

## PENERAPAN ALGORITMA C4.5 TERHADAP DIAGNOSA PENYAKIT DEMAM TIFOID BERBASIS MOBILE

Ihsan<sup>1</sup>, Rusda Wajhillah<sup>2\*</sup>)

Program Studi Sistem Informasi  
Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK Nusamandiri)  
Jl. Veteran II No. 20A Kota Sukabumi  
[Ihsan1605@nusamandiri.ac.id](mailto:Ihsan1605@nusamandiri.ac.id)<sup>1</sup>, [rusda.rwh@bsi.ac.id](mailto:rusda.rwh@bsi.ac.id)<sup>2</sup>

### Abstract

*The use of gadgets in today is not something new anymore. Gadgets can allow a person to perform activities of information. This is used to create mobile-based applications in the use of artificial intelligence in the world of health. Here is a tool used eclips and database use sqlite3 with android OS usage targets. Typhoid is an acute systemic disease caused by salmonella typhimurium. In the world, the incidence of typhoid fever is estimated to reach 16 million cases annually. More than 600,000 people die every year due to this disease. Based on the data processing of typhoid fever Waluran Health Centers using the algorithm C4.5 obtained 15 rule/decision trees were applied in the development of mobile-based applications as the basis for the decision. Utilization data mining in this application is a manifestation of artificial intelligence in mobile-based information system in particular. The same practice will help facilitate the gadget users to diagnose the disease at an early stage of typhoid fever using existing symptoms.*

**Keywords:** *Tifoid Fever, C4.5 Algorithm*

### I. PENDAHULUAN

Demam tifoid merupakan penyakit sistemik akut yang disebabkan oleh *salmonella typhimurium*. Insidensi demam tifoid diperkirakan mencapai 16 juta kasus setiap tahunnya. Lebih dari 600.000 orang meninggal setiap tahun karena penyakit ini. Di Indonesia, demam tifoid atau lebih dikenal sebagai penyakit tifus merupakan penyakit endemik dan menjadi masalah kesehatan yang serius. Insidensi rata-rata mencapai 650 kasus per 100.000 penduduk di Indonesia, dengan mortalitas rata-rata bervariasi dari 3,1 - 10,4% [10].

Penyakit ini dapat menyerang tubuh melalui makanan atau minuman yang menyebabkan infeksi usus halus [9]. Banyak terjadi pada musim penghujan terutama di daerah dengan tingkat sanitasi rendah dan daerah banjir [12].

Data mining adalah proses yang menggunakan statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar.

Algoritma C4.5 yaitu metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon

keputusan yang merepresentasikan aturan. Algoritma C4.5 juga dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari record pada kategori tertentu [6]. Pada algoritma ini dilakukan penghitungan *entropy* dan *gain information* untuk memperoleh node akar dan node lainnya. Algoritma C 4.5 dapat digunakan untuk menentukan penyakit demam tifoid mengacu pada data primer yang ada.

Salah satu penerapannya adalah dalam dunia medis. Untuk penyakit-penyakit khusus diperlukan keahlian seorang dokter untuk melakukan diagnosa dan pemeriksaan. Namun masalahnya dokter yang dimaksud tidak setiap waktu ada ditempat dan tidak semua orang bisa langsung memeriksakan kesehatannya ke dokter maka dari itu akan dibuatkan sebuah perangkat lunak berbasis *mobile* yang bisa menangani permasalahan tersebut.

### II. TINJAUAN PUSTAKA

#### Data Mining

*Data Mining* didefinisikan sebagai sebuah proses untuk menemukan hubungan, pola dan tren baru yang bermakna dengan menyaring data yang sangat besar, yang tersimpan dalam penyimpanan,

menggunakan teknik pengenalan pola seperti teknik statistik dan matematika [4].

Pengolahan data mining terdiri dari beberapa metode pengolahan [4], yaitu:

*Predictive Modelling* merupakan pengolahan data mining dengan melakukan prediksi/peramalan. Tujuan metode ini untuk membangun model prediksi suatu nilai yang mempunyai ciri-ciri tertentu.

*Association* (Asosiasi) merupakan teknik dalam data mining yang mempelajari hubungan antar data.

*Clustering* (Klastering) merupakan teknik untuk mengelompokkan data ke dalam suatu kelompok tertentu.

