

**PENERAPAN *PARTICLE SWARM OPTIMIZATION* UNTUK
SELEKSI ATRIBUT PADA METODE *SUPPORT VECTOR
MACHINE* UNTUK PREDIKSI PENYAKIT DIABETES**



TESIS

FRISMA HANDAYANNA

14000369

**PROGRAM PASCA SARJANA MAGISTER ILMU KOMPUTER
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER**

NUSA MANDIRI

JAKARTA

2012

**PENERAPAN *PARTICLE SWARM OPTIMIZATION* UNTUK
SELEKSI ATRIBUT PADA METODE *SUPPORT VECTOR
MACHINE* UNTUK PREDIKSI PENYAKIT DIABETES**



TESIS

Diajukan sebagai salah satu syarat untuk memperoleh gelar

Magister Ilmu Komputer (M.Kom)

FRISMA HANDAYANNA

14000369

**PROGRAM PASCA SARJANA MAGISTER ILMU KOMPUTER
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER**

NUSA MANDIRI

JAKARTA

2012

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini :

Nama : Frisma Handayanna

NIM : 14000369

Program Studi : Magister Ilmu Komputer

Jenjang : Strata Dua (S2)

Konsentrasi : Management Information System

Penerapan *Particle Swarm Optimization* Untuk Seleksi Atribut Pada Metode *Support Vector Machine* Untuk Prediksi Penyakit Diabetes

Dengan ini menyatakan bahwa tesis yang telah saya buat dengan judul: “Penerapan *Particle Swarm Optimization* Untuk Seleksi Atribut Pada Metode *Support Vector Machine* Untuk Prediksi Penyakit Diabetes” adalah hasil karya sendiri, dan semua sumber baik yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar dan tesis belum pernah diterbitkan atau dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa tesis yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri dicabut/dibatalkan.

Jakarta, 29 Agustus 2012

Yang menyatakan,

Materai Rp. 6.000,-

Friska Handayanna

HALAMAN PENGESAHAN

Tesis ini diajukan oleh :

Nama : Frisma Handayanna
 NIM : 14000369
 Program Studi : Magister Ilmu Komputer
 Jenjang : Strata Dua (S2)
 Konsentrasi : Management Information System
 Judul Tesis : “Penerapan *Particle Swarm Optimization* Untuk Seleksi Atribut
 Pada Metode *Support Vector Machine* Untuk Prediksi Penyakit Diabetes”

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Ilmu Komputer (M.Kom) pada Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri).

Jakarta, 14 September 2012
 Pascasarjana Magister Ilmu Komputer
 STMIK Nusa Mandiri
 Direktur

Prof. Dr. Ir. Kaman Nainggolan, MS

D E W A N P E N G U J I

Penguji I : Ir. Dana Indra Sensuse, MLIS, Ph.D

Penguji II : Windu Gata, M.Kom

Penguji III / Pembimbing : Romi Satria Wahono, M.Eng

KATA PENGANTAR

Puji syukur alhamdulillah, penulis panjatkan kehadiran Allah, SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan tesis ini tepat pada waktunya. Dimana tesis ini penulis sajikan dalam bentuk buku sederhana. Adapun judul tesis, yang penulis ambil sebagai berikut: ” Penerapan *Particle Swarm Optimization* Untuk Seleksi Atribut Pada Metode *Support Vector Machine* Untuk Prediksi Penyakit Diabetes”.

Tujuan penulisan ini dibuat sebagai salah satu syarat untuk mendapatkan gelar Magister Ilmu Komputer (M.Kom) pada Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri).

Tesis ini diambil berdasarkan hasil penelitian atau riset mengenai Penerapan *Particle Swarm Optimization* Untuk Seleksi Atribut Pada Metode *Support Vector Machine* Untuk Prediksi Penyakit Diabetes. Penulis juga mencari dan menganalisa berbagai macam sumber referensi, baik dalam bentuk jurnal ilmiah, buku-buku literatur, internet, dan lain-lain yang berkaitan dengan pembahasan pada tesis ini.

