

# SURVEI PADA PENGGUNAAN TEKNIK DATA MINING PADA BIDANG KESEHATAN DI INDONESIA

**Siti Khomsah**

Program Studi Teknik Informatika  
Fakultas Ilmu Komputer Universitas Alma Ata Yogyakarta  
email: [sitikhomsah5@gmail.com](mailto:sitikhomsah5@gmail.com)

## Abstrak

Layanan kesehatan adalah salah satu sektor pelayanan masyarakat yang berkembang pesat saat ini, sehingga menghasilkan tumpukan data medical record pasien dalam jumlah besar. Tumpukan data tersebut dapat memberikan pengetahuan yang berharga jika diolah dengan cara yang tepat. Knowledge discovery in database (KDD) merupakan rangkaian proses untuk menggali pola pengetahuan yang tersembunyi dalam sekumpulan data yang besar. Teknik data mining adalah satu proses dalam KDD.

Penelitian pada penggunaan teknik data mining dibidang kesehatan khususnya di Indonesia terus berkembang. Data mining dapat diterapkan untuk menemukan pola pengetahuan dari profil pasien dan data riwayat kesehatannya. Pengetahuan yang diperoleh dapat digunakan analisis dan pengambilan keputusan, antara lain untuk memprediksi jenis penyakit, mengetahui pola penyebaran penyakit, dan melihat efektifitas pengobatan. Makalah ini memaparkan berbagai penelitian penggunaan teknik data mining untuk analisis dan prediksi penyakit manusia. Teknik data mining yang dibahas antara lain asosiasi, klasifikasi, clustering, dan prediksi.

Kata kunci : kesehatan, data mining, asosiasi, prediksi, clustering, klasifikasi.

## Abstract

Healthcare is one of the fastest growing fields today, and it is produce enermous quantity data of patients medical record. The data set can provide usefull knowledge if processed in an appropriate way. The discovery of knowledge in the database (KDD) is a series of processes for recording hidden patterns of knowledge in large data sets. Data mining technique is one of a process in KDD.

Researchs of using data mining in health domain, especially in Indonesia continues to grow. Data mining have capabilities to find patterns of knowledge from the patient profile and theirs history of medical record. The knowledge that gained can be used to know the type of disease, the pattern of disease spread, and to know the effectiveness of treatment. This paper describes various studies of using of data mining techniques for the analytic and prediction of human disease. This paper features data mining techniques such as association,clustering, calssification, and prediction.

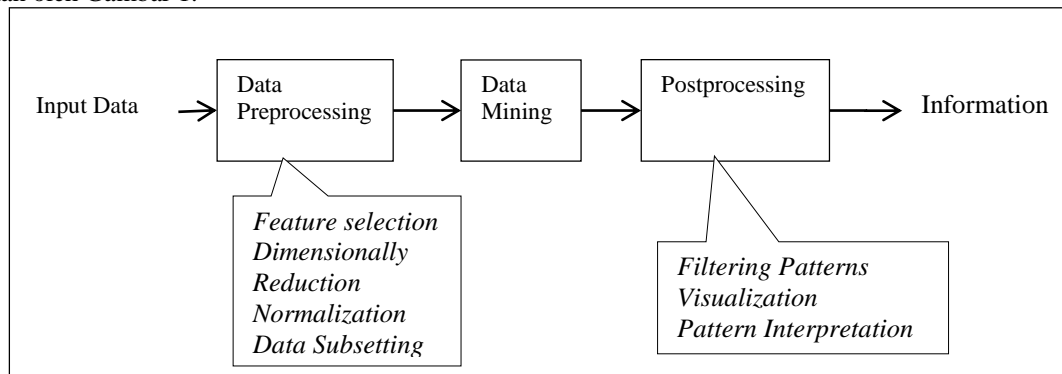
Keyword : healthcare, data mining, association,clustering, calssification,prediction

## 1. PENDAHULUAN

Pemerintah Indonesia memberikan perhatian khusus pada bidang kesehatan dengan mengeluarkan kebijakan untuk penerapan teknologi informasi dan komunikasi pada layanan kesehatan di Indonesia. Kebijakan tersebut salah satunya tertuang dalam rencana strategis 2015 Kementerian Kesehatan, diantaranya rencana strategis penerapan sistem informasi kesehatan dan integrasi data rekam medis rumah sakit rujukan nasional dan regional [1]. Penerapan sistem informasi kesehatan dan integrasi data rekam medis antar rumah sakit akan menghasilkan database rekam medis yang besar. Bagi peneliti, database yang melimpah ini membuka peluang baru pada penelitian analisa data pasien menggunakan berbagai metode yang berkembang saat ini, salah satunya adalah data mining.

