

OPTIMASI ALGORITMA SUPPORT VECTOR MACHINE DENGAN PARTICLE SWARM OPTIMIZATION DALAM MENDETEKSI KETEPATAN WAKTU KELULUSAN MAHASISWA : STUDI KASUS POLTEK LP3I JAKARTA “KAMPUS DEPOK”

Endang Supriyadi ¹⁾, Dana Indra Sensuse ²⁾

¹⁾ Program Pascasarjana Magister Ilmu Komputer

Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri)

Jl. Salemba Raya No. 5 Jakarta Pusat

<http://www.nusamandiri.ac.id>, endangs2013@gmail.com

²⁾ Fakultas Ilmu Komputer

Universitas Indonesia, Kampus UI Depok

dana@cs.ui.ac.id

Abstrak- Deteksi ketepatan kelulusan mahasiswa sangat diperlukan untuk mengantisipasi mahasiswa yang memiliki kinerja yang sangat buruk dalam mengikuti proses pembelajaran dalam suatu lembaga pendidikan karena tujuan semua lembaga pendidikan adalah menciptakan mahasiswa yang berkualitas. Hal tersebut dapat dilihat pada hasil alumni yang telah menyelesaikan pendidikan tersebut. POLTEK LP3I Jakarta Kampus Depok salah satu lembaga pendidikan vokasi yang didirikan untuk menciptakan manusia yang memiliki kemampuan / Skill yang dibutuhkan oleh perusahaan sehingga konsep yang ditawarkan adalah pendidikan yang memiliki Link and Match. Kompetitor yang memiliki tujuan sama merupakan salah satu tantangan yang harus dihadapi oleh lembaga tersebut sehingga diperlukan suatu solusi untuk mengatasi hal tersebut. Salah satu solusinya adalah pendeteksian ketepatan waktu kelulusan mahasiswa. Hal tersebut dapat dilakukan dengan cara menggunakan teknik data mining. Salah satu teknik data mining yang digunakan adalah support vector machines (SVM). Metode support vector machine mampu mengatasi masalah yang berdimensi tinggi, mengatasi masalah klasifikasi dan regresi dengan linier ataupun nonlinier kernel yang dapat menjadi satu kemampuan algoritma pembelajaran untuk klasifikasi serta regresi, namun support vector machine memiliki masalah dalam parameter yang sesuai. Untuk mengatasi masalah tersebut diperlukan metode algoritma particle swarm optimization (PSO) untuk pemilihan parameter yang sesuai pada metode support vector machine. Beberapa eksperimen dilakukan untuk mendapatkan akurasi yang optimal. Eksperimen dengan menggunakan metode support vector machine dan algoritma particle swarm optimization yang digunakan untuk melakukan optimasi parameter C, populasi dan ϵ . Data training yang digunakan data yang dibatasi dengan waktu dan tanpa waktu atau data single year dan multi years. Hasil eksperimen menunjukkan bahwa data training single year memiliki akurasi yang lebih tinggi dibandingkan dengan data multi years yaitu sebesar 80.34 % dan AUC = 0.641 dengan waktu eksekusi lebih cepat selama 00:01:14 sedangkan data multi years memiliki akurasi dan AUC adalah sebesar 77.30 % dan AUC = 0.580 serta memiliki waktu eksekusi selama 00:13:43. Hasil eksperimen tersebut menunjukkan pengujian data set single year menggunakan penerapan algoritma particle swarm optimization pada support vector machine menunjukkan hasil yang lebih akurat untuk deteksi ketepatan kelulusan mahasiswa dan merekomendasikan pengklasifikasian berdasarkan pertahun ajaran.

Kata kunci: Deteksi, Ketepatan Waktu Kelulusan, Support Vector Machine, Algoritma Particle Swarm Optimization..

I. PENDAHULUAN

Tujuan utama lembaga pendidikan diantaranya adalah pendidikan berkualitas bagi mahasiswanya [1] oleh karena itu *keterlambatan mahasiswa lulus* merupakan juga salah satu hal yang mengakibatkan penurunan kualitas pendidikan yang diberikan oleh lembaga pendidikan terhadap peserta didiknya. Kualitas perguruan tinggi diukur berdasarkan 7 standard utama penilaian diantaranya adalah mahasiswa dan lulusannya dengan komponen penilaian IPK dan masa studi [2].

Perkembangan peminatan terhadap perguruan tinggi khususnya politeknik khususnya diploma III di wilayah kopertis III dari periode tahun ajaran 2008/2009 hingga tahun 2011/2012 sangat cukup tinggi yaitu sebesar 8.767 mahasiswa [3]. Jumlah mahasiswa baru TA. 2009/2010 s/d 2012/2013 mengalami kenaikan dari 229 menjadi 285 orang. Sedangkan mahasiswa yang mengikuti studi hingga selesai ditahun yang sama mengalami penurunan 229 menjadi 128, 194 menjadi 143 dst [4]. Salah satu solusi permasalahan tersebut, LP3I khususnya kampus depok harus dapat meningkatkan kualitas pendidikannya yaitu dengan cara mengantisipasi *keterlambatan kelulusan mahasiswanya*.

SVM merupakan salah satu teknik data mining yang dapat digunakan untuk berbagai macam tujuan diantaranya pada bidang bahasa yang dilakukan oleh Saraswati [5], komputer yang dilakukan oleh Jacobus [6], perbankan yang dilakukan oleh Satsiou [7], bidang kedokteran yang dilakukan oleh Yenaeng [8].

PSO merupakan salah satu metode optimasi yang dapat digunakan untuk meningkatkan akurasi (Jing Wang, 2014) sehingga pada penelitian ini dilakukan eksperimen dengan metode SVM-PSO untuk mendeteksi ketepatan waktu kelulusan mahasiswa khususnya mahasiswa LP3I Jakarta kampus "Depok".

II. LANDASAN TEORI

2.1. Mahasiswa

Adalah peserta didik pada jenjang Pendidikan Tinggi (UU No. 12 tahun 2012) [9] sedangkan menurut Azwar bahwa mahasiswa adalah elite masyarakat yang mempunyai intelektualitas yang kompleks dibandingkan dengan kelompok seusianya, dibawah maupun diatasnya yang bukan mahasiswa. Ciri intelektualitas tersebut adalah kemampuan mahasiswa menghadapi, memahami dan mencari cara pemecahan masalah secara lebih sistematis [10].

2.2. Kelulusan Mahasiswa

Menurut (UU No. 12 Thn. 2012) [9] kelulusan mahasiswa adalah "mahasiswa dapat menyelesaikan program Pendidikan sesuai dengan kecepatan belajar masing-masing dan tidak melebihi ketentuan batas waktu yang ditetapkan oleh Perguruan Tinggi".

2.3. Data Mining

Ada beberapa definisi yang diberikan oleh para ahli mengenai data mining diantaranya adalah bahwa data mining merupakan proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis [11]. *Data mining*, sering juga disebut *knowledge discovery in database (KDD)*, adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan pola keteraturan, pola hubungan dalam set data berukuran besar [12] dan keluaran dari data mining ini dapat dijadikan untuk memperbaiki pengambilan keputusan di masa depan.

Berikut ini tahapan proses penemuan pengetahuan dijelaskan menurut [13] adalah sebagai berikut :

1. *Cleaning*
Pembersihan data yang tidak konsisten dan noise.
2. *Integration*
Penggabungan data dari berbagai sumber baik tabel maupun database.
3. *Selection*
Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database.

4. Transformation

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining.

