**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNOLOGI INFORMASI**

**INSTITUT TEKNOLOGI SEPULUH NOPEMBER**

**USULAN TUGAS AKHIR**

# **IDENTITAS PENGUSUL**

Nama : **Ririn Uktarini**

NRP : **5108100192**

Dosen Wali : **Sarwosri, S.Kom, M.Kom**

1. **JUDUL TUGAS AKHIR**

***“Implementasi Metode Radial Basis Function Network Berbasis Time Variant Multi-Objective Particle***

***Swarm Optimization untuk Diagnosis Penyakit”***

1. **LATAR BELAKANG**

Diagnosis penyakit adalah hal yang perlu ditangani dengan benar. Diagnosis awal adalah cara awal untuk mencegah penyebaran penyakit semakin menjalar dalam tubuh. Setelah dilakukan diagnosis penyakit maka akan dilakukan permulaan pengobatan. Dalam setiap pengobatan penyakit akan berbeda dalam mengobatinya. Untuk itulah pentingnya melakuka diagnosis dengan benar. Salah satu metode yang digunakan untuk membantu dalam melakukan diagnosis penyakit adalah menggunakan *Artificial Neural Network (ANN).*

Metode *Artificial Neural Network (ANN)* adalah suatu metode komputasi yang meniru sistem jaringan syaraf biologis. Metode ini dapat memecahkan masalah klasifikasi karena adanya proses *learning*. *Radial Basis Function Network (RBFN)* adalah salah satu varian *ANN.*

*Radial Basis Function Network* berbasis *Time Variant Multi-Objective Particle Swarm Optimization (RBFN-TVMOPSO)* yang menggabungkan antara metode *PSO* sebagai seleksi fitur dan *RBFN* untuk *clustering*. Metode ini digunakan untuk membantu diagnosis penyakit. Ada dua keuntungan utama dari metode ini. Pertama, metode ini dapat diterapkan untuk semua masalah dunia nyata dan kedua, menunjukkan kinerja yang lebih baik dalam hal kesalahan (error), spesifisitas sensitifitas, dan akurasi untuk data set [1].

1. **TUJUAN TUGAS AKHIR**

Tugas akhir ini bertujuan mendesain dan mengimplementasikan sistem perangkat lunak dengan metode yang menggabungkan *Radial Basis Function Network (RBFN)* dan *Time Variant Multi-Objective Particle Swarm Optimization (TVMOPSO)* untuk diagnosis penyakit.

1. **MANFAAT TUGAS AKHIR**

Tugas akhir ini dikerjakan dengan harapan dapat memberikan manfaat yang besar

pada bidang kedokteran agar bisa mendiagnosis penyakit sebelum penyakit menyebar keseluruh organ tubuh dan menyebabkan kematian.

1. **RUMUSAN MASALAH**

Rumusan masalah yang diangkat untuk mengimplementasikan metode *Radial Basis Function Network berbasis Time Variant Multi-Objective Particle Swarm Optimization* ini adalah :

1. Bagaimana mengimplementasikan modul-modul dalam metode *RBF* *network* serbadasarkan *TVMOPSO,* yang meliputi:

a. *Radial Basis Function Network (RBFN)*

b. *Multi-Objective Optimization (MOO)*

c. *Time Variant Multi-Objective Particle Swarm Optimization (TVMOPSO)*

1. Bagaimana mengevaluasi pendiagnosisan penyakit menggunakn metode *RBFN* berbasis *TVMOPSO*
2. **BATASAN MASALAH**

Permasalahan yang dibahas dalam Tugas Akhir ini memiliki beberapa batasan, yaitu sebagai berikut :

1. Sistem perangkat lunak dibangun dengan menggunakan perangkat lunak MATLAB R2008a.
2. Dataset yang digunakan adalah data medis asli berupa data set penyakit kanker, diabetes, dan hepatitis. Ketiga data tersebut didapatkan dari UCI *Repository of Machine Learning Databases.*
3. **RINGKASAN TUGAS AKHIR**