Data mining merupakan salah satu tahapan dalam proses penemuan pola pengetahuan dalam database yang besar atau lebih dikenal dengan istilah *knowledge discovery in database* (KDD) [2]. Istilah *knowledge discovery in database* diciptakan oleh Piatetsky-Shapiro pertama kali pada konferensi KDD 1991 [2]. KDD terdiri dari beberapa tahapan dan data mining adalah salah satu tahapannya yang berupa algoritma penemuan yang menghasilkan pola [2]. Data mining merupakan bagian integral KDD yang merupakan proses keseluruhan dari mulai mengubah data mentah menjadi informasi yang bermanfaat [3]. Istilah data mining tidak dapat dipisahkan dengan istilah KDD sehingga pada perkembangannya, istilah data mining lebih sering digunakan untuk menyebut proyek KDD.

KDD merupakan perpaduan berbagai bidang ilmu yaitu kecerdasan buatan, *machine learning*, statistik, basisdata, pengenalan pola, visualisasi data, dan komputasional tingkat tinggi [2]. Tahapan proses KDD ditunjukkan oleh Gambar 1.



Gambar 1. Proses *knowledge discovery* dalam database(Tan dkk.( 2006))

Komponen dalam proses KDD pada Gambar 1 dapat dijelaskan sebagai berikut [4]:

1. Input data adalah dataset yang akan diekstrak pola pengetahuannya. Data ini dapat disimpan dalam berbagai macam format (file, spreadsheet atau tabel) dalam tempat penyimpanan terpusat atau disimpan terdistribusi lintas sumber.
2. *Data preprocessing* bertujuan mengubah data input ke dalam format yang siap di analisis. Tahapan pada *preprocessing* meliputi penggabungan data dari berbagai sumber data, *data cleaning* untuk menghilangkan noise data dan duplikasi data, dan seleksi fitur yang relevan dengan tujuan data mining, *dimensionally reduction* untuk sampel analisis, normalisasi data ke dalam bentuk yang mudah dianalisis, dan data subsetting.
3. *Postprocessing* adalah menentukan hanya hasil yang valid dan berguna saja yang akan disatukan dalam sistem pengambilan keputusan. Pada tahapan ini diantaranya adalah proses *filtering Patterns*, *visualization*, dan *pattern interpretation*.

Keberhasilan ekstraksi pola dari dataset sehingga menghasilkan pengetahuan yang bermanfaat tergantung pada keahlian dan ketrampilan sumber daya manusia yang menggunakan data mining [5]. Menurut Tomar dan Agarwal, CRISP-DM (CRoss Industry Standard Process for Data Mining) menyediakan standar framework proses data mining dimana tersebut dibagi dalam 6 tahap. Tahap pertama adalah pemahaman proses bisnis sedangkan data dalam proses bisnis dikumpulkan dan dianalisis pada tahap kedua. *Pre-processing* data dilakukan pada tahap ke-tiga dan pemodelan dilakukan pada tahap keempat. Tahap kelima mengevaluasi model dan tahap keenam adalah *deployment* model yang ditafsirkan.

## 2. TEKNIK DATA MINING

Teknik data mining dibagi dalam beberapa kelompok sesuai tujuannya yaitu klasifikasi, *clustering*, dan asosiasi.

### 2.1. Klasifikasi

Klasifikasi adalah proses pembelajaran fungsi target  $f$  yang memetakan setiap sekumpulan himpunan atribut  $x$  ke dalam kelas label  $y$  yang belum diketahui sebelumnya seperti yang diilustrasikan Gambar 3.2.



Gambar 2.1 Klasifikasi sebagai pemetaan input himpunan atribut  $x$  ke dalam kelas label  $y$  (Khomseh, 2016)

Data input untuk klasifikasi adalah sekumpulan *record* dikelompokkan sebagai tupel  $(x,y)$  dimana  $x$  adalah atribut dan  $y$  adalah kelas. Atribut dataset dapat berbentuk nilai diskrit atau kontinyu. Sedangkan kelas label harus berbentuk diskrit. Atribut kontinyu adalah atribut yang mengandung nilai angka riil, misalnya suhu, tinggi dan bobot. Sedangkan, atribut diskrit berupa kategori, misalnya kategori *mammal*, *reptile*, *frog*. Atribut diskrit dapat juga berupa nilai biner (benar atau salah). Model klasifikasi dibangun

dengan *data training*. Model yang dibangun digunakan untuk menentukan kelas label yang belum diketahui (Tan dkk (2006)).

