

# **Aplikasi Data Mining Untuk Mendeteksi Penyakit Hepatitis Menggunakan Metode Support Vector Machine (SVM) Polynomial (Studi Kasus : Data Pasien Hati India)**

**Lusy Octoria Sitanggang<sup>\*1)</sup>, Nurdin Bahtiar<sup>\*2)</sup>**

**\*\*Jurusan Ilmu Komputer/Informatika, Fakultas Sains dan Matematika,  
Universitas Diponegoro**

**<sup>1)</sup>sitangganglusi@gmail.com, <sup>2)</sup>nurdinbahtiar@gmail.com**

## **Abstrak**

*Perkembangan teknologi informasi saat ini telah merambah ke berbagai sektor termasuk sektor kesehatan. Dalam sektor kesehatan, perkembangan ilmu kedokteran mengalami kemajuan yang sangat pesat yang ditandai dengan ditemukannya penyakit-penyakit baru yang belum teridentifikasi sebelumnya. Salah satu penyakit yang berkembang saat ini yaitu penyakit pada organ hati. Salah satunya adalah penyakit Hepatitis. Diagnosa awal penyakit ini setelah memperhatikan gejala adalah melakukan tes fungsi hati yang biasa disebut LFT (Liver Function Test). Dengan beberapa atribut dari hasil pemeriksaan LFT tersebut akan mudah digunakan untuk menganalisis penyakit tersebut. Salah satu teknologi kecerdasan buatan yang dapat digunakan untuk menganalisis penyakit tersebut adalah machine learning. Machine Learning telah banyak digunakan dalam bidang medis yaitu untuk menganalisa dataset medis. Salah satu metode machine learning adalah Support Vector Machine (SVM). Ciri dari metode ini adalah menemukan fungsi pemisah (klasifier) yang optimal yang bisa memisahkan dua set data dari dua kelas yang berbeda. Data yang digunakan pada penelitian ini didapat dari UCI (Universitas California Invene) Machine Learning Repository yang berjumlah 579 data pasien. Dalam dataset tersebut, terdapat 11 atribut yang akan digunakan untuk mendiagnosis penyakit dengan menggunakan metode support vector machine polynomial. Dengan menggunakan Cross validation, menggunakan pengujian pada 10 atribut data Pasien Hati India memiliki rata-rata nilai akurasi sebesar 87.65%.*

**Kata Kunci :** *Support Vector Machine, Hepatitis, Machine Learning, Cross Validation, polynomial*

## **Abstract**

*The development of information technology has now penetrated into various sectors including the health sector. In the health sector, the development of medical science has progressed very rapidly marked by the discovery of new diseases that have not been identified before. One of the diseases that develops now is the disease in the liver. One of them is hepatitis. The initial diagnosis of this disease after paying attention to symptoms is to perform a liver function test which is commonly called the LFT (Liver Function Test). With some attributes from the results of the LFT examination it will be easy to use to analyze the disease. One of the artificial intelligence technologies that can be used to analyze the disease is machine learning. Machine Learning has been widely used in the medical field to analyze medical datasets. One machine learning method is Support Vector Machine (SVM). The characteristic of this method is finding the optimal separator function (classifier) that can separate two data sets from two different classes. The data used in this research was obtained from UCI (University of California Invene) Machine*

*Learning Repository, which amounted to 579 patient data. In the dataset, there are 11 attributes that will be used to diagnose the disease using the support vector machine polynomial method. By using the dataset, there are 11 attributes that will be used to diagnose the disease using the support vector machine polynomial method. By using Cross validation, using tests on 10 attributes of Indian Heart Patients has an average accuracy value of 87.65%.*

**Keywords** : *Support Vector Machine, Hepatitis, Machine Learning, Cross Validation, polynomial*

## 1. PENDAHULUAN

Pada era globalisasi saat ini, Teknologi Informasi turut berkembang sejalan dengan perkembangan peradaban manusia. Perkembangan teknologi informasi ini meliputi perkembangan infrastruktur, seperti *hardware, software*, teknologi penyimpanan data (*storage*), dan teknologi. Saat ini perkembangan teknologi informasi telah merambah ke berbagai sektor termasuk kesehatan.