Dalam implementasi tugas akhir ini digunakan tida dataset dari UCI *Repository of Machine Learning Databases.* Data set tersebut adalah data penyakit kanker, diabetes, dan hepatitis. Untuk kanker diklasifikasikan menjadi tumor benign atau malignant. Data ini diambil dari “Breast Cancer Wisconsin”, data asli dari koleksi *University of Wisconsin Hospital,* Madison dari Dr. William H. Wolberg[2]. 458 (65%) data setnya benign, 241 (34%) adalah malignant. Terdapat sembilan atribut/inputs (*clump thickness, uniformity of cell size and shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli andmitoses*) dan keluarannya adalah *benign* atau *malignant*. Sedangkan untuk data set penyakit diabetes adalah diagnosis dari individu Pima India berdasarkan data pribadi dan pemeriksaan kesehatannya. Data set ini terdapat delapan atribut/input (*no. of pregnant, plasma glucose concentration, diastolic blood pressure, triceps skin fold thickness, serum insulin, BMI,* dan *age*) dan kelas dua output (*diabetes positive* atau *diabetes negative*). Dan dataset terakhir, yaitu penyakit hepatitis. Data ini kompleks dan *noisy* karena banyak data yang *missing (*167 data *missing).* Atribut dari data set ini ads sembilan (age, sex, steroid, antivirals, fatigue, malaise, anorexia, liver big, liver ﬁlm, spleen palpable, spiders, ascites, varices, bilirubin, alk. phosphate, SGOT, albumin, protime dan histology) dan outputnya adalah *live* atau *die.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Data set | Attribute | Class | Samples | Input | Hidden | Output |
| Brest Cancer | 9 | 2 | 699 | 9 | 2 | 1 |
| Diabetes | 8 | 2 | 769 | 8 | 2 | 1 |
| Hepatitis | 19 | 2 | 155 | 19 | 2 | 1 |

Tabel 1 deskripsi data set

1. *Radial Basis Function Network*

*Radial Basis Function Network (RBFN)* adalah salah satu varian ANN yang paling penting dalam machine learning. RBFN merupakan feed-forward network dengan satu lapisan tersembunyi (hidden layer), yang disebut Radial Basis Functions (RBFs). Output dari RBFs adalah menampilkan nilai maksimum pada titik pusat dan menurunkan nilai output sebagai input yang meninggalkan pusat. Biasanya, fungsi Gaussian digunakan untuk fungsi aktivasi. RBFN terdiri dari tiga lapisan (layer): lapisan input (input layer), lapisan tersembunyi (hidden layer) dan lapisan output (output layer).

Pada lapisan input (input layer), jumlah neuron adalah sama dengan jumlah dimensi masukan. Neuron lapisan input (input layer) akan mengirimkan data ke lapisan tersembunyi (hidden layer), dan menghitung nilai RBFs yang diterima dari lapisan input (input layer). Nilai-nilai ini akan dikirim ke lapisan output (output layer) yang menghitung nilai-nilai penjumlahan linear dari tersembunyi neuron (hidden layer). Dalam studi ini, fungsi Gaussian digunakan sebagai RBF.

1

Dimana:

* adalah RBFs ke-j,
* x=(x1, x2,…, xd)T adalah input vektor,
* cj= ( c1j, c2j,…, cdj)T adalah vektor pusat ke-j,
* adalah luas parameter.

Output dari RBF network y adalah penjumlahan linear dari RBFs. Diberikan persamaan sebagai berikut:

2



Gambar 1 Struktur RBFN

Dimana:

* y adalah output dari RBF
* p adalah jumlah neuron pada lapisan tersembunyi (hidden layer)
* wj adalah bobot dari neuron ke-j ke lapisan output (output layer)

Untuk membangun RBFN, jumlah neuron pada lapisan tersembunyi (hidden layer) harus diatur, cj (pusat); j (lebar); dan wj (bobotnya) harus diestimasi. Dalam proses pembelajaran RBF, strukturnya akan ditentukan berdasarkan pengetahuan sebelumnya (*prior knowlege*) atau pengalaman ahli (*expert knowlege*). Parameter diestimasi menggunakan *clustering* atau *least mean squared method*. Di sisi lain, ada pendekatan di mana struktur jaringan dan parameternya diestimasi menggunakan perhitungan evolusioner .