Klasifikasi terdiri dari dua tahap yaitu training dan testing. Tahap training bertujuan menghasilkan aturan- aturan IF- THEN berdasarkan sekumpulan data pelatihan. Aturan- aturan yang dihasilkan proses training akan menjadi model klasifikasi. Tahap testing yaitu memprediksi setiap data testing berdasarkan aturan IF- THEN dari data pelatihan. Contoh klasifikasi calon pendonor darah apakah layak (Ya) atau tidak layak (Tidak), **IF** Kel\_Hct=3 **AND** Kel\_Usia=4 **AND** Kel\_Tda=3 **AND** Jkel=1 **AND** Kel\_Hb=1 **AND** Kel\_Bb=3 **THEN** Ya [6].

## 2.2. Clustering

*Clustering* merupakan metode data mining untuk mengelompokkan data yang memiliki kemiripan (*similarity*). Sekumpulan data atau objek akan dikelompokkan menjadi beberapa kelompok berdasarkan kemiripan masing sehingga dalam satu kelompok hanya terdiri dari data atau objek semirip mungkin [7]. Jumlah kelompok dari proses *clustering* dapat ditentukan (satu, dua, atau tiga). Contoh algoritma *clustering* k-means, c-means, fuzzy c-means [7][8]. K atau c menunjukkan jumlah kelompok yang ingin dibentuk. Pada metode k-means, hanya menjadi anggota satu kelompok berdasarkan jarak kemiripan data terhadap pusat kelompok. Sedangkan pada metode c-means misalnya fuzzy c-means, setiap ada dapat menjadi anggota semua kelompok berdasarkan derajat keanggotaannya [8].

## 2.3. Asosiasi

Asosiasi adalah metode *data mining* untuk mengetahui hubungan antara satu item dengan item lainnya dalam database berukuran besar. Hubungan yang tidak terlihat tersebut dapat diekstraksi menjadi pola yang direpresentasikan dalam bentuk himpunan aturan *IF -THEN*. Metode ini menghitung kelayakan aturan dengan menggunakan nilai *support count* dan nilai *confidence* (Tan dkk (2006)). Aturan asosiasi didefinisikan sebagai implikasi  $X \rightarrow Y$ , dimana  $X$  dan  $Y$  adalah disjoint *itemsets* atau  $X \cap Y = \phi$ . Aturan asosiasi yang kuat ditunjukkan oleh nilai *support* dan *confidence*-nya. Dimana *support* menunjukkan seberapa sering aturan berlaku untuk himpunan data yang diteliti, sedangkan *confidence* adalah seberapa sering *item Y* muncul dalam transaksi yang mengandung  $X$ . Bentuk formal definisi *support* seperti persamaan (3) dan *confidence* pada Persamaan (4).

$$Support(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N} \quad (3)$$

$$Confidence(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{X} \quad (4)$$

Nilai *Support* digunakan untuk mengukur tingkat pentingnya suatu aturan. Aturan dengan *support* yang rendah akan banyak dan tidak bermakna atau tidak menguntungkan sehingga aturan dengan *support* rendah dapat dihilangkan. *Confidence* adalah ukuran tingkat kepercayaan terhadap sebuah aturan. Pada aturan  $X \rightarrow Y$ , nilai *confidence* menunjukkan seberapa besar kemungkinan munculnya  $Y$  jika terdapat  $X$ . Artinya seberapa yakin kemunculan item  $X$  pasti akan diikuti juga dengan pembelian item  $Y$ . Seberapa penting suatu aturan ditunjukkan dengan nilai *lift*-nya [9]. Nilai *lift* diperoleh dengan persamaan (5).

$$Lift = \frac{confidence\ rule}{(support\ antecedent)(support\ consequent)} \quad (5)$$

Formula data mining aturan asosiasi pada umumnya dinyatakan sebagai berikut:

- Tentukan himpunan transaksi  $T$ .
- Temukan semua aturan yang mempunyai  $support \geq minsup$  dan  $confidence \geq minconf$  dimana  $minsup$  dan  $minconf$  adalah ambang batas nilai minimal *support* dan nilai minimal *confidence* yang ditentukan.

Secara umum jumlah aturan yang dibangkitkan dari dataset yang mempunyai  $d$  item dituliskan seperti persamaan (6).

$$R = 3^d - 2^{d+1} + 1. \quad (6)$$

Algoritma asosiasi membutuhkan komputasi yang besar sehingga perlu strategi pembangkitan aturan dengan membagi dalam dua proses yaitu:

- Menemukan semua *itemset* yang memenuhi ambang batas (*minsup*) yang ditentukan. *Itemset- itemset* ini disebut dengan *frequent itemsets*.
- Membangkitkan aturan, yaitu mencari aturan yang mempunyai nilai *confidence* yang tinggi.

## 3. PENERAPAN DATA MINING DIBIDANG KESEHATAN DI INDONESIA

Teknik data mining dibagi menjadi beberapa kelompok sesuai dengan tujuan aplikasi. Di sektor kesehatan, teknik data mining dapat digunakan untuk analisa deskripsi atau prediksi suatu penyakit pasien.

