

# EDIKUR 1

*by* Edi Kurniawan

---

**Submission date:** 03-Jul-2018 02:12PM (UTC+0700)

**Submission ID:** 980141508

**File name:** Akurasi\_Penggalian\_Informasi\_Pada\_Data\_Rekam\_Medis\_ICD-10.pdf (303.87K)

**Word count:** 2860

**Character count:** 17102

## Akurasi Penggalan Informasi Pada Data Rekam Medis ICD-10

### Accuracy of Information Extraction on ICD-10 Medical Record Data

2  
Edy Kurniawan

Program Studi Teknik Elektro, Fakultas Teknik Universitas Muhammadiyah Ponorogo  
Jl. Budi Utomo. No. 10 Ponorogo Telp. (0325) 481124, www.umpo.ac.id  
email : [edy@umpo.ac.id](mailto:edy@umpo.ac.id)

#### Abstrak

Banyaknya variabel dalam data rekam medis menyebabkan proses input data oleh tenaga administrasi sering mengalami keterlambatan. Akurasi data sangat diperlukan bagi berbagai pihak yang akan menggunakan data tersebut untuk mencari informasi yang terkandung didalamnya. Sering terjadi adanya data rusak diakibatkan tidak lengkapnya dalam proses input data. Jika data tersebut langsung dianalisa tanpa melewati proses pembersihan, maka akan menyebabkan informasi yang dianalisa tidak bisa dipertanggung jawabkan. Sebagai langkah untuk mengolah atau menggali informasi yang ada didalam tumpukan data rekam medis maka diperlukan pembersihan data dari noise. Hal ini dimaksudkan untuk menjaga keakuratan informasi yang nantinya dihasilkan. Penyelarasan struktur atribut dalam variabel data rekam medis dilakukan satu persatu. Sebab disetiap variabel tersebut mempunyai kode sendiri-sendiri yang tidak bisa dikerjakan secara bersama-sama. Untuk memudahkan dalam pembersihan data maka perlu dilakukan pembersihan variabel yang akan di proses saja, sedangkan variabel yang tidak diperlukan dihilangkan terlebih dahulu. Dari kumpulan data rekam medis yang akan di gali informasinya mempunyai cacat data yang cukup tinggi. Dari total data yang ada hampir 74% data mengalami kerusakan. Kerusakan ini dikarenakan proses inputan pertama yang tidak lengkap maupun adanya singkatan-singkatan, sehingga menyebabkan informasi tidak bisa dipertanggung jawabkan.

**Kata Kunci** : Rekam Medis, Data Mining, Analisa

#### Abstract

The number of variables in the medical record data causes the data input process by administrative personnel often experience delays. Accuracy of data is needed for various parties who will use the data to find information contained therein. Frequently the occurrence of damaged data due to incomplete data input process. If the data is directly analyzed without going through the cleaning process, it will cause the analyzed information can not be justified. As a step to process or dig the information in the stack of medical records data, it is necessary to clean the data from noise. This is intended to maintain the accuracy of information that will be generated. The alignment of the attribute structure in the medical records data variable is done one by one. Because in each variable has its own code that can not be done together. To facilitate the cleaning of data it is necessary to clean the variables that will be processed only, while the

*unnecessary variables are removed first. From a collection of medical record data to be dig the information has a data defect is quite high. Of the total data available almost 74% of data is damaged. This damage is due to incomplete first inputan process or the abbreviations, thus causing information can not be justified.*

**Key-word** : Medical Record, Data Mining, Analysis

## PENDAHULUAN

Tumpukan data rekam medis yang ada di Rumah sakit maupun di dinas kesehatan saat ini mempunyai 2 metode. Pertama adalah rekam medis yang diarsipkan berupa catatan fisik (lembaran kertas), kedua adalah rekam medis yang didokumentasikan dengan data digital. Data-data tersebut sudah ada acuan yang resmi, yaitu ICD dengan berbagai versinya. Acuan rekam medis menggunakan kode ICD tersebut tidak selalu dijadikan pedoman utama. Namun sering kali pedoman tersebut hanya sekedar aturan yang belum sepenuhnya ditaati