2. *Multi-Objective Optimization (MOO)*

*Multiple Objective Optimization* (MOO) yakni *object* yang ingin dicapai lebih dari satu. Dalam MOO biasanya terdapat lebih dari satu solusi yang optimum untuk semua objek, konsekuensinya terdapat beberapa solusi yang optimal, dikenal dengan solusi *pareto-optimal*. *Pareto optimal set* adalah sebuah kumpulan solusi non dominan yang berhubungan satu sama lain ketika berpindah dari solusi pareto satu ke solusi lain. Kumpulan solusi *pareto-optimal* seringkali mengacu kepada solusi tunggal karena dapat diaplikasikan berdasarkan pada masalah yang terdapat dalam kehidupan nyata. *Pareto optimal set* mempunyai ukuran yang bervariasi, namun ukuran dari *Pareto set* biasanya bertambah seiring dengan bertambahnya jumlah dari *objevtives* (obyek), pada sebagian besar kasus, jumlah solusi *pareto-optimal* sangat banyak.

Masalah umum MOO terdiri dari sejumlah tujuan untuk dioptimalkan secara bersamaan dan berhubungan dengan sejumlah persamaan (*equality constraints*) dan kendala ketidakpersamaan (*inequality constraints*) . Hal ini dapat dirumuskan sebagai berikut:

Maksimum/minimum :

Subyek untuk 3

Dimana:

* fi adalah fungsi objek ke-i,
* gj adalah persamaan ke-j,
* hk adalah kendala ketidakpersamaan ke-k
* x adalah vektor keputusan yang merepresentasikan solusi
* M,N, dan K adalah jumlah dari objek, persamaan, dan ketidak persamaan kendala

Untuk masalah MOO, setiap dua solusi x1 dan x­2 bisa memilih salah satu dari dua kemungkinan yang ada – satu mendominasi yang lain atau tidak mendominasi yang lain. Dalam masalah minimalisasi, sebuah x1 dan x­2 mendominasi jika keduanya memenuhi hal berikut:

4

Operator ≤ dan < bisa disebut sebagai “kurang dari sama dengan dari” dan “kurang dari”, secara berurutan. Solusi yang tidak didominasi oleh yang lain disebut solusi *Pareto-optimal*. Secara umum, Solusi *Pareto-optimal* disebut sebagai *Pareto-optimal front.* Himpunan non-dominan diperlukan untuk didekatkan pada *true Pareto front* dan didistribusikan secara tidak *uniform*.

2. *Time Variant Multi-Objective Particle Swarm Optimization*

*Particle Swarm Optimazation (PSO)* adalah populasi berdasarkan teknik optimasi *stochastic* dikembangkan oleh Kennedy dan Eberhat (1995), terinspirasi dari kumpulan burung atau sekelompok ikan, dimana setiap individu diperlakukan sebagai partikel yang kecil dalam n-ruang dimensi dengan vektor posisi dan kecepatan vektor partikel i direpresentasikan sebagai Xi(t)=(Xi1(t), Xi2(t),…, Xin(t)) dan Vi(t)=(Vi1(t), Vi2(t),…, Vin(t)) Partikel digerakkan berdasarkan persamaan berikut:

5

6

Dimana:

* c1 dan c2 adalah koefisien kecepatan,
* vektor Pi = (Pi1, Pi2,…., Pin) adalah posisi terbaik sebelumnya (posisi diberikan dengan nilai fitness terbaik) dari partikel i diketahui sebagai personal posisi terbaik (**pbest**),
* vektor Pg =( Pg1, Pg2,…, Pgn) adalah posisi terbaik diantara personal posisi terbaik dari partikel dalam populasi dan diketahui sebagai global posisi terbaik (**gbest**)

Parameter r1 dan r2 adalah dua bilangan random terdistribusi uniform (0,1). Secara umum nilai Vid dibatasi dalam interval [-Vmax, Vmax]. Bobot inersia w, dikenalkan oleh Shi dan Eberhast umumnya digunakan untuk percepatan kecepatan algoritma secara konvergen.