### 3.1. Prediksi Penyakit

Algoritma data mining digunakan untuk klasifikasi jenis penyakit, menentukan diagnosa penyakit, dan sistem pendukung keputusan. Penelitian tentang prediksi penyakit yang dikaji dalam paper ini dirangkum pada Tabel 1.

**Tabel 1.** Penerapan algoritma untuk prediksi penyakit

| No. | Jenis Prediksi Penyakit       | Algoritma   |
|-----|-------------------------------|---|
| 1.  | Jantung                       | C.45 , Adaboost[19], Particle Swarm Optimation[27], Neural Network[21] , CBR[54], RBF[33], Naïve Bayes[36][48]                                    |
| 2.  | Karies Gigi                   | Fuzzy C-Means[8][50]  |
| 3.  | Diabetes Melitus (DM) Tipe II | ID3[34], Decision tree J48[60], Naïve Bayes, SVM, Boosted[16], NN MLP[65], SMO ZeroR.[52]   |
| 4.  | Diabetes Mellitus (DM) Tipe I | NN[65], Boosting, ELM[49], SMO ZeroR.[52]   |
| 5.  | Hepatitis/Liver               | Naïve Bayes SVM[38], Fuzzy Dessionion Tree[39], C.45[63][64], QUEST (Quick, Unbiased, And Efficient Statistical Tree), Rule Based Classifier [28] |
| 6.  | Kanker Payudara               | SVM[15], Gain Ratio[45]   |
| 7.  | Kanker Servic                 | SVM, PSO Naïve Bayes[41][46]  |
| 8.  | Kanker Ovarium                | Assosiation Rule (apriori)[9]   |
| 9.  | Demam Berdarah (Dengue)       | SVM K-means, C.45, SVM[58][61], ID3, J48[56]  |
| 10. | Level Kesehatan Pasien        | C.45, Decision Tree [37]  |
| 11. | Parkinson                     | Naïve Bayes RIPPER J45 [42]   |
| 12. | Demam Biasa                   | K-NN[22]  |
| 13. | Hemodialisa                   | C.45[44]  |
| 14. | Penyakit Mata                 | Naïve Bayes[31][62]   |
| 15. | Stroke                        | Improves PSO SVM[47], ECLAT[51], C.45   |
| 16. | TBC (Tuberculosis)            | SVM Naïve Bayes NN Logistik Regression[40], K-NN[35], C.45[32]  |
| 17. | Alzheimer                     | NN[18]  |
| 18. | Thalasemia                    | K-Means[7]  |

### 3.2. Analisa Pola Data Kesehatan

Data mining dapat juga diterapkan untuk menganalisa pola hubungan antara penyakit dan faktor penyebab, pola penyebaran penyakit, perilaku pasien, dan evaluasi kasus medis. Beberapa penelitian tentang analisa pola data terangkum pada Tabel 2.

**Tabel 2.** Penerapan algoritma untuk analisa pola data kesehatan

| No. | Analisa                           | Algoritma       |
|-----|-----------------------------------|-----------------|
| 1.  | Pola pembelian obat               | Apriori[29]     |
| 2.  | Pola penyebaran ISPA              | Apriori[24][59] |
| 3.  | Pola distribusi Penyakit Disentri | Apriori[53]     |
| 4.  | Pola penyakit akibat rokok        | Apriori[57]     |
| 5.  | Kelompok penyakit pasien          | K-Means[43]     |

## 5. KESIMPULAN

Penerapan teknologi informasi untuk mengelola data rekam medis pasien membawa dampak pada melimpahnya database pasien. Penelitian data mining dibidang kesehatan oleh para di Indonesia mulai berkembang meskipun belum banyak penelitian yang dapat terimplementasi langsung di rumah sakit dengan baik. Domain penyakit yang paling banyak menjadi perhatian peneliti antara lain DM, Jantung, Hepatitis, Kanker Payudara, TB, dan Demam Berdarah. Penelitian-penelitian yang dikaji dalam paper ini belum mempunyai rantai keberlanjutan antar satu penelitian dengan lainnya meskipun domainnya sama sehingga perkembangan penelitian satu domain tidak tumbuh cepat. Penelitian menggunakan dataset yang berbeda

meskipun penyakit yang diteliti sama sehingga meskipun algoritma yang sama diterapkan pada domain penyakit yang sama, akurasinya berbeda, dampaknya yaitu efektifitas algoritma untuk prediksi suatu penyakit terlihat bias. Kedepannya akan sangat baik jika dibangun satu sistem dataset penyakit yang dapat digunakan oleh para peneliti secara luas.

