

Volume 5 Nomor 2 Mei 2020

# INFORMASI INTERAKTIF

JURNAL INFORMATIKA DAN TEKNOLOGI INFORMASI

**PROGRAM STUDI TEKNIK INFORMATIKA – FAKULTAS TEKNIK -UNIVERSITAS JANABADRA**

SURVEI DAN ANALISIS PENGGUNAAN INTERNET DI UNIVERSITAS JANABADRA

**Ryan Ari Setyawan, Taofiq Krisdiyanto**

EVALUASI ANTARMUKA WEBSITE SMKN DI TANJUNGPANDAN BELITUNG MENGGUNAKAN *USABILITY TESTING*

**Brama Wahyu Prabowo, Bambang Soedijono, Sudarmawan**

PENERAPAN FRAMEWORK COBIT 4.1 DAN BSC PADA AUDIT SISTEM INFORMASI INSTLASI RAWAT INAP RUMAH SAKIT

**Patmawati Hasan, Elvis Pawan, Sariaty H. Y. Bei, Rosiyati M. H. Thamrin**

PENGUKURAN TINGKAT KEMATANGAN TATA KELOLA TEKNOLOGI INFORMASI PADA INSTANSI PEMERINTAHAN XYZ MENGGUNAKAN COBIT 4.1

**Irfan Purwanto, Wing Wahyu Winarno, Asro Nasiri**

EVALUASI TINGKAT KEMATANGAN TEKNOLOGI INFORMASI MENGGUNAKAN FRAMEWORK COBIT 4.1 PADA INSTANSI PEMERINTAHAN ABC

**Joni Saputra, Bambang Soedijono, M. Rudyanto Arief**

*CLUSTERING DATA* NILAI ADAPTIF SISWA MENGGUNAKAN ALGORITMA K-MEANS

**Khoironi, Kusrini, M. Rudyanto Arief**

PENILAIAN TERHADAP TINGKAT KEMATANGAN DALAM PENINGKATAN EFISIENSI BIAYA IT DAN KONTRIBUSINYA PADA KEUNTUNGAN BISNIS (STUDI KASUS: HOTEL TICKLE YOGYAKARTA)

**Selviana Yunita**

PREDIKSI RISIKO KEMATIAN PASIEN STROKE PERDARAHAN DENGAN MENGGUNAKAN TEKNIK KLASIFIKASI *DATA MINING*

**Indarto, Ema Utami, Suwanto Raharjo**



INFORMASI  
INTERAKTIF

Vol. 5

No. 2

Hal. 39 - 91

Yogyakarta  
Mei 2020

ISSN  
2527-5240

**DEWAN EDITORIAL**

- Penerbit** : Program Studi Teknik Informatika Fakultas Teknik Universitas Janabadra
- Ketua Penyunting  
(Editor in Chief)** : Fatsyahrina Fitriastuti, S.Si., M.T. (Universitas Janabadra)
- Penyunting (Editor)** : 1. Selo, S.T., M.T., M.Sc., Ph.D. (Universitas Gajah Mada)  
2. Dr. Kusriani, S.Kom., M.Kom. (Universitas Amikom Yogyakarta)  
3. Jemmy Edwin B, S.Kom., M.Eng. (Universitas Janabadra)  
4. Ryan Ari Setyawan, S.Kom., M.Eng. (Universitas Janabadra)  
5. Yumarlin MZ, S.Kom., M.Pd., M.Kom. (Universitas Janabadra)
- Alamat Redaksi** : Program Studi Teknik Informatika Fakultas Teknik  
Universitas Janabadra  
Jl. Tentara Rakyat Mataram No. 55-57  
Yogyakarta 55231  
Telp./Fax : (0274) 543676  
E-mail: [informasi.interaktif@janabadra.ac.id](mailto:informasi.interaktif@janabadra.ac.id)  
Website : <http://e-journal.janabadra.ac.id/>
- Frekuensi Terbit** : 3 kali setahun

**JURNAL INFORMASI INTERAKTIF** merupakan media komunikasi hasil penelitian, studi kasus, dan ulasan ilmiah bagi ilmuwan dan praktisi dibidang Teknik Informatika. Diterbitkan oleh Program Studi Teknik Informatika Fakultas Teknik Universitas Janabadra di Yogyakarta, tiga kali setahun pada bulan Januari, Mei dan September.