TVMOPSO bersifat adaptif dengan memungkinkan parameter pentingnya ( bobot inersia dan koefisien akselerasi ) untuk menyesuaikan dengan iterasi. Sifat ini sangat membantu dalam hal mencari ruang yang lebih efisien. Parameter keragaman digunakan untuk memastikan keragaman di antara solusi dari yang *front non-dominant* di waktu yang sama secara konvergensi ke *front pareto optimal*. TVMOPSO memperluas algoritma dari satu tujuan PSO untuk menangani masalah MOO. Ini menggabungkan mekanisme perhitungan jarak kepadatan (*distance crowded*) ke dalam algoritma PSO, khususnya pada pemilihan global posisi terbaik (**gbest**), dalam metode penghapusan dari sebuah arsip eksternal solusi non-dominan dan menyesuaikan parameter PSO. Mekanisme jarak kepadatan bersama dengan operator mutasi mempertahankan keragaman solusi yang non-dominan di arsip eksternal.

Algoritma dari TVMOPSO adalah sebagai berikutnya:

Step 1. inisialisasi P, V, menghitung **pbest** dan **gbest**

Step 2. Inisialisai t=0

Step 3. Menyimpan vektor yang non-dominan yang ditemukan dari P dalam A

Step 4. Mengulang

* + - * 1. menghitung nilai jarak kerapatan setiap solusi yang tidak dominan dalam A
        2. Menurutkan solusi non-dominan dalam A berdasarkan nilai jarak kerapatan secara descending
        3. Mengulang

Pilih secara random global dari P dari bagian yang paling atas dari A yang sudah di urutkan dan simpan posisinya untuk **gbest**

Menyesuaikan parameter PSO (bobot inersia dan koefisien percepatan) pada iterasi ke t. Nilai dari wt diijinkan turun secara liniar dengan iterasi t, wt terdiri sebagai berikut:

7

Dimana

- max t adalah jumlah maksimal dari iterasi dan t adalah jumlah iterasi.

- nilai c1 turun dari inisial nilai c1i ke c1f sementara nilai c2 menjadi dinaikkan dari c2i ke c2.

Nilai persamaan berikut:

8

9

Modifikasi setiap titik menggunakan formula PSO dan **gbest**

10

11

Dimana:

* + - adalah bobot inersia di iterasi t,
    - adalah kecepatan baru partikel i
    - adalah posisi baru dari partikel i
    - dan adalah koefisien kecepatan pada iterasi t
    - dan adalah bilangan random pada range [0,1]
    - adalah posisi terbaik yang diraih partike i
    - adalah petunjuk global terbaik dari setiap solusi non-dominan.

Jika posisi sekarang diluar batas maka dilakukan *upper* atau *lower bound* dan kecepatannya dikalikan dengan -1 agar pencarian pada arah yang sebaliknya.

Jika (t<(max\*Pmut)) maka mutasi Pi+1 dimana maxt adalah jumlah maksimal dari iterasi dan Pmut adalah peluang mutasi

Evaluasi populasi Pi+1

1. Sampai jumlah populasi adalah N
2. Memasukkan semua nilai solusi non-dominan dalam P ke A. Jika termasuk non-dominan simpan solusi. Semua solusi dominan dalam arsip oleh solusi baru di hapus dari arsip. Jika arsip penuh maka solusi *direplace* menggunakan step dibawah ini:

Menghitung jarak kepadatan dalam A setiap solusi non-dominan dalam arsip A

Mengurutkan solusi non-dominan dalam A berdasarkan nilai jarak kepadatan secara *descending*

Memilih secara random partikel di posisi terbawah yang terdiri dari partikel yang paling padat dalam arsip maka *replace* itu dengan solusi baru.