## 6. DAFTAR PUSTAKA

- [1] *Rencana Strategis Kementerian Kesehatan Tahun 2015-2019*, [www.depkes.go.id/resources/download/info-publik/Renstra-2015.pdf](http://www.depkes.go.id/resources/download/info-publik/Renstra-2015.pdf), diakses 19 Mei 2017.
- [2] Fayyad, U., Piatetsky-Shapiro, G., dan Smyth, P., *From Data Mining to Knowledge Discovery in Databases*, 1996, American Association for Artificial Intelligence, 0738-4602-1996.
- [3] Tan, P., Steinbach, M., dan Kumar, V., 2006, *Data Mining Concept and Technique*, Morgan Kaufman Publisher, San Francisco.
- [4] Khomsah, S., dan Winarko, E., *Pemanfaatan Algoritma Wit-Tree Dan Hits Untuk Klasifikasi Tingkat Keberhasilan Pemberdayaan Keluarga Miskin (Studi Kasus Kabupaten Bantul)*, 2015, Tesis, Universitas Gadjah Mada.
- [5] Tomar, D. dan Agarwal, S., 2013, *A survey on Data Mining approaches for Healthcare*, International Journal of Bio-Science and Bio-Technology Vol.5, No.5 , pp. 241-266, DOI: 10.14257/ijbsbt.2013.5.5.25.
- [6] Yunus, M., Dahlan, H.S., dan Santoso, P.B, 2014, *SPK Pemilihan Calon Pendorong Darah Potensial dengan Algoritma C4.5 dan Fuzzy Tahani*, Jurnal EECIS Vol. 8, No. 1, p:47-54.
- [7] Sulastri, H. dan Gufroni, A.I., 2017, *Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia*, Jurnal Teknologi dan Sistem Informasi, Vol. 03 No. 02, ISSN (Print): 2460-3465, ISSN (Online) 2476-8812, <https://doi.org/10.25077/TEKNOSI.v3i2.2017.299-305>.
- [8] Havens, T.C., Bezdek, J.C., Leckie, C., Hall, L.O., dan Palaniswami, M., 2012, *Fuzzy c-Means Algorithms for Very Large Data*, IEEE Transactions On Fuzzy Systems, Vol. 20, No. 6.
- [9] Kuswardani, D., Widyanto, M.R., dan Trihandini, I., 2011, *Metode Association Rule Untuk Analisis Citra CT Organ Pasien Kanker Ovarium*, Jurnal Ilmiah KURSUSOR Vol. 6, No. 2, ISSN 0216 – 0544, p: 111-120.
- [10] Rismawan, T., Irawan, A.W., Prabowo, W., dan Kusumadesi, S., 2008, *Sistem Pendukung Keputusan Berbasis Pocket PC Sebagai Penentu Status Gizi Menggunakan Metode Knn (K-Nearest Neighbor)*, Teknoin, Volume 13, Nomor 2, ISSN: 0853-8697.
- [11] Kusumadewi, S., 2009, *Aplikasi Informatika Medis Untuk Penatalaksanaan Diabetes Melitus Secara Terpadu*, Seminar Nasional Aplikasi Teknologi Informasi 2009 (SNATI 2009) ISSN: 1907-5022
- [12] Kuswardani, D., Rahmat Widyanto, M., Trihandini, I., 2011, *Metode Association Rule Untuk Analisis Citra Ct Organ Pasien Kanker Ovarium*, Jurnal Ilmiah Kursor, Vol. 6, No. 2, Juli 2011 ISSN 0216 - 0544
- [13] Bayu Adhi Tama, B., Rodiyatul F. S., Hermansyah, 2011, *An Early Detection Method of Type-2 Diabetes Mellitus in Public Hospital*, TELKOMNIKA, Vol.9, No.2, August 2011, pp. 287~294 ISSN: 1693-6930
- [14] Rizal, A., 2012, *Wrapper Features Subset Selection Pada Ekstraksi Ciri Sinyal EKG Menggunakan Metode Dekomposisi Paket Wavelet*, Jurnal Informatika, Vol. 8, No.2, p: 105 – 114.
- [15] Ayu Novianti, F., dan Wulan Purnami, S., 2012, *Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan Support Vector Machine (SVM) Berdasarkan Hasil Mamografi*, JURNAL SAINS DAN SENI ITS Vol. 1, No. 1, (Sept. 2012) ISSN: 2301-928X
- [16] Wahyu Fernanda, J., dan W. Otok, B., 2012, *Boosting Neural Network dan Boosting Cart Pada Klasifikasi Diabetes Militus Tipe II*, Jurnal Matematika Vol. 2 No. 2, Desember 2012. ISSN : 1693-1394
- [17] Dillak, R. Y., Bintiri, M.B., Pangestuty, D.M., 2012, *Pemanfaatan Algoritma Jaringan Syaraf Tiruan*

