IMPLEMENTASI ALGORITMA *SUPPORT VECTOR MACHINE* UNTUK KLASIFIKASI PASIEN PENDERITA DIABETES MELLITUS

RISMA PUSPITA

H1091201060

SKRIPSI

PROGRAM STUDI STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS TANJUNGPURA

PONTIANAK

2025

IMPLEMENTASI ALGORITMA *SUPPORT VECTOR MACHINE* UNTUK KLASIFIKASI PASIEN PENDERITA DIABETES MELLITUS

RISMA PUSPITA
H1091201060

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Statistika pada Program Studi Statistika

PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS TANJUNGPURA
PONTIANAK
2025

LEMBAR PENGESAHAN

Judul Tugas Akhir	:	Implementasi Algoritma *Support Vector Machine* untuk Klasifikasi Pasien Penderita Diabetes Mellitus
Nama Mahasiswa	:	Risma Puspita
Jurusan/Program Studi	:	Matematika/Statistika
Tanggal Lulus	:	02 Juni 2025
SK Pembimbing	:	No. 2579/UN22.8/TD.06/2024/Tanggal 2 September 2024
SK Penguji	:	No. 721/UN22.8/TD.06/2025/Tanggal 3 Maret 2025

Dosen Pembimbing

Pembimbing I	Pembimbing II
Hendra Perdana, M.Sc	Shantika Martha, M.Si
NIP 198810102019031020	NIP 198403082008122003

Dosen Penguji

Ketua Penguji	Anggota Penguji
Dr. Evy Sulistianingsih. M.Sc	Dr. Evi Noviani. M.Si., CFrA
NIP 198502172008122006	NIP 198402292006042001

Hendra Perdana, M.Sc	Shantika Martha, M.Si
NIP 198810102019031020	NIP 198403082008122003

Mengesahkan
Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Tanjungpura

Prof. Dr. Gusrizal, M.Si
NIP 197108022000031001

PERNYATAAN INTEGRITAS AKADEMIK

Saya yang bertanda tangan di bawah ini :

Nama	: Risma Puspita
NIM	: H1091201060
Program Studi/Jurusan	: Statistika/Matematika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa dokumen ilmiah Tugas Akhir yang disajikan ini tidak mengandung unsur pelanggaran integritas akademik sesuai Peraturan Menteri Pendidikan, Kebudayaan, Riset, Dan Teknologi Republik Indonesia No 39 Tahun 2021. Apabila ada kemudian hari dokumen ilmiah Tugas Akhir ini mengandung unsur pelangaran integritas akademik sesuai ketentuan perundangan tersebut, maka saya bersedia menerima sanksi akademik dan/atau sanksi hukum yang berlaku.

Demikian pernyataan ini untuk digunakan sebagaimana mestinya.

Pontianak, Juni 2025

Risma Puspita
NIM. H1091201060

Implementasi Algoritma *Support Vector Machine* untuk Klasifikasi Pasien Penderita Diabetes Mellitus

Abstrak

Diabetes Mellitus merupakan kondisi di mana tubuh gagal mengontrol kadar glukosa dalam darah. Penyakit ini terjadi karena tubuh tidak mampu menghasilkan insulin dalam jumlah yang cukup, sehingga menyebabkan peningkatan kadar gula darah tetap tinggi dan dapat menyebabkan dampak serius bagi kesehatan pasien. Dalam upaya memahami dan menganalisis masalah Diabetes Mellitus secara lebih mendalam sistem klasifikasi yang akurat untuk mendeteksi penyakit ini, diperlukan pendekatan yang tepat, salah satunya melalui metode berbasis Machine Learning. Untuk mengidentifikasi faktor-faktor yang mempengaruhi klasifikasi pasien diabetes mellitus serta menentukan tingkat akurasi algoritma Support Vector Machine (SVM) dalam proses klasifikasi maka dari itu penelitian ini menggunakan dua jenis kernel yaitu kernel lienar dan kernel Radial Basic Function (RBF). Data yang digunakan merupakan data sekunder dari 982 pasien di Puskesmas Gang Sehat, dengan variabel yang dianalisis meliputi jenis kelamin, usia, tekanan darah, serta Indeks Massa Tubuh. Hasil menunjukkan bahwa kernel RBF memiliki akurasi lebih tinggi 80,2% dibandingkan kernel linear 77,8%, sehingga kernel RBF lebih efektif dalam menangani pola data dalam penelitian ini.

