**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNOLOGI INFORMASI**

**INSTITUT TEKNOLOGI SEPULUH NOPEMBER**

**USULAN TUGAS AKHIR**

# IDENTITAS PENGUSUL

**NAMA : Astris Dyah Perwita**

**NRP : 5110100178**

**DOSEN WALI : Victor Hariadi, S.Si, M.Kom**

**DOSEN PEMBIMBING : 1.Dr. Chastine Fatichah, S.Kom, M.Kom  
 2. Rully Soelaiman, S.Kom, M.Kom**

# JUDUL TUGAS AKHIR

“Penerapan Metode SVM dan *Simulated Annealing* Pada Klasifikasi Penyakit Hepatitis”

# LATAR BELAKANG

Viral hepatitis adalah salah satu masalah kesehatan dunia yang sangat penting. Selama ini diagnosis hepatitis diterapkan selama tes darah yang dilakukan secara rutin atau saat mendonasikan darah. Dokter ahli biasanya menyimpulkan diagnosis seseorang dengan cara mengevaluasi hasil tes terkini seorang pasien atau membandingkan kondisi yang sama dengan pasien yang lain. Oleh karena itu, cukup sulit untuk seorang dokter untuk mendiagnosis hepatitis. Berdasarkan alasan-alasan di atas, kini banyak sistem cerdas dan teknik penggalian data yang dipelajari dan dikembangkan untuk mendiagnosis penyakit hepatitis secara otomatis.

SVM adalah salah satu algoritma yang dapat digunakan untuk mendiagnosis hepatitis karena memiliki kemampuan prediksi yang relatif baik. Namun SVM memiliki beberapa kelemahan antara lain harus memperhatikan aspek-aspek seperti jenis kernel dan parameter. parameter yang ada dalam SVM juga perlu diperhatikan agar dapat menghasilkan hasil yang optimal contohnya parameter *C* dan . Algoritma SA merupakan salah satu algoritma yang dapat mengoptimasi metode SVM dalam menentukan parameter. SA akan menghasilkan parameter terbaik untuk metode SVM sehingga menghasilkan akurasi yang lebih baik pula.

Untuk mengevaluasi metode SVM-SA, tugas akhir ini akan membandingkan metode SVM-SA dengan metode klasifikasi lain. Hasil eksperimen yang didapatkan pada studi ini diharapkan mempunyai hasil akurasi yang lebih tinggi daripada metode klasifikasi yang lain.

# RUMUSAN MASALAH

Permasalahan yang dibahas dalam tugas akhir ini meliputi hal-hal sebagai berikut.

* 1. Bagaimana proses penggabungan metode SVM dan SA sehingga menghasilkan sistem cerdas yang bekerja dengan baik?
  2. Bagaimana metode SVM-SA dapat digunakan untuk mengklasifikasi penyakit hepatitis?
  3. Bagaimana menganalisis kinerja metode SVM-SA?

# BATASAN MASALAH

Batasan masalah yang ada pada tugas akhir ini antara lain

* 1. Dataset yang digunakan adalah dataset yang didapatkan dari <http://archive.ics.uci.edu/ml/datasets/hepatitis>.
  2. Metode yang digunakan adalah SVM digabung dengan algoritma SA
  3. Parameter yang dioptimasi adalah parameter C dan

# TUJUAN PEMBUATAN TUGAS AKHIR

Tujuan dari tugas akhir ini antara lain

* 1. Membangun sebuah metode gabungan antara SVM dan SA
  2. Membuktikan bahwa metode SVM-SA dapat mengklasifikasi penyakit hepatitis
  3. Membuktikan bahwa kinerja metode SVM-SA dapat menghasilkan akurasi yang lebih baik daripada metode klasifikasi lain terhadap diagnosis penyakit hepatitis.

# MANFAAT TUGAS AKHIR

Manfaat yang diharapkan dari tugas ini antara lain

* 1. Mengaplikasikan kakas bantu untuk mempermudah diagnosis penyakit hepatitis
  2. Mendapatkan metode gabungan yang akurasinya lebih baik daripada metode yang selama ini digunakan.

