

Diagnosa Penyakit Demam Tifoid Dengan Model Algoritma C4.5

Diagnosis of Typhoid Fever Using the C4.5 . Algorithm Model

Ramadhan Hermansyah¹, Deni Arifianto^{2*}, Dewi Lusiana³

¹Mahasiswa Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember
Email: ramadhanhermansyah@gmail.com

²Dosen Fakultas Teknik, Universitas Muhammadiyah Jember* Koresponden Author
Email: deniarifianto@unmuhjember.ac.id

³Dosen Fakultas Teknik, Universitas Muhammadiyah Jember
Email: dewilusiana2011@gmail.com

Abstrak

Demam tifoid adalah suatu penyakit yang disebabkan oleh bakteri yang bernama salmonella typhii. kasus penyakit ini diperhitungkan menyentuh angka 16 juta kasus pertahunnya. Orang meninggal lebih dari 600.000 jiwa setiap tahunnya karena penyakit tifus. Perkembangan teknologi informasi selalu mengalami perkembangan yang pesat setiap tahunnya. Dampak dari perkembangan teknologi informasi yaitu munculnya banyak fasilitas-fasilitas yang memudahkan, sehingga secara langsung berdampak kepada kegiatan organisasi. Algoritma C4.5 adalah suatu algoritma yang menghasilkan pohon keputusan yang mudah dijelaskan, memiliki tingkat akurasi yang bisa diterima, dan efisiensi yang cukup dalam menangani suatu atribut bertipe diskret dan numerik. Karena kelebihan itu algoritma ini lebih sering digunakan. Adapun hasil yang diperoleh dari penerapan metode C4.5 pada penelitian ini yaitu aplikasi dapat digunakan sebagai sistem diagnosa demam tifoid lebih awal dan aplikasi dapat diterima oleh masyarakat serta mendapatkan tingkat nilai akurasi sebesar 90%.

Kata Kunci: *decision tree , C4.5, Cross Validation, Klasifikasi*

Abstract

Typhoid fever is a disease caused by a bacterium called salmonella typhii. The number of cases of this disease is calculated to reach 16 million cases per year. More than 600,000 people die each year from typhus. The development of information technology always experiences rapid development every year. The impact of the development of information technology is the emergence of many facilities that facilitate, so that it directly affects the activities of the organization. The C4.5 algorithm is an algorithm that produces a decision tree that is easy to explain, has an acceptable level of accuracy, and has sufficient efficiency in handling discrete and numeric type attributes. Because of its advantages, this algorithm is used more often. The results obtained from the application of the C4.5 method in this study are that the application can be used as an early typhoid fever diagnosis system and the application can be accepted by the community and get an accuracy value of 90%.

Keywords: *decision tree , C4.5, Cross Validation, Clasification.*

1. PENDAHULUAN

Demam tifoid adalah merupakan suatu penyakit yang disebabkan oleh bakteri yang bernama *salmonella typhii*. Kasus penyakit ini diperhitungkan menyentuh angka 16 juta kasus pertahunnya. Banyak Orang meninggal lebih dari 600.000 jiwa setiap tahunnya karena penyakit tifus. Di Indonesia saja, penyakit endemik ini sering menjadi masalah kesehatan yang serius dan penyakit ini sering lebih sering disebut penyakit tifus ini oleh masyarakat Indonesia. Penyakit ini dapat menyerang tubuh melalui makanan maupun minuman yang menyebabkan infeksi usus halus (Tolle, 2008). Banyak masyarakat yang masih menganggap demam tifoid itu hanya demam biasa dan tidak mau melakukan pemeriksaan secara lebih intensif untuk lebih mengetahui lagi tentang gejala-gejala yang dirasakan. Kurangnya pengetahuan yang ada pada masyarakat mengenai gejala dan cara penanganan penyakit demam tifoid merupakan salah satu alasan penyebab tingginya angka kematian akibat penyakit demam tifoid dan juga merupakan suatu permasalahan pada kasus demam tifoid ini.

