

1 Pendahuluan

1.1 Latar Belakang

Berkembangnya penggunaan komputer dalam bidang manajemen data menyebabkan akumulasi data dalam jumlah sangat besar di beberapa organisasi. Apalagi dengan berkembangnya persepsi bahwa analisa terhadap data yang besar ini akan mengubah data pasif menjadi informasi yang berguna. Salah satu cara untuk melakukan hal itu adalah dengan menggunakan metode Data Mining atau *Knowledge Discovery in Databases*. Di dalam konsep Data Mining terdapat berbagai cara dan metode untuk mengekstrak informasi dari data yang besar.

Klasifikasi adalah salah satu metode untuk melakukan ekstraksi informasi. Data input atau yang biasa disebut dengan *training set*, terdiri dari banyak *record*, yang tiap *record*-nya mempunyai beberapa atribut. Setiap *record* ini juga mempunyai sebuah label kelas. Tujuan dari klasifikasi ini adalah untuk menganalisa data *input* dan membuat model yang akurat untuk setiap kelasnya berdasarkan data yang ada. Model kelas tersebut juga digunakan untuk mengklasifikasikan data tes lain/data tes baru untuk ditentukan label kelasnya.

Terdapat banyak algoritma yang dapat digunakan untuk melakukan klasifikasi data mining, salah satunya dengan Jaringan Syaraf Tiruan (JST). JST merupakan suatu arsitektur jaringan untuk memodelkan cara kerja sistem syaraf manusia (otak) dalam melaksanakan tugas tertentu. Pemodelan ini didasari oleh kemampuan otak manusia dalam mengorganisasi sel-sel penyusunnya atau *neuron*, sehingga memiliki kemampuan untuk melaksanakan tugas-tugas tertentu khususnya pengenalan pola dengan efektivitas jaringan sangat tinggi [9]. Dalam mencari arsitektur yang optimal bukanlah hal yang mudah dalam penggunaan JST. Jadi salah satu kelemahan JST adalah penentuan arsitektur yang optimal, yang dimaksud arsitektur adalah penentuan struktur dan bobot-bobot koneksi dalam JST.

Evolutionary Algorithms (EAs) adalah algoritma-algoritma optimasi yang berbasis evolusi dalam dunia nyata. Oleh karena itu EAs dapat digunakan dalam optimasi pencarian arsitektur yang optimal dari JST. Algoritma EAs yang digunakan dalam tugas akhir ini adalah *Evolution Strategies* (ES). Pemilihan algoritma ES disebabkan kecepatan proses ES lebih baik dibandingkan dengan *Genetic Algorithm*[9].

Setiap atribut data memiliki pengaruh yang berbeda-beda dalam pengklasifikasian data. Hal ini tergantung pada seberapa besar nilai keinformatifan atau kontribusinya suatu atribut dalam pengklasifikasian data. Sehingga diperlukan *feature selection* terhadap atribut yang akan dijadikan sebagai input dalam pengklasifikasian data menggunakan JST.

Pada tugas akhir penulis mencoba menganalisis metode klasifikasi JST yang dipadukan dengan ES dan melakukan proses *feature selection* pada saat *preprocessing*. Metode JST yang digunakan adalah Feedforward Networks dengan *Supervised Learning*.

1.2 Perumusan Masalah

Dapat dirumuskan beberapa masalah yang dapat diangkat melalui penelitian Tugas Akhir ini, yaitu :

1. Bagaimana menerapkan proses *feature selection* dengan *information gain*.
2. Bagaimana membuat suatu model *classifier* dengan menerapkan metode JST dan ES?
3. Bagaimana menganalisa sistem klasifikasi berdasarkan parameter-parameter masukannya pada metode ES (seperti ukuran populasi struktur, ukuran populasi bobot serta ukuran *mutation step size*) untuk mendapatkan struktur dan bobot yang optimal pada JST.

Batasan masalah dari penelitian Tugas Akhir ini yaitu :

1. Jenis JST yang digunakan adalah *Feedforward Networks* dengan *Supervised Learning*.
2. Arsitektur JST yang digunakan adalah satu *hidden layer*.
3. Proses *discretization* menggunakan dua nilai diskret pada saat *preprocessing*.
4. Terdapat dua data yang digunakan adalah data *Pima Indians Diabetes* dan *Breast Cancer* yang diambil dari www.ics.uci.edu [5]
5. Aplikasi yang dibangun bersifat *stand alone*.
6. Menggunakan MATLAB.

1.3 Tujuan

Tujuan dari Tugas Akhir ini adalah sebagai berikut :

1. Menerapkan *information gain* pada suatu data set.
2. Mengimplementasikan ES dalam menentukan struktur dan bobot-bobot yang optimal pada JST dalam membangun sebuah classifier.
3. Melakukan analisis terhadap performansi sistem dan parameter-parameter masukan ES untuk mendapatkan struktur dan bobot yang optimal pada JST.

1.4 Metodologi Penyelesaian Masalah

Metodologi yang digunakan dalam Tugas Akhir ini sebagai berikut :

1. Studi Literatur
 - a. Pencarian referensi dan sumber-sumber yang berhubungan dengan JST, ES dan *Feature Selection*.
 - b. Pendalaman materi JST, ES dan *Feature Selection*.
 - c. Pengumpulan data dari www.ics.uci.edu
2. Perancangan Sistem
Pada tahap ini dilakukan perancangan sistem yang akan diimplementasikan dengan menentukan parameter-parameter yang akan digunakan dalam pembangunan sistem.
3. Implementasi dan Training
Pada tahap ini dilakukan implementasi ES dalam mengoptimasi dan melatih JST sesuai dengan perancangan sistem.
4. Testing dan Analisis Hasil
Pada tahap ini dilakukan *testing* terhadap JST yang dihasilkan dari proses Implementasi dan Training, kemudian dilakukan analisis terhadap hasil *testing* dalam mengklasifikasikan data.
5. Penyusunan Laporan
Penyusunan laporan dalam bentuk buku Tugas Akhir dengan mengikuti kaidah penulisan yang berlaku dan berdasarkan hasil penelitian.

1.5 Sistematika Penulisan

Penulisan Tugas Akhir ini dibagi dalam lima bab, yang terdiri atas:

BAB 1 : Pendahuluan

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan, metodologi penyelesaian masalah dan sistematika penulisan yang digunakan dalam penyusunan Tugas Akhir.

BAB II : Landasan Teori

Bab ini memuat dasar teori yang mendukung dan mendasari penulisan tugas akhir, yaitu mengenai *discretization*, *feature selection*, Jaringan Syaraf Tiruan dan *evolution strategies*.

BAB III : Analisis dan Perancangan Sistem

Bab ini menguraikan gambaran umum sistem serta pemodelan sistem.

BAB IV : Implementasi dan Pengujian Sistem

Bab ini berisi uraian mengenai implementasi dan analisis hasil percobaan yang telah dilakukan.

BAB V : Penutup

Bab terakhir ini menjelaskan kesimpulan umum dari seluruh rangkaian penelitian yang dilakukan dan saran untuk pengembangan selanjutnya.