*Classification* merupakan teknik mengklasifikasikan data.

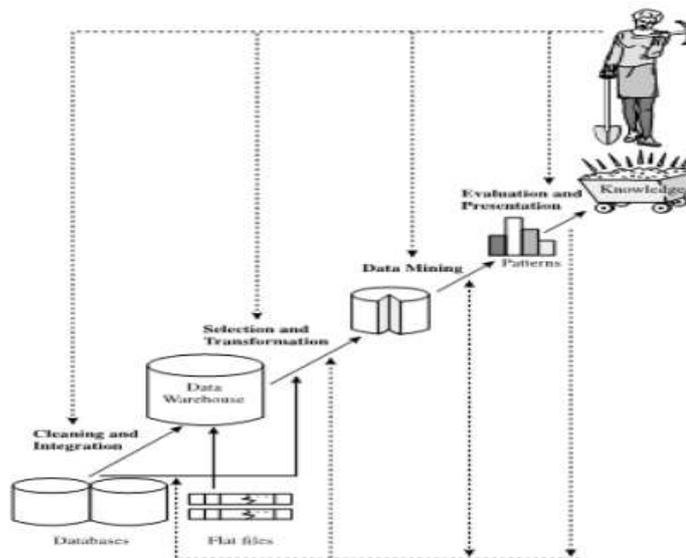
**Algoritma C4.5**

Algoritma C4.5 merupakan kelompok algoritma *Decision Tree* (Pohon Keputusan). Algoritma ini mempunyai input berupa *training samples* dan *samples*. *Training samples* berupa data contoh yang akan digunakan untuk membangun sebuah tree yang telah diuji kebenarannya. Sedangkan *samples* merupakan field-field data yang nantinya akan digunakan sebagai parameter dalam melakukan klasifikasi data [2].

Secara umum alur proses algoritma C4.5 untuk membangun pohon keputusan dalam data mining [13] adalah:

- a) Pilih atribut sebagai simpul akar.
- b) Buat cabang untuk tiap-tiap nilai.
- c) Bagi kasus dalam cabang.

Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.



Sumber: [13]

**Gambar 1.** Tahapan Dalam Data Mining

Pemilihan atribut sebagai simpul, baik simpul akar (*root*) atau simpul internal didasarkan pada nilai Gain tertinggi dari atribut-atribut yang ada. Penghitungan nilai Gain digunakan rumus seperti dalam Persamaan 1.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \dots (1)$$

- S : Himpunan kasus
- A : Atribut
- n : Jumlah partisi atribut A
- |S<sub>i</sub>| : Jumlah kasus pada partisi ke-i
- |S| : Jumlah kasus dalam S

Untuk menghitung nilai *entropy* dapat dilihat pada Persamaan 2.

$$Entropy(S) = \sum_{i=1}^n - pi * \log_2 pi \dots (2)$$

- n : Jumlah partisi S
- pi : Proporsi dari S<sub>i</sub> terhadap S

**III. METODE DAN OBJEK PENELITIAN**

Pada penelitian ini menggunakan data primer yang terdiri dari 142 data pasien yang diperoleh dari puskesmas Waluran Jampang Kulon Kabupaten Sukabumi periode 2015. Meliputi beberapa gejala atau atribut yakni demam tinggi, sakit kepala,

pusing, pegal-pegal, mual, muntah, batuk, dengan nadi antara 80-100 kali permenit, perut kembung dan merasa tak enak, sedangkan diare dan sembelit

silih berganti. Pada tabel 1 ditampilkan sample data yang digunakan pada penelitian ini.