Kata Kunci : Akurasi, *Kernel Linear, Kernel RBF*

Implementation of the Support Vector Machine Algorithm for the Classification of Diabetes Mellitus Patients

Abstract

Diabetes mellitus is a condition in which the body fails to regulate glucose levels in the blood effectively. This disease occurs because the body is unable to produce sufficient insulin, leading to persistently high blood sugar levels that can have serious health consequences for patients. To better understand and analyze the problem of diabetes mellitus, an accurate classification system is essential for detecting this disease, and one effective approach is through machine learning-based methods. This study aims to identify factors influencing the classification of diabetes mellitus patients and evaluate the accuracy of the Support Vector Machine (SVM) algorithm in the classification process. Two types of kernels are used: the linear kernel and the Radial Basis Function (RBF) kernel. The dataset consists of secondary data from 982 patients at Puskesmas Gang Sehat, with variables including gender, age, blood pressure systolic and diastolic, and Body Mass Index. The results indicate that the RBF kernel achieves a higher accuracy of 80.2% compared to the linear kernel, which has an accuracy of 77.8%. This suggests that the RBF kernel is more effective in handling the data patterns in this study.

Keywords : Accuracy, Linear Kernel, RBF Kernel

PRAKATA

Puji Syukur Allah SWT, Tuhan yang Maha Esa atas Rahmat dan karunia-Nya penulis bisa menyelesaikan penulisan Tugas Akhir yang berjudul "Implementasi Algoritma Support Vector Machine untuk Klasifikasi Pasien Penderita Diabetes Mellitus". Adapun penelitian ini dibuat untuk memenuhi syarat penulis untuk memperoleh gelar Sarjana Statistika pada Program Studi Statistika di Fakultas Matematika dan Ilmu Pengetahuan Alam di Universitas Tanjungpura. Penulisan ini dapat diselesaikan tidak terlepas dari bimbingan, dukungan, saran, dan dukungan dari berbagai pihak. Oleh karena itu, penulis berterima kasih kepada :

1. Kedua orangtua penulis Bapak Hariadi dan Ibu Juliana. Terima kasih atas do'a, pengorbanan, cinta, motivasi, kepercayaan, semangat serta dukungan dalam hidup penulis. Kalian berdua sangat bearti dalam kehidupan penulis. Sehat selalu dan semoga Allah memberikan umur yang panjang dan hidup yang lebih lama lagi. Mamak dan bapak harus selalu ada disetiap perjalanan dan pencapaian hidup penulis.
2. Saudara kandung penulis, Rian Firdaus, Riski Febriadi dan Rizmi Asyfa Fadillah serta Kakak Ipar penulis Yeni Jumiarti dan Keponakan penulis Naadhira Syabilla Firdaus yang selalu memberikan dukungan serta motivasi sehingga penulis bisa sampai ke tahap ini. Semoga Allah selalu memberkahi setiap jalan dan diberikan kesehatan selalu.
3. Ibu Mariatul Kiftiah. M.Si dan Ibu Wirda Andani. M.Sc selaku Dosen Pembimbing Akademik yang memberikan motivasi, saran dan masukan selama masa perkuliahan.
4. Bapak Hendra Perdana, M.Sc. selaku dosen pembimbing pertama penulis yang telah memberikan waktu untuk membimbing, memberikan masukan serta motivasi yang bermanfaat bagi penulis dalam menjalankan proses penulisan skripsi.
5. Ibu Shantika Martha, M.Si,. selaku Dosen Pembimbing Kedua yang telah memberikan waktu untuk membimbing memberikan masukan serta motivasi yang bermanfaat bagi penulis dalam menjalankan proses penulisan skripsi.