- [18] *Levenberg Marquadt Untuk Mendeteksi Penyakit Alzheimer*, Seminar Nasional Informatika 2012 UPN "Veteran" Yogyakarta, ISSN: 1979-2328
- [19] Suwondo, A., Asmarajati, D., Surahman, H., 2013, *Algoritma C4.5 Berbasis Adaboost Untuk Prediksi Penyakit Jantung Koroner*, Prosiding Seminar Nasional Teknologi dan Teknopreneur (SNTT) 2013 ISSN: 2338-3887 FASTIKOM UNSIQ Wonosobo
- [20] Parassa, P., 2013, *Sistem Informasi Berbasis Web Dalam Pengendalian Masalah Kesehatan Pada Dinas Kesehatan Provinsi Sulawesi Utara*, Seminar Nasional SESINDO.
- [21] Rifai, B., 2013, *Algoritma Neural Network Untuk Prediksi Penyakit Jantung*, Vol. IX No.1, Maret 2013 Techno Nusa Mandiri
- [22] Redjeki, Sri., 2013, *Identifikasi Penyakit dengan Gejala Awal Demam Menggunakan K-Nearest Neighbor (K-NN)*, Jurnal Buana Informatika, Volume 4, Nomor 1, Januari 2013: 1-10
- [23] Ade Prasetyo, R., 2013, *Aplikasi Data Mining Association Rules Untuk Menampilkan Informasi*
- [24] *Pola Penyebaran Penyakit Ispa Menggunakan Algoritma Apriori (Studi Kasus di Poliklinik Universitas Dian Nuswantoro Semarang)*, Tugas Akhir Program Studi Teknik Informatika - S1 Fakultas Ilmu Komputer Universitas Dian Nuswantoro Semarang
- [25] Aeni Widiastuti, N., Santosa, S., Supriyanto, C., 2014, *Algoritma Klasifikasi Data Mining Naïve Bayes Berbasis Particle*
- [26] Fathan Hidayatullah, A., Dwi Prasetyo, A., Puspita Sari, D., Pratiwi, I., 2014, *Analisis Kualitas Data dan Klasifikasi Data Pasien Kanker*, Seminar Nasional Informatika Medis (SNIMed) V 2014, Magister Teknik Informatika, Universitas Islam Indonesia
- [27] *Swarm Optimization Untuk Deteksi Penyakit Jantung*, Jurnal Pseudocode, Volume 1 Nomor 1, Februari 2014, ISSN 2355 – 5920
- [28] Widodo, P., 2014, *Rule-Based Classifier Untuk Mendeteksi Penyakit Liver*, Bianglala Informatika Vol . II No 1.
- [29] Nur Salamah, E., Ulinnuha, E., *Analisis Pola Pembelian Obat dan Alat Kesehatan di Klinik Ibu dan Anak Graha Amani dengan Menggunakan Algoritma Apriori*, Jurnal INFORM Vol.xx No.xx, Bulan Tahun, ISSN : 2502-3470
- [30] Kurniawan, H., Fujiati, Saleh, A., 2014, *Analisa Pola Data Penyakit Rumah Sakit Dengan Menerapkan Metode Association Rule Menggunakan Algoritma Apriori*, Seminar Nasional Informatika STMIK Potensi Utama 2014
- [31] Setiawan, W. dan Ratnasari S., 2014, *Sistem Pakar Diagnosis Penyakit Mata Menggunakan naïve bayes Classifier*, Seminar Nasional Sains dan Teknologi 2014, Fakultas Teknik Universitas Muhammadiyah Jakarta , ISSN : 2407-1846.
- [32] Amegia Saputra, R., 2014, *Komparasi Algoritma Klasifikasi Data Mining Untuk Memprediksi Penyakit Tuberculosis (Tb): Studi Kasus Puskesmas Karawang Sukabumi*, Seminar Nasional Inovasi Dan Tren (Snit), Proceedings SNIT 2014
- [33] Setiadi, A., 2014, *Komparasi Algoritma Multi Layer Perceptron Dan Radial Basis Function Untuk Diagnosa Penyakit Jantung*, Jurnal Pilar Nusa Mandiri Vol.X No.1, Maret 2014
- [34] Enggar Tyasti, A., Ispriyanti, D., Hoyyi, A., 2015, *Algoritma Iterative Dichotomiser 3 (ID3) Untuk Mengidentifikasi Data Rekam Medis (Studi Kasus Penyakit Diabetes Mellitus Di Balai Kesehatan Kementerian Perindustrian, Jakarta)*, JURNAL GAUSSIAN, Volume 4, Nomor 2, Tahun 2015, Halaman 237 – 246, ISSN: 2339-2541