Dalam sektor kesehatan, perkembangan ilmu kedokteran mengalami kemajuan yang sangat pesat yang ditandai dengan ditemukannya penyakit-penyakit baru yang belum teridentifikasi sebelumnya. Salah satu penyakit yang berkembang saat ini yaitu penyakit pada organ hati. Salah satunya adalah penyakit hepatitis.

Hepatitis adalah kelainan hati berupa peradangan sel hati. Penyakit ini disebabkan adanya gangguan atau kerusakan membran hati. Ada dua faktor penyebabnya yaitu faktor infeksi dan faktor non infeksi. Faktor penyebab infeksi antara lain virus hepatitis dan bakteri yang terdapat pada hati (Lika, 2007).

Peradangan ini ditandai dengan meningkatnya kadar enzim hati. Diagnosis awal penyakit ini setelah memperhatikan gejala adalah melakukan tes fungsi hati yang biasa disebut LFT (*Liver Function Test*) (Green, 2005). Dengan beberapa atribut dari

hasil pemeriksaan tersebut akan mudah digunakan untuk menganalisis penyakit tersebut. Salah satu teknologi kecerdasan buatan yang dapat digunakan untuk menganalisis penyakit tersebut adalah *machine learning*.

*Machine learning* telah banyak digunakan dalam bidang medis untuk menganalisa dataset medis (Munawarah, 2016). Salah satu metode *machine learning* adalah *Support Vector Machine* (SVM). *Support Vector Machine* (SVM) adalah sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi *linier* dalam sebuah ruang fitur (*feature space*) berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan *learning bias* yang berasal dari teori pembelajaran statistik (Sembiring 2007). Ciri dari metode ini adalah menemukan fungsi pemisah (*klasifier*) yang optimal yang bisa memisahkan dua set data dari dua kelas yang berbeda.

Teori yang mendasari SVM sendiri sudah berkembang sejak 1960-an, tetapi baru diperkenalkan oleh Vapnik, Boser, dan Guyon pada tahun 1992 dan sejak itu SVM berkembang dengan pesat. SVM merupakan salah satu teknik yang relative baru dibandingkan dengan teknik lain, tetapi memiliki performansi yang lebih baik di berbagai bidang aplikasi seperti *bioinformatics*, pengenalan tulisan tangan, klasifikasi teks dan lain sebagainya

(Ramana, 2011). SVM juga dikenal sebagai teknik pembelajaran mesin (*machine learning*) paling mutakhir setelah pembelajaran mesin sebelumnya yang dikenal sebagai *Neural Network* (NN)(Kerami, 2004). SVM berada dalam satu kelas dengan ANN dalam hal fungsi dan kondisi permasalahan yang bisa diselesaikan. Keduanya masuk dalam kelas *supervised learning*. Terbukti dalam banyak implementasi, SVM memberi hasil yang lebih baik dari ANN, terutama dalam hal solusi yang dicapai. SVM selalu mencapai solusi yang sama untuk setiap *running* (Santoso, 2007).

Sejalan dengan perkembangan teknologi saat ini, masyarakat sudah mampu mengakses internet salah satunya untuk mengakses *web*. Dilihat dari artikel-artikel ataupun *website* belum ditemukan suatu aplikasi yang dapat mendiagnosis penyakit Hepatitis berbasis *web*. Sehingga, dengan menggunakan metode klasifikasi *Support Vector Machine* (SVM) akan dibangun sebuah aplikasi yang dapat mendiagnosis seseorang terkena penyakit hepatitis dengan berbasis *web*. Aplikasi ini dirancang untuk memberi kemudahan kepada masyarakat atau tenaga medis untuk mengetahui diagnosis penyakit Hepatitis. Aplikasi ini dibangun dengan menggunakan bahasa pemrograman PHP. Bahasa pemrograman ini sangat sederhana untuk dipahami dalam mengelola situs *web* dibandingkan dengan bahasa pemrograman lainnya.

## 2. TINJAUAN PUSTAKA

Bagian ini menyajikan hasil studi pustaka mengenai teori yang berhubungan dengan pelaksanaan dan penyusunan tahap pengembangan.