6. Ibu Dr. Evy Sulistianingsih. M.Sc Selaku Dosen Penguji pertama yang telah memberikan saran dan masukkan untuk memperbaiki skripsi sehingga menjadi lebih baik.
7. Ibu Evi Noviani, M.Si., CFrA Selaku Dosen Penguji Kedua yang telah memberikan saran dan masukkan untuk perbaiki skripsi sehingga menjadi lebih baik.
8. Teruntuk Resti Arsanti, Leona Idilla dan Isna Firdayanti yang telah membersamai penulis dari awal perkuliahan tanpa henti. Terima kasih selalu memberikan semangat dan dorongan yang luar biasa kepada penulis disaat penulis sedih maupun senang.
9. Teruntuk Rika, Syania, Delvina, dan Ainna terima kasih telah membersamai penulis hingga saat ini. Terima kasih atas segala semangat, doa dan motivasi yang telah diberikan serta menjadi bagian penting selama penulis tinggal dan hidup diperantauan.
10. Teruntuk Sisi Rahayu dan Rinanti Putri terima kasih telah membersamai penulis dari SMA hingga saat ini. Terima kasih atas segala semangat, doa, dan motivasi yang telah diberikan kepada penulis.
11. Teruntuk teman-teman asrama Romi, Yuni, Rosi, Widia, Rizkiana, Zaskia terima kasih telah membersamai penulis selama tinggal diasrama, terima kasih atas segala semangat, bantuan, doa dan motivasi yang diberikan selama penulis tinggal diasrama.
12. Teman-teman Statistika Universitas Tanjungpura yang telah membersamai selama perkuliahan seta memberikan semangat dalam penulisan skripsi.
13. Serta orang-orang yang pernah mampir dikehidupan penulis terima kasih atas segala semangat dan motivasi yang diberikan kepada penulis.

Penulis berharap dapat memperoleh kritik dan saran yang bisa membangun dari semua pihak untuk kesempurnaan skripsi ini.

Pontianak, Juni 2025

Risma Puspita

DAFTAR ISI

LEMBAR PENGESAHAN iii
PERNYATAAN INTEGRITAS AKADEMIK iv
Abstrak v
Abstract vi
PRAKATA vii
DAFTAR ISI ix
DAFTAR TABEL xi
DAFTAR GAMBAR xii
DAFTAR LAMPIRAN xiii
DAFTAR SIMBOL xiv
BAB I PENDAHULUAN 1
1. 2 Rumusan Masalah 3
1. 3 Tujuan Penelitian 3
1. 4 Batasan Masalah 3
1. 5 Tinjauan Pustaka 4
1. 6 Metodologi Penelitian 5
BAB II LANDASAN TEORI 6
2. 1 Diabetes Mellitus 6
2. 2 Klasifikasi *Diabetes Mellitus* (DM) 6
2. 3 Kriteria Diabetes Mellitus 7
2. 4 Dat*a Training* dan Data *Testing* 8
2. 5 *Machine Learning* 9
2. 6 Akurasi dan *Confusion Matrix* 9
BAB III SUPPORT VECTOR MACHINE 11
3. 1 *Support Vector Machine* (SVM) 11
3. 2 *Support Vector Machine Linear* 12
3. 3 *Support Vector Machine Non Linear* 16
3. 4 Kelebihan Metode *Support Vector Machine Learning* 17
BAB IV HASIL DAN PEMBAHASAN 18
4. 1 Gambaran Umum Lokasi Penelitian 18
4. 2 Data dan Variabel Penelitian 18
4. 3 Analisis Statistik Deskriptif 20

4. 4 Analisis Metode Support Vector Machine (SVM)..... 22
4.4. 1 Pembentukan SVM *Kernel Linear*..... 22
4.4. 2 Persamaan *Hyperplane Kernel Linear*..... 22
4.4. 3 Pembentukan SVM Kernel RBF..... 25
4.4. 4 Perbandingan Dua Fungsi Kernel SVM..... 27
BAB V PENUTUP..... 29
5. 1 Kesimpulan..... 29
5.2 Saran..... 30
DAFTAR PUSTAKA..... 31
LAMPIRAN..... 33

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix Untuk Klasifikasi Dua Kelas 10
Tabel 3. 1 Fungsi Kernel 16
Tabel 4. 1 Deskriptif Data 18
Tabel 4. 2 Statistik Deskriptif 20
Tabel 4. 3 Parameter Model SVM *Kernel Linear* 24
Tabel 4. 4 Confusion Matrix Kernel Linear 24
Tabel 4. 5 Hasil Perhitungan Nilai pada SVM Model *Default Kernel RBF* 25
Tabel 4. 6 Parameter Model Kernel Radial 27
Tabel 4. 7Confusion Matrix Kernel Radial 27
Tabel 4. 8 Tabel Nilai Akurasi Fungsi Kernel 28