- [35] Sari Wardani,R., and Purwanto, 2015, *Model Diagnosis Tuberkulosis Menggunakan K-Nearest Neighbor Berbasis Seleksi Atribut*, The 2nd University Research Coloquium 2015 ISSN 2407-9189
- [36] Sulaksono, J. dan Darsono, 2015, *Sistem Pakar Penentuan Penyakit Gagal Jantung Menggunakan Metode Naive Bayes Classifier*, Seminar Nasional Teknologi Informasi dan Multimedia 2015 STMIK AMIKOM, Yogyakarta, ISSN : 2302-3805, p:3.6-19.
- [37] Himawan, H., S. Simanjuntak, O., Triawan, A., 2015, *Diagnosa Tingkat Kesehatan Pasien Menggunakan Metode Decision Tree*, Seminar Nasional Informatika 2015 (semnasIF 2015) ISSN: 1979-2328
- [38] Wulansari Fridayanthie, E., 2015, *Analisa Data Mining Untuk Prediksi Penyakit Hepatitis Dengan Menggunakan Metode Naive Bayes Dan Support Vector Machine*, JURNAL KHATULISTIWA INFORMATIKA, VOL. 3, NO. 1 JUNI 2015
- [39] Latu Handarko, J., Alamsyah, 2015, *Implementasi Fuzzy Decision Tree untuk Mendiagnosa Penyakit Hepatitis*, UNNES Journal of Mathematics, UJM 4 (2) (2015)
- [40] Hasim Iswanto, M., Erna Permanasari, A., Adi Nugroho, H.,2015, *Pemanfaatan Teknik Data Mining Untuk Diagnosis Penyakit Tuberculosis (Tbc)*, Seminar Nasional Teknologi Informasi dan Multimedia 2015 ISSN : 2302-3805
- [41] Ery Kurniawati, Y., Erna Permanasari, A., Fauziati, S., 2016,*Comparative Study on Data Mining Classification Methods for Cervical Cancer Prediction Using Pap Smear Results*, 2016 1st International Conference on Biomedical Engineering (IBIOMED), Yogyakarta, Indonesia
- [42] Astuti1, T., Ferinanto, T., 2016, *Diagnosis Penyakit Parkinson Berdasarkan Kombinasi Algoritme Data Mining Dan Seleksi Fitur*, Seminar Nasional APTIKOM (SEMNASITIKOM), Lombok, Mataram, 28-29 Oktober 2016
- [43] Khrisna Wardhani, A., 2016,*Implementasi Algoritma K-Means Untuk Pengelompokkan Penyakit Pasien Pada Puskesmas Kajen Pekalongan*, JURNAL TRANSFORMATIKA, Volume 14, Nomor 1, Juli 2016
- [44] Ardian Lazuardi, M., 2016, *Implementasi Data Mining Untuk Memprediksi Pasien Haemodialisa Menggunakan Algoritma C.45*, Skripsi Prodi Teknik Informatika - S1 Fik Udinus 2016
- [45] Aisyah Farahdiba, B., dan Sulisty Nugroho, Y., 2016, *Klasifikasi Kanker Payudara Menggunakan Algoritma Gain Ratio*, Jurnal Teknik Elektro Vol. 8 No. 2 Juli - Desember 2016 ISSN 1411 – 0059
- [46] Hidayatulloh, T., Herliana, A., Arifin, T., 2016, *Klasifikasi Sel Tunggal Pap Smear Berdasarkan Analisis Fitur Berbasis Naïve Bayes Classifier Dan Particle Swarm Optimization*, SWABUMI VOL IV No. 2, September 2016 ISSN 2355-990X
- [47] Cholissodin,I., Farisuddin, F., Santoso, E., 2016, *Klasifikasi Tingkat Resiko Stroke Menggunakan Improved Particle Swarm Optimization dan Support Vector Machine*, Konferensi Nasional Sistem & Informasi 2016 STT Ibnu Sina Batam
- [48] Rohman, A., 2016, *Komporasi Metode Klasifikasi Data Mining Untuk Prediksi Penyakit Jantung*, Jurnal Neo Teknik Vol. 2 No.2, Desember 2016, hal. 21-28
- [49] Junifer Pangaribuan, J., 2016, *Mendiagnosis Penyakit Diabetes Melitus Dengan Menggunakan Metode Extreme Learning Machine*, Jurnal ISD Vol.2 No.2 Juli - Desember 2016 ISSN : 2528-5114
- [50] Astuti, N., Soesanto,O., Kartini, D., 2017, *Algoritma Fuzzy C-Means (FCM) Untuk Penentuan Nilai Center Radial Basis Function (RBF) Pada Klasifikasi Data Penyakit Karies Gigi*, Jurnal Elektronik Nasional Teknologi dan Ilmu Komputer (JENTIK)
- [51] Fernando,R., Anggraini. L., Nazir. A., 2017, *Analisa Keterkaitan Risk Factor Stroke dengan Jenis Stroke yang Diderita Menggunakan Algoritma ECLAT*, Seminar Nasional Teknologi Informasi,