5. Data mining

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. Pattern evaluation

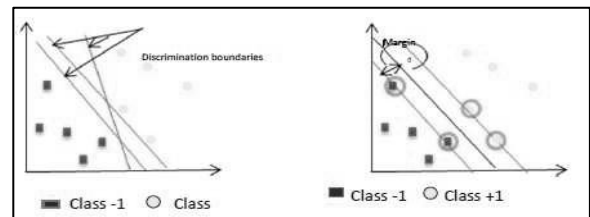
Hasil data yang telah ditemukan berupa pola – pola tertentu sesuai dengan hipotesa di evaluasi jika tidak sesuai dikembalikan untuk diperbaiki proses data miningnya atau Hasil dapat langsung dijadikan hasil akhir yang mungkin bermanfaat.

7. Presentation

Presentasi pola yang ditemukan untuk menghasilkan aksi dan memformulasikan keputusan atau aksi dari hasil analisa yang didapat.

2.4. Algoritma Support Vector Machine (SVM)

SVM adalah usaha mencari *hyperplane* terbaik yang berfungsi sebagai pemisah dua buah kelas pada *input space* [14].



Gambar 1. SVM mencari hyperplane terbaik [14]

Input data dapat berupa linear dan non linear. Jika input data berupa linear maka pemisahan *hyperplane* dapat diberikan dalam persamaan :

$$f(X) = wTx + b \quad (2.1)$$

dimana w adalah n -dimensi bobot vektor dan b adalah pengali skala atau nilai bias. Persamaan ini menemukan maksimum margin untuk memisahkan kelas dari kelas positif dari kelas negatif. Fungsi keputusan ditunjukkan dalam persamaan. Contoh untuk data linear terpisah ditunjukkan pada Gambar 2.2 :

$$y_i(w \cdot x_i + b) \geq 1 \quad i=1, \dots, k \quad (2.2)$$

Pemilihan Parameter pada Support Vector Machine

Untuk mendapatkan tingkat kinerja yang tinggi, beberapa parameter dari algoritma SVM harus diperbaiki, termasuk [15]:

1. Pemilihan Fungsi Kernel
2. Kinerja SVM tergantung pada pilihan fungsi kernel, besaran parameter kernel dan penentuan parameter C . Fungsi kernel yang berbeda memperoleh tingkat keberhasilan yang berbeda untuk berbagai jenis data aplikasi. Ketika nilai penentuan parameter C yang dipilih terlalu besar atau terlalu kecil, generalisasi SVM mungkin berkurang. Jika parameter kernel dan penentuan

parameter yang tepat dipilih, kinerja SVM akan optimal.

3. Parameter kernel(s);
4. Parameter regularisasi (C , ν , ϵ) untuk *tradeoff* antara kompleksitas model dan akurasi mode

2.5. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) banyak digunakan untuk memecahkan masalah optimasi serta sebagai masalah seleksi fitur [16]. Optimasi adalah proses menyesuaikan kepada masukan atau karakteristik perangkat, proses matematis, atau percobaan untuk menemukan output minimum atau maksimum hasil. Input terdiri dari variabel, proses atau fungsi dikenal sebagai fungsi biaya, fungsi tujuan atau kemampuan fungsi dan output adalah biaya atau tujuan, jika proses adalah sebuah percobaan, kemudian variabel adalah masukan fisik untuk percobaan [17]. Dalam teknik *Particle Swarm Optimization* (PSO) terdapat beberapa cara untuk melakukan pengoptimasian diantaranya meningkatkan bobot atribut (*attribute weight*) terhadap semua atribut atau variabel yang dipakai, menseleksi atribut (*attribute selection*) dan *feature selection*. *Particle Swarm Optimization* (PSO) adalah suatu teknik optimasi yang sangat sederhana untuk menerapkan dan memodifikasi beberapa parameter [18].

Untuk menemukan solusi yang optimal, masing-masing partikel bergerak ke arah posisi sebelumnya terbaik (*pbest*) dan terbaik posisi global (*gbest*). Kecepatan dan posisi partikel dapat diperbarui sebagai berikut persamaan:

$$v_{ij}(t+1) = w * v_{ij}(t) + c_1 * rand_1 * (pbest_{ij}(t) - p_{ij}(t)) + c_2 * rand_2 * (gbest_{ij}(t) - p_{ij}(t))$$

$$p_{ij}(t+1) = p_{ij}(t) + \beta * v_{ij}(t+1)$$

Dimana:

- t = menunjukkan *counter iterasi*
- V_{ij} = kecepatan partikel i pada dimensi ke-j (nilainya terbatas antara [- v_{max}, v_{max}],
- p = posisi partikel i pada j dimensi (nilainya terbatas [-p_{max}, p_{max}])
- pbest_{ij} = posisi pbest partikel i pada dimensi ke-j
- gbest_{ij} = posisi *gbest* dari dimensi ke-j
- w = berat inersia (menyeimbangkan eksplorasi global dan lokal eksploitasi)
- rand1 dan rand2 = fungsi acak di rentang [0, 1]
- β = faktor kendala untuk mengontrol kecepatan berat (nilainya ke 1)
- c1 dan c2 adalah faktor pembelajaran pribadi dan sosial (nilainya ke 2).

Pengujian dan Evaluasi serta Validasi Algoritma Data Mining

1. Pengujian K-fold *Cross-validation*

K-Fold *Cross Validation* adalah teknik validasi yang membagi data ke dalam k bagian dan kemudian

masing-masing bagian akan dilakukan proses klasifikasi. Dengan menggunakan K-Fold *Cross Validation* akan dilakukan percobaan sebanyak k. Tiap percobaan akan menggunakan satu data testing dan k-1 bagian akan menjadi data training, kemudian data testing itu akan ditukar dengan satu buah data training sehingga untuk tiap percobaan akan didapatkan data testing yang berbeda-beda. Data training adalah data yang akan dipakai dalam melakukan pembelajaran sedangkan data testing adalah data yang belum pernah dipakai sebagai pembelajaran dan akan berfungsi sebagai data pengujian kebenaran atau keakurasian hasil pembelajaran [19].

2. Evaluasi & Validasi Metode

a. Confusion Matrik

Confusion Matrix adalah alat (*tools*) visualisasi yang biasa digunakan pada supervised learning. Tiap kolom pada matriks adalah contoh kelas prediksi, sedangkan tiap baris mewakili kejadian di kelas yang sebenarnya [20]. *Confusion matrix* berisi informasi aktual (*actual*) dan prediksi (*predicted*) pada sistem klasifikasi.

b. ROC (*Receiver Operating Characteristic*)

Kurva ROC menunjukkan akurasi dan membandingkan klasifikasi secara visual. ROC mengekspresikan *confusion matrix*. ROC adalah grafik dua dimensi dengan *false positives* sebagai garis horizontal dan *true positives* untuk mengukur perbedaan performansi metode yang digunakan. ROC Curve adalah cara lain untuk menguji kinerja pengklasifikasian [20].