DAFTAR GAMBAR

Gambar 1. 1 *Flowchart Support Vector Machine* 5
Gambar 3. 1 (a) dan (b). SVM berusaha menemukan hyperplane terbaik yang memisahkan kelas -1 dan +1 12

DAFTAR LAMPIRAN

Lampiran 1 Surat Permohonan Data 33

DAFTAR SIMBOL

n : banyak nya sampel data

N : jumlah populasi

e : Presentase Tingkat kesalahan pengambilan sampel

$\boldsymbol{w}$: Vector bobot

b : Bias atau *error*

R : Vector bilangan real

$\|.\|$: Panjang *Vector*

a_i : Nilai *Lagrange Multiplier* pada data ke-i

x_i : data uji yang akan dilakukan prediksi

y_i : data uji yang akan dilakukan prediksi

SV : Jumlah *Support Vector*

ξi : Variabel *Slank*

ϕ : Fungsi *Kernel*

c : Nilai Parameter konstanta C*ost*

γ : Nilai Parameter konstanta *Gamma*

d : Nilai Parameter konstanta *Degree*

D : Himpunan Semesta

BAB I
PENDAHULUAN

1. 1. Latar Belakang

Kesehatan merupakan aspek penting dalam kehidupan, banyak permasalahan yang terjadi dalam peningkatan taraf kesehatan masyarakat yang berhubungan dengan gaya hidup. Gaya hidup seseorang kurang sehat dapat mengakibatkan munculnya gangguan kesehatan. Gangguan kesehatan yang kerap terjadi pada masyarakat salah satunya adalah Diabetes Mellitus (DM) yang merupakan penyakit yang disebabkan kadar gula darah yang tinggi. Hal ini menjadi tantangan yang berat pada sistem pelayanan Kesehatan di Indonesia, seperti peningkatan jumlah penderita DM, tingginya biaya pengobatan, keterbatasan akses pelayanan kesehatan, serta kurangnya kesadaran masyarakat dalam menerapkan pola hidup sehat (Zahtamal et al., 2007).

Dari berbagai penelitian di Indonesia yang dilakukan oleh pusat-pusat diabetes, sekitar tahun 1980-an prevalensi DM pada penduduk usia 15 tahun ke atas sebesar 1,5-2,3% dengan prevalensi di daerah rural/pedesaan lebih rendah dibandingkan perkotaan. Pada Survei Kesehatan Rumah Tangga (SKRT) 2001 mendapatkan prevalensi DM pada penduduk usia 25-64 tahun di Jawa dan Bali sebesar 7.5%. Prevalensi dari penderita DM cenderung pada perempuan dibandingkan dengan laki-laki dan terjadi peningkatan prevalensi penyakit DM sesuai dengan pertambahan umur namun mulai umur 65 tahun cenderung menurun.

Diabetes Mellitus (DM) didefinisikan sebagai suatu penyakit atau gangguan metabolisme kronis dengan multi etiologi yaitu penyakit dengan penyebab lebih dari satu yang ditandai dengan tingginya kadar gula darah disertai dengan gangguan metabolisme karbohidrat, lipid, dan protein sebagai akibat insufisiensi yaitu ketidakcukupan fungsi atau performa suatu organ, sistem tubuh atau zat tertentu dalam memenuhi kebututuhan fisiologi

tubuh fungsi insulin. Insufisiensi fungsi insulin dapat disebabkan oleh atau defisiensi yaitu kekurangan suatu elemen zat atau fungsi penting dalam tubuh produksi insulin oleh sel-sel beta Langerhans kelenjar pankreas, atau disebabkan oleh kurang responsifnya sel-sel tubuh terhadap insulin (WHO, 1999). Diabetes Mellitus merupakan salah satu penyakit tidak menular yang saat ini masih mengancam keberlangsungan hidup manusia. Gejala umum penyakit Diabetes Mellitus yaitu *polyphagia* (tingginya ritme rasa lapar), *polydipsia* (rasa haus yang tidak berhenti) dan *palyura* (keinginan buang air kecil yang terlalu tinggi) (Aljumah, Ahmad, dan Siddique, 2012).