Komunikasi dan Industri (SNTIKI) 9 ISSN (Printed) : 2579-7271 Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau ISSN (Online) : 2579-5406 Pekanbaru, 18-19 Mei 2017

- [52] Ferdian Salim, M., Sugeng, 2017, *Analisis Rekam Medis Pasien Diabetes Mellitus Melalui Implementasi Teknik Data Mining di RSUP Dr. Sardjito Yogyakarta*, Jkesvo (Jurnal Kesehatan Vokasional) Vol. 2 No 2 – November 2017 ISSN 2541-0644
- [53] Kelana Aji, S., Purwanti K.D., *Aplikasi Data Mining Distribusi Penyakit Diare Dengan Metode Association Rule Menggunakan Algoritma A-Priori (Studi Kasus Puskesmas Tambakboyo)*, Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
- [54] Wahyudi, E., Hartati, S., 2017, *Case-Based Reasoning* untuk Diagnosis Penyakit Jantung, IJCCS, Vol.11, No.1, January 2017, pp. 1~10, ISSN: 1978-1520 1
- [55] Yenty Yuliana, O., Intan, R., Irwanto, L., *Data Mining Application For Analyzing Patient Track Record Using Decision Tree Induction Approach*, The 5th International Conference on Information & Communication Technology and Systems
- [56] Ziyah Azkiya1, Z., Indriani, F., Kartika Chandra, H., 2017, *Deteksi Penyakit Dengue Hemorrhagic Fever dengan Pendekatan One Class Classification*, Journal of Information Systems Engineering and Business Intelligence Vol. 3, No. 2, October 2017 e-ISSN 2443-2555
- [57] Fince Tinus Waruwu1, Efori Buulolo1, Eferoni Ndruru1, 2017, *Implementasi Algoritma Apriori Pada Analisa Pola Data Penyakit Manusia Yang Disebabkan Oleh Rokok*, Konferensi Nasional Teknologi Informasi dan Komputer, Volume I, Nomor 1, Oktober 2017 ISSN 2597-4610
- [58] Arrasyid Hasibuan, C., Abdul Mukid, M., Prahutama, A., 2017, *Klasifikasi Diagnosa Penyakit Demam Berdarah Dengue (DBD) Menggunakan Support Vector Machine (SVM) Berbasis GUI Matlab*, Jurnal Gaussian, Volume 6, Nomor 2, Tahun 2017, Halaman 171-180 ISSN: 2339-2541
- [59] Mustakim, A., *Implementasi Data Mining Untuk Identifikasi Pola Penyakit ISPA Dengan Menggunakan Algoritma Apriori (Studi Kasus Di Uptd Puskesmas Bae Kabupaten Kudus)*, Program Studi Sistem Informasi, Fakultas Ilmu Komputer Universitas Dian Nuswantoro Semarang.
- [60] Rodiyatul FS, Adhi Tama, B., *Implementasi Teknik Data Mining Didalam Analisis Penyakit Diabetes Mellitus Tipe II Menggunakan Decision Tree*, Fakultas Ilmu Komputer, Universitas Sriwijaya
- [61] Arrasyid Hasibuan, C., Abdul Mukid, M., Prahutama, A., 2017, *Klasifikasi Diagnosa Penyakit Demam Berdarah Dengue (DBD) Menggunakan Support Vector Machine (Svm) Berbasis Gui Matlab Jurnal Gaussian*, Volume 6, Nomor 2, Tahun 2017, Halaman 171-180 ISSN: 2339-2541
- [62] Taufiq Budiman,Z., Witanti, W., Nursantika, D., 2017, *Klasifikasi Kecenderungan Penyakit Mata Di Jawa Barat Dengan Association Rule Dan Naïve Bayes Classifier*, Seminar Nasional Teknologi Informasi dan Multimedia STMIK AMIKOM Yogyakarta ISSN : 2302-3805
- [63] Wisti Dwi Septiani, 2017, *Komparasi Metode Klasifikasi Data Mining Algoritma C4.5 Dan Naive Bayes Untuk Prediksi Penyakit Hepatitis*, Jurnal Pilar Nusa Mandiri Volume 13 No.1, Maret 2017 ISSN 1978 – 1946
- [64] Ma'arif, F., Arifin, T., 2017, *Optimasi Fitur Menggunakan Backward Elimination Dan Algoritma SVM Untuk Klasifikasi Kanker Payudara*, JURNAL INFORMATIKA, Vol.4 No.1 April 2017, pp. 46~53, ISSN: 2355-6579
- [65] Honainah, *Optimasi Neural Network Untuk Mendeteksi Penyakit Diabetes Dengan Algoritma Genetika*, Jurusan Teknik Informatika, STT Nurul Jadid Probolinggo