### 2.1 PENYAKIT HEPATITIS

Menurut Green (2005), Hati adalah organ yang bekerja paling keras di dalam tubuh manusia yang berbentuk seperti bola yang berada tepat di bawah tulang rusuk di sisi kanan perut. Hati sangat berperan penting dalam tubuh manusia untuk mencerna makanan, menyingkirkan tubuh dari zat beracun dan menyimpan energi dalam tubuh. Salah satu penyakit yang dapat menyerang hati adalah penyakit Hepatitis. Cara terbaik untuk memeriksa hepatitis adalah dengan tes darah. Tes darah akan menunjukkan hasil dari fungsi hati dengan mengukur: Alanineaminotransferase (ALT)/SGPT, aspartateaminotransferase (AST)/SGOT, alkalinephosphatase (ALP), Bilirubin, Albumin dan total Protein (TP). (Ramana, 2011)

### 2.2 DATA MINING

*Data Mining* adalah proses pengambilan informasi dari sejumlah data yang besar dimana informasi yang akan didapatkan adalah informasi penting yang tidak bisa didapatkan dengan cara manual, melainkan dengan mencari pola atau kecenderungan tertentu (Chakrabarti, et al., 2006).

*Data mining* memiliki tahap-tahap yang berasal dari suatu data menjadi suatu pengetahuan yang berguna. Tahap- tahap *data mining* yaitu (Han & Kamber, 2006):

#### 1. *Data Cleaning*

Merupakan tahap pembuangan data data yang tidak relevan dan data yang tidak konsisten. Pembuangan data merupakan pengecekan data apa saja yang hilang, tidak valid dan data yang tidak sesuai dengan hipotesa.

#### 2. *Data Integration*

Merupakan proses penambahan data dengan yang relevan dan penggabungan data.

3. *Data Selection*

Pemilihan data yang relevan sebagai dasar analisis untuk melakukan proses *mining*.

4. *Data Transformation*

Proses merubah data ke format tertentu sehingga sesuai untuk melakukan proses *mining*.

5. *Data Mining*

Proses mencari pola atau informasi dengan metode atau algoritma *data mining* tertentu.

6. *Pattern evaluation*

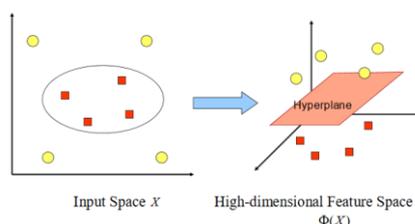
Mengidentifikasi pola yang menarik dari hasil *mining*. Dan dari pola ini dievaluasi apakah sesuai dengan hipotesa atau tidak.

7. *Knowledge presentation*

Menampilkan pola informasi yang dihasilkan dari proses *mining*, dan menampilkan hasil ke bentuk yang mudah dipahami.

2.3 **SUPPORT VECTOR MACHINE (SVM)**

SVM pertama kali diperkenalkan oleh Vapnik tahun 1992 sebagai rangkaian harmonis konsep-konsep unggulan dalam bidang *pattern recognition*. Sebagai salah satu metode *pattern recognition*, usia SVM terbilang masih relatif muda. *Support Vector Machine* (SVM) melakukan suatu teknik untuk menemukan fungsi pemisah yang bisa memisahkan dua set data dari dua kelas yang berbeda (Vapnik, 2013).



**Gambar 2.1 Usaha SVM dalam menemukan hyperplane terbaik secara nonlinear**

Konsep SVM dapat dijelaskan secara sederhana sebagai usaha mencari *hyperplane-hyperplane* terbaik yang berfungsi sebagai pemisah dua *class* pada *input space*. Gambar 2.2a memperlihatkan beberapa pola yang merupakan anggota dari dua buah *class* : -1 dan +1. Pola yang tergabung pada *class* -1 disimbolkan dengan warna merah (kotak), sedangkan pola pada *class* +1, disimbolkan dengan warna kuning (lingkaran). Problem klasifikasi dapat diterjemahkan dengan usaha menemukan garis (*hyperplane*) yang memisahkan antara kedua kelompok tersebut.