**Tabel 1.** Sample Data Pasien Penyakit Demam Tifoid

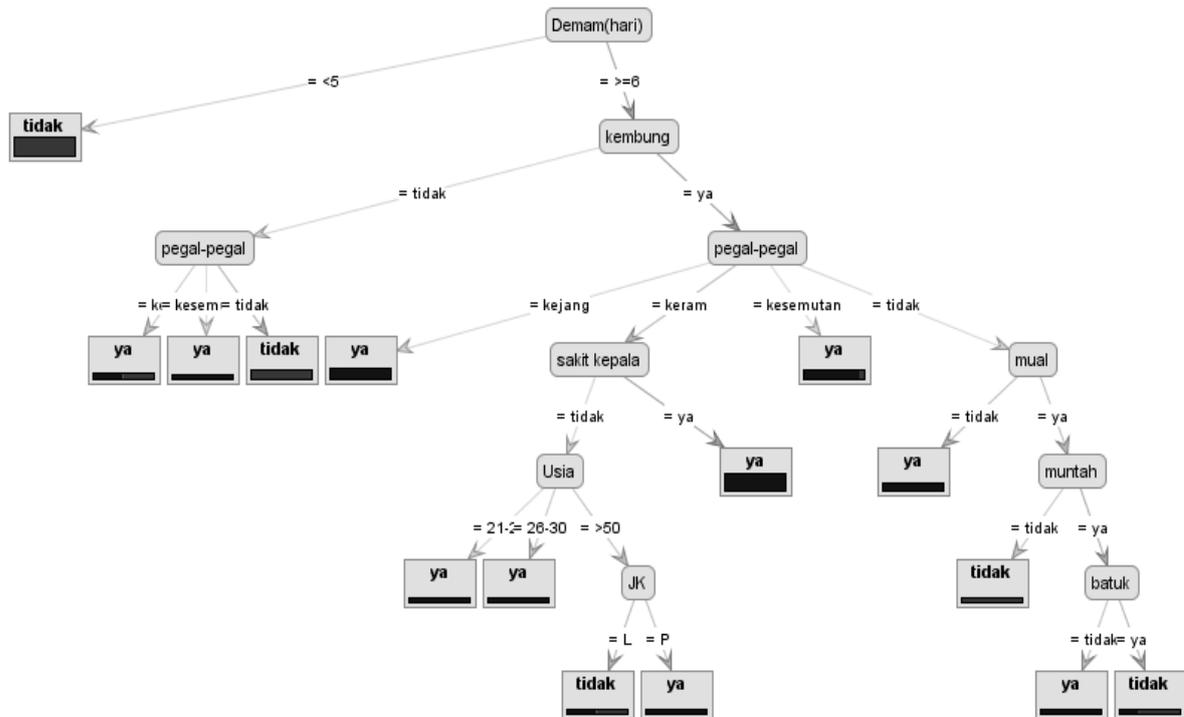
Usia	JK	Demam (hari)	Sakit kepala	Pegal-pegal	Mual	Muntah	Batuk	Denyut nadi/menit	Kembung	Diare	Kelas
16-20	P	>=6	ya	keram	ya	ya	ya	80-100	ya	ya	ya
>50	P	<5	tidak	tidak	tidak	tidak	ya	<80	tidak	ya	tidak
<=10	L	>=6	ya	keram	ya	ya	ya	80-100	ya	ya	ya
26-30	L	>=6	ya	keram	ya	ya	tidak	80-100	ya	ya	ya
26-30	L	<5	tidak	tidak	ya	ya	tidak	<80	ya	tidak	tidak
>50	P	<5	tidak	tidak	tidak	tidak	ya	80-100	tidak	ya	tidak
16-20	L	>=6	ya	tidak	ya	ya	ya	80-100	tidak	ya	tidak
>50	L	>=6	tidak	keram	tidak	tidak	ya	80-100	ya	ya	tidak
11-15	P	>=6	tidak	tidak	ya	tidak	tidak	80-100	ya	ya	tidak
26-30	P	>=6	tidak	keram	tidak	tidak	ya	80-100	tidak	tidak	tidak
26-30	L	<5	tidak	tidak	tidak	ya	tidak	<80	ya	ya	tidak
26-30	L	<5	tidak	keram	tidak	tidak	ya	<80	tidak	ya	tidak
>50	L	>=6	ya	kesemutan	tidak	ya	tidak	80-100	ya	tidak	ya
>50	P	>=6	ya	kejang	tidak	tidak	ya	80-100	ya	ya	ya
11-15	L	>=6	ya	keram	ya	ya	ya	80-100	ya	ya	ya
26-30	L	>=6	ya	keram	ya	ya	tidak	80-100	ya	tidak	ya
26-30	L	>=6	tidak	keram	ya	ya	ya	80-100	ya	tidak	ya
<=10	P	>=6	ya	keram	ya	ya	ya	80-100	ya	tidak	ya
26-30	P	>=6	ya	keram	ya	ya	tidak	80-100	ya	tidak	ya
26-30	L	>=6	ya	keram	ya	ya	tidak	80-100	ya	tidak	ya
>50	P	>=6	ya	keram	ya	ya	ya	80-100	ya	tidak	ya
11-15	L	>=6	ya	kejang	tidak	tidak	ya	80-100	ya	tidak	ya
>50	L	>=6	tidak	kejang	tidak	tidak	ya	80-100	ya	tidak	ya