# TINJAUAN PUSTAKA

**SVM**

*Support Vector Machine* (SVM) yang pertama kali diperkenalkan oleh Vapnik pada tahun 1995 merupakan salah satu sistem cerdas yang dibangun berdasarkan teori statistik. SVM merupakan salah satu metode klasifikasi yang paling efektif. Tujuan utama SVM adalah untuk mendapatkan keputusan yang memiliki jarak margin paling jauh dari titik data pelatihan yang terdekat. Prinsip dasar klasifikasi SVM yang pertama adalah memetakan masukan yang berupa vektor ke dalam ruang fitur (*feature space*) secara linear maupun non-linear dan sesuai dengan jenis kernel yang dipakai.

Diberikan data pelatihan {(*xi , yi* )|*yi* =1 or − 1*, i* =1*, . . . , N.*} untuk klasifikasi dua-kelas (*N* adalah jumlah sampel pelatihan, *xi* adalah masukan atau *input* vektor, dan *yi* adalah label kelas), penampang pemisah yang terbentuk dari SVM ditunjukkan dalam bentuk persamaan

(1)

K adalah fungsi kernel yang merupakan ruang fitur, b adalah nilai bias, adalah nilai yang didapat dari permasalahan *quadratic programming* yang nantinya akan menghasilkan nilai margin *hyperplane* yang maksimum, dibatasi dengan . Performa klasifikasi SVM tergantung pada kombinasi beberapa parameter antara lain kapasitas parameter *C*, tipe kernel *K*, dan beberapa parameter lain yang berkoresponden. [[1](#Sar11)]

Pada teori-teori optimasi SVM biasanya melibatkan beberapa parameter. Salah satunya adalah parameter *C* yang mengontrol hasil keluaran agar margin maksimal dan kesalahan minimal. Parameter lain terbentuk pada pemetaan non-linear ruang fitur. Parameter-parameter ini dikenal sebagai *kernel parameters.* Secara sederhana, proses yang dilakukan biasanya memungkinkan parameter *C* dikenal sebagai *kernel parameter* sehingga semua parameter dapat diperlakukan sebagai satu kerangka. [[2](#Cha02)]

**SA**

*Simulated annealing* adalah teknik untuk menemukan solusi yang tepat pada permasalahan optimasi dengan cara mencoba variasi acak terhadap keadaan saat ini. Teknik ini merupakan hasil generalisasi dari metode Monte Carlo yang memeriksa *equation of state* dan *frozen state of n-body system*. Masalah yang membutuhkan pendekatan SA adalah masalah-masalah optimasi kombinatorial yang mempunyai ruang pencarian solusi yang sangat besar.

*Pseudocode* algoritma ini dapat dijabarkan sebagai berikut

[[1](#Sar11)]

s s0; p P(s)



// keadaan awal, presisi

sbest  s; pbest p



//keadaan awal untuk solusi terbaik

k 0; kmax Constant value



MaxScore Constant value



//hitung evaluasi

while k < kmax dan p <= MaxScore

{

//selama waktu tersisa dan hasil blm maksimal

Snew Neighbor(s)



//pilih *neighbor*

Pnew P(snew)



//hitung presisi

If exp(pnew - p) > Random() then

S snew; pbest pnew



If pnew >pbest

Sbest snew; pbest pnew



K k+1



}

Return sbest, pbest

**Dataset**

Dataset hepatitis yang diproses pada tugas akhir ini didapatkan dari UCI Respository of Machine Learning Databases (<http://archive.ics.uci.edu/ml/datasets/hepatitis>).Dataset hepatitis akan menentukan kondisi pasien akan meninggal atau tidak. Dataset ini digunakan untuk memprediksi keberadaan penyakit hepatitis berdasarkan hasil berbagai tes kesehatan yang dilakukan pasien. Dataset ini berisi 155 sampel, 32 (20,6%) sampel meninggal dan 123(79.4%) sampel hidup. Terdapat 19 atribut, 13 atribut berupa atribut biner, dan 6 atribut berupa 6-8 nilai diskrit. Beberapa atribut ini memiliki nilai yang hilang atau tidak memenuhi kriteria. Oleh karena itu, setelah dilakukan pengurangan terhadap atribut yang tidak memenuhi syarat, terdapat 13 sampel hidup 67 sampel meninggal.

# RINGKASAN ISI TUGAS AKHIR

Secara garis besar, proses yang dilakukan pada tugas akhir ini ditunjukkan pada Gambar 1



Gambar 1. Bagan proses pengerjaan tugas akhir

1. **Praproses dataset**

Praproses dataset dilakukan untuk mengolah data agar dapat diproses dengan baik selama implementasi metode. Praproses dilakukan dengan cara merepresentasikan data-data tersebut dalam bentuk vektor dan mengubah nilai kategorial menjadi nilai angka.