## DAFTAR ISI

	<i>halaman</i>
Survei Dan Analisis Penggunaan Internet di Universitas Janabadra <b>Ryan Ari Setyawan, Taofiq Krisdiyanto</b>	39 - 46
Evaluasi Antarmuka Website SMKN di Tanjungpandan Belitung Menggunakan <i>Usability Testing</i> <b>Brama Wahyu Prabowo, Bambang Soedijono, Sudarmawan</b>	46 - 53
Penerapan Framework Cobit 4.1 dan BSC pada Audit Sistem Informasi Instalasi Rawat Inap Rumah Sakit <b>Patmawati Hasan, Elvis Pawan, Sariaty H. Y. Bei, Rosiyati M. H. Thamrin</b>	54 – 61
Pengukuran Tingkat Kematangan Tata Kelola Teknologi Informasi Pada Instansi Pemerintahan XYZ Menggunakan COBIT 4.1 <b>Irfan Purwanto, Wing Wahyu Winarno, Asro Nasiri</b>	62 - 69
Evaluasi Tingkat Kematangan Teknologi Informasi Menggunakan Framework COBIT 4.1 pada Instansi Pemerintahan ABC <b>Joni Saputra, Bambang Soedijono, M. Rudyanto Arief</b>	70 -75
Clustering Data Nilai Adaptif Siswa Menggunakan Algoritma K-Means <b>Khoironi, Kusri, M. Rudyanto Arief</b>	76 - 79
Penilaian Terhadap Tingkat Kematangan Dalam Peningkatan Efisiensi Biaya IT dan Kontribusinya pada Keuntungan Bisnis (Studi Kasus: Hotel Tickle Yogyakarta) <b>Selviana Yunita</b>	80 - 85
Prediksi Risiko Kematian Pasien Stroke Perdarahan Dengan Menggunakan Teknik Klasifikasi Data Mining <b>Indarto, Ema Utami, Suwanto Raharjo</b>	86 - 91

## **PENGANTAR REDAKSI**

Puji syukur kami panjatkan kehadiran Allah Tuhan Yang Maha Kuasa atas terbitnya JURNAL INFORMASI INTERAKTIF Volume 5, Nomor 2, Edisi Mei 2020. Pada edisi kali ini memuat 8 (delapan) tulisan hasil penelitian dalam bidang teknik informatika.

Harapan kami semoga naskah yang tersaji dalam JURNAL INFORMASI INTERAKTIF edisi Januari tahun 2020 dapat menambah pengetahuan dan wawasan di bidangnya masing-masing dan bagi penulis, jurnal ini diharapkan menjadi salah satu wadah untuk berbagi hasil-hasil penelitian yang telah dilakukan kepada seluruh akademisi maupun masyarakat pada umumnya.

Redaksi

## PREDIKSI RISIKO KEMATIAN PASIEN STROKE PERDARAHAN DENGAN MENGGUNAKAN TEKNIK KLASIFIKASI DATA MINING

Indarto<sup>1</sup>, Ema Utami<sup>2</sup>, Suwanto Raharjo<sup>3</sup>

<sup>1)2)</sup>Magister Teknik Informatika, Fakultas Teknik Informatika,  
Universitas Amikom Yogyakarta

<sup>3)</sup> Teknik Informatika, Fakultas Teknologi Industri, Institut Sains & Teknologi AKPRIND Yogyakarta

Email : <sup>1)</sup>indarglo2019@gmail.com, <sup>2)</sup>ema.u@amikom.ac.id, <sup>3)</sup>wa2n@nrar.net

### ABSTRACT

*Stroke is a major health problem in Indonesia and the world, and the cause of disability and death. Hemorrhagic stroke has higher mortality rate compared to ischemic stroke. The objective of this study is to create a mortality risk prediction model by using Logistic Regression, Support Vector Machine, Random Forest, and Decision Tree C4.5 algorithms based on the data demographics and the clinical data in patients with hemorrhagic stroke. Prediction of outcome of patients of stroke help the physician to determine prognosis, targeted treatments and prepare patients and families. 538 subjects obtained from Hospital Stroke Registry in Yogyakarta. This study uses 10 fold cross validation to evaluate a model. The performance of Decision Tree C4.5 is higher than Logistic Regression, Support Vector Machine, and Random Forest. Prediction accuracy of Decision Tree C4.5 is 90.5%. The use of data mining algorithms able to predict the mortality and functional outcome of patients with hemorrhagic stroke.*