Sumber: Data rekam medis Puskesmas Waluran

IV. HASIL DAN PEMBAHASAN

4.1 Pohon Keputusan

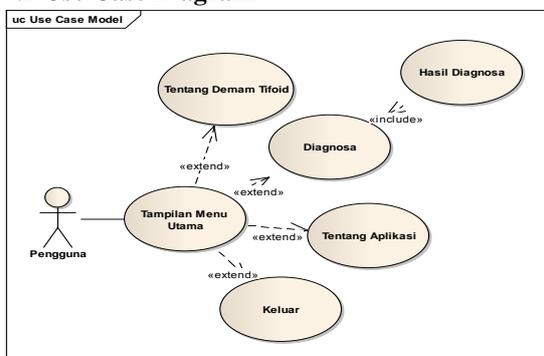
Setelah didapatkan hasil perhitungan entropy dan gain, serta aturan-aturan atau rule maka pohon keputusan yang terbentuk dapat dilihat seperti gambar di bawah ini:



Gambar 2. Pohon Pakar

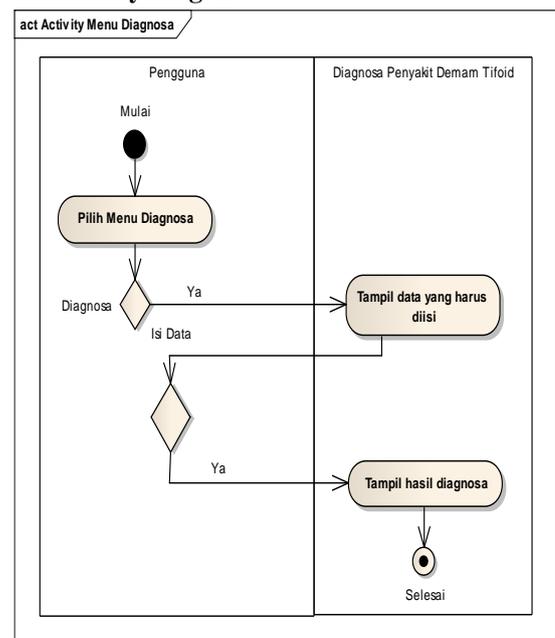
Penerapan algoritma C4.5 untuk diagnosa penyakit demam tifoid berbasis mobile, dirancang agar pengguna dapat mendiagnosa secara dini tentang penyakit demam tifoid melalui perangkat mobile khususnya Android. Berikut ini spesifikasi kebutuhan (*system requirement*) dari penerapan algoritma C4.5 untuk diagnosa penyakit demam tifoid berbasis mobile.

