}|\text{Kista}) \cdot P(\text{Kista})}{P(\text{Keputihan}|\text{Kista}) \cdot P(\text{Kista}) + P(\text{Keputihan}|\text{Kanker}) \cdot P(\text{Kanker})}$$

$$P(\text{Kista}|\text{Keputihan}) = \frac{\frac{7}{8} \times \frac{8}{50}}{\frac{7}{8} \times \frac{8}{50} + \frac{5}{6} \times \frac{6}{50}} = 0.58$$

- (2) Probabilitas kanker jika diketahui gejala keputihan adalah

$$P(\text{Kanker}|\text{Keputihan}) = \frac{P(\text{Keputihan}|\text{kanker}) \cdot P(\text{kanker})}{P(\text{Keputihan}|\text{Kista}) \cdot P(\text{Kista}) + P(\text{keputihan}|\text{Kanker}) \cdot P(\text{Kanker})}$$

$$P(\text{Kanker}|\text{Keputihan}) = \frac{\frac{5}{6} \times \frac{6}{50}}{\frac{7}{8} \times \frac{8}{50} + \frac{5}{6} \times \frac{6}{50}} = 0.42$$

Dalam kasus kista indung telur dan kanker nilai probabilitas **0.58** dan **0.42** mengandung makna bahwa probabilitas penyakit tersebut mencakup dari 50 orang pasien sehingga dapat disimpulkan bahwa kemungkinan terkena penyakit kista indung telur.

2.4 Infeksi Saluran Pernafasan

Infeksi pada saluran napas merupakan penyakit yang umum terjadi pada masyarakat. Dalam sebuah penelitian yang menguji hubungan antara jumlah kasus sistem pernafasan dan faktor lingkungan yang mempengaruhi kasus ini seperti tingkat polusi udara perkotaan, penggunaan lahan saat ini, jumlah populasi, dan jumlah kendaraan. Data dikumpulkan di enam kabupaten Izmir, Turki untuk tahun antara 2007 dan 2011. Temuan analisis menunjukkan bahwa tingkat peningkatan jumlah

penduduk, kendaraan bermotor, perubahan penting dalam penggunaan lahan saat ini (daerah terbuka dan hijau, kawasan industri dan pemukiman) dan tingkat penurunan kualitas udara di daerah studi. Hasil penelitian menunjukkan bahwa ada hubungan signifikan secara statistik antara jumlah kasus sistem pernafasan dan faktor lingkungan perkotaan (Tarhan, Ozcan, & Ozkan, 2015).

Infeksi saluran napas berdasarkan wilayah infeksiya terbagi menjadi infeksi saluran napas atas dan infeksi saluran napas bawah. Infeksi saluran napas atas meliputi rhinitis, sinusitis, faringitis, laringitis, epiglottitis, tonsilitis, otitis. Sedangkan infeksi saluran napas bawah meliputi infeksi pada bronkhus, alveoli seperti bronkhitis, bronkiolitis, pneumonia. Infeksi saluran napas atas bila tidak diatasi dengan baik dapat berkembang menyebabkan infeksi saluran nafas bawah. Infeksi saluran nafas atas yang paling banyak terjadi serta perlunya penanganan dengan baik karena dampak komplikasinya yang membahayakan adalah otitis, sinusitis, dan faringitis (Mumpuni & Romiyanti, 2016).

Secara umum penyebab dari infeksi saluran napas adalah berbagai mikroorganisme, namun yang terbanyak akibat infeksi virus dan bakteri. Infeksi saluran napas dapat terjadi sepanjang tahun, meskipun beberapa infeksi lebih mudah terjadi pada musim hujan. Faktor-faktor yang mempengaruhi penyebaran infeksi saluran napas antara lain faktor lingkungan, perilaku masyarakat yang kurang baik terhadap kesehatan diri maupun publik, serta rendahnya gizi. Faktor lingkungan meliputi belum terpenuhinya sanitasi dasar seperti air bersih, jamban, pengelolaan

sampah, limbah, pemukiman sehat hingga pencemaran air dan udara. Perilaku masyarakat yang kurang baik tercermin dari belum terbiasanya cuci tangan, membuang sampah dan meludah di sembarang tempat. Kesadaran untuk mengisolasi diri dengan cara menutup mulut dan hidung pada saat bersin ataupun menggunakan masker pada saat mengalami flu supaya tidak menulari orang lain masih rendah. Pengetahuan dan pemahaman tentang infeksi ini menjadi penting di samping karena penyebarannya sangat luas yaitu melanda bayi, anak-anak dan dewasa, komplikasinya yang membahayakan serta menyebabkan hilangnya hari kerja ataupun hari sekolah, bahkan berakibat kematian khususnya pneumonia.