**Keywords:** mortality risk prediction, data mining, hemorrhagic stroke.

### 1. PENDAHULUAN

Penyakit tidak menular seperti penyakit kardiovaskular, kanker, diabetes dan penyakit pernafasan kronik menjadi penyebab kematian utama. Penyakit tidak menular menjadi penyebab lebih dari 70% kematian di seluruh dunia [1]. Sedangkan menurut data World Health Organization tahun 2016, penyakit kardiovaskular menjadi penyebab kematian 31% dari semua kematian di dunia [2]. Pada tahun 2017 stroke menjadi penyebab kematian ketiga di dunia [3]. Riset Kesehatan Dasar yang dilakukan pemerintah Indonesia pada tahun 2007, 2013 dan 2018 menunjukkan peningkatan penyakit tidak menular. Di Indonesia, stroke menjadi penyebab kematian pertama di rumah sakit [4], [5], [6].

Angka kejadian stroke perdarahan lebih sedikit dibandingkan stroke infark tetapi stroke perdarahan mempunyai angka kematian yang lebih tinggi dibanding stroke infark [2], [7], [8]. Salah satu tipe stroke perdarahan yaitu perdarahan intracerebral disebabkan oleh robeknya pembuluh darah di otak yang mengakibatkan perdarahan di parenkim otak. Perdarahan Intracerebral sekitar 9 – 27% stroke di dunia [9]. Hampir 50% pasien dengan ICH Hipertensi meninggal [1].

Saat ini masih jarang penelitian yang dilakukan di Indonesia untuk memprediksi risiko kematian pada pasien stroke perdarahan. Karena tingginya angka kematian pada pasien stroke perdarahan maka kami melakukan penelitian ini. Tujuan penelitian ini adalah memprediksi risiko kematian pasien stroke perdarahan dengan menggunakan data mining. Prediksi risiko kematian akan membantu dokter sejak awal dalam menentukan prognosis, menentukan target terapi dan memberikan edukasi serta mempersiapkan pasien dan keluarga [10].

I. C. Hostettler menggunakan algoritme decision tree untuk memprediksi kematian, luaran fungsional dan ketergantungan pada *ventriculoperitoneal shunt*. Akurasi prediksi kelangsungan hidup pada hari pertama adalah 75.2%, akurasi prediksi untuk luaran fungsional adalah 71.1% pada data latih dan 66.7% pada data uji [11]. W. Y. Lin menggunakan tiga algoritme data mining yaitu Regresi Logistik, *Random Forest*, and *Support Vector Machine* untuk memprediksi luaran fungsional. Performa Regresi Logistik dan *Random Forest* lebih tinggi dibanding *Support Vector Machine* [12]. H. L. Wang mendapatkan bahwa *Random Forest* adalah algoritma terbaik untuk memprediksi luaran fungsional stroke. Akurasi

prediksi luaran fungsional pada bulan pertama adalah 83.1% dan untuk bulan ke enam sebesar 83.9% [13].

## 2. TINJAUAN PUSTAKA

Data mining adalah suatu proses mencari hubungan, pola dan kecenderungan yang mempunyai makna dengan memilah - milah data yang besar yang tersimpan di repositori menggunakan teknologi pengenalan pola, matematika dan teknik statistik [14], [15]. Pada intinya data mining membuat model yang menyajikan suatu pola yang ringkas dengan menggunakan data yang lalu dan menerapkan model itu dengan data yang baru [16]. Obermeyer menyatakan bahwa algoritme akan mengubah bidang kesehatan. Ia percaya bahwa perhatian akan bergeser dari metode statistik ke metode *machine learning* [17].

### 2.1 Regresi Logistik

Regresi Logistik (RL) bertujuan untuk mengetahui hubungan antara variabel bebas dengan satu atau lebih variabel terikat dengan menggunakan angka kemungkinan sebagai nilai prediksi dari variabel terikat. Regresi logistik multinomial digunakan pada saat variabel terikat adalah kategori dengan lebih dari dua kategori [18].