1. Memodifikasi **pbest** dari setiap partikel P jika posisi sekarang dominan *pbest,* replace *pbest* dengan posisi sekarang
2. Menaikkan iterasi counter t

Step 5 sampai jumlah maksimal iterasi tercapai



Gambar 2 prosedur RBFN-TVMOPSO

Metode ini, RBFN direpresentasikan sebagai vektor dengan dimensi D terdiri dari koneksi (bobot dan pusat) dari jaringan. Dimensi parikel adalah:

12

Dimana:

* + - I adalah jumlah neuron input dari RBFN
    - H adalah jumlah neuron tersembunyi (*hidden layer*) dari RBFN
    - O adalah jumlah neuron outpur dari RBFN

Pusat dari RBFN adalah lapisan tersembunyi yang diinisialisasi dari algoritma *k-means clustering* dan bobot RBFN adaah inisialisasi dari algoritma Least Mean Squared (LSM) atau nilai random. Pada algoritma RBFN-TVMOPSO ini, partikel mewakili lapisan satu tersembunyi (*hidden layer*) RBFN dan PSO terdiri dari populasi satu lapisan tersembunyi jaringan

RBFN-TVMOPSO dievaluasi menggunakan akurasi berdasarkan *Mean Square Error (MSE)* persamaannya sebagai berikut,

13

Dimana:

* adalah output *network*

Dan menggunakan kompleksitas berdasarkan *sum of square weight*

14

Dimana :

* j=1,….Q adalah bobot dalam jaringan
* adalah jumlah bobot total

1. **METODOLOGI**

Metodologi yang akan dilakukan dalam Tugas Akhir ini memiliki beberapa tahapan, diantaranya sebagai berikut:

1. Penyusunan Proposal Tugas Akhir

Tahap awal untuk memulai pengerjaan Tugas Akhir adalah penyusunan Proposal Tugas Akhir. Pada proposal ini, penulis mengajukan gagasan pembuatan sistem untuk mendiagnosis penyakit.

1. Studi Literatur

Pada tahap ini dilakukan pencarian, pengumpulan, penyaringan, pembelajaran dan pemahaman literatur yang berhubungan dengan algoritma metode *RBFN-TVMPSO* serta pengumpulan dataset. Literatur yang digunakan dalam pengerjaan Tugas Akhir ini sebagian besar berasal dari internet berupa makalah ilmiah, tesis, artikel, materi kuliah, serta beberapa buku referensi.

1. Implementasi

Implementasi merupakan tahap untuk membangun sistem tersebut.

1. Pengujian dan Evaluasi

Pada tahap ini dilakukan uji coba terhadap sistem yang telah dibuat, mengamati kinerja sistem yang baru dibuat, serta mengidentifikasi kendala yang mungkin timbul.

5. Penyusunan Buku Tugas Akhir

Tahap terakhir merupakan penyusunan laporan yang memuat dokumentasi mengenai pembuatan serta hasil dari implementasi perangkat lunak yang telah dibuat.

1. **JADWAL PEMBUATAN TUGAS AKHIR**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No** | **Tahapan** | **Bulan** | | | | | | | | | | | | | | | |
| **Oktober** | | | | **November** | | | | **Desember** | | | | **Januari** | | | |
| 1 | Penyusunan Proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | Studi Literatur |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | Implementasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| a. modul untuk RBFN |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| b.modul untuk MO |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | c. modul untuk TVMOPSO |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | Pengujian dan Evaluasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | Penyusunan Buku Tugas Akhir |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

1. **DAFTAR PUSTAKA**
2. Sultan Noman Qasem, Siti Mariyam Shamsuddin. Radial Basis Function Network Based on time Variant Multi-Objective Particle Swarm Optimization for Medical Diseases Diagnosis. 2011.
3. Umi Fitria P, Tri Santoso, Prima Kristalina. Simulasi Coverage Pada Wireless Sensor Network dengan Menggunakan Algoritma Genetika Pareto.

**LEMBAR PENGESAHAN**

###### Surabaya, 17 Maret 2012

Menyetujui,

|  |  |
| --- | --- |
| Dosen Pembimbing I,  **(Isye Arieshanti, S.Kom, M.Phil)**  **( NIP. 197804122006042001 )** | Dosen Pembimbing II,  **(Ahmad Saikhu, S.SI., M.T)**  **(NIP. 197107182006041001)** |