PENERAPAN DATA MINING UNTUK PENGELOMPOKAN HASIL DIAGNOSA PENYAKIT PASIEN PENGGUNA BPJS KESEHATAN (STUDI KASUS PADA RSUD UNDATA PALU)

**Implementation of Data Mining For Grouping Results Diagnosis of
Patient Disease of Indonesia's Health Insurance Scheme Users (BPJS
Kesehatan) (Case Study On Regional General Hospital Undata Palu)**

Amriana¹, Yuri Yudhaswana Joeфри², Farah Nabila Meidji³
^{1,2,3}Jurusan Teknologi Informasi, Fakultas Teknik Universitas Tadulako;
e-mail: 1ayazih08@yahoo.com, 2yuri.yudhaswana@untad.ac.id

Received 12 / 03 / 2018, Revised 13 / 04 / 2018, Accepted 22 / 04 / 2018

Abstract

This research was conducted to process medical record data in RSUD Undata of central Sulawesi province, for some BPJS Kesehatan insurance member. Medical record contain information about identity and medical history by patient in Hospital or community health center (Puskesmas). Medical records have disease information by patients encoded according to WHO standart. The code is called ICD (International Classification of Disease) and this research use C4.5 Algorithm as Classification method to process patient medical record which then uses address attributes, gender, age and disease diagnosis (ICD-10). Of the five attributes are groupings then processed into group of age, regional and icd. The result of this studi can find patterns of disease tendency that most suffered by people in a region.

Keywords— *Rekam Medis, Algoritma C4.5, Kode Penyakit internasional (ICD-10);*

1. Pendahuluan

Salah satu rumah sakit rujukan BPJS kesehatan di Kota Palu adalah RSUD Undata yang menghasilkan dan mengumpulkan banyak data rekam medis setiap harinya. Tumpukan data rekam medis digunakan untuk kebutuhan operasional, bahkan tidak jarang juga tumpukan data rekam medis dibiarkan begitu saja tanpa diberdayakan.

Pemanfaatan data secara umum yang digunakan oleh rumah sakit dan juga dinas kesehatan untuk melakukan kebijakan-kebijakan yang berhubungan dengan penyuluhan kesehatan kepada masyarakat hanya sebatas memberikan data jumlah pasien yang berobat dengan penyakit yang di deritanya beserta laporan kepulauan pasien tersebut. Mengenai pola kecenderungan penyakit yang di derita oleh masyarakat dalam sekelompok wilayah belum digali secara maksimal untuk dijadikan acuan dalam upaya penyuluhan kesehatan kepada masyarakat sehingga bisa tepat sasaran.

Hasil dari olahan data ini dapat digunakan oleh RSUD Undata dan Dinas Kesehatan Kota Palu untuk melakukan penyuluhan, pencegahan ataupun peningkatan mutu RSUD Undata dalam melayani pasien pengguna BPJS kesehatan. Oleh karena itu penulis tertarik untuk melakukan penelitian mengenai “Penerapan Data Mining untuk Pengelompokan Hasil Diagnosa Penyakit Pasien Pengguna BPJS kesehatan”.

1.1. Algoritma C4.5

Algoritma C4.5 diperkenalkan oleh Quinlan (1996) sebagai versi perbaikan dari ID3. Dalam ID3, induksi *decision tree* hanya bisa dilakukan pada fitur bertipe kategorikal (nominal atau ordinal), sedangkan tipe numerik (interval atau rasio) tidak dapat digunakan. Perbaikan yang membedakan algoritma C4.5 dan ID3 adalah dapat menangani fitur dengan tipe numerik, melakukan pemotongan (*prunning*) *decision tree* [2]. Proses pada pohon keputusan adalah mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi *rule*. [3] Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Secara umum Algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua kasus memiliki kelas yang sama

Untuk memilih atribut sebagai akar, didasarkan pada nilai *Gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *Gain* digunakan rumus seperti tertera dalam persamaan 1

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

Keterangan :

S = Himpunan kasus

A = Atribut

N = Jumlah partisi atribut A

|S_i| = Jumlah kasus pada partisi ke i

|S| = Jumlah kasus dalam S

Sementara itu, perhitungan nilai *entropy* dapat dilihat pada persamaan 2

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

Keterangan :

S = Himpunan kasus

A = Atribut

N = Jumlah partisi atribut A

P_i = Proporsi dari S_i terhadap S

1.2. Data Mining

Data mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika.[4]

Data mining, sering juga disebut sebagai *knowledge discovery in database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar.[5]

2. Metode Penelitian

2.1. Bahan Penelitian

Dalam penelitian ini, bahan penelitian yang akan digunakan adalah data pasien pengguna BPJS kesehatan pada tahun 2015 di RSUD Undata Provinsi Sulawesi Tengah.