Terdapat beberapa metode kernel dalam *Support Vector Machine* (SVM) :

**Tabel 2.1 Jenis Metode Kernel SVM**

No	Jenis Kernel	Definisi
1	Linear	$K(x_i, x_j) = x_i \cdot x_j$
2	Polynomial	$K(x_i, x_j) = (x_i \cdot x_j + 1)^d$
3	Gaussian atau Radial Basis Function (RBF)	$K(x_i, x_j) = \exp \left( -\frac{\ x_i - x_j\ ^2}{2\sigma^2} \right)$
4	Tangent Hyperbolic (Sigmoid)	$K(x_i, x_j) = \tanh (\alpha x_i \cdot x_j + \beta)$

Adapun langkah-langkah dalam klasifikasi SVM adalah sebagai berikut :

1. Menginisiasi awal untuk nilai  $\alpha$ , C, epsilon, gamma, lamda, degree (d). 2.1
2. Memasukkan data latih
3. Menentukan *dot product* setiap data dengan memasukkan fungsi kernel (K). Untuk rumus fungsi kernel Linear adalah sebagai berikut : 2.1  
 $K(x_i, x_j) = x_i \cdot x_j$
4. Menghitung matriks dengan rumus : 2.2  
 $D_{ij} = y_i y_j (K((x_i \cdot x_j) + \lambda^2)$
5. Mencari nilai error dengan rumus : 2.3  
 $E_i = \sum_{j=1}^i \alpha_j D_{ij}$

6. Menghitung nilai delta alpha dengan rumus :  

$$\delta\alpha_i = \min\{\max[\gamma(1-E_i), -\alpha_i], C-\alpha_i\} \dots\dots\dots 2.4$$
7. Menghitung nilai  $\alpha$  baru dengan rumus :  

$$\alpha_i = \alpha_i + \delta\alpha_i \dots\dots\dots 2.5$$
8. Mencari nilai bias dengan rumus :  

$$b = \frac{1}{2}(\langle w \cdot x^{-1} \rangle + \langle w \cdot x^{+1} \rangle) \dots\dots\dots 2.6$$
9. Setelah mendapat nilai  $\alpha$ ,  $w$  dan  $b$  maka dapat dilakukan pengujian. Untuk menentukan fungsi keputusan dengan rumus :  

$$f(x) = w \cdot x + b \text{ atau } f(x) = \sum_{i=1}^m \alpha_i y_i K(x, x_i) + b \dots\dots\dots 2.7$$

2	(-)	FP (False Positif)	TN (True Negatif)
---	-----	--------------------	-------------------

Hasil yang diperoleh dalam *confusion matrix* digunakan untuk mengukur performa dari kinerja *classifier* dalam bentuk akurasi dan parameter evaluasi lainnya.

**2.4 CROSS VALIDATION**

*Cross-validation* (CV) adalah metode statistik yang dapat digunakan untuk mengevaluasi kinerja model atau algoritma dimana data dipisahkan menjadi dua subset yaitu data proses pembelajaran dan data validasi / evaluasi. *K-Fold Cross Validation* melakukan pembagian data dengan membagi dataset secara terstruktur menjadi k himpunan bagian (*subset*). Nilai K pada penelitian biasanya menggunakan nilai 10 karena 10-fold lebih baik dari metode *leave-one-out* (Kohavi, 1995).

**2.5 CONFUSION MATRIX**

*Confusion Matrix* adalah konsep dari *machine learning*, yang berisi informasi tentang klasifikasi aktual dan prediksi yang dilakukan oleh sistem klasifikasi. Sebuah *confusion matrix* memiliki dua dimensi, satu dimensi di indeks oleh aktual kelas objek, yang lain diindeks oleh kelas yang *classifier* memprediksi. Pada tabel 2.2 menyajikan bentuk dasar dari *confusion matrix* untuk klasifikasi N=2 (Deng, et al., 2016).