Performance keakurasian AUC dapat diklasifikasikan menjadi lima kelompok yaitu [20] :

0.90 – 1.00 = *Excellent Clasification*

0.80 – 0.90 = *Good Clasification*

0.70 – 0.80 = *Fair Clasification*

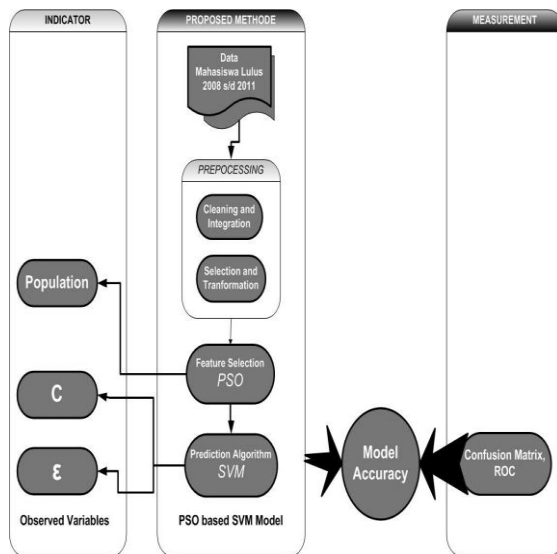
0.60 – 0.70 = *Poor Clasification*

0.50 – 0.60 = *Failure*

2.6. Kerangka Kerja Penelitian

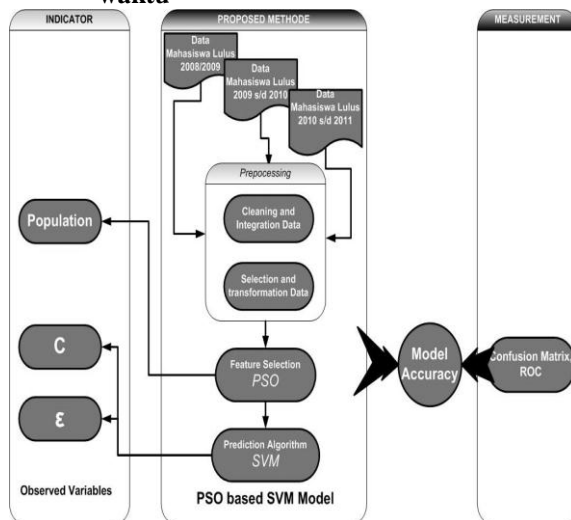
Model kerangka pemikiran yang digunakan adalah *method improvement* (perbaikan metode), yang sering digunakan pada penelitian di bidang sains dan teknik, termasuk bidang *computing* didalamnya. Komponen dari model kerangka pemikiran perbaikan metode (*method improvement*) adalah *Indicators, Proposed Method, Objectives, dan Measurements* [21]. Kerangka pemikiran pada penelitian ini dimulai dari prediksi hasil pemilihan umum. Maka dengan ini penulis mencoba membuat sebuah *soft computing* dengan menggunakan *support vector machine* dengan teknik optimasi *Particle Swarm Optimization* (PSO).

2.6.1. Model prediksi tanpa perbedaan waktu



Gambar 2. Model Prediksi Data *Multi Years*

2.6.2. Model prediksi berdasarkan perbedaan waktu



Gambar 3. Model Prediksi Data *Single Year*

2.7. Hipotesis

Hipotesis dari penelitian ini adalah:

2.7.1. Hipotesis Usulan Model Prediksi Tanpa Perbedaan Waktu

1. Diduga Support Vector Machine (SVM) mampu dalam menyelesaikan masalah dalam klasifikasi ketepatan waktu kelulusan mahasiswa tanpa memperhatikan perbedaan waktu.
2. Diduga seleksi atribut Particle Swarm Optimization akan mampu meningkatkan akurasi Support Vector Machine (SVM) dalam mendeteksi ketepatan waktu kelulusan mahasiswa tanpa memperhatikan perbedaan waktu.

2.7.2. Hipotesis Usulan Model Prediksi Berdasarkan Perbedaan Waktu

1. Diduga Support Vector Machine (SVM) mampu dalam menyelesaikan masalah dalam klasifikasi

ketepatan waktu kelulusan mahasiswa yang dibedakan berdasarkan perbedaan waktu.

2. Perbedaan antar waktu tersebut dapat dideteksi oleh tingkat akurasi yang diukur oleh Confusion Matrix dan ROC (Receiver Operating Characteristic).

III. METODOLOGI PENELITIAN

3.1. Desain Penelitian

Pengertian penelitian dalam akademik yaitu digunakan untuk mengacu pada aktivitas yang rajin dan penyelidikan sistematis atau investigasi di suatu daerah, dengan tujuan menemukan atau merevisi fakta, teori, aplikasi dan tujuannya adalah untuk menemukan dan menyebarkan pengetahuan baru [22].

Menurut [23] ada empat metode penelitian yang umum digunakan yaitu tindakan penelitian, eksperimen, studi kasus dan survei.

Dalam penelitian ini penulis menggunakan metode penelitian eksperimen dengan beberapa langkah yang dilakukan adalah sebagai berikut :

3.2. Pengumpulan Data

Teknik pengumpulan data ialah teknik atau cara-cara yang dapat digunakan untuk menggunakan data [24]. Dalam pengumpulan data terdapat sumber data, sumber data yang terhimpun langsung oleh peneliti disebut dengan sumber primer, sedangkan apabila melalui tangan kedua disebut sumber sekunder. Data yang dikumpulkan penulis merupakan data sekunder karena diperoleh dari database mahasiswa yang terkumpul pada file-file yang terpisah dalam format microsoft excel pada bagian pendidikan yang merupakan salah satu struktur dalam fungsi organisasi di LP3I Jakarta khususnya kampus Depok.

Dataset yang dikumpulkan adalah dataset yang telah melalui proses preprocessing berupa data mahasiswa 3 periode tahun ajaran mulai dari tahun ajaran 2008-2009 s/d 2010-2011. Tahun ajaran 2008-2009 sebanyak 160 mahasiswa, tahun ajaran 2009-2010 sebanyak 152, tahun ajaran 2010-2011 sebanyak 208 mahasiswa sehingga total jumlah mahasiswa sebanyak 520 mahasiswa yang terdiri dari atribut mahasiswa umur, prodi (program studi), kons (konsentrasi), ips (indeks prestasi semester), dan status kelulusan “tepat waktu” dan “Telat”. Contoh sampel dataset seperti yang terlihat pada gambar berikut ini :

Tabel 1. Dataset akademik

UMUR	PRODI	KONS	IPS_1	IPS_2	IPS_3	IPS_4	STATUS_LULUS
17	1	11	3.48	3.26	3.38	3.27	1
17	3	31	3.07	2.95	3.04	3.06	1
17	4	41	2.87	2.81	2.94	3.11	1
17	1	11	0.75	2.16	2.50	0.30	0
18	1	11	1.81	1.76	2.04	1.15	0
18	1	12	2.64	2.52	2.94	2.90	0
18	1	12	2.28	2.22	2.27	2.83	0

3.3. Pengolahan data awal

Jumlah data awal yang diperoleh dari pengumpulan data yaitu sebanyak **520** data, namun tidak semua data dapat digunakan dan tidak semua atribut digunakan karena harus melalui beberapa tahap pengolahan awal data (preparation data). Untuk mendapatkan data yang berkualitas, beberapa teknik yang dilakukan sebagai berikut [25] :

1. Data validation

Adalah untuk mengidentifikasi dan menghapus data yang ganjil (*outlier/noise*), data yang tidak konsisten, dan data yang tidak lengkap (*missing value*).

2. Data integration and transformation.

Adalah untuk meningkatkan akurasi dan efisiensi algoritma. Data yang digunakan dalam penulisan ini bernilai kategorikal. Data ditransformasikan kedalam software Rapidminer.

3. Data size reduction and discretization

Adalah untuk memperoleh data set dengan jumlah atribut dan record yang lebih sedikit tetapi bersifat informatif.