<sup>1</sup> Tujuan dan kegunaan rekam medis adalah menunjang tercapainya tertib administrasi dalam rangka upaya peningkatan pelayanan kesehatan. Tanpa didukung suatu sistem pengelolaan rekam medis yang baik dan benar, maka tertib admnistrasi tidak akan berhasil. Kegunaan rekam medis antara lain adalah : [1]

- a. Aspek Administrasi, rekam medis mempunyai nilai administrasi, karena isinya menyangkut tindakan berdasarkan wewenang dan tanggung jawab sebagai tenaga medis dan perawat dalam mencapai tujuan pelayanan kesehatan,
- b. Aspek Medis, Catatan tersebut dipergunakan sebagai dasar untuk merencanakan pengobatan/perawatan yang harus diberikan kepada pasien. Contoh adalah, tindakan yang diberikan dokter kepada pasien atas dasar riwayat penyakit yang dideritanya (sesuai dengan catatan di rekam medis).
- c. Aspek Hukum, menyangkut masalah jaminan kepastian hukum atas dasar

keadilan, dalam rangka usaha menegakkan hukum serta penyediaan bahan tanda bukti untuk menegakkan keadilan.

- d. Aspek penelitian, berkas rekam medis mempunyai nilai penelitian, karena isinya menyangkut data/informasi yang dapat digunakan sebagai aspek penelitian.

Ada beberapa variabel yang menjadi keharusan dalam pencatatan rekam medis. Dari variabel-variabel tersebut harus diisi dengan benar sesuai dengan kondisi pasien. Dari informasi di dalam rekam medis mengandung paling tidak ada 23 variabel. Dari 23 tersebut adalah merupakan informasi pasien disaat mereka masuk ke Rumah sakit maupun disaat mereka perawatan sampai kondisi pulang

## RUMUSAN MASALAH

Data rekam medis yang dikodekan dalam bentuk ICD oleh perakam medis sebuah rumah sakit membantu dalam administrasi data rumah sakit yang bersangkutan. Namun tidak sebatas pelengkap administrasi saja, yang lebih penting adalah sebagai sumber pelacakan riwayat penyakit yang diderita oleh seorang pasien yang akan berobat di rumah sakit tersebut. Disamping itu sering difungsikan juga untuk melakukan pelacakan apabila terjadi kejadian yang aneh pada pasien.

Kumpulan data yang begitu banyak tersebut sebenarnya masih terkandung informasi yang terpendam. Hampir

sebagian besar tumpukan data tersebut hanya sebagai pelengkap administrasi institusi (Rumah Sakit). Kumpulan data rekam medis berupa kode ICD ini dapat diklasifikasikan menjadi empat atribut yang bisa memberikan informasi penting tentang kecenderungan masyarakat menderita sebuah penyakit tertentu. Jenis kelamin, Wilayah, dan umur merupakan field utama yang dijadikan atribut pengolahan data ICD di sebuah Rumah Sakit.

Atribut-atribut tersebut merupakan bagian utama yang terkandung dalam rekam medis. Setiap rumah sakit atau balai pengobatan, informasi tentang rekam medis tersebut bisa dipastikan ada dalam administrasi pasien.

#### **TUJUAN DAN MANFAAT**

Tumpukan data rekam medis yang setiap hari selalu bertambah merupakan gudang informasi yang bisa digali lebih jauh. Untuk mencari informasi yang benar maka diperlukan data yang bersih dan dapat diolah sesuai kebutuhan. Hasil penggalan informasi dari gudang data tersebut bisa dimanfaatkan oleh berbagai pihak (rumah sakit, dinas kesehatan).