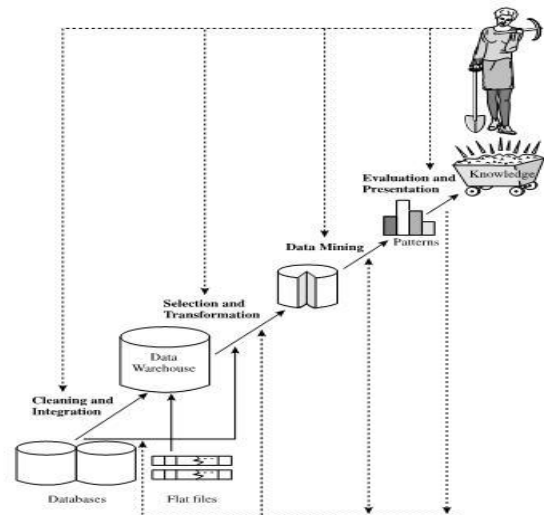
Sistem pakar adalah sistem yang berusaha menerapkan pengetahuan manusia pada komputer, sehingga nantinya komputer dapat menyelesaikan persoalan tersebut layaknya seorang ahli (Sri Kusumadewi, 2003). Dalam bidang kedokteran sistem pakar diterapkan untuk mendiagnosa penyakit. Pada Penelitian sebelumnya yang telah dilakukan oleh Furqon dan Rahayudi (2018) dengan judul penelitian “Klasifikasi Penyakit Typhoid Fever (TF) dan Dengue Haemorrhagic Fever (DHF) dengan Menerapkan Algoritma Decision tree C4.5”. Maka peneliti hendak mengusulkan suatu program yang dapat dipakai untuk mengelompokkan seorang pasien tersebut masuk kedalam kategori terkena penyakit demam tifoid maupun tidak berdasarkan gejala yang dirasakan para pasien.

2. TINJAUAN PUSTAKA

A. Data Mining

Data Mining merupakan suatu metode yang digunakan untuk mendapatkan pola, hubungan dan tren baru yang berfungsi untuk

menyaring data yang sangat besar, yang tersimpan pada penyimpanan, memakai teknik pengenalan pola seperti Teknik statistik dan matematika (Kursini dan Luthfi, 2009). Data Mining merupakan beberapa tahap yang dibagi dari suatu rangkaian proses. Tahapan tersebut bisa digambarkan sebagai berikut :



Gambar 1. Pengertian Data Mining

Sumber: Website Pengertian Data Mining

B. Demam Tifoid

Pengertian Demam *tifoid* adalah infeksi akut saluran cerna yang disebabkan oleh *Salmonella typhi*. Demam paratifoid adalah penyakit sejenis yang disebabkan oleh *Salmonella Paratyphi A, B, dan C*. Gejala dan tanda penyakit tersebut hampir sama, namun manifestasi paratifoid lebih ringan (Widoyono, 2008).

C. Algoritma C4.5

Algoritma C4.5 cukup sering digunakan dalam mengklasifikasi suatu data yang memiliki atribut numerik dan kategorial. Proses dari hasil klasifikasi algoritma C4.5 bisa dipakai dalam menebak nilai atribut berjenis diskret dari record yang baru. Metode algoritma C4.5 adalah perkembangan algoritma ID3, dimana peningkatannya agar bisa melakukan beberapa hal, seperti bisa menangani data hilang, dan dapat menangani data *peruning* dan kontinu.

D. PHP

Menurut Arief (2011) bahasa pemrograman PHP merupakan suatu bahasa pemrograman

server-side-scripting yang dapat digabungkan dengan HTML dalam pembuatan halaman web yang lebih dinamis. *Server-side-scripting* merupakan bagian dari bahasa pemrograman PHP yang dimana perintah-perintah dan sintaks akan dieksekusi di *server* kemudian hasil dari sintaks dan perintah-perintah tersebut akan dikirimkan ke *browser* dengan format HTML.

E. MySQL

Menurut Arief (2011) “MySQL merupakan *Database server* yang banyak digunakan dalam membangun suatu aplikasi web yang database digunakan sebagai sumber dan pengolahan datanya”.

3. METODOLOGI PENELITIAN

A. Identifikasi Masalah

Perumusan masalah merupakan tahapan awal di dalam penelitian ini yang dijadikan objek penelitian. Melihat kondisi aktual di lapangan terlebih dahulu merupakan cara melakukan perumusan masalah. Setelah dirumuskannya masalah tahapan berikutnya yaitu dengan menentukan tujuan penelitian. Tujuan penelitian merupakan poin penting yang akan dihasilkan dari penyelesaian permasalahan yang diteliti.

B. Tinjauan Pustaka

Tahapan ini digunakan untuk mendapatkan literatur pendukung penelitian. Pada tahap penulis menjelaskan mengenai metode yang digunakan dan menjelaskan mengenai webiste atau situs yang dipelajari maupun dikunjungi, teori-teori, dan tools yang berhubungan dengan diagnosa penyakit demam *tifoid*.