Tabel 2. Tabel Atribut dataset akademik

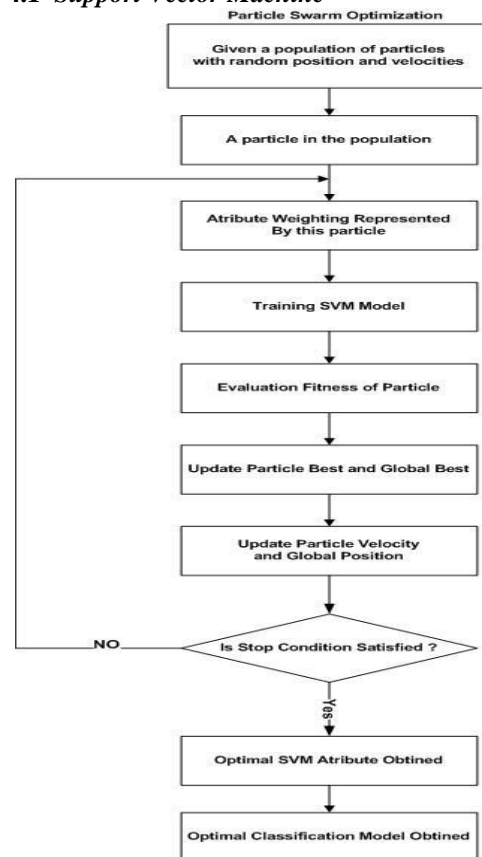
No.	Atribut	Keterangan
1	Umur	Umur mahasiswa sesuai tahun ajaran
2	Prodi	[1] Manajemen Informatika
		[2] Administrasi Bisnis
		[3] Komputerisasi Akuntansi
		[4] Hubungan Masyarakat
3	Kons	[1] Informatika Komputer
		[2] Komputer Desain & Multimedia
		[3] Administrasi Perkantoran
		[4] Sekretaris
		[5] Komputer Akuntansi
		[6] Financial Syari'ah Banking
		[7] Marketing Bisnis Komunikasi
4	IPS	Indeks Prestasi Semester (IPS_1 s/d IPS4
5	Status	0 = Telat Waktu 1 = Lulus tepat waktu

3.4. Metode yang diusulkan

Pada tahap modeling ini dilakukan pemrosesan data training sehingga akan membahas metode algoritma yang diuji dengan memasukan data mahasiswa kemudian dianalisa dan dikomparasi. Berikut ini bentuk gambaran metode algoritma yang akan diuji.

IV. PEMBAHASAN

4.1 Support Vector Machine



4.1.1 Pengujian Data Single Year

Penentuan nilai *training cycles* dalam penelitian ini dilakukan dengan cara uji coba memasukkan nilai *C* dan *Epsilon* pada data 1 tahun (*Single Year*) yaitu data persatu tahun ajaran (periode tahun ajaran 2008/2009 atau 2009/2010 atau 2010/2011). Berikut ini tabel hasil uji coba yang telah dilakukan :

Tabel 3 Eksperimen penentuan nilai *training cycle* SVM

C	ε	Periode TA. 2008/2009			Periode TA. 2009/2010			Periode TA. 2010/2011		
		ACC %	AUC	Exe. Time	ACC %	AUC	Exe. Time	ACC %	AUC	Exe. Time
1.0	0.0	83.75	0.812	0s	78.67	0.543	1s	77.86	0.669	0s
0.0	1.0	75.00	0.500	0s	79.33	0.500	0s	72.28	0.500	0s
0.5	0.5	78.75	0.648	2s	76.67	0.434	1s	66.36	0.500	1s
0.6	0.6	80.00	0.681	2s	80.00	0.455	1s	58.35	0.562	2s
0.7	0.7	66.25	0.510	9s	81.33	0.528	1s	59.12	0.500	4s
0.8	0.8	75.00	0.690	5s	73.33	0.481	2s	61.58	0.525	5s
0.9	0.9	61.88	0.523	11s	72.67	0.532	2s	56.14	0.542	7s

Hasil terbaik pada eksperimen SVM diatas adalah **TA. 2008/2009** dengan nilai $C=1.0$ dan $Epsilon=0.0$ menghasilkan *accuracy* sebesar 83.75 dan AUCnya 0.812 waktu eksekusi selama 0 detik.

4.1.2 Hasil Perhitungan Rata-rata Data Single Year

Tabel 4. Hasil Rerata Data Single Year

C	ϵ	Periode TA. 2008/2009 s/d 2010/2011		
		\bar{X} Accuracy	\bar{X} AUC	\bar{X} Exec Time
1.0	0.0	80.09 %	0.67	0 s
0.0	1.0	75.54 %	0.50	0 s
0.5	0.5	73.93 %	0.53	1 s
0.6	0.6	72.78 %	0.57	2 s
0.7	0.7	68.90 %	0.51	5 s
0.8	0.8	69.97 %	0.57	4 s
0.9	0.9	63.56 %	0.53	7 s

Dari hasil perhitungan rata-rata data single year (Periode TA. 2008/2009 s/d 2010/2011) dari tabel 4 dapat dikumpulkan hasilnya sebagai mana terlihat pada tabel 3 . Sehingga dapat diketahui bahwa nilai rata-rata akurasi tertinggi adalah sebesar 80.09 % dan $AUC = 0.67$ dengan waktu eksekusi selama 0 detik.

4.1.3 Pengujian Data Multi Years

Penentuan nilai *training cycles* dalam penelitian ini dilakukan juga dengan cara uji coba memasukkan nilai C dan *Epsilon* pada data *multi years*, yaitu data dari periode tahun ajaran 2008/2009 s/d 2010/2011. Dengan hasil uji coba seperti yang terlihat pada gambar tabel berikut ini :

Tabel 5. Hasil Pengujian Data Multi Years

C	ϵ	Multi Years		
		Acc	AUC	Exec Time
1.0	0.0	75.62%	0.668	1s
0.0	1.0	75.42%	0.500	0s
0.5	0.5	66.36%	0.500	13s
0.6	0.6	66.75%	0.551	3s
0.7	0.7	66.82%	0.500	6s
0.8	0.8	67.67%	0.565	21s
0.9	0.9	51.90%	0.510	21s

Dari hasil pengujian / eksperimen yang telah dilakukan terhadap data *multi years* (periode tahun ajaran 2008/2009 hingga tahun ajaran 2010/2011) yang terkumpul pada tabel 5, dapat disimpulkan bahwa data *multi years* dengan nilai parameter $C = 1.0$ dan $epsilon = 0.0$ memiliki akurasi yang paling tinggi yaitu sebesar 75.62 % dan AUC (Area Under

Curve) sebesar 0.668 dengan waktu eksekusi selama 1 detik.

4.2 Support Vector Machine Berbasis Particle Swarm Optimization

4.2.1 Pengujian Data Single Year

Penentuan nilai *training cycles* dalam penelitian ini dilakukan dengan cara uji coba memasukkan nilai C dan Epsilon pada data 1 tahun (*Single Year*) yaitu data persatu tahun ajaran (periode tahun ajaran 2008/2009 atau 2009/2010 atau 2010/2011) dengan menggunakan metode SVM-PSO. Berikut ini tabel hasil uji coba yang telah dilakukan :

Tabel 6. Hasil Pengujian Penentuan nilai training cycle SVM-PSO data single year

C	ϵ	Pop Size	Periode TA. 2008/2009			Periode TA. 2009/2010			Periode TA. 2010/2011		
			ACC %	AUC	Ex.T	ACC %	AUC	Ex.T	ACC %	AUC	Ex.T
0.0	1.0	5	75.00	0.500	18s	79.33	0.500	17s	73.53	0.478	48s
0.5	0.5	5	83.75	0.762	6:34	82.00	0.648	1:56	72.87	0.525	1:14
0.6	0.6	5	83.75	0.827	9:46	80.67	0.478	2:26	72.90	0.561	1:33
0.7	0.7	5	81.88	0.758	12:02	80.67	0.560	1:46	72.90	0.511	47s
0.8	0.8	5	80.00	0.681	12:49	80.67	0.624	3:35	72.28	0.489	2:58
0.9	0.9	5	80.00	0.723	19:07	80.67	0.487	4:22	73.46	0.472	1:27

Dari hasil pengujian / eksperimen yang telah dilakukan terhadap data single year (periode tahun ajaran 2008/2009 atau 2009/2010 atau 2010/2011) yang terkumpul pada tabel 6, dapat diketahui bahwa data single year khususnya data periode tahun ajaran 2008/2009 menempati peringkat akurasi yang paling tinggi yaitu sebesar 85,62% dan AUC (Area Under Curve) sebesar 0.792 dengan nilai parameter $C = 1.0$ dan $epsilon = 0.0$ dengan waktu eksekusi selama 28 detik.