Penulis menyadari bahwa tanpa bimbingan dan dukungan dari semua pihak dalam pembuatan tesis ini, maka penulis tidak dapat menyelesaikan tesis ini tepat pada waktunya. Untuk itu ijinkanlah pada kesempatan ini untuk mengucapkan ucapan terimakasih yang sebesar-besarnya kepada:

1. Allah SWT yang selalu mencurahkan nikmat dan rahmatNya pada saya sehingga saya dapat menyelesaikan tesis ini tepat pada waktunya.
2. Bapak Romi Satrio Wahono, M.Eng selaku pembimbing tesis yang telah menyediakan waktu, pikiran dan tenaga dalam membimbing penulis sampai penulis dapat menyelesaikan tesis ini.
3. Isnaini Firdaus dan Elmiwati (Orang Tua), Edward Supiko, ST (kakak), dan Frisma Handayanny, S.Kom (Kembaran) yang selalu memberikan dukungan moral dan spiritual kepada penulis.
4. Mohammad Badrul, Rinawati, dan Titin Kristiana yang selalu memberi semangat, dukungan dan perjuangannya dalam mengerjakan tesis ini.

4. Hilda Amalia, M.Kom yang telah memberikan bantuan dan dukungannya.
5. Hendra Widiarta, S.Kom yang telah memberikan bantuan dan dukungannya.
6. Seluruh staf pengajar (dosen) Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri yang telah memberikan pelajaran yang berarti bagi penulis selama menempuh studi.
7. Seluruh staf dan karyawan Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri yang telah melayani penulis dengan baik selama kuliah maupun selama penulis menyusun tesis.
8. Rekan-rekan angkatan periode V Pascasarjana Magister Ilmu Komputer STMIK Nusa Mandiri.

Serta semua pihak yang terlalu banyak untuk penulis sebutkan satu-persatu sehingga terwujudnya penulisan tesis ini. Penulis menyadari bahwa penulisan tesis ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan karya ilmiah yang penulis hasilkan untuk yang akan datang.

Akhir kata semoga tesis ini dapat bermanfaat bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, Agustus 2012

Frisma Handayanna

Penulis

**SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA
ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini, saya :

Nama : Frisma Handayanna
NIM : 14000369
Program Studi : Magister Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Management Information System
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, dengan ini menyetujui untuk memberikan ijin kepada pihak Program Pascasarjana Magister Ilmu Komputer Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri (STMIK Nusa Mandiri) Hak **Bebas Royalti Non-Eksklusif (Non-exclusive Royalti-Free Right)** atas karya ilmiah kami yang berjudul : Penerapan *Particle Swarm Optimization* Untuk Seleksi Atribut Pada Metode *Support Vector Machine* Untuk Prediksi Penyakit Diabetes, beserta perangkat yang diperlukan (apabila ada).

Dengan **Hak Bebas Royalti Non-Eksklusif** ini pihak STMIK Nusa Mandiri berhak menyimpan, mengalih-media atau bentuk-kan, mengelolanya dalam pangkalan data (database), mendistribusikannya dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak STMIK Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 29 Agustus 2012
Yang menyatakan,

Materai Rp. 6.000,-

Frisma handayanna

ABSTRAK

Nama : Frisma Handayanna
NIM : 14000369
Program Studi : Magister Ilmu Komputer
Jenjang : Strata Dua(S2)
Konsentrasi : Management Information System
Judul Tesis :” Penerapan *Particle Swarm Optimization* Untuk Seleksi Atribut Pada Metode *Support Vector Machine* Untuk Prediksi Penyakit Diabetes”

Penyakit diabetes adalah salah satu penyakit yang dapat menyebabkan komplikasi bahkan dapat menyebabkan kematian. Saat ini penyakit diabetes semakin lama semakin meningkat jumlah penderitanya. Banyak penelitian yang menggunakan metode *support vector machines* dalam memprediksi penyakit diabetes tetapi nilai akurasi yang dihasilkan masih kurang akurat. Dalam penelitian ini dibuatkan model algoritma *support vector machines* dan model algoritma *support vector machines* berbasis *Particle Swarm Optimization* untuk mendapatkan rule dalam memprediksi penyakit diabetes dan memberikan nilai akurasi yang lebih akurat. Setelah dilakukan pengujian dengan dua model yaitu Algoritma *support vector machines* dan *support vector machines* berbasis *Particle Swarm Optimization* maka hasil yang didapat adalah algoritma sehingga didapat pengujian dengan menggunakan *support vector machines* dimana didapat nilai *accuracy* adalah 74.21 % dan nilai AUC adalah 0.758, sedangkan pengujian dengan menggunakan *support vector machines* berbasis *Particle Swarm Optimization* didapatkan nilai *accuracy* 77.36% dan nilai AUC adalah 0.765 dengan tingkat diagnosa *good classification*. Sehingga kedua metode tersebut memiliki perbedaan tingkat akurasi yaitu sebesar 3.15% dan perbedaan nilai AUC sebesar 0,017.