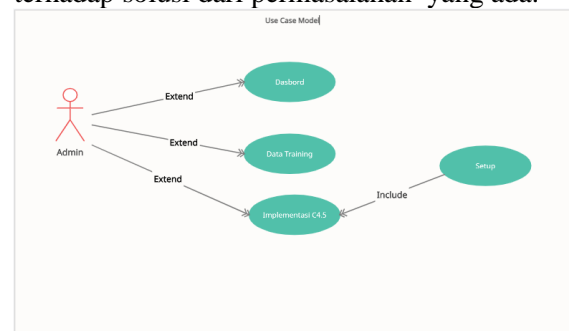
C. Metode Pengumpulan Data

Terdapat dua metode cara dalam pengumpulan data pada penelitian ini yang pertama yaitu dengan melakukan Pengamatan suatu tempat dan literatur. Pengamatan dilakukan demi mendapatkan dan mengumpulkan data yang dibutuhkan. Pengamatan ini dilakukan di RS Al Huda Banyuwangi.

D. Desain Sistem

Di tahapan ini penulis akan membentuk suatu perancangan sistem yang akan diusulkan. Ada beberapa urutan dalam perancangan sistem yang akan penulis lakukan pada penelitian ini, berikut urutannya:

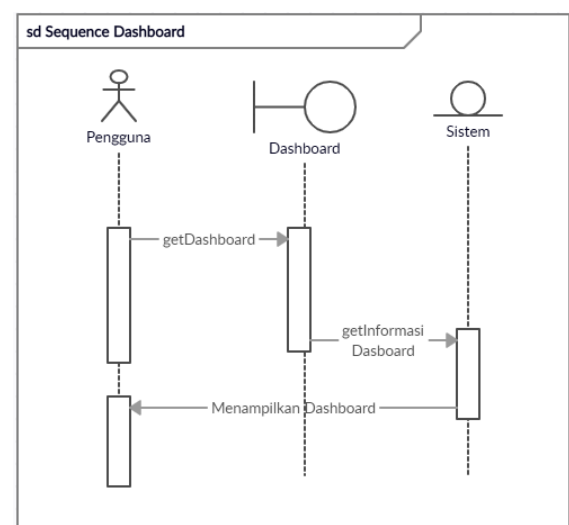
- Pertama yaitu merancang tampilan user agar terlihat menarik dan mudah dimengerti,
- kedua merancang basis data untuk sistem supaya lebih mudah dalam manajemen file yang ada, Dan yang terakhir yaitu merancang coding program dari suatu informasi. Perangkat pemodelan seperti Unified Modeling Language (UML) dapat digunakan sebagai tahapan penguasaan pikiran dan perancangan sistem terhadap solusi dari permasalahan yang ada.