Tabel 2.2 Contoh Confusion Matrix 2 Kelas

No	Aktual	Prediksi	
		(+)	(-)
1	(+)	TP (True Positif)	FN (False Negatif)

**3. METODOLOGI PENELITIAN**

Bab ini merupakan tahapan dari proses pembangunan perangkat lunak, yaitu deskripsi perangkat lunak (analisis penilaian dan kebutuhan fungsional), pemodelan dan desain sistem. Tahapan ini juga merupakan fase dari pengembangan perangkat lunak menggunakan model proses Waterfall.

**3.1 PENGUMPULAN DATA**

Proses ini merupakan pengumpulan data dan informasi terkait data pasien penyakit hati yang akan digunakan dalam penelitian. Data yang digunakan pada penelitian ini didapat dari UCI (*Universitas California Invene*) *Machine Learning Repository* (Janosi & Steinbrunn, 2011) yaitu *Indian Liver Dataset Patient* atau Data Pasien Hati India. Data Pasien Hati India merupakan catatan data-data pasien hati di berbagai rumah sakit India. Data tersebut dikumpulkan dari timur laut Andhra Pradesh, India (Classifai, 2017). Data terdiri dari 11 atribut yang meliputi 10 atribut masukan dan 1 atribut kelas.

Terdapat 583 jumlah total data dengan 416 data untuk kelas sakit dan 167 data untuk kelas sehat. Untuk data lengkap Pasien Hati India dapat dilihat di <http://archive.ics.uci.edu/ml/datasets/ILPD+%28Indian+Liver+Patient+Dataset%29>.

**3.2 PREPROCESSING DATA**

Sebelum melakukan proses data mining perlu dilakukan *preprocessing* untuk memastikan data yang akan diolah di data

mining adalah data yang baik. Data yang kualitasnya kurang baik adalah data yang masih bersifat *incomplete*, *noisy*, dan *inconsistent*. *Incomplete* merupakan tidak lengkapnya nilai suatu atribut, tidak lengkapnya atribut-atribut yang penting, atau hanya mempunyai data yang merupakan rekapitulasi. *Noisy* merupakan data masih mengandung error atau merupakan value yang tidak wajar. *Inconsistent* merupakan data masih mengandung nilai yang saling bertentangan (Larose, 2005). Beberapa tahapan yang dilakukan dalam *preprocessing* data adalah sebagai berikut :

### 1. Pembersihan Data (*Data Cleaning*)

Proses ini digunakan untuk membuang data yang tidak konsisten dan bersifat noise dari data yang terdapat pada dataset. Pada data Pasien Hati India terdapat 1 tabel dengan nilai yang kosong yaitu pada tabel ratio. Pada tabel ratio terdapat 4 nilai data yang kosong sehingga dilakukan pembuangan data terhadap data tersebut. Maka dari 583 data pasien yang digunakan sebagai data pelatihan hanya berjumlah 579 dengan 11 atribut yang sudah lengkap.

### 2. Seleksi Data (*Data Selection*)

Proses ini digunakan untuk Pemilihan data yang relevan sebagai dasar analisis untuk melakukan proses *mining*. Pada data Pasien Hati India terdapat 11 atribut, 10 atribut masukan dan 1 atribut kelas. Tidak ada seleksi atribut yang dilakukan pada data Pasien Hati India. Sehingga, untuk mendiagnosis seseorang terkena penyakit Hepatitis dilakukan dengan menggunakan 10 atribut pada data Pasien Hati India. Atribut-atribut tersebut yaitu: usia, jenis kelamin,

total bilirubin, direct bilirubin, alkaline, alamine, aspartate, total proteins, albumin, dan ratio.

### 3. Transformasi Data (*Data Transformation*)

Proses ini digunakan untuk merubah data ke format tertentu sehingga sesuai untuk melakukan proses *mining*. Terdapat 11 atribut pada data Pasien Hati India yaitu 10 atribut masukan dan 1 atribut kelas. Pada data tersebut akan dilakukan normalisasi data. Normalisasi adalah proses transformasi dimana sebuah atribut numerik diskalakan dalam range yang lebih kecil seperti -1.0 sampai 1.0, atau 0.0 sampai 1.0. Proses normalisasi pada studi kasus ini, atribut numerik diskalakan dalam range 0.0 sampai 1.0. Teknik yang dilakukan untuk normalisasi data dengan menggunakan *min-max normalization*. *Min-max normalization* memetakan sebuah *value v* dari atribut *A* menjadi *v'* ke dalam range [*new\_minA*, *new\_maxA*] dengan rumus:

$$v' = \frac{v - \text{minA}}{\text{maxA} - \text{minA}} (\text{new\_maxA} - \text{new\_minA}) + \text{new\_minA} \dots \dots \dots 3.1$$