4.2.2 Pengujian Data Multi Years

Penentuan nilai *training cycles* dalam penelitian ini dilakukan juga dengan cara uji coba memasukkan nilai C dan *Epsilon* pada data multi years, yaitu data dari periode tahun ajaran 2008/2009 s/d 2010/2011. Dengan hasil uji coba seperti yang terlihat pada gambar tabel berikut ini :

Tabel 7. hasil uji coba penentuan nilai training Data Single Year

C	(ϵ)	Multi Years		
		Acc %	AUC	Exec Times
1.0	0.0	77.30	0.580	13:43

0.0	1.0	75.42	0.500	37s
0.5	0.5	77.10	0.622	10:07
0.6	0.6	77.09	0.583	7:34
0.7	0.7	76.26	0.580	12:01
0.8	0.8	75.42	0.573	9:48
0.9	0.9	75.62	0.567	9:58

Dari hasil pengujian / eksperimen yang telah dilakukan terhadap *data multi years* (periode tahun ajaran 2008/2009 hingga tahun ajaran 2010/2011) yang terkumpul pada tabel 7, dapat disimpulkan bahwa *data multi years* dengan nilai parameter $C = 1.0$ dan $\epsilon = 0.0$ memiliki akurasi yang paling tinggi yaitu sebesar 77.30 % dan AUC (*Area Under Curve*) sebesar 0.580 dengan waktu eksekusi selama 13:43 detik.

4.3 Hasil Optimasi Akurasi Pengujian Data Single Year & Multi Years

4.3.1 Perbandingan SVM dan SVM berbasis PSO pada Rerata Data Terpilih (*Rerata Single Year*)

Tabel 8. Perbandingan Metode SVM dan SVM-PSO Menggunakan Data Rerata Single Year

C	ϵ	SVM			Pop Size	SVM - PSO		
		Acc %	AUC	Exec Time		Acc%	AUC	Exec Time
1.0	0.0	80.09	0.67	0 s	5	80.34	0.641	00:01:14
0.0	1.0	75.54	0.50	0 s	5	75.95	0.493	00:08:12
0.5	0.5	73.93	0.53	1 s	5	79.54	0.645	00:02:50
0.6	0.6	72.78	0.57	2 s	5	79.11	0.622	00:04:04
0.7	0.7	68.90	0.51	5 s	5	78.48	0.610	00:04:36
0.8	0.8	69.97	0.57	4 s	5	77.65	0.598	00:05:28
0.9	0.9	63.56	0.53	7 s	5	78.04	0.561	00:03:50

Dari hasil pengujian / eksperimen dengan menggunakan metode SVM dan SVM-PSO yang telah dilakukan terhadap *data single year* kemudian dihitung reratanya maka jika dibandingkan dengan hasil uji coba tersebut, seperti yang terlihat pada tabel 8, dapat diketahui bahwa rerata *data single year* dengan menggunakan metode SVM-PSO memiliki akurasi yang paling tinggi yaitu sebesar 80.34 % dan AUC (*Area Under Curve*) sebesar 0.641 dengan nilai parameter $C = 1.0$ dan $\epsilon = 0.0$ dengan waktu eksekusi selama 1 menit 14 detik. Sedangkan rerata *data single year* dengan menggunakan metode SVM memiliki akurasi yang lebih rendah yaitu sebesar 80.09 % akan tetapi memiliki AUC =0.67 lebih tinggi serta eksekusi waktu lebih cepat dibandingkan dengan metode SVM-PSO.

4.3.2 Perbandingan SVM dan SVM berbasis PSO pada data Multi Years

Tabel 9. Perbandingan Metode SVM dan SVM-PSO Menggunakan Data Multi Years

c	ϵ	SVM			Pop Size	SVM-PSO		
		Acc%	AUC	Exec. Time		ACC %	AUC	Exec. Times
1.0	0.0	75.62	0.668	1s	5	77.30	0.580	13:43
0.0	1.0	75.42	0.500	0s	5	75.42	0.500	37s
0.5	0.5	66.36	0.500	13s	5	77.10	0.622	10:07
0.6	0.6	66.75	0.551	3s	5	77.09	0.583	7:34
0.7	0.7	66.82	0.500	6s	5	76.26	0.580	12:01
0.8	0.8	67.67	0.565	21s	5	75.42	0.573	9:48
0.9	0.9	51.90	0.510	21s	5	75.62	0.567	9:58

Dari hasil pengujian / eksperimen dengan menggunakan metode SVM dan SVM-PSO yang telah dilakukan terhadap *data multi years* maka jika dibandingkan dengan hasil uji coba tersebut, seperti yang terlihat pada tabel 9, dapat diketahui bahwa *data multi years* dengan menggunakan metode SVM-PSO memiliki akurasi yang paling tinggi yaitu sebesar 77.30 % dan AUC (*Area Under Curve*) sebesar 0.580 dengan nilai parameter $C = 1.0$ dan $\epsilon = 0.0$ dengan waktu eksekusi selama 13 menit 43 detik. Sedangkan rerata *data single year* dengan menggunakan metode SVM memiliki akurasi yang lebih rendah yaitu sebesar 75.62 % akan tetapi memiliki AUC =0.668 lebih tinggi serta eksekusi waktu lebih cepat dibandingkan dengan metode SVM-PSO.

4.3.3 Perbandingan Data Terpilih Single Year dan Multi Years Metode SVM-PSO

Tabel 10. Perbandingan Data Terpilih Single Year dan Multi Years

C	ϵ	Pop Size	\bar{X} Single Year			Multi Years		
			ACC %	AUC	Exec Times	Acc%	AUC	Execution Times
1.0	0.0	5	80.34	0.641	0:01:14	77.30	0.580	13:43
0.0	1.0	5	75.95	0.493	0:08:12	75.42	0.500	37s
0.5	0.5	5	79.54	0.645	0:02:50	77.10	0.622	10:07
0.6	0.6	5	79.11	0.622	0:04:04	77.09	0.583	7:34
0.7	0.7	5	78.48	0.610	0:04:36	76.26	0.580	12:01
0.8	0.8	5	77.65	0.598	0:05:28	75.42	0.573	9:48
0.9	0.9	5	78.04	0.561	0:03:50	75.62	0.567	9:58

Dari hasil pengujian / eksperimen yang telah dilakukan terhadap *data single year* kemudian dihitung reratanya maka jika dibandingkan dengan

hasil uji coba dengan data *multi years*, seperti yang terlihat pada tabel 10, dapat diketahui bahwa rerata data *single year* memiliki akurasi yang paling tinggi yaitu sebesar **80.34 %** dan *AUC (Area Under Curve)* sebesar **0.641** serta rata-rata waktu eksekusi selama **00:01:14** dengan nilai parameter $C = 1.0$ dan $\epsilon = 0.0$.

Sedangkan data training *multi years* khususnya 3 periode tahun ajaran memiliki akurasi dan *AUC* yang lebih rendah yaitu sebesar **77.30 %** dan **0.580** dan waktu eksekusi selama **00:13:43**.