1. **Normalisasi data**

Normalisasi data dilakukan agar data yang diproses tidak memiliki jarak terlalu jauh antara satu data dengan yang lainnya. Tujuan utama dari pengskalaan adalah untuk menghindari dominasi atribut yang memiliki skala besar terhadap atribut yang memiliki skala yang lebih kecil. Keuntungan lainnya adalah untuk menghindari kesulitan pada proses penghitungan.

1. **Implementasi algoritma SVM-SA**

Implementasi algoritma diterapkan 2 metode yang telah diajukan yaitu SVM dan SA. Implementasi dari SVM-SA ini dapat dijelaskan dalam bagan seperti pada Gambar 2.



Gambar 2. Bagan implementasi SVM-SA

1. **Evaluasi dan pengujian**

Pada tahap evaluasi dan pengujian, hasil implementasi akan dihitung akurasinya menggunakan *confussion matrix*. Dari metode evaluasi ini akan didapatkan tingkat kebenaran diagnosia yang direpresentasikan sebagai *sensitivity, specificity,* dan akurasi.

1. **Pembandingan akurasi**

Proses ini akan membandingkan akurasi metode SVM-SA dengan akurasi metode klasifikasi lain.

# METODOLOGI

## Penyusunan proposal tugas akhir

Proposal tugas akhir ditulis untuk mengajukan ide atas pengerjaan tugas akhir. Proposal tugas akhir juga mengandung proyeksi hasil dari ide tugas akhir yang diajukan

## Studi literatur

Pada proses ini dilakukan studi lebih lanjut terhadap konsep-konsep yang terdapat pada jurnal, buku, artikel, dan literatur lain yang menunjang. Studi dilakukan untuk mendalami konsep algoritma baik SVM maupun SA, karakteristik penyakit hepatitis, dan konsep lain yang berguna untuk menyelesaikan permasalahan yang muncul pada proses pengerjaan tugas akhir ini.

## Implementasi algoritma

## Implementasi merupakan tahap untuk membangun sistem tersebut. Terdapat dua macam metodeserta algoritma yang akan diimplementasikan yaitu metode SVM dan algoritma SA. Implementasi diproses menggunakan Phyton dan Matlab.

## Pengujian dan evaluasi

Performa dari algoritma yang diterapkan akan dievaluasi menggunakan *confussion matrix*. Dari metode evaluasi ini akan didapatkan tingkat kebenaran diagnosia yang direpresentasikan sebagai *sensitivity, specificity,* dan akurasi.

## Penyusunan Buku Tugas Akhir

Pada tahap ini dilakukan penyusunan laporan yang menjelaskan dasar teori dan metode yang digunakan dalam tugas akhir ini serta hasil dari implementasi aplikasi perangkat lunak yang telah dibuat. Sistematika penulisan buku tugas akhir secara garis besar antara lain:

1. Pendahuluan
   1. Latar Belakang
   2. Rumusan Masalah
   3. Batasan Tugas Akhir
   4. Tujuan
   5. Metodologi
   6. Sistematika Penulisan
2. Tinjauan Pustaka
3. Desain dan Implementasi
4. Pengujian dan Evaluasi
5. Kesimpulan dan Saran
6. Daftar Pustaka

# JADWAL KEGIATAN

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Tahapan | Tahun | | | | | | | | | | | | | | | | | | | | |
| September | | | | Oktober | | | | November | | | | Desember | | | | | Januari | | | |
| Penyusunan Proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Studi literature |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Implementasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Pengujian dan evaluasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Penyusunan buku |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

# 12. DAFTAR PUSTAKA

x

|  |  |
| --- | --- |
| [1] | Javad Salimi Sartakhti, Mohammad Hossein Zangooei, and Kourosh Mozafari, "Hepatitis disease diagnosis using a novel hybrid method based in support vector machine and simulated annealing (SVM-SA)," 2011. |
| [2] | Olivier Chapelle, Vladimir Vapnik, Olivier Bousquet, and Sayan Mukherjee, "Choosing Multiple Parameters for Support Vector Machines," in *Machine Learning*, Nello Cristianini, Ed. Netherlands: Kluwer Academic Publishers, 2002, ch. 2, pp. 131-159. |
| [3] | Zheng Jiang et al., "Support Vector Machine-Based Feature Selection for Classification of Liver Fbrosis Grade in Chronic Hepatitis C," 2006. |

x