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (1)$$

Dimana x adalah nilai probabilitas dari  $0 \leq \pi(x) \leq 1$ . Dengan mentransformasikan  $\pi(x)$  dengan transformasi logit  $g(x)$ , dimana :

$$g(x) = \ln \left( \frac{\pi(x)}{1 - \pi(x)} \right) \quad (2)$$

Maka didapatkan bentuk logit :

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (3)$$

### 2.2 Support Vector Machine

*Support Vector Machine* (SVM) adalah metode klasifikasi yang menggunakan data terpilih untuk membentuk model. Data - data

yang berkontribusi membentuk model tersebut disebut *support vector*. Konsep SVM adalah usaha memperoleh batas keputusan atau *hyperplane* terbaik yang bertujuan untuk memisahkan dua kelas data pada *input space* [19].

### 2.3 Random Forest

*Random Forest* (RF) merupakan gabungan pohon keputusan sedemikian rupa sehingga setiap pohon bergantung pada nilai - nilai *vector* acak yang disampling secara *independen* dan sebaran yang sama. Kemampuan RF terletak pada seleksi fitur yang dilakukan dengan acak untuk memilah-milah setiap simpul atau node untuk menghasilkan tingkat kesalahan yang relative rendah [20].

### 2.4 Decision Tree C4.5

*Decision tree* mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan [15], [20], [21]. Algoritma C4.5 adalah pengembangan dari algoritma ID3 dan menjadi *benchmark* bagi algoritma *supervised learning* yang baru [21].

Metode *Decision tree* C4.5 digunakan untuk data kategorial maupun numerik. *Decision tree* C4.5 mengevaluasi semua atribut menggunakan ukuran *impurity Gain Ratio*. Data dalam *Decision tree* dinyatakan dalam bentuk tabel dengan atribut dan *record*. Proses pada *Decision Tree* adalah mengubah bentuk data menjadi model pohon, mengubah model pohon menjadi *rule* dan menyederhanakan *rule* [15], [20].

Untuk menghitung gain digunakan rumus sebagai berikut :

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{S_i}{S} * \text{Entropy}(S_i) \quad (4)$$

Keterangan :

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

Sedangkan untuk menghitung entropy adalah :

$$\text{Entropy (S)} = \sum_{i=1}^n -p_i * \log_2 p_i \quad (5)$$

S : himpunan kasus

n : jumlah partisi S

$p_i$  : proporsi dari  $S_i$  terhadap S

## 2.5 Stroke

Stroke adalah suatu kondisi terhentinya aliran darah yang disebabkan oleh sumbatan pembuluh darah atau perdarahan yang sifatnya mendadak dan megakibatkan gangguan fungsi otak [22]. Perubahan aliran darah di otak menyebabkan gangguan pasokan oksigen dan zat makanan ke otak yang mengakibatkan kematian sel saraf. Stroke ada 2 jenis, yaitu stroke perdarahan dan stroke sumbatan. Stroke sumbatan terjadi bila pembuluh darah otak mengalami sumbatan. Sedangkan stroke perdarahan terjadi karena pecahnya pembuluh darah di otak.

Gejala dan tanda yang muncul berbeda - beda tergantung bagian otak yang terpengaruh. Gejala dan tanda yang muncul antara lain kelemahan tubuh sisi kanan atau kiri, kelemahan lengan tangan atau tungkai kaki, wajah perot, gangguan bicara, pusing berputar, sakit kepala, penurunan kesadaran dan lain-lain [5].

Stroke terjadi karena adanya faktor risiko. Faktor risiko stroke adalah tekanan darah tinggi, kencing manis, gangguan kadar lemak tubuh atau dislipidemi, merokok, obesitas, fibrilasi atrium, usia tua, jenis kelamin, ras, riwayat keluarga dan lain-lain. Semakin banyak faktor risiko pada seseorang maka semakin besar risiko terjadinya stroke.

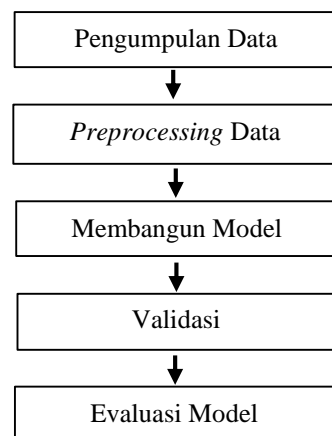
## 3. METODOLOGI PENELITIAN

Pada bagian metodologi menjelaskan cara pengumpulan data, pemilihan atribut penelitian, algoritme data mining yang digunakan untuk membangun model dan kriteria evaluasi seperti yang ditunjukkan pada gambar 1.