Kata kunci:

Diabetes, Algoritma *Support Vector Machine*, Seleksi Atribut, *Particle Swarm Optimization*

ABSTRACT

Nama : Frisma Handayanna
NIM : 14000369
Program Studi : Magister Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : Management Information System
Judul Tesis : "Application of Particle Swarm Optimization for Feature Selection In Support Vector Machine Method To Predict Diabetes"

Diabetes is a disease that can cause complications even can causes death. At this time diabetes has increased the number of its patient. A lot of research using the method of support vector machines to predict diabetes but the value of the resulting accuracy is less accurate. In this research made model algorithm support vector machines and model algorithm support vector machines based on Particle Swarm Optimization to get the rule to predict the disease diabetes and give a more accurate value of the accuracy. After the testing with two models that support vector machines algorithms and support vector machines based on Particle Swarm Optimization and so test the results are by using support vector machines are get accuracy values 74.21% and AUC values was 0.758, while testing by support vector machines based particle swarm optimization are get value accuracy 77.36 % and value AUC is 0.765 to level diagnose good classification. The two this method having the different levels of accuracy is as much as 3.15 % and the difference in value AUC of 0,017.

Keywords:

Diabetes, Support Vector Machine Algorithm, Feature Selection, Particle Swarm Optimization

DAFTAR ISI

	Halaman
HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
HALAMAN PERNYATAAN ORISINALITAS	iii
HALAMAN PENGESAHAN	iv
KATA PENGANTAR	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS.....	vii
ABSTAK.....	viii
ABSTACT.....	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang Penulisan.....	1
1.2 Identifikasi Masalah.....	3
1.3 Problem Statement.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Ruang Lingkup Penelitian.....	4
1.7 Sistematika Penulisan	4
BAB II LANDASAN TEORI	6
2.1 Tinjauan Pustaka	6
2.1.1 Diabetes.....	6
2.1.2 Data Mining	8
2.1.3 Support Vector Machine	12
2.1.4 Feature Selection.....	16
2.1.5 Particle Swarm Optimization	17

2.1.6 Pengujian Evaluasi dan Validasi Metode Klasifikasi Data Mining.....	18
2.1 Tinjauan Studi	22
2.3 Kerangka Pemikiran	24
BAB III METODOLOGI PENELITIAN.....	27
3.1 Desain Penelitian	27
3.2 Pengumpulan Data.....	28
3.3 Pengolahan Data Awal.....	29
3.3 Model yang Diusulan.....	30
3.3 Eksperimen dan Pengujian Metode	31
3.6 Evaluasi dan Validasi Hasil	31
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	33
4.1 Hasil Eksperimen dan Pengujian Metode	33
4.1.1 Support Vector Machine	33
4.1.2 Support Vector Machine berbasis Berbasis Particle Swarm Optimization.....	33
4.2 Evaluasi dan Validasi Hasil	34
4.2.1 Hasil Pengujian Support Vector Machine.....	34
4.2.2 Hasil Pengujian Support Vector Machine Machine Berbasis Particle Swarm Optimization	37
4.2.3 Analisis Hasil Pengujian Support Vector Machine	40
4.3 Pembahasan.....	41
4.3 Implikasi Penelitian	42
BAB V PENUTUP	43
5.1 Kesimpulan	43
5.2 Saran....	43
DAFTAR LAMPIRAN.....	44
LAMPIRAN	

DAFTAR TABEL

Tabel 2.1	AND Problem.....	15
Tabel 2.2	Model <i>Confusion Matrix</i> (Vercellis, 2009).....	19
Tabel 2.3	<i>State of the art</i> prediksi penyakit menggunakan <i>Support Vector Machine</i>	24
Tabel 3.1	Atribut dan data penyakit diabetes	28
Tabel 3.2	Atribut yang digunakan.....	29
Tabel 3.3	Spesifikasi <i>hardware</i> dan <i>software</i>	31
Tabel 4.1	Eksperiment penentuan nilai <i>training cycle</i> SVM	33
Tabel 4.2	Eksperiment penentuan nilai <i>training cycle</i> SVM berbasis PSO	33
Tabel 4.3	Model <i>Confusion Matrix</i> untuk Metode <i>Support Vector Machine</i>	35
Tabel 4.4	Nilai <i>accuracy</i> , <i>sensitivity</i> , <i>specificity</i> , <i>ppv</i> , dan <i>npv</i> untuk Metode <i>Support Vector Machine</i>	36
Tabel 4.5	Model <i>Confusion Matrix</i> untuk Metode <i>Support Vector Machine</i> berbasis <i>Particle Swarm Optimization</i>	38
Tabel 4.6	Nilai <i>accuracy</i> , <i>sensitivity</i> , <i>specificity</i> , <i>ppv</i> , dan <i>npv</i> untuk Metode <i>Support Vector Machine</i> berbasis <i>Particle Swarm Optimization</i>	39
Tabel 4.7	Pengujian algoritma SVM dan SVM berbasis PSO	40