Diabetes Mellitus merupakan masalah penyakit dimana tubuh pasien tidak dapat mengontrol tingkat konsentrasi glukosa pada darahnya. Penderita tidak dapat menghasilkan insulin pada jumlah cukup sehingga kadar gula dalam darah cukup tinggi, jika kondisi ini berlanjut maka akan menimbulkan bahaya bagi tubuh dan untuk itu, diperlukan metode yang mampu mengklasifikasikan pasien berdasarkan data kesehatan yang tersedia. Salah satu metode machine learning yang efektif adalah Support Vector Machine (SVM). Oleh karena itu, penelitian ini bertujuan untuk mengimplementasikan metode SVM dalam membangun sistem klasifikasi untuk membedakan pasien penderita dan non-penderita Diabetes Mellitus, serta mengukur akurasi dari berbagai kernel yang digunakan dalam model SVM tersebut

Support Vector Machine (SVM) adalah salah satu metode atau algoritma untuk melakukan klasifikasi maupun prediksi. Prinsip kerja dari metode ini adalah mencari ruang pemisah yang paling optimal dari suatu dataset dalam kelas yang berbeda (Sukmawati et al., 2013). SVM merupakan algoritma klasifikasi yang diatur pada *machine learning*. Dalam metode SVM, proses klasifikasi dilakukan dengan pendekatan *supervised learning* untuk mencari nilai *hyperplane* yang dimiliki margin paling optimal. SVM banyak diminati karena memiliki algoritma yang dapat menghasilkan akurasi yang bagus dengan daya komputasi yang sedikit. Karakteristik SVM secara garis besar adalah sebagai berikut :

1. SVM mampu mengenali pola data dengan melakukan transformasi dimensi ruang input ke ruang dimensi yang lebih tinggi untuk melakukan optimasi.
2. Menerapkan strategi *structural risk minimization* yang mencari *hyperplane* terbaik yang memiliki nilai margin paling optimal.
3. SVM merupakan *linear classifier* yang membagi dataset menjadi 2 bagian secara linear dengan menggunakan garis *hyperplane*.

Berdasarkan latar belakang yang ditulis maka penelitian ini menggunakan *support vector machine* untuk memecahkan klasifikasi berbagai algoritma pada *machine learning.*

1. 2 Rumusan Masalah

Berdasarkan latar belakang, maka diambil rumusan masalah sebagai berikut :

1. Bagaimana implementasi SVM untuk klasifikasi pasien penderita Diabetes Mellitus.
2. Bagaimana tingkat akurasi pada model SVM dalam klasifikasi penyakit Diabetes Mellitus.

1. 3 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan penelitian ini sebagai berikut :

1. Mengimplementasikan SVM kedalam klasifikasi pasien penderita Diabetes Mellitus.
2. Menentukan tingkat akurasi pada model SVM dalam klasifikasi penyakit Diabetes Mellitus.

1. 4 Batasan Masalah

Berdasarkan tujuan penelitian, maka batasan masalah penelitian ini sebagai berikut :

1. Data yang digunakan merupakan data pasien Diabetes Mellitus dan *non*-Diabetes Mellitus di Puskesmas Gang Sehat dengan variabel sebagai berikut: Jenis Kelamin, Umur, Hasil IMT, Sistole dan Diastole
2. *Fungsi Kernel* yang digunakan dalam penelitian ini adalah *Kernel RBF* dan *Kernel Linear*.

1. 5 Tinjauan Pustaka

Junus, Tarno dkk. (2023) membahas penerapan metode *support vector machine* dan *random forest* untuk mendeteksi risiko diabetes mellitus secara dini. Algoritma klasifikasi yang digunakan adalah SVM dan *Random Forest*, dengan analisis kinerja dilakukan pada pengklasifikasian data diabetes mellitus. Data yang digunakan dalam penelitian adalah data sekunder yang di ambil dari UCI *Machine Learning Repository* Bangladesh dengan 16 variabel independen dan 1 variabel terikat. Variabel terikat mengkategorikan hasil tes menjadi kelas diabetes mellitus positif dan negatif. Hasil analisis menunjukkan bahwa *Random Forest* memiliki kinerja terbaik dengan akurasi 98,08% sedangkan hasil analisis menggunakan metode *Support Vector Machine* menghasilkan nilai akurasi sebesar 91% .