Dengan :

- v'* : nilai atribut setelah normalisasi
- v* : nilai atribut sebelum normalisasi
- minA*: nilai minimum dari semua nilai dalam atribut
- maxA*: nilai maksimum dari semua nilai dalam atribut
- new\_maxA* : nilai maksimum dari semua nilai dalam atribut setelah normalisasi
- new\_minA* : nilai minimum dari semua nilai dalam atribut setelah normalisasi

### 3.3 PROSES KLASIFIKASI

Proses klasifikasi adalah proses pengelompokan berdasarkan ciri-ciri persamaan dan perbedaan. Pada penelitian ini penulis menggunakan metode *Support Vector Machine* (SVM) yang digunakan untuk mengklasifikasi dan memprediksi diagnosis penyakit hepatitis. Metode *Support Vector Machine* yang digunakan adalah metode *Support Vector Machine Polynomial*. Pada proses klasifikasi menggunakan *Support Vector Machine* ini hanya memiliki 2 label yaitu 1 adalah label positif dan -1 adalah label negatif. Label positif merupakan seseorang yang terdiagnosis penyakit hepatitis sedangkan label negatif merupakan seseorang yang tidak terdiagnosis penyakit hepatitis. Untuk klasifikasi data menggunakan persamaan-persamaan pada langkah-langkah klasifikasi SVM.

### 4. PEMBAHASAN

Bab ini membahas mengenai fase *implementation*, dan fase *testing* pada Aplikasi Data Mining untuk Mendeteksi Penyakit Hepatitis menggunakan metode *Support Vector Machine (SVM) Polynomial*.

#### 4.1 IMPLEMENTATION

Fase *Implementation* menjelaskan implementasi dari hasil analisis dan desain dengan pengkodean untuk menghasilkan Aplikasi Deteksi Penyakit Hepatitis. Pada fase ini menghasilkan implementasi *source code*, dan implementasi antarmuka.

#### 4.2 PENGUJIAN KLASIFIKASI POLYNOMIAL

Pengujian klasifikasi ini menggunakan *K-Fold Cross Validation*. *K-Fold Cross Validation* melakukan pembagian data dengan membagi dataset secara terstruktur menjadi k himpunan bagian (*subset*). Pembagian data menggunakan *K-Fold Cross Validation* membagi data menjadi K

partisi dengan setiap partisinya memiliki banyak data yang sama. Dataset yang berjumlah 579 ini akan dilakukan proses evaluasi dengan K = 10. Dengan 90% data untuk data latih dan 10% data untuk data uji. Pengujian klasifikasi metode SVM *Polynomial* dengan menggunakan 10 atribut data. 10 atribut data tersebut merupakan atribut masukan yang ada pada data Pasien Hati India yaitu usia, jenis kelamin, total bilirubin, direct bilirubin, alkaline, alamine, aspartate, total proteins, albumin, dan ratio.

Adapun hasil pengujian dengan K-Fold Validation adalah sebagai berikut :

Tabel 4.1 Hasil Pengujian Klasifikasi SVM

K	SVM Linear	SVM Polynomial
1	71.93%	71.93%
2	70.17%	70.17%
3	73.68%	73.68%
4	73.68%	73.68%
5	73.68%	73.68%
6	64.91%	64.91%
7	66.67%	66.67%
8	66.67%	66.67%
9	71.92%	71.92%
10	74.24%	74.24%

Maka dapat kita simpulkan bahwa rata-rata nilai akurasi metode SVM Polynomial dengan 10 atribut:

$$\begin{aligned}
 \text{Rata-Rata Akurasi} &= \frac{89.47+89.47+87.72+91.22+87.72+87.72+91.22+85.96+84.21+81.82}{10} \\
 &= \frac{876.53}{10} \\
 &= 87.65 \%
 \end{aligned}$$