DAFTAR GAMBAR

Gambar 2.1	Tahapan Data Mining	9
Gambar 2.2	SVM Berusaha Menemukan Hyperplane Terbaik Yang Memisahkan Kedua Class Negatif Dan Positif	13
Gambar 2.3	Taksonomi masalah pengurangan dimensi.....	16
Gambar 2.4	K-fold Cross-validation	19
Gambar 2.5	Grafik ROC (discrete dan continuous).....	22
Gambar 2.6	Kerangka Pemikiran.....	26
Gambar 3.1	Metode yang diusulkan	30
Gambar 4.1	Desain Model Validasi	34
Gambar 4.2	Model Pengujian validasi Support Vector Machine	35
Gambar 4.3	Kurva ROC dengan Metode <i>Support Vector Machines</i>	39
Gambar 4.4	Model pengujian validasi <i>support vector machine</i> berbasis algoritma <i>Particle Swarm Optimization</i>	37
Gambar 4.5	Kurva ROC dengan <i>support vector machine</i> berbasis <i>Particle Swarm Optimization</i>	39
Gambar 4.6	Atribute weight dengan <i>support vector machine</i> berbasis <i>Particle Swarm Optimization</i>	40
Gambar 4.7	Kurva ROC dengan <i>support vector machine</i> dan <i>support vector machine</i> berbasis <i>Particle Swarm</i> <i>Optimization</i>	41

DAFTAR LAMPIRAN

Lampiran 1	Data penyakit diabetes	47
------------	------------------------------	----

BAB I

PENDAHULUAN

1.1 Latar Belakang Penulisan

Perkiraan terakhir populasi penderita penyakit diabetes menunjukkan 171 juta orang di dunia pada tahun 2000 dan diperkirakan akan meningkat menjadi 366 juta pada 2030 (Report WHO, 2006). Penyakit diabetes disebabkan oleh peningkatan kadar glukosa dalam darah, apabila kadar glukosa darah meningkat dalam jangka waktu yang lama maka akan menyebabkan komplikasi seperti gagal ginjal, kebutaan dan serangan jantung (Jayalskshmi & Santhakumaran, 2010). Kontrol glukosa darah merupakan hal terpenting dalam praktek medis penyakit diabetes dan penyakit kritis lainnya (Iancu, Iancu, & Sfredel, 2010). Kelainan darah diabetes dan gula lain disebabkan oleh apa yang kita makan dan bagaimana cara kita hidup (Mason, 2005).

Penyakit diabetes merupakan salah satu penyakit yang mematikan, faktor resiko tinggi dalam keluarga yang menyebabkan penyakit diabetes antara lain dikarenakan orang gemuk yang tidak melakukan latihan fisik, dan orang-orang yang memiliki gaya hidup yang tidak sehat dan makanan yang berlebih dari apa yang dibutuhkan oleh tubuh (Nuwangi, Oruthotaarachchi, Tilakaratna, & Caldera, 2010). Sehingga untuk menghindari penyakit diabetes diupayakan kita memiliki gaya hidup yang sehat serta tidak makan berlebihan dari apa yang diperlukan oleh tubuh.

Penyakit diabetes perlu diprediksi dengan akurat karena penyakit diabetes merupakan penyakit sosial yang serius dan bisa terkena orang dalam jumlah besar, serta menyebabkan komplikasi dan melibatkan biaya yang tinggi serta dapat meningkatkan keadaan sakit melalui penyakit diabetes terutama pada anak-anak dan anak muda (Iancu, Mota, & Iancu, 2008).

