

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis multivariat adalah analisis statistika yang digunakan pada data yang memiliki lebih dari dua variabel secara bersamaan. Secara umum, Analisis multivariat berhubungan dengan metode-metode statistik yang melakukan analisis terhadap lebih dari dua variabel secara bersama-sama (simultan) pada setiap objek atau orang (Santoso, 2010, hlm. 7). Analisis multivariat digunakan karena pada kenyataannya masalah yang terjadi di lapangan tidak dapat diselesaikan dengan hanya melibatkan satu atau dua variabel saja, akan tetapi melibatkan banyak variabel.

Teknik analisis multivariat diklasifikasikan menjadi menjadi dua yaitu analisis dependensi dan analisis interdependensi. Analisis dependensi merupakan analisis untuk mengetahui hubungan antara variabel terikat dengan dua atau lebih variabel bebas. Contoh analisis dependensi, yaitu analisis varians (ANOVA), analisis varians multivariat (MANOVA), analisis kovarians (ANCOVA), analisis regresi berganda, dan analisis diskriman. Sedangkan analisis interdependensi adalah analisis untuk mengetahui hubungan antar variabel bebas. Contoh analisis interdependensi, yaitu analisis faktor, analisis kluster, MSD (*Multidimensional Scalling*), dan CA (*Categorical Analysis*) (Nuningsih, 2010, hlm.1).

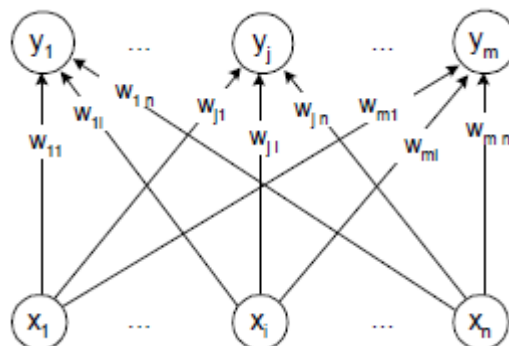
Analisis kluster merupakan salah satu teknik analisis multivariat yang bertujuan untuk mengalokasikan sekelompok objek pada suatu kelompok-kelompok yang saling bebas, yang disebut sebagai kluster, sehingga objek-objek didalam satu kelompok yang sama mirip satu sama lain, sedangkan objek-objek di dalam kelompok yang berbeda tidak mirip (Rachmatin, 2014, hlm. 134).

Nuningsih (2010, hlm.2) menjelaskan bahwa prinsip dasar dalam analisis kluster adalah mengelompokkan objek pada suatu kluster yang memiliki kemiripan sangat besar dengan objek lain dalam kluster yang sama, tetapi sangat tidak mirip dengan objek lain pada kluster yang berbeda. Hal ini berarti bahwa kluster yang baik adalah kluster yang mempunyai homogenitas (kesamaan) yang tinggi antar anggota dalam satu kluster (*within-cluster*) dan heterogenitas

(perbedaan) yang tinggi antara klaster yang satu dengan klaster yang lainnya (*between-klaster*) (Santoso, 2014, hlm.113).

Ada bermacam-macam metode untuk melakukan analisis klaster mulai dari metode yang sederhana diantaranya metode hierarki dan non-hierarki, hingga metode yang kompleks yaitu dengan menggunakan kecerdasan buatan, seperti jaringan syaraf tiruan (*artificial neural network*). Metode jaringan syaraf tiruan untuk melakukan *clustering* adalah metode jaringan syaraf (*neural network*) yang menggunakan *unsupervised learning* (pembelajaran tak terawasi), yaitu *Self-Organizing Map*.

Metode *Self-Organizing Map* (SOM) merupakan suatu metode pada Jaringan Syaraf Tiruan yang diperkenalkan oleh Teuvo Kohonen pada tahun 1980-an, sebagai salah satu bentuk topologi dari *Unsupervised Artificial Neural Network* dimana dalam proses pelatihannya tidak memerlukan pengawasan atau target keluaran (Anis&Isnanto, 2014, hlm. 48). Leleury dan Patty (2013, hlm.152) lebih lanjut menjelaskan bahwa pada jaringan SOM terdiri dari dua lapisan (*layer*), yaitu lapisan *input* dan lapisan *output*. Setiap neuron dalam lapisan *input* terhubung dengan setiap neuron pada lapisan *output*. Setiap neuron dalam lapisan *output* merepresentatifkan kelas (klaster) dari *input* yang diberikan. Setiap neuron *output* mempunyai bobot untuk masing-masing neuron *input*. Selama proses *clustering*, klaster (neuron pada lapisan *output*) yang memiliki jarak paling dekat dengan pola input akan terpilih sebagai pemenang dan beserta neuron tetangganya akan memperbaiki bobotnya. Berikut adalah struktur jaringan SOM (Siang, 2004, hlm. 142).