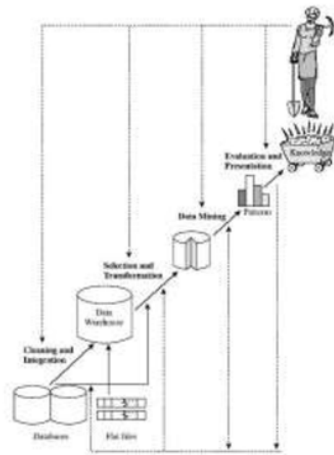
Pada penelitian ini dilakukan pembersihan data rekam medis supaya bisa digunakan untuk mencari keputusan. Jika data bersih dan benar serta tidak ada error maka akan memberikan informasi yang valid kepada pengguna (Rumah sakit, Dinas kesehatan, dan pasien)

#### **KAJIAN PUSTAKA**

Meningkatnya transaksi yang disimpan dengan sistem basis data sekarang ini, maka dibutuhkan proses untuk menangani data tersebut. Proses untuk menangani data tersebut dikenal dengan Knowledge Discovery in Databases (KDD). Knowledge Discovery

in Databases (KDD) merupakan proses nontrivial dalam mengekstraksi data implisit, yang belum diketahui sebelumnya dan berpotensi menjadi informasi yang berguna. Nontrivial karena beberapa pencarian atau inferensi yang dilibatkan bukan merupakan hasil komputasi secara langsung terhadap kuantitas yang telah didefinisikan sebelumnya, seperti komputasi nilai rata-rata sekumpulan bilangan. Pola yang ditemukan harus benar terhadap data baru pada suatu tingkat kepastian tertentu. Pola-pola tersebut harus dapat menjadi gambaran tentang suatu pengetahuan yang secara potensial berguna dan menguntungkan bagi pengguna atau tugas tertentu. Akhirnya, pola-pola tersebut juga harus dapat dipahami dan dimengerti, walaupun terdapat kemungkinan tidak dapat secara langsung dan harus melewati beberapa proses dahulu. Pada aplikasinya, sebenarnya data mining merupakan bagian dari proses KDD. Sebagai komponen dalam KDD, data mining terutama berkaitan dengan ekstraksi dan penghitungan pola-pola dari data yang ditelaah [3].

Data mining adalah suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database dalam jumlah besar. Machine learning adalah suatu area dalam artificial intelligence(AI) atau kecerdasan buatan yang berhubungan dengan pengembangan teknik-teknik yang bisa diprogramkan dan belajar dari data masa lalu.



Gambar 1. Alur Proses KDD [1]

Dalam Alur proses KDD pada gambar 2.1 dapat diuraikan sebagai berikut :

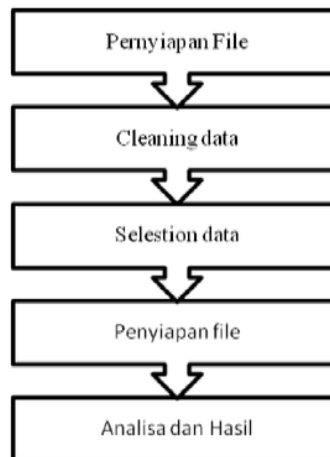
- **Data Cleaning** mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi) [3].
- **Data Integration** merupakan proses menghubungkan antar data yang saling terkait untuk dapat memperkaya informasi yang dihasilkan. Hasil integrasi data sering diwujudkan dalam sebuah data warehouse karena dengan data warehouse, data dikonsolidasikan dengan struktur khusus yang efisien. Selain itu data warehouse juga memungkinkan tipe analisa seperti OLAP.
- **Data Selection** menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada subset variabel atau sampel data, dimana penemuan (discovery) akan dilakukan analisa dan menghasilkan informasi yang berharga.
- **Data Transformation** merupakan proses transformasi data untuk menentukan kualitas dari hasil data mining, sehingga data diubah menjadi bentuk sesuai untuk dilakukan pengolahan.

2

- Data Mining merupakan proses penggalian informasi, ada beberapa teknik data mining yang sudah umum dipakai.
- Pattern Evaluation and Knowledge Presentation dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Presentasi pola yang ditemukan untuk menghasilkan aksi tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisa yang didapat [2].

## METODOLOGI

Untuk melakukan klasifikasi variabel yang ada dalam data rekam medis diperlukan metode yang bisa memudahkan pengguna informasi data rekam medis yang masih asli. Berikut metode yang digunakan untuk membersihkan data-data dan file yang mempunyai noise. Sebab dengan adanya noise data tersebut akan menyebabkan informasi yang diolah bisa tidak lebih akurat.