4.2 Use Case Diagram



Gambar 3. Use case Diagram

4.3 Activity Diagram

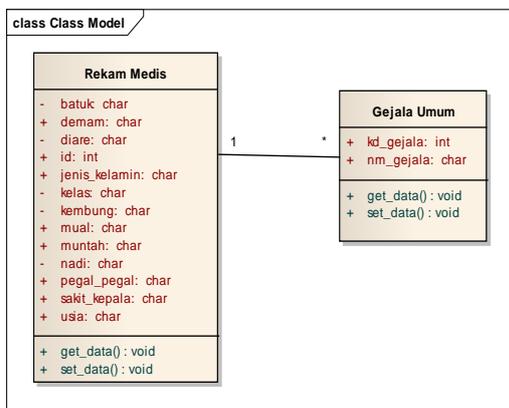


Gambar 4. Activity Diagram

Deskripsi Use Case Diagram Diagnosa

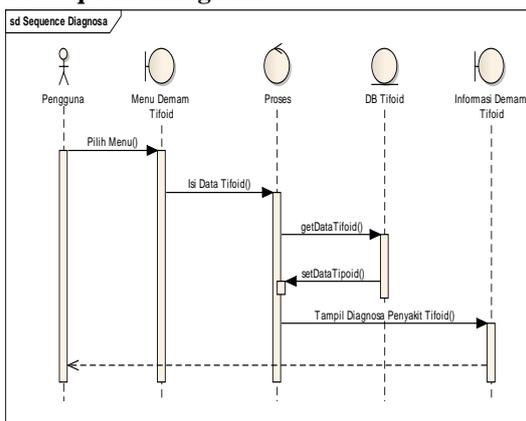
Use Case Name	Diagnosa
Requirment	A2
Goal	Pengguna dapat melakukan Diagnosa
Pre-condition	Pengguna memilih menu utama
Post-condition	Tampil Diagnosa
Failed end condition	Pengguna tidak memilih diagnosa
Primary Actor	Pengguna
Main Flow/Basic Path	Pengguna memilih menu diagnosa
Invariant	-

4.4 Class Diagram



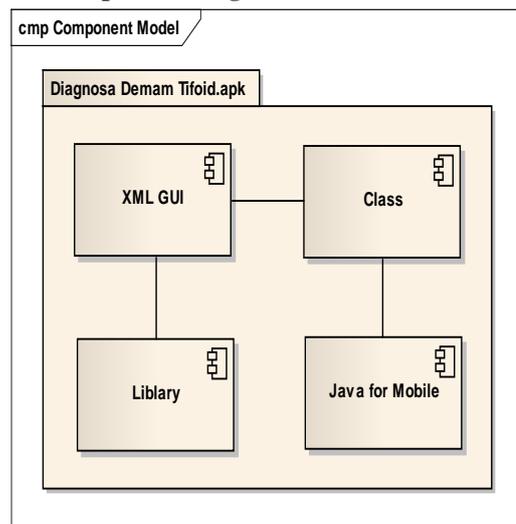
Gambar 5. Class Diagram

4.5 Sequence Diagram



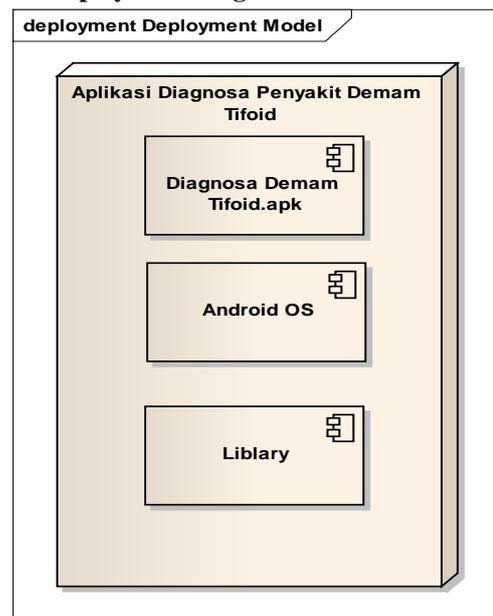
Gambar 6. Sequence Diagram

4.6 Component Diagram



Gambar 7. Component Diagram

4.7 Deployment Diagram



Gambar 8. Deployment Diagram

4.9 User Interface

Setelah proses analisa dan perancangan selesai, berikut adalah implementasi berupa tampilan antar muka dari aplikasi berbasis Mobile untuk diagnosa penyakit demam tifoid ditunjukkan pada Gambar 9.

Gambar 9. Interface aplikasi

## V. KESIMPULAN

Dengan adanya aplikasi ini bisa memudahkan untuk mendiagnosa penyakit demam tifoid sejak dini dari gejala-gejala yang dialami oleh penderita sekalipun seorang penderita tersebut tidak mengerti dengan ilmu kesehatan terutama dalam bidang demam tifoid, dengan adanya aplikasi ini diharapkan dapat meminimalisir biaya periksa yang selama ini terkesan mahal apalagi melakukan pemeriksaan kepada dokter spesialis. Dimanapun dan kapanpun selama pasien menggunakan aplikasi ini akan mampu melakukan diagnose sendiri, *Artificial intelligence* yang khusus mendiagnosa penyakit demam tipoid ini menggunakan algoritma c4.5 yang sesuai dengan penyeleksian gejala, jadi tingkat akurasi dari penyeleksian gejala itu sangat tepat. Aplikasi ini akan sangat membantu pengguna terutama dalam memeriksakan kesehatannya khususnya dalam demam tifoid tanpa terkendala dengan biaya yang harus dikeluarkan. Tidak selalu harus mencari dokter spesialis demam tifoid dikarnakan sudah ada aplikasi yang bisa mendiagnosa demam tifoid ini.