### 3.1 Pengumpulan Data

Data diperoleh dari Stroke Register Rumah Sakit di Yogyakarta dari tanggal 1 Januari 2017 sampai 31 Desember 2018. Kami minta pertimbangan dokter spesialis saraf dalam memilih data pasien maupun atribut yang akan

diolah. Subyek penelitian dibagi menjadi dua kelas atau label, yaitu kelas yang meninggal dan kelas yang hidup.



Gambar 1. Diagram Alur Penelitian

### 3.2 Preprocessing Data

Data penelitian meliputi data demografik dan data klinis pasien. Data demografi terdiri dari atribut umur, jenis kelamin. Sedangkan data klinis terdiri dari atribut onset stroke, jenis ulangan, tipe stroke yang sudah dikonfirmasi dengan pemeriksaan radiologi, lama rawat inap atau *length of stay* (LOS), gejala dan tanda stroke yaitu penurunan kesadaran, afasia, kelemahan anggota gerak, disartria, wajah perot dan faktor risiko stroke yaitu hipertensi, dislipidemia, *Ischemic Heart Disease* (IHD), Atrial Fibrillation (AF) serta komplikasi seperti Infeksi Saluran Kencing (ISK), pneumonia, Gastrointestinal (GI) *bleeding* dan dekubitus.

### 3.3 Algoritme Data Mining

Semua algoritme data mining yang dilakukan pada penelitian ini diolah dengan menggunakan perangkat lunak Waikato Environment for Knowledge Analysis (WEKA) versi 3.8, Machine Learning Group, University of Waikato, Hamilton, New Zealand. Piranti lunak WEKA berisi kumpulan metode *machine learning* dan *preprocessing* [23], [24].

Penelitian ini menggunakan *ten fold cross validation* dan *confusion matrix* untuk menguji model. Sampel asli dipartisi secara acak menjadi sepuluh subsampel dengan ukuran yang kira-kira sama. Satu dari sepuluh subsampel digunakan sebagai set data uji untuk menguji model, dan sembilan subsampel sisanya digunakan sebagai set data latih. Teknik validasi

silang kemudian diulang sepuluh kali dengan satu dari sepuluh sub-sampel yang digunakan secara berurutan untuk setiap validasi. Sepuluh hasil dari setiap validasi yang diulang kemudian dirata-rata untuk menghasilkan estimasi akhir. Secara umum *ten fold cross validation* direkomendasikan untuk memperkirakan akurasi sebagai akibat dari bias dan varietas yang relatif rendah [13], [21].

*Confusion matrix* adalah matriks klasifikasi yang salah dan benar yang dibuat oleh algoritma. Kolom sebagai klasifikasi yang diprediksi dan baris sebagai klasifikasi yang benar [14]. Dalam *confusion matrix*, *true negative* (TN) adalah jumlah sampel negatif yang dengan benar diklasifikasikan sebagai negatif, *true positive* (TP) adalah jumlah sampel positif yang diklasifikasikan yang dengan benar sebagai positif. Jadi TN dan TP adalah sampel yang diklasifikasikan dengan benar. *False negative* (FN) adalah jumlah sampel positif yang salah diklasifikasikan sebagai negatif dan *false positive* (FP) adalah jumlah sampel negatif yang salah diklasifikasikan sebagai positif [23]. Akurasi adalah ukuran kinerja yang umumnya dikaitkan dengan algoritma pembelajaran mesin dan didefinisikan sebagai

$$\text{Akurasi} = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

$$\text{Presisi} = \frac{TP}{TP + FP} \quad (7)$$

$$\text{Recall} = \frac{P}{TP + FN} \quad (8)$$

$$F\text{-measure} = \frac{2 \times \text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (9)$$

## 4. HASIL DAN PEMBAHASAN

### 4.1 Hasil

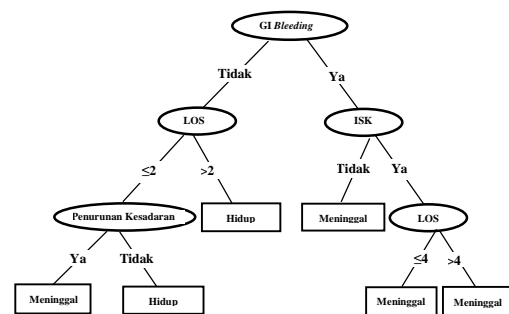
Hasil performa dari beberapa model klasifikasi menunjukkan *Decision tree* C4.5 mendapatkan akurasi tertinggi, yaitu 90.5% diikuti oleh RF sebesar 89.6%, RL sebesar 88.8% dan SVM sebesar 86.6% (Tabel 1).