Gambar 2. Sistematika Penelitian



**HASIL DAN PEMBAHASAN**

Dari 23 atribut yang ada di data rekam medis dengan menggunakan ICD 10 perlu di rekonstruksi ulang sesuai kebutuhan analisa. Hal ini dilakukan supaya memudahkan pihak pengguna dalam menganalisa dari data-data rekam medis tersebut. Untuk itu perlu adanya seleksi-seleksi dan pembersihan atribut yang ada. Struktur dasar data rekam medis tersebut bisa di jabarkan dalam tabel 1.

Tabel 1. Variabel di dalam data rekam medis

Nama Field	Tipe data	Keterangan
Regestrasi	Number Not Null	Nomer regestrasi
Jenis Kelamin	Varchar	
Umur	Number	
Tanggal Masuk	Date	
Tanggal Pulang	Date	
Ruang Rawat Inap	Varchar	
Alamat	Varchar	
Diagnosa	Vachar	Diagnosa penyakit
Kode ICD	Vachar	Kode penyakit hasil diagnosa
Symbol	Vachar	

Type Ex	Vachar	
Identifikasi	Vachar	
Tinta Biru	Vachar	
Dokter yang merawat	Vachar	
Alergi Obat	Vachar	
Jenis Penyakit Menular	Vachar	
Cara masuk	Vachar	
Status Pulang	Vachar	
Kategori PX	Vachar	
Kelengkapan DMK Dokter	Vachar	
Infonent Concent	Vachar	
DPJP	Vachar	
Keterangan	Vachar	

Atribut tersebut tidak semuanya belum tentu digunakan untuk analisa sebuah masalah. Sebagai peneliti hanya membutuhkan beberapa saja sebagai dasar melakukan penelitiannya.

**Pembersihan Data**

Artibut yang digunakan untuk pengolahan data pada penelitian ini diperlukan pembersihan dari noise yang ada di dalamnya, tidak sesuainya isi dari atribut tersebut akan mempersulit pengolahan data dan bahkan akan membuat hasil dari analisa menjadi kurang sempurna. Empat atribut (jenis kelamin, umur, kode ICD, Alamat) yang

perlu dilakukan pembersihan adalah pada "Alamat". Penulisan pada atribut alamat yang dilakukan oleh perekam medis mempunyai keterangan yang sangat detil (contoh : Jl. Budi utomo no. 102 Siman). Dari contoh yang ada diperlukan pengelompokan berdasarkan kecamatan, sehingga dari alamat yang tertulis di rekam medis tersebut dikelompokkan menjadi nama kecamatan.

**Pengelompokan Atribut Alamat**

Pengelompokan atribut "Alamat" tersebut menjadi kelompok kecamatan dimaksudkan untuk memudahkan dalam pembacaan analisa, sehingga hasil yang ada nantinya sudah memberikan informasi berdasarkan lokasi/wilayah kecamatan yang bersangkutan. Pengolahan dan pengelompokan ini dilakukan pada sistem database terpisah dengan database yang ada di rumah sakit, hal ini karena tidak ada izin interkoneksi dengan database rumah sakit. Metode yang digunakan untuk pengolahan awal pada indentitas alamat menjadi kelompok kecamatan ini menggunakan program php dan mysql sebagai databasenya.



Gambar 3. Alur pengelompokan alamat

Alamat yang digunakan pada rekam medis adalah keterangan alamat lengkap "Jl. Budi Utomo No. 102, Siman", hal ini akan menyulitkan pembacaan hasil dari pengolahan awal sistem aplikasi. Pembacaan yang sempurna akan menyebabkan akurasi dari analisa menjadi menurun. Untuk alamat dilakukan standarisasi

penamaan berdasarkan pada kecamatan dari pasien yang bersangkutan. Dari "Jl. Budi Utomo No. 102, Siman" di standarkan menjadi nama kecamatan "Siman".