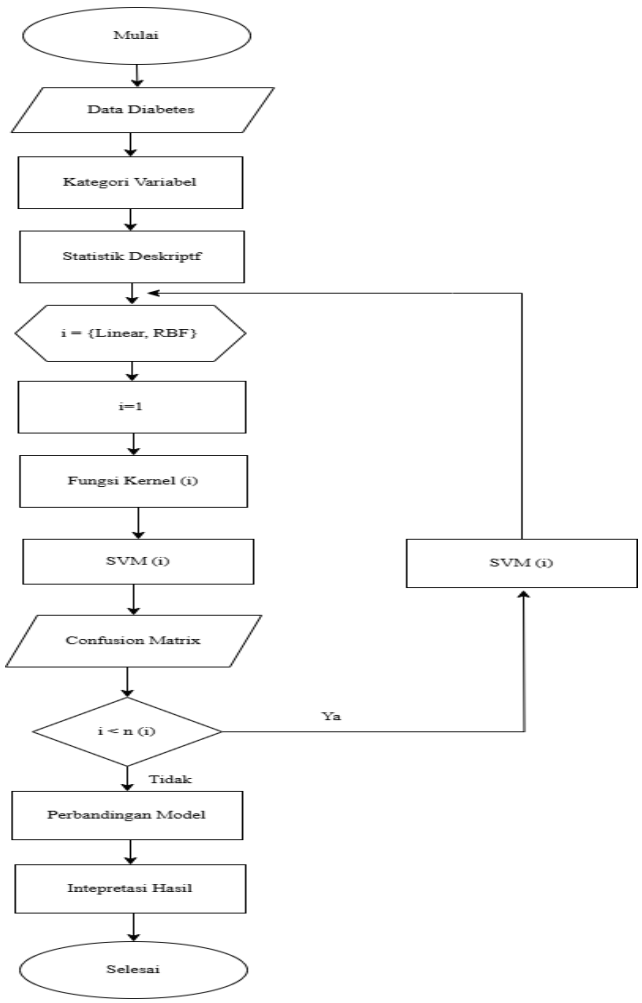
Lumbanraja, dkk. (2022) membahas implemetasi metode SVM dalam klasifikasi penderita diabetes mellitus. Penelitian ini bertujuan untuk mengembangkan model prediksi menggunakan R *Shiny* serta mengevaluasi kinerja klasifikasi metode SVM. Dataset yang digunakan adalah diabetes *130-US Hospital For Years 1999-2008*. Bersumber dari UCI *Machine Learning Repository* yang terdiri dari 34 variabel dan 84900 *records* teknik pembagian dan pengujian dataset dilakukan menggunakan SVM dengan Kernel Gaussian. Hasil penelitian menunjukkan bahwa model prediksi berbasis R *shiny* yang bertujuan untuk mempermudah pengguna dalam mengetahui prediksi dan mendapatkan hasil akurasi yaitu 82,76% dari *kernel gaussian*.

Prasatya dkk. (2020) membahas tentang penerapan metode K-Means dan C4.5 untuk prediksi diabetes. Penelitian ini melakukan prediksi HbA1C dengan mengkombinasikan metode K-Means dengan C4.5. Data yang digunakan dalam

penelitian ini yaitu Diabetes *130-US Hospitals For Years 1999-2008 Dataset*. Hasil dari prediksi yang didapatkan dilakukan dengan melalui proses validasi untuk mengetahui tingkat akurasi yang menggunakan *K-Fold Cross Validation.* Nilai yang didapat sebesar 72%

1. 6 Metodologi Penelitian

Penelitian ini menggunakan data sekunder dari pasien Diabetes Mellitus yang diperoleh dari Puskesmas Gang Sehat. Data yang dikumpulkan terdiri dari pasien yang terdiagnosis Diabetes Mellitus serta pasien yang tidak terdiagnosis Diabetes Mellitus. Teknik analisis data yang diterapkan dalam penelitian ini menggunakan metode *Support Vector Machine* (SVM), sebuah teknik *machine learning* yang digunakan untuk klasifikasi.



Gambar 1. 1 *Flowchart Support Vector Machine*