4.3.4 Perbandingan SVM dan SVM berbasis PSO dengan Data Terpilih (*Single Year*)

Tabel 11. hasil uji coba penentuan nilai training Cycles data terpilih single year

C	ϵ	SVM			Pop Size	SVM - PSO		
		Acc%	AUC	Exec Time		Acc%	AUC	Exec Time
1.0	0.0	83.75	0.812	0s	5	85.62	0.792	28s
0.0	1.0	75.00	0.500	0s	5	75.00	0.500	18s
0.5	0.5	78.75	0.648	2s	5	83.75	0.762	6:34
0.6	0.6	80.00	0.681	2s	5	83.75	0.827	9:46
0.7	0.7	66.25	0.510	9s	5	81.88	0.758	12:02
0.8	0.8	75.00	0.690	5s	5	80.00	0.681	12:49
0.9	0.9	61.88	0.523	11s	5	80.00	0.723	19:07

Dari hasil pengujian / eksperimen yang telah dilakukan terhadap data terpilih yaitu data training periode tahun ajaran 2008/2009 yang terkumpul pada tabel 11, dapat disimpulkan bahwa pengujian data training dengan metode SVM-PSO dapat meningkatkan akurasi dan nilai *AUC* yang lebih tinggi dibandingkan dengan metode SVM.

Dari hasil eksperimen tersebut parameter $C = 1.0$ dan $\epsilon = 0.0$ serta *population size* = 5 memiliki akurasi dan *AUC* yang paling tinggi yaitu sebesar 85.62% dan *AUC (Area Under Curve)* sebesar 0.792. Sedangkan peningkatan akurasi dan *AUC* dari kedua metode tersebut yaitu sebesar 1.87 % sedangkan *AUC* mengalami penurunan 0.02 serta waktu eksekusi lebih lama yaitu selama 28 detik.

4.4 Pembahasan

4.4.1 Evaluasi dan Validasi Hasil

Pengujian model untuk mendeteksi ketepatan waktu kelulusan mahasiswa khususnya mahasiswa POLTEK LP3I Jakarta Kampus Depok menggunakan metode SVM dan SVM-PSO dilakukan menggunakan rapidmaner 5.3. Validasi model menggunakan *Cross Validation* dengan nilai parameter 10.

4.4.1.1 Pengujian Model Metode SVM

4.4.1.1.1 Confusion Matrix

Tabel 12. Model Confusion Matrix dengan Metode SVM

Accuracy 83,75%+/-4,15% (mikro:83,75%)			
	True 1.0	True 0.0	Class Precision
Pred 1.0	118	24	83,10%
Pred 0.0	2	16	88,89%
Class recall	98,33%	40,00%	

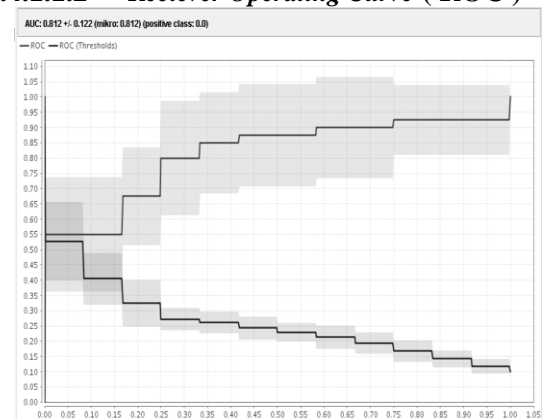
Dari hasil pengujian model dengan menggunakan metode SVM terhadap data *single year* dan *multi years* yang menghasilkan akurasi tertinggi diantara hasil pengujian lainnya termuat dalam table 12, dapat diketahui bahwa dari jumlah data training (data periode tahun ajaran 2008/2009) yaitu sebanyak 160 data terdiri dari 118 diklasifikasikan lulus tepat waktu (1) hasilnya sesuai prediksi yang dilakukan dengan metode SVM, 24 data diprediksi lulus tepat waktu tetapi hasil prediksi telat (0) sedangkan 2 data diprediksi telat ternyata lulus tepat waktu (1) dan 16 data diprediksi telat hasilnya sesuai.

Berdasarkan hasil tersebut dapat diketahui pula perhitungan nilai *accuracy*, *sensitivity*, *specificity*, *ppv*, *npv* yang dikumpulkan dalam tabel 12 sebagai berikut :

Tabel 13. Hasil Perhitungan Accuracy, Sensitivity, Specificity, PPV dan NPV dengan Metode SVM untuk data terpilih.

	Nilai (%)
<i>Accuracy</i>	83.75
<i>Sensitivity</i>	83.09
<i>Specificity</i>	88.89
PPV	98.33
NPV	40.00

4.4.1.1.2 Reciever Operating Curve (ROC)



Gambar 5. Kurva ROC Metode SVM dengan Data Terpilih

Dari gambar 5, dapat diketahui bahwa pengujian data training terpilih dengan akurasi tertinggi menghasilkan kurva ROC dengan nilai *AUC* sebesar 0.812. Hal ini berarti bahwa klasifikasi data

training tersebut memiliki hasil diagnosa *good classification*.

4.4.1.2 Pengujian Model Support Vector Machine berbasis PSO

4.4.1.2.1 Confusion Matrix

Tabel 14. Model *Confusion Matrix* dengan Metode SVM-PSO

Accuracy 85,62%+/-6,28% (mikro:85,62%)			
	True 1.0	True 0.0	Class Precision
Pred 1.0	119	22	83,10%
Pred 0.0	1	18	88,89%
Class recall	99,17%	45,00%	

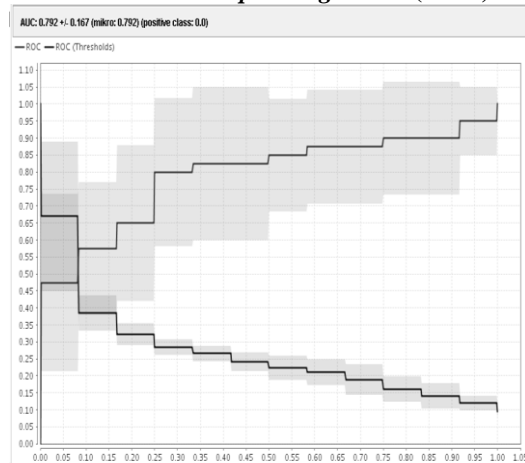
Dari hasil pengujian model dengan menggunakan metode SVM dengan optimasi PSO terhadap data single year khususnya data terpilih yang menghasilkan akurasi tertinggi diantara hasil pengujian lainnya termuat dalam tabel 14, dapat diketahui bahwa dari jumlah data training (data periode tahun ajaran 2008/2009) yaitu sebanyak 160 data terdiri dari 119 diklasifikasikan *lulus tepat waktu* (1) hasilnya sesuai prediksi yang dilakukan dengan metode SVM-PSO, 22 data diprediksi *lulus tepat waktu* tetapi hasil prediksi *telat* (0) sedangkan 1 data diprediksi *telat* ternyata *lulus tepat waktu* (1) dan 18 data diprediksi *telat* hasilnya sesuai.

Berdasarkan hasil tersebut dapat diketahui pula perhitungan nilai *accuracy*, *sensitivity*, *specificity*, *ppv*, *npv* yang terkumpul dalam tabel berikut ini :

Tabel 15. Hasil Perhitungan *Accuracy*, *Sensitivity*, *Specificity*, *PPV* dan *NPV* dengan Metode SVM – PSO untuk data terpilih.