### Saran

Dari pembahasan ini ada beberapa saran untuk pengembangan kedepannya:

- Kedepannya diharapkan adanya input gejala meskipun selama ini untuk gejala-gejala yang

ada pada demam tifoid tidak pernah berubah, tapi kemungkinan dengan ilmu pengetahuan yang semakin maju itu bisa bertambah maka diharapkan adanya input gejala supaya aplikasi ini lebih dinamis dan tentunya ada konektifitas antara administrator aplikasi sama pengguna aplikasi.

- Dalam penyusunan penelitian ini metode yang digunakan adalah algoritma c4.5, untuk kedepannya diharapkan ada yang menggunakan metode lain selain metode algoritma c4.5 atau bahkan bisa optimasi penggabungan metode supaya aplikasi yang dibuat semakin tinggi tingkat akurasinya.
- Untuk pengembangan diharapkan tidak hanya dikembangkan untuk OS Android saja, bisa juga digunakan untuk semua OS misalnya IOS, Windows dan Blackberry yang banyak digunakan pada gadget dimasa ini.
- Demikian saran yang dibuat dalam penyusunan penelitian ini. Semoga semua yang disarankan dapat segera dengan cepat direalisasikan supaya aplikasi yang dibuat bisa lebih disempurnakan dan tepat guna.

## VI. DAFTAR PUSTAKA

- Aprilla C, D., Baskoro, D. A., Ambarwati, L., & Wicaksana, I. W. 2013. Belajar Data Mining dengan Rapid Miner. Jakarta: Open Content model.
- Ginting, Selvia Lorena Br, Wendi Zarman, Ida Hamidah. 2014. Analisa dan Penerapan Algoritma C4.5 dalam Data Mining untuk Memprediksi Masa Studi Mahasiswa Berdasarkan Data Nilai Akademik. ISSN: 1997-911X. Yogyakarta: Porsiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2014.
- Kamagi, David Hartanto dan Seng Hansun. 2014. Implementasi Data Mining dengan Algoritma C4.5 Memprediksi Tingkat Kelulusan Mahasiswa. ISSN: 2085-4552. Tangerang: UTLIMATICS, Vol. VI, No. 1, Juni 2014.
- Kursini dan Emha Taufiq Luthfi. 2009. Algoritma Data Mining. Yogyakarta: C.V ANDI OFFSET.
- Solihah, Nur Baiti. 2010. Identifikasi Penyakit Typus Dengan Analisa Citra Darah Menggunakan Jaringan Syaraf Tiruan. Malang: Jurnal Neutrino Vol. 3 No.1, Oktober 2010.

- [6]. Suhartono, Vincent et al. 2009. Multimedia Pembelajaran Reproduksi Pada Manusia. ISSN. 1414-9999. Jurnal Teknologi Informasi, Volume 5 Nomor 2, Oktober 2009. 695-710.
- [7]. Susanti, R et al. 2012. Aktivitas Reactive Oxygen Species Makrofag Akibat Stimulasi Gel Lidah Buaya Pada Infeksi Salmonella typhimurium. ISSN 0215-9945. Jurnal MIPA 35 (1) (2012).
- [8]. Suwondo, Adi dkk.2013. Algoritma c4.5 berbasis adaboost untuk Prediksi penyakit jantung coroner. ISSN: 2338-3887. Wonosobo: Prosiding Seminar Nasional Teknologi dan Teknopreneur (SNTT) 2013.
- [9]. Tolle, H.2008. Pengantar Sistem Pakar.Diambil dari: <http://www.hermantolle.com/blog/file-sharing/>. (28 Mei 2015)
- [10]. WHO. 2015. Typhoid fever – Uganda.Diambil dari:<http://www.who.int/csr/don/17-march-2015-uganda/en/>. (05 Mei 2015)
- [11]. Yunus, Mahmud, Harry Soekotjo Dahlan dan Purnomo Budi Santoso. 2014. SPK Pemilihan Calon Pendorong Darah Potensial dengan Algoritma C4.5 dan Fuzzy Tahani. Malang: Jurnal EECCIS Vol. 8, No. 1, Juni 2014.
- [12]. Yuwono, Bambang. 2010. Pengembangan sistem pakar pada perangkat mobile