Gambar 2. Use case Diagram

Sumber: Hasil Pembuatan Sequence Diagram

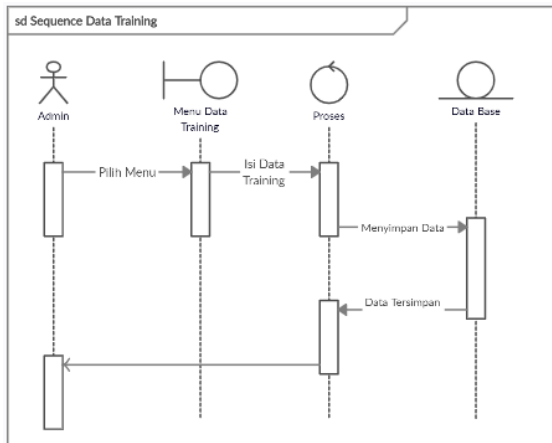
Di situs <https://app.creately.com/>



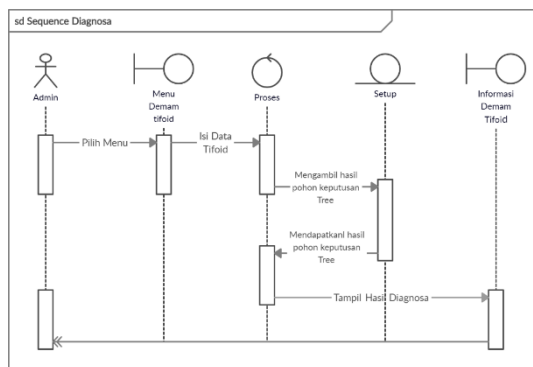
Gambar 3. Sequence Dashboard

Sumber: Hasil Pembuatan Sequence Diagram

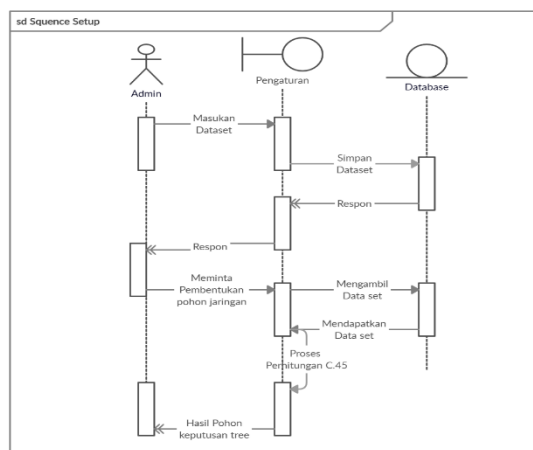
Di situs <https://app.creately.com/>



Gambar 4. Sequence Diagram Data Training
 Sumber: Hasil Pembuatan Sequence Diagram
 Di situs <https://app.creately.com/>



Gambar 5. Sequence Diagram Diagnosa
 Sumber: Hasil Pembuatan Sequence Diagram
 Di situs <https://app.creately.com/>

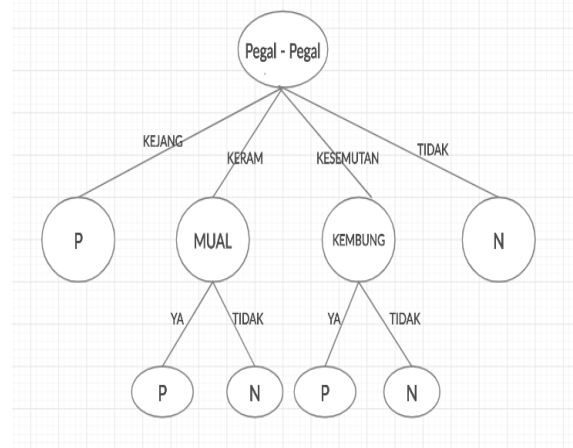


Gambar 6. Sequence Diagram Setup

Sumber: Hasil Pembuatan Sequence Diagram
 Di situs <https://app.creately.com/>

E. Implementasi Sistem Dan Testing

Agar aplikasi yang sedang dikembangkan bebas dari kesalahan ataupun *bug*, dilakukanlah tahap Implementasi dan testing pada aplikasi tersebut. Ditahapan ini dilakukan sebuah evaluasi pada hasil penelitian yang sudah dilakukan. Evaluasi mencakup hasil dan manfaat dengan membandingkan hasil yang didapatkan dengan kebutuhan pengguna. Dari hasil perhitungan metode C4.5 diatas maka diperoleh pohon keputusan seperti berikut:



Gambar 7 Pohon keputusan
 Sumber: Hasil perhitungan

F. Gambaran Data Set

Data penelitian yang akan digunakan yaitu data dari RS Al Huda Banyuwangi. Data ini berupa data gejala demam tifoid yang terdiri dari 10 atribut yaitu Usia, Jenis Kelamin, Demam, Sakit kepala, Pegal-pegal, Mual, Muntah, Batuk, Kembung, Diare dan 2 output yaitu Positif dan Negatif. Data ini di ambil pada periode 2019 sampai Januari 2020 yaitu sebanyak 100 data.

G. Skenario Uji

Pengujian yang akan dilakukan menggunakan K-fold cross validation. pada pengujian yang menggunakan K-fold cross validation data terlebih dahulu di bagi menjadi data training dan data testing. Dimana dengan jumlah data yang berbeda yang habis dibagi 2. Data akan di bagi menjadi k sama dengan 2, 4, 5, dan 10.