Tabel 1. Performa Klasifikasi

Performa	RL	SVM	RF	C4.5
Akurasi	88.8%	86.6%	89.6%	90.5%
Presisi	88.6%	86.4%	89.4%	90.4%
Recall	88.8%	86.6%	89.6%	90.5%
F-measure	88.5%	85.9%	89.4%	90.4%

Sedangkan waktu yang dibutuhkan untuk membangun model adalah LR 0.03 detik, SVM 0.05 detik, RF 0.14 detik dan *Decision tree* C4.5 0 detik

### 4.2 Pembahasan



Gambar 2. Pohon Keputusan C4.5

*Intra Cerebral Hemorrhage* (ICH) score digunakan secara luas di praktek kedokteran untuk menentukan risiko kematian pada stroke perdarahan. Faktor penentunya adalah nilai rendah Glasgow Coma Scale, usia tua, lokasi perdarahan dan volume perdarahan otak. Pada gambar Pohon Keputusan C4.5 menunjukkan bahwa *GI Bleeding* atau perdarahan saluran cerna menjadi faktor penentu pertama terjadinya risiko kematian pada pasien stroke perdarahan. Dengan demikian penggunaan data mining untuk mengolah data stroke perdarahan mampu mengidentifikasi faktor penentu yang sebelumnya kurang mendapatkan perhatian. Sehingga temuan ini dapat meningkatkan penatalaksanaan pasien dengan stroke perdarahan.

Pada penelitian ini, nilai tertinggi akurasi dihasilkan oleh algoritma *Decision tree* C4.5. Akurasi yang dihasilkan lebih tinggi dibandingkan hasil penelitian yang dilakukan oleh I. C. Hostettler. Kemungkinan disebabkan oleh set data pada penelitian Hostettler mempunyai atribut yang kosong atau *missing value* [11]. Sedangkan dibandingkan dengan penelitian H. L. Wang yang menyatakan



*Random Forest* merupakan model prediksi terbaik untuk prediksi luaran fungsional stroke perdarahan, maka dalam penelitian ini *Random Forest* mempunyai nilai tertinggi kedua sesudah *Decision tree C4.5*.

Secara umum *Decision trees* adalah metode yang handal dan efektif untuk membuat keputusan yang memberikan akurasi yang tinggi dengan representasi yang sederhana untuk menghasilkan pengetahuan dan selama ini sudah dipergunakan di berbagai bagian dari pengambilan keputusan di bidang kesehatan [25]. Sehingga pada penelitian ini *Decision tree C4.5* dan *Random Forest* menghasilkan akurasi yang lebih baik dibandingkan model prediksi yang lain.

## 5. KESIMPULAN

Hasil dari penelitian ini menunjukkan bahwa data mining dapat memprediksi risiko kematian pada pasien stroke perdarahan dan menemukan faktor penentu risiko kematian pada pasien stroke perdarahan. Faktor penentu risiko kematian pada pasien stroke perdarahan adalah perdarahan saluran cerna. Prediksi risiko kematian pada pasien stroke perdarahan dengan menggunakan data mining masih jarang dilakukan dan sangat besar kemungkinan untuk digali sehingga menghasilkan model prediksi terbaik.

Penelitian selanjutnya diharapkan dapat menggunakan jumlah data pasien yang lebih banyak dan dapat menggunakan proses data reduction untuk mempercepat proses data mining.