Gambar 1. 1 Struktur Jaringan Syaraf Tiruan pada SOM

Metode *Self-Organizing Map* (SOM) menghasilkan akurasi yang lebih baik dalam mengelompokkan objek ke dalam kelompok yang cocok dari pada

algoritma *K-Means* dan metode hirarki serta menunjukkan hasil yang baik saat menggunakan himpunan data yang kecil daripada himpunan data yang besar (Syaripudin, 2013, hlm.146). Selain itu, pada penelitian yang dilakukan oleh Yiheng Chen dan kawan-kawan (Suryaningsih, 2015, hlm.1) dengan membandingkan algoritma SOM dan algoritma *K-Means* dalam kasus *clustering* dokumen (data sampel yang digunakan adalah 420 artikel dari topik yang berbeda), teridentifikasi bahwa secara keseluruhan kinerja algoritma SOM lebih bagus daripada algoritma *K-Means* untuk kasus *clustering* dokumen. SOM tidak sensitif terhadap inisialisasi awal, ditunjukkan dengan *f-measure* yang stabil selama dua puluh kali percobaan. Sedangkan *K-Means* tidak stabil terhadap inisialisasi awal.

Proses *clustering* dengan menggunakan metode *Self Organizing Map* memerlukan waktu yang cukup lama serta dapat terjadi kesalahan dalam perhitungannya apabila dilakukan secara manual. Oleh karena itu, perhitungan *Self Organizing Map* memerlukan bantuan *software* yang juga dapat menerjemahkan algoritma *Self Organizing Map* ke dalam bahasa pemrograman sehingga perhitungan *Self Organizing Map* dapat dikerjakan dengan program yang telah dibuat sesuai algoritma. Wiharto (2013, hlm. 73) menjelaskan bahwa R adalah suatu perangkat lunak (*software*) yang digunakan untuk manipulasi data, perhitungan, simulasi, penayangan grafik, dan sekaligus sebagai bahasa pemrograman yang bersifat *interpreter* (penerjemah). R diturunkan dari bahasa S, suatu bahasa pemrograman yang dikembangkan di Laboratorium Bell. Oleh karena R berlisensi *open source* maka dapat diperoleh dan diedarkan secara cuma-cuma di bawah lisensi publik GNU. Lebih lanjut Zuur,*et. al* (Wiharto, 2013, hlm. 74) menjelaskan bahwa R merupakan bahasa pemrograman komputer yang memungkinkan pengguna untuk memrogramkan algoritma dan menggunakan alat yang telah dikembangkan melalui R oleh pengguna lainnya. Oleh karena hal tersebut, maka penulis tertarik untuk membuat skripsi yang berjudul “**Program Aplikasi Pengelompokan Objek dengan Metode *Self Organizing Map* Menggunakan Bahasa R**”.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara mengkontruksi program aplikasi untuk pengelompokan objek dengan metode *Self Organizing Map* menggunakan bahasa R?
2. Bagaimana hasil kontruksi program aplikasi untuk pengelompokan objek dengan metode *self organizing map* menggunakan bahasa pemograman R?
3. Bagaimana hasil pengaplikasian program untuk pengelompokan objek berdasarkan data Indeks Pembangunan Manusia Provinsi Aceh Tahun 2013?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan yang hendak dicapai dalam penelitian ini adalah:

1. Mengkontruksi program aplikasi untuk pengelompokan objek dengan metode *Self Organizing Map* menggunakan bahasa R.
2. Untuk mengetahui hasil kontruksi program aplikasi untuk pengelompokan objek dengan metode *self organizing map* menggunakan bahasa pemograman R.
3. Untuk mengetahui hasil pengaplikasian program pengelompokan objek berdasarkan data Indeks Pembangunan Manusia Provinsi Aceh Tahun 2013.

1.4 Manfaat Penelitian

Adapun manfaat penelitian yang diharapkan adalah sebagai berikut.

1.4.1 Manfaat Teoritis

Secara teoritis manfaat penelitian tugas akhir ini adalah untuk menambah ilmu dan wawasan baru mengenai pengelompokan objek dengan menggunakan metode *Self-Organizing Map* (SOM) dalam analisis kluster dan menambah wawasan mengenai bahasa pemograman R.

1.4.2 Manfaat Teoritis

Bagi penulis, penelitian ini diharapkan dapat menambah wawasan dan mengembangkan kemampuan dalam pembuatan program aplikasi untuk pengelompokan objek menggunakan metode *Self Organizing Map* dengan bahasa pemrograman R. Sedangkan bagi masyarakat umum, penelitian ini diharapkan dapat menjadi bahan pertimbangan dan referensi untuk penelitian selanjutnya yang berkenaan dengan analisis kluster.

1.5 Sistematika Penulisan

Isi dari skripsi ini akan terdiri dari beberapa bab dengan sistematika sebagai berikut:

BAB I PENDAHULUAN

Mengemukakan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penulisan, manfaat penulisan, dan sistematika penulisan.

BAB II KAJIAN PUSTAKA

Mengemukakan beberapa materi yang mendasari mengenai analisis kluster, metode *Self Organizing Map (SOM)*, dan bahasa pemrograman R.

BAB III METODE PENELITIAN

Mengemukakan metodologi penelitian, pembahasan tentang metode *Self Organizing Map(SOM)*, dan langkah-langkah pemrograman *Self Organizing Map* ke dalam bahasa R.

BAB IV PEMBAHASAN

Mengemukakan data yang dipakai dalam penelitian, pengujian asumsi analisis kluster terhadap data, pengaplikasian program aplikasi ke dalam data, dan validasi hasil program aplikasi yang dibandingkan dengan perhitungan manual.

BAB V PENUTUP

Kesimpulan dan saran dari hasil pembahasan materi.