	Nilai (%)
<i>Accuracy</i>	85.62
<i>Sensitivity</i>	84.40
<i>Specificity</i>	94.74
PPV	99.17
NPV	45.00

4.4.1.2.2 Reciever Operating Curve (ROC)



Gambar 6. Visualisasi Kurva ROC Metode SVM berbasis PSO

Dari gambar 6, dapat diketahui bahwa pengujian data training terpilih dengan akurasi tertinggi menghasilkan kurva ROC dengan nilai AUC sebesar 0.792. Hal ini berarti bahwa klasifikasi data training tersebut memiliki hasil diagnosa *fair classification*.

4.5 Desain dan Implementasi

4.5.1 Desain Aplikasi

Desain aplikasi menggunakan aplikasi Matlab 2014 dengan menggunakan data testing mahasiswa TA 2012/2013 yang belum diketahui status kelulusannya.

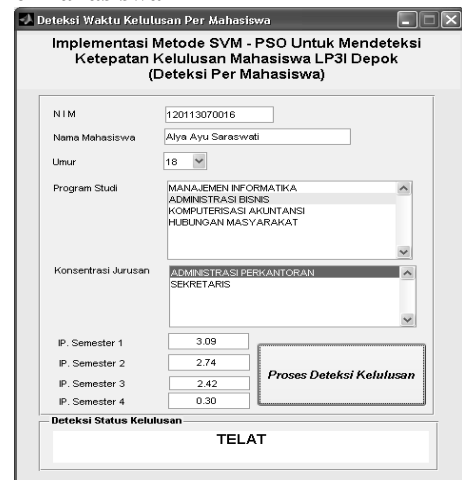
Alur proses aplikasi dimulai dari data akademik berupa file text dengan melalui tahap preprocessing (*Validation, Integration & Transformation, Discretization & Reduction*). Setelah dilakukan proses preprocessing kemudian dilakukan proses optimalisasi populasi untuk mendapatkan hasil atribut yang telah terpilih dengan menggunakan metode *particle swarm optimization(PSO)* kemudian atribut tersebut digunakan untuk memprediksi status kelulusan mahasiswa.

4.5.2 Implementasi Model SVM berbasis PSO



Gambar 7. Menu Utama Aplikasi

4.5.3 Screenshot Deteksi Ketepatan Kelulusan Permahasiswa



Gambar 8. Deteksi Permahasiswa

4.5.4 Screenshot Deteksi Ketepatan Kelulusan Seluruh Mahasiswa

The screenshot shows a window titled "Prediksi Ketepatan Waktu Kelulusan Mahasiswa" with a subtitle "Implementasi Metode SVM - PSO Untuk Mendeteksi Ketepatan Kelulusan Mahasiswa LP3I Depok (Prediksi Seluruh Mahasiswa)". It displays a table with columns: No, Umur, Program Studi, Konsentrasi, Jurusan, Indeks Prestasi 1, Indeks Prestasi 2, Indeks Prestasi 3, Indeks Prestasi 4, and Status Kelulusan. The accuracy is shown as 63.3333 %.

Gambar 9. Deteksi seluruh mahasiswa

4.6 Hasil Pengujian Prototipe Aplikasi Deteksi Ketepatan Waktu Kelulusan

Untuk memastikan bahwa perangkat lunak yang dibuat memiliki standar minimal kualitas, maka salah satu metode untuk pengukuran kualitas perangkat lunak penulis menggunakan metode skala likert [27]. Interval dari kriteria penilaian rata-rata adalah sebagai berikut :

1. Sangat Buruk (SBR) = 1,00 – 1,79
2. Buruk (BR) = 1,80 – 2,59
3. Cukup Baik (CB) = 2,60 – 3,39
4. Baik (B) = 3,40 – 4,19
5. Sangat Baik (SB) = 4,20 – 5,00

Ada 8 buah kriteria yang digunakan penulis untuk mengukur kualitas sebuah perangkat lunak secara kuantitatif. Penulis mencoba untuk menganalisa seluruh jawaban yang diberikan oleh 11 responden yang terdiri dari staff pendidikan & staff IT dan pusat data serta dosen komputer. Mengenai penilaian mereka terhadap kualitas prototipe aplikasi deteksi ketepatan kelulusan mahasiswa khususnya mahasiswa LP3I Jakarta Kampus "Depok", berikut ini hasil analisa penulis terhadap seluruh jawaban yang diberikan oleh responden yang dikumpulkan dalam bentuk tabel:

Tabel 16. Kategori Pengujian & hasil Perhitungan

No.	Kategori	Skor	Mean	Ket
1	<i>Operability</i>	46	4.18	B
2	<i>Auditability</i>	136	4.12	B
3	<i>Simplicity</i>	48	4.36	SB
4	<i>Accuracy</i>	45	4.09	B
5	<i>Execution Efficiency</i>	47	4.27	SB
6	<i>Error Tolerance</i>	45	4.09	B
7	<i>Completeness</i>	42	3.82	B
8	<i>Training</i>	49	4.45	SB
	<i>Rata-rata</i>	458	4.16	B

Dari semua jawaban responden dapat disimpulkan bahwa prototipe aplikasi deteksi ketepatan waktu kelulusan mahasiswa khususnya mahasiswa LP3I Jakarta kampus Depok memenuhi jaminan kualitas aplikasi dengan nilai *baik* karena

nilai skor rata-rata sebesar **4.16** yang berada pada interval antara **3,40 - 4,19**.

Artinya bahwa seluruh responden setuju bahwa prototipe aplikasi deteksi waktu kelulusan memenuhi standar jaminan kualitas aplikasi baik dari segi *kemudahan pengoperasiannya (operability)*, *pemenuhan standard (Auditability)*, *kemudahan untuk dipahami (simplicity)*, *akurasi komputasi (accuracy)*, *kinerja eksekusi (execution efficiency)*, *toleransi kesalahan (error tolerance)*, *Kelengkapan (completeness)*, *kemudahan pembelajaran(training)*.

4.7 Implikasi Penelitian

Implikasi penelitian merupakan suatu penjelasan tentang tindak lanjut penelitian yang terkait dengan aspek sistem, aspek manajerial maupun aspek penelitian lanjutan. Implikasi penelitian ini pada:

4.7.1 Aspek Sistem

a) Hasil pengujian dataset akademik LP3I Jakarta Kampus Depok baik data berupa single year dan multi years menunjukkan bahwa metode particle swarm optimization (PSO) dapat meningkatkan akurasi metode SVM dalam mengklasifikasikan mahasiswa yang tepat lulus dan telat. Sehingga diharapkan metode SVM-PSO dalam wujud prototipe aplikasi deteksi ketepatan kelulusan mahasiswa dapat membantu penanganan terhadap mahasiswa yang bermasalah dengan kinerja akademiknya. Prototipe aplikasi deteksi ketepatan kelulusan mahasiswa dapat dikembangkan dengan menyesuaikan sistem akademik yang sudah ada sehingga dapat meningkatkan jaminan kualitas aplikasi yang ada sebelumnya.

4.7.2 Aspek Manajerial

a) Metode SVM-PSO yang diwujudkan dalam bentuk prototipe deteksi ketepatan waktu kelulusan mahasiswa dapat membantu pihak manajemen dalam hal ini adalah bagian akademik untuk dapat mengambil tindakan preventif terhadap mahasiswa yang mengalami keterlambatan kelulusan. Jika hal tersebut dilakukan dengan baik oleh manajemen diharapkan akan dapat meminimalisir penurunan angka mahasiswa yang telat waktu dan meningkatkan angka kelulusan mahasiswa sesuai dengan waktu yang sudah ditentukan oleh lembaga tersebut.

b) Prototipe deteksi ketepatan waktu kelulusan akan dapat membantu pihak staff IT dan pusat data dalam hal pengembangan sistem akademik yang telah ada sebelumnya.