Tabel 2. Konversi alaman menjadi kecamatan

Alamat	Kecamatan
Jl. Budi Utomo No. 102, Siman	Siman
Mangunsuman, siman	Siman
Jl. Batoro Katong No. 122, Babadan	Babadan
Kertosari, Babadan	Babadan

Data yang ada di triwulan pertama pada rumah sakit aisiyah memuat banyak informasi dari wilayah pasien. Tidak hanya dari kota ponorogo saja, namun ada beberapa yang berasal dari kota disekitar ponorogo (pacitan, magetan, madiun, trenggalek, wonogiri). Pada penelitian ini dilakukan pengolahan data pasien yang berasal dari kota ponorogo saja. Adapun ponorogo sendiri mempunyai 21 kecamatan yang terbagi seperti pada tabel 3.

Tabel 3. Hasil pengelompokan kecamatan

No	Kecamatan
1	Ponorogo / Kota
2	Siman
3	Sumoroto
4	Badegan
5	Sukorejo
6	Jambon
7	Balong
No	Kecamatan
8	Bungkal
9	Slahung
10	Sawoo
11	Sambit
12	Sokoo

**Edy Kurniawan**

Akurasi Penggalan Informasi Pada Data Rekam Medis ICD-10

13	Pudak
14	Pulung
15	Ngebel
16	Jenangan
17	Babadan
18	Ngrayun
19	Sampung
20	Mlarak
21	Jetis

Hasil dari pengelompokan kecamatan yang ada di kabupaten ponorogo tersebut diklasifikasikan berdasarkan wilayah perkotaan, pedesaan, pegunungan. Sebagai acuan pengelompokan tersebut adalah :

Tabel 4. Kelompok wilayah

Kelompok wilayah	Kecamatan
Perkotaan	Ponorogo, Siman, Babadan
Pedesaan	Jenangan, Sumoroto, Sampung, Jambon, Badegan, Sukorejo, Kauman, Slahung, Balong, Bungkal, Jetis, Mlarak,
Pegunungan	Ngrayun, Pudak, Pulung, Ngebel, Sokoo, Sawoo

**Pengelompokan Atribut Umur**

Atribut umur mempunyai range sangat panjang, dari 0 tahun sampai tak terhingga, hal ini karena berhubungan langsung dengan pasien yang terdaftar di rumah sakit yang bersangkutan. Untuk mencari hasil dari analisa menjadi lebih akurat dan dengan banyaknya batasan umur yang ada pada database tersebut akan dianggap

sebagai noise data, maka diperlukan pengelompokan umur.

Tabel 5. Pengelompokan Umur berdasarkan discretize

Umur	Identitas Pengelompokan Umur
≤ 9 th 6 bln	-inf – 9.6
9 th 6 bln – 19 th 2 bln	9.6 – 19.2
19 th 2 bln – 28 th 8 bln	19.2 – 28.8
38 th 4 bln – 48 th	38.4 – 48
48 th – 57 th 6 bln	48 – 57.6
57 th 6 bln – 67 th 2 bln	57.6 – 67.2
67 th 2 bln – 76 th 8 bln	67.2 – 76.8
76 th 8 bln – 86 th 4 bln	76.8 – 86.4
≥ 86 th 4 bln	86.4 – inf

Teknik pengelompokan umur yang kedua adalah mengelompokkan umur berdasarkan jenjang (bayi dan anak-anak, muda dan dewasa, tua). Kategori dari pengelompokan umur berdasarkan jenjang adalah sebagai berikut :

Tabel 6. <sup>1</sup> Pengelompokan umur berdasarkan jenjang menurut WHO

Umur	Jenjang
15 ≤	Bayi dan Anak-anak
15 s/d 50	Muda dan Dewasa
≥ 50	Tua



**Seleksi Data**

Setelah data dilakukan proses cleaning dan integration, data diseleksi dan diambil atribut yang menghasilkan informasi berharga. Atribut data tersebut antara lain : (jenis kelamin, umur, ICD, kecamatan). Data yang diambil merupakan data riwayat pasien saat berobat di rumah sakit yang bersangkutan (RS. Aisyah Ponorogo) pada triwulan pertama di tahun 2011. Syarat lain adalah bahwa data atribut haruslah lebih besar dari 0 dan tidak boleh memiliki nilai kosong atau null.