2.5 PHP (*Hypertext PreProcessor*)

PHP mula mula dikembangkan oleh Rasmus Lerdorf. Istilah PHP mengacu pada *Hypertext Preprocessor*. PHP merupakan Bahasa berbentuk script yang ditempatkan dalam server dan diproses di server. Hasilnya akan dikirim client, tempat pemakai menggunakan browser. PHP dikenal sebagai Bahasa scripting, yang menyatu dengan tag tag HTML, dieksekusi di server, dan digunakan untuk membuat halaman web yang dinamis seperti *Active Server Pages (ASP)* atau *Java Server Pages (JSP)*. PHP adalah Bahasa pemrograman web yang terintegrasi dengan HTML dan berada pada server, PHP banyak dipakai untuk membuat situs web dinamis (Anhar, 2010).

Kelebihan PHP dari Bahasa pemrograman lainnya, yaitu :

- a. Sebuah Bahasa script yang tidak melakukan sebuah kompilasi dalam penggunaannya.
- b. Web server yang mendukung PHP dapat ditemukan dimana mana dari apache, IIS, Lighttpd, hingga Xitami dengan konfigurasi yang relative mudah.
- c. Dalam sisi pengembangan lebih mudah, karena banyaknya milis milis dan developer ang siap membantu pengembangan.
- d. Dalam sisi pemahaman, PHP adalah Bahasa scripting yang paling mudah karena memiliki referensi banyak.

2.6 MySQL

Tujuan awal didirikannya program MySQL adalah untuk mengembangkan aplikasi WEB yang digunakan salah satu client MySQL AB. Pada saat itu MySQL AB adalah sebuah perusahaan konsultan database dan pengembangan software. MySQL versi 1.0 dirilis pada Mei 1996 dan penggunaannya hanya terbatas dikalangan internal saja. Pada bulan Oktober 1996, MySQL versi 3.11.0 dirilis kemasyarakat luas dibawah lisensi “terbuka tapi terbatas”. Dengan lisensi ini, maka siapapun boleh melihat program aslinya dan menggunakan server MySQL secara gratis untuk kegiatan – kegiatan non komersial. Tetapi, untuk kegiatan komersial, maka harus membayar lisensi tersebut.

Pada bulan juni 2000, MySQL AB mengumumkan bahwa mulai versi 3.23.19 diterapkan sebagai General Public License (GPL). Dengan

lisensi ini, maka siapapun boleh melihat program aslinya dan menggunkan program executablenya secara open source atau gratis.

Beberapa keunggulan dari MySQL adalah :

- a. Mampu menangani jutaan user dalam waktu bersamaan
- b. Mampu menanggung lebih dari 50.000.000 record
- c. Sangat cepat mengeksekusi perintah
- d. Memiliki user privilege yang mudah dan efisien

Sebuah website yang interaktif dan dinamis, tentunya akan membutuhkan penyimpanan data yang fleksibel dan cepat untuk diakses. Salah satu database untuk server adalah MySQL, jenis database ini sangat populer dan digunakan pada banyak website di internet sebagai bank data. MySQL menggunakan SQL dan bersifat free (gratis atau tidak perlu membayar untuk menggunakannya). Selain itu, MySQL dapat berjalan di berbagai platform, antara lain Linux dan Windows (Anhar, 2010).

2.7 Penelitian Sebelumnya

Berikut Penelitian-penelitian sebelumnya yang berkaitan dengan penelitian ini, dapat dilihat pada tabel 2.2 :

Tabel 2. 2 Penelitian sebelumnya yang berkaitan

No	Judul	Penulis	Tahun	Deskripsi
1	An Approach to Develop Expert Systems in Medical diagnosis using Machine Learning Algorithms (Asthma) and a Performance Study	Prasadl, Khrisna Prasad, Y Sagar	2011	Dalam penelitian ini menjelaskan bahwa dalam merancang sistem pakar untuk diagnosa asma dapat dilakukan dengan dua pendekatan yaitu 1) melalui kuisisioner 2) melalui data klinis. Sebuah keuntungan menggunakan sistem pakar adalah untuk diagnosa kemungkinan yang berbeda