H. Kisi-kisi kuisioner

Keterangan pilihan jawaban:
 STS : Sangat Tidak Setuju
 TS : Tidak Setuju
 N : Netral
 S : Setuju
 SS : Sangat Setuju

Tabel 1. Pertanyaan kuisioner untuk responden

No	Keterangan	STS	TS	N	S	SS
1	Kebutuhan untuk menggunakan aplikasi tersebut sangat tinggi					
2	Informasi yang dihasilkan oleh aplikasi apa mudah dipahami					
3	Aplikasi cukup ringan untuk dijalankan					
4	Parameter gejala pada aplikasi sudah cukup					
5	Aplikasi sangat mudah diakses dari semua komputer					
6	Aplikasi dapat menghasilkan laporan yang fleksibel					
7	Aplikasi yang ada mudah dipahami					
8	Aplikasi yang ada mudah digunakan					
9	Aplikasi tidak pernah mengalami error					
10	Tata letak tampilan/display mudah dikenali/dilihat					

Sumber: Penulis

Dikarenakan pada kuisioner ini saya menggunakan pernyataan yang favorable atau pernyataan item positif, maka untuk itu saya akan memberikan skor pada tiap-tiap jawab.

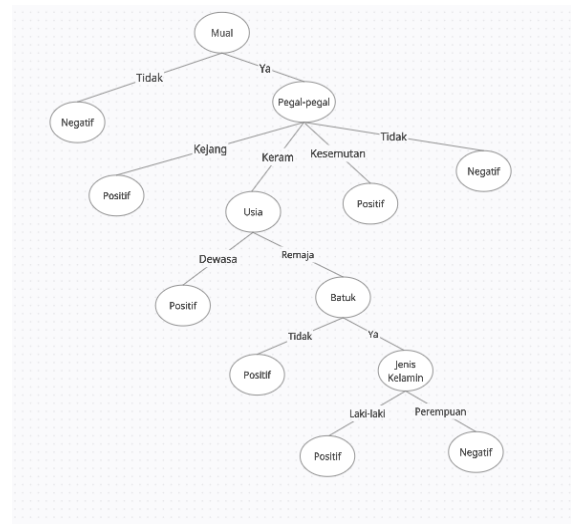
Berikut skor untuk tiap-tiap jawaban:

STS : Sangat Tidak Setuju (Skor 1)
 TS : Tidak Setuju (Skor 2)
 N : Netral (Skor 3)
 S : Setuju (Skor 4)
 SS : Sangat Setuju (Skor 5)

Untuk mengetahui hasil kuisioner penulis akan menyebarkan kepada 50 orang responden. Hasil akan diambil dari suara terbanyak, dan jika mendapatkan hasil suara yang seimbang, maka peneliti akan menambah jumlah responden

4. PEMBAHASAN

Pada hasil pengujian ini yang akan perlihatkan adalah hasil terbaik dari cross validation, didapatkan akurasi terbaik pada 5-fold percobaan ke 2 sebesar 90,00%. Dengan presisi dimana nilai Prediksi positif memiliki hasil 100%, nilai Prediksi negatif 80%. Pada K-fold 5 percobaan 2 terdapat 75 data training yang akan digunakan sebagai acuan untuk dimasukan kedalam Aplikasi. Berikut Tabel perhitungan manual untuk data K-fold 5 percobaan 2.



Gambar 8 .Pohon Keputusan

Sumber: Hasil Perhitungan

Tabel 2 Hasil akurasi tiap fold

No	Fold	Hasil akurasi
1	Fold 2 Percobaan ke 1	84%
2	Fold 2 Percobaan ke 2	84%
3	Fold 4 Percobaan ke 1	80%
4	Fold 4 Percobaan ke 2	76%
5	Fold 4 Percobaan ke 3	84%
6	Fold 4 Percobaan ke 4	88%
7	Fold 5 Percobaan ke 1	85%
8	Fold 5 Percobaan ke 2	90%
9	Fold 5 Percobaan ke 3	75%
10	Fold 5 Percobaan ke 4	70%
11	Fold 5 Percobaan ke 5	85%
12	Fold 10 Percobaan ke 1	80%
13	Fold 10 Percobaan ke 2	80%
14	Fold 10 Percobaan ke 3	90%
15	Fold 10 Percobaan ke 4	90%
16	Fold 10 Percobaan ke 5	80%
17	Fold 10 Percobaan ke 6	80%
18	Fold 10 Percobaan ke 7	80%
19	Fold 10 Percobaan ke 8	80%
20	Fold 10 Percobaan ke 9	90%
21	Fold 10 Percobaan ke 10	90%

Sumber: Hasil Perhitungan

Dari 10 pertanyaan didapatkan 8 pertanyaan dengan hasil kategori “Sangat

Setuju”. Dapat disimpulkan dari hasil kuisioner diatas jika aplikasi diagnosa dapat diterima masyarakat dalam mengetahui penyakit tifoid lebih awal