Tabel 7. Atribut yang di bisa digunakan utuk analisa

Nama Field	Tipe data	Keterangan
Jenis Kelamin	Varchar	
Umur	Number	
Alamat	Varchar	
Kode ICD	Vachar	Kode penyakit hasil diagnosa



Gambar 7. Hasil pembersihan data

**KESIMPULAN DAN SARAN**

Data rekam medis yang diinput oleh **2**haga administrasi Rumah sakit belum

tentu semua inputannya adalah benar. Dari hasil penelitian didapat 74% adalah data noise, dimana data tersebut mempunyai cacat informasi, sehingga jika digunakan untuk penggalan informasi yang ada didalamnya akan menyebabkan informasi tersebut tidak valid.

Sebagai langkah untuk menghilangkan noise dalam analisa data diperlukan penyesuaian atribut di dalam variabel data rekam medis. Penyesuaian ini terutama terletak pada informasi alamat serta informasi kode ICD yang ada. Sebab sering atribut dalam variable tersebut disingkat maupun tidak tertulis secara lengkap.

Sebagai saran adalah jika akan melakukan penggalan data yang terkandung dalam tumpukan data rekam medis, sebaiknya dilakukan pembersihan data dari noise. Kedua adalah, bagi input data maka sebaiknya memasukkan data sesuai dengan standart yang telah diatur oleh WHO maupun Rumah sakit bersangkutan.

**1** **DAFTAR REFERENSI**

ICD-10 (International Classification of Diseases) [http://www.who.int/classifications/icd/en/] diakses Mei 2017

**1** Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques", Second Edition, 2006

Ian H. Witten and Eibe Frank, " Data Mining Practical Machine Learning Tools and Techniques", Morgan Kaufmann Publishers is an imprint of Elsevier., San Francisco, (2005)

Berkhin, Pavel. Survey of Clustering Data Mining Techniques. Accrue Software, Inc. 2002.

**1** Krzysztof J. Cios, Witod Pedrycz, "Data Mining A Knowledge Discovery Approach", 2007

Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, A. Discovering datamining: From

**Edy Kurniawan**

Akurasi Penggalan Informasi Pada Data Rekam Medis ICD-10

---

concept to implementation. New  
Jersey: Prentice Hall. (1997).

Daniel T. Larose, "Discovering Knowledge  
in Data", 2005

# EDIKUR 1

## ORIGINALITY REPORT

19%

SIMILARITY INDEX

14%

INTERNET SOURCES

5%

PUBLICATIONS

7%

STUDENT PAPERS

## PRIMARY SOURCES

1	<a href="http://digilib.its.ac.id">digilib.its.ac.id</a> Internet Source	7%
2	Besty Afriandini. "Analisis Respon Dinamik Gedung Bertingkat Tujuh", Techno (Jurnal Fakultas Teknik, Universitas Muhammadiyah Purwokerto), 2017 Publication	4%
3	<a href="http://docslide.us">docslide.us</a> Internet Source	2%
4	<a href="http://www.lambat.web.id">www.lambat.web.id</a> Internet Source	2%
5	<a href="http://rudyanto62.blogspot.com">rudyanto62.blogspot.com</a> Internet Source	2%
6	Yeti Rusmiati Hasanah, Haryanto Haryanto. "Pengaruh Penambahan Filler Kalsium Karbonat (Caco3) dan Clay Terhadap Sifat Mekanik dan Biodegradable Plastik dari Limbah Tapioka", Techno (Jurnal Fakultas Teknik, Universitas Muhammadiyah Purwokerto), 2017 Publication	1%

---

7

Submitted to Universitas Dian Nuswantoro

Student Paper

1%

---

8

intern-only.blogspot.com

Internet Source

1%

---

---

Exclude quotes      On

Exclude matches      < 20 words

Exclude bibliography      On