Penelitian terlebih dahulu mengenai penyakit diabetes telah banyak dilakukan seperti penelitian yang dilakukan oleh T.Jayalskshmi dan A.Santhakumaran yang berjudul *Impact of Preprocessing for Diagnosis of Diabetes Mellitus Using Artificial Neural Networks* menggunakan Neural Network untuk diagnosa penyakit diabetes mellitus dengan menggunakan metode prosedur pra-pengolahan dan nilai-nilai yang hilang mempengaruhi set data selama klasifikasi dan hasilnya dampak dari nilai yang hilang teknik dan pra-pengolahan teknik yang membuktikan bahwa beberapa kombinasi nilai-nilai yang hilang dan pra-pengolahan akurasi yang sangat meningkat (Jayalskshmi & Santhakumaran, 2010).

Penelitian selanjutnya dilakukan oleh Jianchao Han, Juan C. Rodriguze, Mohsen Beheshti dengan judul *Diabetes Data Analysis and Prediction Model Discovery Using RapidMiner* menggunakan model ID3 dan Decison Tree, dengan metode Decision tree yang dilakukan adalah memisahkan sumber data set ke dalam subset berdasarkan uji nilai atribut dan dengan ID3 yaitu dengan pendekatan di mana pohon keputusan yang secara rekursif dibangun dengan top-down dan dimulai dengan dengan memulai serangkaian pelatihan tuple dan label kelas yang terkait, dimana hasilnya akurasi ID3 lebih tinggi dari pada Decison Tree (Han, Rodriguze, & Beheshti, 2008).

Selanjutnya penelitian yang dilakukan oleh He Chunjian dan Zhang Cuilian Zhao Yan dengan judul *A New SVM Merged into Data Information*, dengan metode kernel fungsi dimana beberapa kernel dilatih dan kernel terbaik tampil di set validasi kemudian dipilih untuk pengujian dan kinerjanya dievaluasi pada set tes dan menunjukkan bahwa pendekatan secara efektif dapat meningkatkan klasifikasi akurasi (Chunjiang & Yan, 2009).

Neural Network mempunyai kelebihan dalam hal kemampuan generalisasi tergantung pada seberapa baik Neural Network meminimalkan resiko empiris namun Neural Network mempunyai kelemahan dimana menggunakan data pelatihan cukup besar (Vapnik, 1999). Decison tree dan ID3 mempunyai

kelebihan untuk keputusan pengklasifikasi memiliki akurasi yang baik namun memiliki kelemahan karena perlu mengumpulkan lebih banyak data (Han, Rodriguze, & Beheshti, 2008). SVM adalah kasus khusus dari keluarga algoritma yang kita sebut sebagai regularized metode klasifikasi linier dan metode yang kuat untuk minimalisasi resiko (Weiss, Indurkha, & Zhang, 2010). Dan kelebihan SVM lainnya adalah dapat meminimalkan kesalahan melalui memaksimalkan margin dengan misahkan antara hyper-plane dan satu set data bahkan dengan jumlah sample yang kecil (Chunjiang & Yan, 2009).

Namun demikian masalah aplikasi tertentu, tidak semua fitur ini sama-sama penting dan kinerja yang lebih baik dapat dicapai dengan membuang beberapa fitur dengan begitu fitur dalam SVM memiliki pengaruh penting dalam akurasi klasifikasi (Zhao, Fu, Ji, Tang, & Zhou, 2011). Dataset yang tidak penting, fitur yang banyak atau sangat berhubungan secara signifikan akan mengurangi tingkat akurasi klasifikasi dengan menghapus fitur ini, dengan begitu tingkat akurasi efisiensi dan klasifikasi dapat diperoleh (Lin a, Shiue b, & Chen, 2009).

PSO sangat menarik untuk pemilihan fitur dimana kawanan partikel akan menemukan kombinasi fitur terbaik pada saat pencarian ruang masalah dan PSO dapat menemukan solusi yang optimal dengan cepat (Parimala & Nallaswamy, 2012). PSO banyak digunakan untuk memecahkan masalah optimasi, serta sebagai masalah seleksi fitur (Liu, Wang, Chen, Dong, Zhu, & Wang, 2011). Dalam teknik 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*.

1.2 Problem Statement

Support Vector Machine dapat menyelesaikan masalah *decision tree* khususnya sampel data yang kecil yang ada di pima diabetes, tetapi *Support Vector Machine* memiliki kelemahan pada sulitnya pemilihan fitur yang sesuai dan optimal pada bobot atribut yang digunakan sehingga menyebabkan tingkat akurasi prediksi menjadi rendah.