Tabel 3 Hasil kuisioner pada setiap pertanyaan

Pertanyaan	Nilai	Kategori
1	80%	Sangat Setuju
2	82,6%	Sangat Setuju
3	62%	Setuju
4	83,6%	Sangat Setuju
5	53,6%	Netral
6	75,6%	Setuju
7	93,2%	Sangat Setuju
8	92%	Sangat Setuju
9	54%	Netral
10	88%	Sangat Setuju

Sumber: Hasil Perhitungan

5. KESIMPULAN

Berdasarkan hasil penelitian tentang diagnosa penyakit *Typhoid Fever (TF)* dengan menerapkan algoritma decision tree C4.5 dapat disimpulkan bahwa :

1. Dari 100 record data dengan 10 kali percobaan dimana nilai akurasi tertinggi didapatkan pada k-5 pada percobaan 2, dimana akurasi yang didapatkan sebesar 90%.
2. Sedangkan hasil untuk kuisioner, didapatkan hasil kategori “Sangat Setuju”. Dapat disimpulkan bahwa diagnosa penyakit demam *tidoid* menggunakan aplikasi dapat diterima oleh masyarakat.

6. DAFTAR PUSTAKA

- Adeyemo, O. O. Adeyeye, T. O & Ogunbiyi, O. 2015. *Comparative Study of ID3/C4.5 Decision tree and Multilayer Perceptron Algorithms for the Prediction ofTyphoid Fever* . Afr J. of Comp & ICTs. Vol 8, No. 1. ISSN 2006-1781:103-112
- Alexander F. K. Sibero, 2011, *Kitab Suci Web Programming*, Media Kom, Yogyakarta
- Arief, M.Rudianto. 2011. *Pemrograman Web Dinamis Menggunakan* Php dan Mysql. Yogyakarta: ANDI
- Astuti, O. (2013). *Demam Tifoid*. Fakultas Kedokteran Universitas Muhammadiyah Surakarta.
- Betha Sidik., 2012, *Pemrograman Web dengan PHP*, Informatika,Bandung.
- Depkes RI. 2006. *Pedoman Penyelenggaraan dan Prosedur Rekam Medis Rumah Sakit di Indonesia*. Jakarta: Depkes RI.
- Depertemen Kesehatan RI. 2008. *Standar Pelayanan Minimal Rumah Sakit*. Jakarta: Departemen Kesehatan RI.
- Gorunescu, F., 2011, *Data mining concepts models and technique*. berlin: Springer.
- Harahap, N., 2011. ‘Karakteristik Penderita Demam Tifoid Rawat Inap Di RSUD Deli Serdang Lubuk Pakam’. FK Universitas Sumatera Utara, <<http://Repository.USU.2011.pdf>>. Diakses tanggal 10 April 2017.
- Kursini Dan Luthfi, E. T., 2009, *Algoritma Data Mining*. Yogyakarta: C.V Andi Offset.
- Moertini, S.,2003, *Towards The Use Of C4.5 Algorithm For Classifying Banking Dataset., Integral*. Vol. 8 (2), pp. 105-115
- Nugroho, (2006), *Membuat Aplikasi Sistem Pakar dengan PHP dan My SQL dengan PHP dan MySQL dengan Editor Dreamweaver*, Ardana Media, Yogyakarta.
- Putra, Ade. 2012. *Hubungan Antara Tingkat Pengetahuan Ibu Tentang Demam Tifoid Terhadap Kebiasaan Jajan Anak Sekolah Dasar*. Semarang: Program Pendidikan Sarjana

kedokteran Fakultas Kedokteran
Universitas Diponegoro

- Tolle, H., 2008. Pengantar Sistem Pakar.
Diambil dari:
<http://www.hermantolle.com/blog/file-sharing/>. (03 Januari 2020)
- Wajhillah, Rusda. 2014. *Optimasi Algoritma Klasifikasi C4.5 Berbasis Particle Swarm Optimization untuk Prediksi Penyakit Jantung*. SWABUMI, Vol 1.No.1
- WHO. 2015. Typhoid fever – Uganda.
Diambil dari:
<http://www.who.int/csr/don/17-march-2015-uganda/en/>. (05 Januari 2020)
- Widodo, J., 2006, *Demam Tifoid, Buku Ajar Ilmu Penyakit Dalam*, Pusat Penerbitan Departemen Ilmu Penyakit Dalam FKIU, Jakarta.
- Yuwono, B., 2010. Pengembangan sistem pakar pada perangkat mobile.