				dalam situasi klinis tertentu tidak harus dilakukan oleh spesialis (<i>human expert</i>), tetapi secara otomatis disediakan oleh sistem dan sistem melakukan akuisisi pengalaman dari pakar (<i>human expert</i>) (Prasadl, Prasad, & Sagar, 2011).
2	Expert Systems for Self-Diagnosing of Eye Diseases Using Naive Bayes	Rahmat Kurniawan, Novi Yanti, Mohd Zakree Ahmad Nazri, Zulvandri	2014	Penelitian ini menggunakan konsep Case Base Reasoning (CBR). Model CBR digunakan untuk menyelesaikan masalah dan melakukan <i>generate</i> hasil yang didasarkan pada <i>history</i> diagnosa penyakit mata. Beberapa proses dalam CBR diantaranya <i>retrieve</i> , <i>reuse</i> , <i>revise</i> , dan <i>retain</i> . Data pada penelitian ini menggunakan 12 penyakit mata. Persentase kesesuaian antara diagnosis sistem pakar dan pakar sebenarnya (<i>human expert</i>) adalah sebesar 82% (Kurniawan, Yanti, Nazri, & Zulvandri, 2014).
3	Komparasi Algoritma Klasifikasi DataMining Untuk Memprediksi Penyakit Tuberkulosis (TB): Studi Kasus Puskesmas Karawang Sukabumi	Rizal Amegia Saputra	2014	Dalam penelitian ini dilakukan komparasi algoritma <i>C4.5</i> , <i>naive baye</i> , <i>neural network</i> , dan <i>logistic regression</i> yang diaplikasikan terhadap data pasien yang dinyatakan positif tuberkulosis dan negative tuberkulosis. Dengan mengukur kinerja dari keempat algoritma tersebut menggunakan metode pengujian <i>Confusion Matrix</i> dan <i>Kurva ROC</i> . Penelitian ini menyimpulkan dengan membandingkan empat metode data mining dari algoritma tersebut, hasil evaluasi dan validasi diketahui bahwa <i>Naive Bayes</i> memiliki nilai akurasi dan <i>AUC</i> paling tinggi diantara metode lainnya yang dikomparasikan, diikuti oleh algoritma <i>C45</i> , <i>Neural Network</i> ,

				dan <i>Logistic Regression</i> memiliki akurasi paling rendah (Saputra, 2014).
4	Fault Diagnosis of Welded Joints Through Vibration Signals Using Naive Bayes Algorithm	Girish Kumar M, Hemanth K, Gandadhar N, Nemantha Kumar, Prasad Khrisna	2014	Penelitian ini menyajikan tentang pemeriksaan sambungan las memainkan peran penting dalam menemukan kelemahan internal dan kerusakan pada percobaan. Penelitian ini menggunakan getaran sinyal yang diperoleh dari <i>accelometer</i> dalam percobaan las untuk bebas kerusakan sambungan las dan imbas kerusakan dari sambungan las untuk kesalahan diagnosa melalui <i>machine learning</i> menggunakan Decission Tree (J48) dan Naive Bayes yang digunakan sebagai classifier. Hasil dari algoritma naive bayes digunakan untuk pengetahuan dan menggolongkan kondisi sambungan las, diperoleh akurasi 85,56%. Dengan demikian model <i>naive bayes classifier</i> pada kenyataanya dapat digunakan untuk diagnosa kualitas kondisi sambungan las (Girish, Kumar, Gangad, Kumar, & Krishna, 2014).
5	Classification of Complete Blood Count and Haemoglobin Typing Data by a C4.5 Decision Tree, a Naive Bayes Classifier and Multilayer Preceptron for Thalassaemia Screening	Damrongrit Setsirichok, Theera Piroonratana, Wawanyu Wongseeree, Touchpong Usavanarong, Nuttawut Paulkhaolarn, Chompunut Kanjanakorn, Monchan Sirikong, Chaini Limwongse, Nachol Chaiyaratana	2012	Penelitian ini menyajikan klasifikasi karakteristik darah dengan <i>C4.5</i> , <i>Naive Bayes Classifier</i> , <i>Multilayer Perceptron</i> untuk <i>thalassemia screening</i> . Data yang digunakan adalah 18 kelas kelainan <i>thalassemia</i> . Penelitian ini menerapkan <i>Naive Bayes Classifier</i> dan <i>Multilayer Perceptron</i> . Hasil pengujian akurasi menunjukkan 99,39% untuk <i>Naive Bayes Classifier</i> . Dan 99,71% untuk <i>Multilayer Perceptron</i> (Setsirichok, et al., 2012).

6	The Relationship Between Respiratory System Case and Environmental Urban Factors	Cigdem Tarhan, Nur Sinem Ozcan, Sevim Pelin Ozkan	2015	Tujuan utama dalam penelitian ini adalah untuk menguji hubungan antara jumlah kasus sistem pernafasan dan faktor lingkungan yang mempengaruhi kasus ini seperti tingkat polusi udara perkotaan, penggunaan lahan saat ini, jumlah populasi, dan jumlah kendaraan. Data dikumpulkan di enam kabupaten Izmir, Turki untuk tahun antara 2007 dan 2011. Temuan analisis menunjukkan bahwa tingkat peningkatan jumlah penduduk, kendaraan bermotor, perubahan penting dalam penggunaan lahan saat ini (daerah terbuka dan hijau, kawasan industri dan pemukiman) dan tingkat penurunan kualitas udara di daerah studi. Hasil penelitian menunjukkan bahwa ada hubungan signifikan secara statistik antara jumlah kasus sistem pernafasan dan faktor lingkungan perkotaan (Tarhan, Ozcan, & Ozkan, 2015).
---	--	---	------	---