## DAFTAR PUSTAKA

- [1] World Health Organization, Noncommunicable Diseases Progress Monitor 2020, WHO, Switzerland, 2020
- [2] World Health Organization, Noncommunicable Diseases Country Profiles 2018., WHO, Switzerland, 2018.
- [3] IHME, Finding from the Global Burde of Disease Study 2017, Institute for Health Metrics and Evaluation (IHME), United State of America, 2018
- [4] Kementerian Kesehatan Republik Indonesia, Profil Kesehatan Indonesia Tahun 2018, Kementerian Kesehatan RI, Jakarta, 2018
- [5] R. T. Pinzon, *Awas Stroke*, Betha Grafika, Yogyakarta, 2016
- [6] Y. Roohani, N. Saji, P. Madhyastha, "Predicting Language Recovery after Stroke with Convolutional Networks on Sticked MRI," Arxiv : 1811v1 (cs.CV) 26 Nov 2018
- [7] K. K. Andersen, et. al., "Hemorrhagic and Ischemic Strokes Compared Stroke Severity, Mortality, and Risk Factors," *AHA Journals. Amerika Serikat*, 2009
- [8] I. Setyopranoto, et. al., "Prevalence of Stroke and Associated Risk Factors in Sleman District of Yogyakarta Special Region, Indonesia," *Hindawi Stroke Research and Treatment*, Volume 2019.
- [9] Steiner, et al., "European Stroke Organisation (ESO) Guidelines For The Management Of Spontaneous Intracerebral Hemorrhage," *World Stroke Organization*, 2014.
- [10] M. D. Smith, N. Shobha, et. al., "A Risk Score for In-Hospital Death in Patients Admitted With Ischemic or Hemorrhagic Stroke," *Journal of the American Heart Association* 2013.
- [11] I. C. Hostettler, C. Murol, J. K. Richter, J. Schmid, et al, " Decision Tree Analysis in Subaracnoid Hemorrhage : Prediction of Outcome Parameters During The Course of Aneurysmal Subarachnoid Hemorrhage Using Decision Tree Analysis" *J Neurosurg* 129 : 1499-1510, 2018
- [12] W. Y. Lin, C. H. Chen, Y. J. Tseng, Y. T. Tsai, et al, "Predicting Post-Stroke Activities of Daily Living Through A Machine Learning Based Approach on Initiating Rehabilitation" *International Journal of Medical Informatics* 111 : 159-164, 2018
- [13] H. L. Wang, W. Y. Hsu, M. H. Lee, H. H. Weng, et al. , " Automatic Machine-Learning-Based Outcome Prediction in Patients With Primary Intracerebral Hemorrhage" *Fontiers in Neurology*, Volume 10, article 910, August 2019
- [14] D.T. Larose, *Discovering Knowledge in Data : An Introduction to Data mining*, John Wiley & Sons Inc, Hoboken, New Jersey, pp 204, 2005
- [15] Kusriani, E.T. Luthfi, *Algoritma Data Mining*, Penerbit Andi, Yogyakarta, 2009.
- [16] Hornick, F. Mark, "Java Data Mining : Strategy, Standard and Practice : A Practical Guide for Architecture, Design, and Implementation," *Morgan Kaufmann Publisher*, San Frannisco 2007.
- [17] Obermeyer, "Predicting the Future-Big Data, Machine Learning and Clinical Medicine," *New England Journal Medicine* 375;13, Sept 29, 2016
- [18] D. W. Hosmer, S. Lemeshow, *Applied Logistic Regression*, A Wiley-Interscience Publication, United States of America, 2000
- [19] E. Prasetyo, *Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab*, Penerbit Andi, Yogyakarta, 2014
- [20] Suyanto, *Machine Learning Tingkat Dasar dan Lanjut*, Penerbit Informatika, Bandung, 2018

- [21] J. Han, M. Kamber, Data Mining Concepts and Techniques, Morgan Kaufmann Publisher, USA, 2012
- [22] A. Khosravi, H. Amirifard, H., “Frequency and Causes of Mortality in Patients with Stroke Referred to Zahedan Hospital in 2016,” International Journal of Research in Medical Sciences, Vol 6 (3) : 743-746, March 2018
- [23] R. R. Bouckaert, E. Frank, M. Hall, R. Kirkby, P. Reutemann, A. Seewald, D. Scuse, Weka Manual for Version 3-8-3, The University of Waikato, September 4, 2018
- [24] I. H. Witte, F. Eibbe, M. A. Hall, Data Mining Practical Machine Learning Tools and Techniques, Morgan Kaufmann Publisher, USA, 2011
- [25] V. Podgorelec, P. Kokol, B. Stiglic, I. Rozman, “Decision Trees : an Overview and Their Use in Medicine,” Journal of Medical Systems, Kluwer Academic/Plenum Press, vol 26, Num 5, pp 445 – 463, October 2002