4.7.3 Penelitian Lanjutan

a) Penelitian berikutnya dalam hal pengujian metode terhadap data akademik untuk mengklasifikasikan kinerja mahasiswa dapat dilakukan dengan menggunakan metode lainnya seperti metode SVM –GA atau membuat suatu perbandingan kinerja akurasi antara metode SVM – PSO dan SVM-GA sehingga akan dapat diketahui secara pasti metode yang memiliki akurasi yang lebih unggul dari keduanya dan dapat diterapkan untuk mengklasifikasikan dataset akademik.

b) Dengan menggunakan metode yang sama dapat dilakukan penelitian berikutnya berkaitan dengan klasifikasi mahasiswa magang atau kerja hal tersebut sangat penting karena sesuai dengan slogan yang dimiliki oleh LP3I yaitu **“Tepat & Cepat Kerja!”** Selain itu juga metode SVM dengan metode optimasinya dapat juga diterapkan tidak hanya pada domain dataset akademik saja akan tetapi dapat pula diterapkan pada dataset lainnya seperti : bidang perbankan, ketahanan, kedokteran, geofisika dll.

V. KESIMPULAN

Data *single year* khususnya data tahun ajaran 2008/2009 memiliki akurasi yang lebih tinggi daripada data lainnya sehingga dapat disimpulkan bahwa metode SVM-PSO sangat efektif jika menggunakan dataset dengan volume data kecil akan tetapi memiliki atribut terpilih secara selektif dan parameter yang sesuai.

Penerapan metode SVM – PSO yang diwujudkan dalam prototipe aplikasi ketepatan waktu kelulusan mahasiswa akan sangat membantu untuk meningkatkan jumlah kelulusan ditahun yang akan datang. Sehingga dapat meningkatkan kualitas lembaga pendidikan tersebut ke jenjang akreditasi yang lebih baik dari jenjang akreditasi sebelumnya.

DAFTAR REFERENSI

- [1] Ahmed, "Data Mining : A Prediction for Student's Performance Using Classification Method," *World Journal of Computer Application and Technology* 2, pp. (2) 43-47, 2014.
- [2] BAN-PT, *Akreditasi Institusi Perguruan Tinggi - Buku III Pedoman Penyusunan Borang.*, 2011.
- [3] Kopertis3. (2014) Peraturan Perundangan. [Online]. <http://www.dikti.go.id/id/peraturan-perundangan/>
- [4] Kabag Marketing & C N P, "Laporan Perkembangan Mahasiswa LP3i Depok TA. 2009 s/d 2013," Depok, 2014.
- [5] N. W. S Saraswati, "Text mining dengan metode naïve bayes classifier dan support vector machines untuk sentiment analysis," 2014.
- [6] Agustinus Jacobus, "Penerapan Metode Support Vector Machine pada Sistem Deteksi Intrusi secara Real-time," *IJCCS*, pp. Vol.8, No.1, January 2014, pp. 13~24, 2014.
- [7] ana Satsiou, "Genetic Algorithms for the Optimization of Support Vector Machines in Credit Risk Rating," 2002.
- [8] S. Yenaeng, "Automatic Medical Case Study Essay Scoring by Support Vector Machine and Genetic Algorithms," *IJIET*, pp. Vol. 4, No. 2., 2014.
- [9] Undang-Undang R.I nomor 12 tahun 2012. (2015, January) www.academia.edu. [Online]. http://www.academia.edu/2160186/Undang-Undang_R.I_nomor_12_tahun_2012_tentang_Pendidikan_Tinggi
- [10] S Azwar, *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Pelajar, 2004.
- [11] F.A. Hermawati, *Data Mining*. Yogyakarta: Andi Offset, 2013.
- [12] B. Santosa, *Data Mining Teknik Pemanfaat Data Untuk Keperluan Bisnis*. Yogyakarta : Graha Ilmu, 2007.
- [13] J., & Kamber, M. Han, *Data Mining Concepts and Techniques*. San Fransisco: Mofgan Kaufan Publisher., 2007.
- [14] A. S. Nugroho, "Support Vector Machine: Paradigma Baru Dalam Softcomputing," *Konferensi Nasional Sistem Dan Informatika* , pp. 92-99, 2008.
- [15] O. Maimon, *Data Mining And Knowledge Discovery Handbook*. New York Dordrecht Heidelberg London: Springer, 2010.
- [16] Y., Wang, G., Chen, H., Dong, H., Zhu, X., & Wang, S. Liu, "An Improved Particle Swarm Optimization for Feature Selection," *Journal of Bionic Engineering*, pp. 8(2), 191–200. doi:10.1016/S1672-6529(11)60020-6, 2011.
- [17] R. L., & Haupt, S. E. Haupt, *Practical Genetic Algorithms*. United State of America: A John Wiley & Sons Inc Publication, 2004.
- [18] A. S. H., Hussin, B., Ananta, I. G. P., & Zeniarja, J. Basari, "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization.," *Procedia Engineering*, pp. 53, 453–462. doi:10.1016/j.proeng.2013.02.059, 2013.
- [19] H Ian Witten, Frank Eibe, and A Mark Hall, *Data Mining Machine Learning Tools and Techniques*. Burlington: Morgan Kaufmann Publisher, 2011.
- [20] F. Gorunescu, *Data Mining: Concepts and Techniques*. Verlag berlin Heidelberg: Springer, 2011.
- [21] G. Polancic, "Empirical Research Method Poster," 2007.
- [22] M., Hansson, J., Olsson, B., & Lundell, B. Berndtsson, *A Guide For Students In Computer Science And Information Systems*. London: Springer, 2008.
- [23] C. W. Dawson, *Projects In Computing And Information System A Student's Guide*. England: Addison - Wesley, 2009.
- [24] Riduwan, *Metode dan Teknik Menyusun Tesis*. Bandung: Alfabeta, 2008.
- [25] C Vercellis, *Business Intelligence Data Mining And Optimization For Decision Making*. United Kingdom: A John Wiley And Sons, Ltd., Publication., 2009.

- [26] Frisma Handayana, "Penerapan Particle Swarm Optimization untuk Seleksi Atribut Pada Metode support Vector Machine untuk Prediksi Penyakit Diabetes ," 2012.
- [27] Sugiyono, *Metode Penelitian Kuantitatif Kualitatif dan R & D*. Bandung: Alfabeta, 2011.

Biodata Penulis

Endang Supriyadi, memperoleh gelar Diploma III (D3), Jurusan Manajemen Informatika Universitas Gunadarma Depok, lulus tahun 1998 dan memperoleh gelar Sarjana Agama (S.Ag) Jurusan Tarbiyah Perguruan Tinggi Ilmu Al-qur'an Jakarta (PTIQ) lulus tahun 1999. Memperoleh gelar Magister Komputer

(M.Kom) Program Pasca Sarjana Magister Ilmu Komputer Konsentrasi e-Business STMIK Nusa Mandiri Jakarta, lulus tahun 2015. Saat ini menjadi Dosen di LP3I Jakarta "Kampus Depok", STIAMI Jakarta & Depok, STIMA IMMI Cibubur.

Dana Indra Sensuse, memperoleh gelar Sarjana dari Institut Pertanian Bogor tahun lulus 1985. Gelar MLIS diperolehnya dari Dalhousie University – Kanada pada tahun 1994. Sedangkan gelar Ph.D diraih dari University of Toronto tahun, 2004. Saat ini menjadi tenaga pengajar pada Universitas Indonesia, dengan bidang penelitian E-government, Knowledge Management, dan Sistem Informasi.