### 5. PENUTUP

#### 5.1 KESIMPULAN

Berdasarkan pengerjaan skripsi yang telah dilakukan, kesimpulan yang dapat diambil meliputi:

1. Dihasilkannya suatu Aplikasi Deteksi Penyakit Menggunakan metode *Support Vector Machine (SVM) Polynomial* (Studi Kasus: Data Pasien Hati India). Aplikasi ini dikembangkan dengan Bahasa Pemrograman PHP. DBMS MySQL.
2. Aplikasi ini dapat menampilkan hasil akurasi SVM *Polynomial* menggunakan *K-Fold Validation*.
3. Dengan menggunakan *Cross validation*. menggunakan pengujian pada 10 atribut. data Pasien Hati India memiliki rata-rata nilai akurasi sebesar 87.65%.

## 5.2 SARAN

Untuk pengembangan penelitian selanjutnya. dapat dicoba dengan menggunakan metode SVM Non linear lainnya yaitu RBF dan Sigmoid untuk membandingkan nilai akurasi setiap jenis kernel metode *Support Vector Machine*.

## REFERENSI

- [1] Anhar (2010). *Panduan Menguasai PHP dan MySQL*. Jakarta: Media Kita.
- [2] Barbara Rosario (2007). *Applied Natural Language Processing*.
- [3] Binus. (2017) “10 Fold Cross Validation”.
- [4] Chakrabarti. S.. Ester. M.. Fayyad. U.. Gehrke. J.. Han. J.. Morishita. S.. . . . Wang. W. (2006). *Data Mining Curriculum: A Proposal*.
- [5] Deng. et al.. (2016). *An improved method to construct basic probability assignment based on the confusion matrix for classification problem*. vol 340-341. pp 250-261
- [6] Green. Chris W.(2005). *Hepatitis Virus dan HIV*. Yayasan Spiritika:Jakarta
- [7] Hamilton. H..2012. *Confusion Matrix*.
- [8] Janosi & Steinbrunn. (2011). *UCI Machine Learning Repository*.
- [9] Ladjamudin. Al Bahra Bin. 2005. “*Analisis dan Desain Sistem Informasi*”. Yogyakarta: Graha Ilmu.
- [10] Larose. D. T. (2005). *Discovering Knowledge in Data: An Introduction to Data Mining*. John Willey & Sons. Inc.
- [11] Lika (2017). *Informasi Lengkap Seputar Penyakit Hepatitis yang Wajib kamu ketahui*.
- [12] Medistra. (2008). *Gangguan Faal (Fungsi) Hati Yang Sering Ditanyakan Oleh Penderita*
- [13] Munawarah. Raudlatul (2016). *Implementasi Metode Support Vector Machine untuk Mendiagnosa Penyakit Hepatitis*. Program S-1 Ilmu Komputer. Universitas Lambung Mangkurat:Banjarbaru.
- [14] Nendrabertus. (2015) “*Eksplorasi Data Mining menggunakan RapidMiner*”
- [15] Nugroho. Anto Satriyo. dkk. 2003. “*Support Vector Machine: Teori dan Aplikasi dalam Bioinformatika*”. dalam jurnal Kuliah Umum IlmuKomputer.Com. Jakarta.
- [16] Pressman. R. (2010). *Software Engineering: A Practitioner's Approach 7th Edition*. New York: McGraw Hill.
- [17] Ramana. B.V.. Babu. S.P.. & Venkateswarlu. N.B. (2011). *A Critical Study Of Selected Classification Algorithms For Liver Disease Diagnosis*. International Journal of Database Management System. Vol.3(2). pp 101-114.
- [18] Santosa. Budi. 2007. “*Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*”. Yogyakarta: Penerbit Graha Ilmu
- [19] Sembiring. Krisantus. (2007). *Penerapan Teknik Support Vector Machine untuk pendeteksian Intrusi pada Jaringan*.
- [20] Sommerville (2011). *Software Engineering*. Jakarta: Erlangga.