2.2. Alat Penelitian

Implementasi perangkat keras yang digunakan dengan spesifikasi processor Intel®Core™ i3, RAM 2 GB, Monitor dengan resolusi 1366x768 dan sistem operasi 32 bit.

Untuk implementasi perangkat lunak digunakan bahasa pemrograman Visual Basic 2012 sebagai media pembuatan dan perancangan system. MySQL Connector.Net sebagai media penghubung antara bahasa pemrograman dengan database server

2.3. Lokasi Penelitian

Penelitian ini dilakukan di RSUD UNDATA Provinsi Sulawesi Tengah.

3. Hasil dan Pembahasan

3.1. Input Data Training

DATA MINING IMPLEMENTATION TO CLUSTER DISEASE DIAGNOSIS RESULT OF PATIENT WHO USE "BPJS KESEHATAN"

Farah Nabila Melidi
F.551.12.036

MENU

- Data Training
- Data Testing
- Graph of Mined Data
- Tree of Mined Data

Sunday, 04 June 2017
02:59:15 PM

Close

Address:

Gender:

Age: y/o

Diagnosis:

NO	ADDRESS	GENDER	AGE	AGE CLUSTER	REGIONAL CLUSTER
1	J.A.SULAIMES	Male	65	Tua	Pekitikan
2	J.A.YANI	Wanita	50	Tua	Pekitikan
3	MEL KONSUTANA	Wanita	50	Tua	Pekitikan
4	J.A.MULASABANAN	Wanita	15	Muda dan Dewasa	Pekitikan
5	E.KAMBOLUANGA	Male	60	Tua	Pekitikan

Add Save Edit Delete

Gambar 1. Form Input Data Training

Form ini digunakan untuk proses memasukkan data pasien ke dalam basis data sistem. Data pasien yang dimasukkan antara lain alamat, jenis kelamin, umur dan diagnosa penyakit. Data training rekam medis yang sudah ada merupakan data yang diperoleh dari rumah sakit.

3.2. Data Testing

Gambar 2. *Form* Data Testing

Form ini digunakan untuk menampilkan hasil keluaran dalam bentuk *rule* atau aturan dari pengelompokan/klasifikasi berdasarkan kode penyakit internasional (ICD-10) yang di proses menggunakan algoritma C4.5

4. Kesimpulan

Berdasarkan dari penelitian yang telah dilaksanakan dan sudah diuraikan dalam penerapan data mining untuk pengelompokan diagnosa penyakit pasien pengguna BPJS kesehatan pada RSUD Undata Sulawesi tengah, maka penulis dapat menarik kesimpulan sebagai berikut :

1. Algoritma klasifikasi C4.5 dapat menghasilkan model/pola klasifikasi data rekam medis berdasarkan kode penyakit internasional (ICD-10).
2. Mengolah tumpukan data rekam medis dilakukan menggunakan persamaan 2.1 dan persamaan 2.2 sehingga didapat nilai entropi dan gain tertinggi. Nilai ini kemudian digunakan untuk menghasilkan rule dan pohon keputusan yang kemudian dapat menemukan pola klasifikasi penyakit berdasarkan kode penyakit internasional.

5. Saran

Berdasarkan kesimpulan yang telah diuraikan diatas, maka dapat disarankan untuk penelitian selanjutnya mengenai penerapan data mining untuk pengelompokan hasil diagnosa penyakit pasien

diharapkan mengambil lebih banyak parameter yang ada pada data rekam medis guna meningkatkan akurasi algoritma mining sesuai dengan masalah yang akan diselesaikan.

DAFTAR REFERENSI

- [1] Sutarjo, 2014. *Profil kesehatan indonesia*. Kementrian Kesehatan Republik Indonesia.
- [2] Prasetyo Eko, 2014. *Data mining-mengolah data menjadi informasi menggunakan matlab*. Penerbit Andi. Yogyakarta.
- [3] Zalilia, 2007. *Penerapan data mining untuk IDS*. Institut Teknologi Bandung. Bandung.
- [4] Kusrini; & Luthfi, E.T. 2009. *Algoritma Data Mining*. Penerbit Andi. Yogyakarta:
- [5] Santoso, Budi, 2007. "*Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*". Graha Ilmu. Yogyakarta