1.3 Rumusan masalah

Rumusan masalah yang diangkat pada penelitian ini adalah seberapa besar akurasi metode *Support Vector Machine* yang ditingkatkan dengan *Particle Swarm Optimization* (PSO) dengan cara melakukan seleksi atribut untuk prediksi penyakit diabetes?

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah menerapkan *Particle Swarm Optimization* (PSO) untuk seleksi atribut yang optimal dan membobot atribut dari *dataset* pada metode *Support Vector Machine* (SVM) guna meningkatkan akurasi prediksi penyakit diabetes.

1.5 Manfaat penelitian

- a. Manfaat praktis dari hasil penelitian ini adalah dapat digunakan oleh pihak medis untuk penentuan penyakit dengan lebih baik.
- b. Manfaat kebijakan dari hasil penelitian ini adalah dapat digunakan sebagai bahan pertimbangan dalam penentuan penyakit diabetes.
- c. Manfaat teoritis dari penelitian ini yaitu diharapkan dapat memberikan sumbangan untuk pengembangan teori yang berkaitan dengan penerapan *Particle swarm optimization* pada *Support Vector Machine* untuk meningkatkan akurasi prediksi penyakit diabetes.

1.6 Ruang Lingkup Penelitian

Ruang lingkup pembahasan dalam penelitian ini dibatasi pada perbandingan metode *Support Vector Machine* dengan *Support Vector Machine* berbasis *Particle swarm optimization* (PSO) dengan cara menganalisis atribut menjadi bobot atribut (*attribute weighting*) dalam prediksi penyakit diabetes.

1.7 Sistematika Penulisan

Disajikan dalam lima bab dan masing-masing bab terdiri dari beberapa sub bab yaitu sebagai berikut:

Bab I Pendahuluan

Bab ini membahas tentang latar belakang penulisan, identifikasi permasalahan, rumusan masalah, tujuan penelitian, ruang lingkup penelitian dan hipotesis.

Bab II Landasan Teori

Bab ini membahas tentang landasan teori yang melandasi penelitian.

Bab III Metode Penelitian

Bab ini berisi tentang metode penelitian yang membahas tentang perancangan penelitian, tahap *computing approach* dan pengembangan sistem.

Bab IV Hasil dan Pembahasan

Bab ini berisi tentang hasil dan pembahasan yang menguraikan tentang implementasi sistem, pengukuran serta implikasi penelitian.

Bab V Kesimpulan dan Saran

Bab ini membahas kesimpulan dari penelitian dan saran untuk penelitian selanjutnya.

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Dalam penulisan tesis ini penulis melakukan tinjauan dengan menggunakan buku dan jurnal yang berhubungan dengan tema yang dipilih. Secara lebih detail tinjauan dalam penulisan tesis ini dapat dijelaskan sebagai berikut :

2.1.1 Diabetes

Diabetes adalah epidemi yang paling cepat berkembang di Barat dunia. Satu dari tiga anak-anak Amerika akan tumbuh dan terkena diabetes, 24% orang dewasa Amerika resisten insulin dan 45% orang dewasa di atas usia 60 resisten insulin (Mason, 2005). Diabetes adalah salah satu penyebab utama kematian di banyak negara dan penyebab utama kebutaan, gagal ginjal, dan *nontraumatic* amputasi (Robert, Zgonis, & Driver, 2006).

Faktor penyebab diabetes adalah (Nurrahmani, 2012):

1. Gen diabetes dalam keluarga

Gen merupakan sel pembawa sifat yang dapat diwariskan orang tua kepada keturunannya, dan diabetes merupakan penyakit yang bisa diwariskan.

2. Insulin dan gula darah

Insulin adalah karena ketidakmampuan beta sel-sel di pankreas untuk memproduksi insulin (Mason, 2005). Produksi ini disebabkan oleh tingginya kadar gula dalam darah, sehingga menyebabkan insulin diproduksi semakin tinggi.

3. Kegemukan (Obesitas)

Pada kegemukan atau obesitas sel-sel lemak yang menggemuk yang jumlahnya lebih banyak dari pada keadaan tidak gemuk, sehingga menyebabkan resistensi terhadap insulin dimana gula darah sulit masuk kedalam sel, sehingga gula darah tetap tinggi (hiperglikemi) sehingga terjadilah diabetes, khususnya terjadi pada diabetes tipe 2.