**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNOLOGI INFORMASI**

**INSTITUT TEKNOLOGI SEPULUH NOPEMBER**

**USULAN TUGAS AKHIR**

# **IDENTITAS PENGUSUL**

Nama : **Tesa Eranti Putri**

NRP : **5107 100 012**

Dosen Wali : **Darlis Heru Murti, S.Kom., M.Sc.**

1. **JUDUL TUGAS AKHIR**

***Implementasi Analisis Weblog untuk Rekomendasi Link Favorit dengan Kombinasi Algoritma SOM dan Fuzzy***

1. **LATAR BELAKANG**

Dengan banyaknya situs internet yang ada sekarang, seringkali pengguna internet mengalami kesulitan untuk mengingat alamat *link* tertentu meskipun mereka sering mengunjunginya. Pada akhirnya mereka bergantung pada kecepatan dan ketepatan situs mesin pencari untuk menemukan alamat situs yang diinginkan dan akibatnya memakan waktu dua kali lebih lama.

Sebenarnya, *browser* yang ada sekarang telah menyediakan fitur yang sangat membantu bagi pengguna internet yang mengalami masalah ini. Fitur ini dikenal dengan fitur *speed dial*. *Speed dial* pada *browser* merupakan kumpulan *bookmark* yang ditampilkan pada halaman konten *browser* berupa kotak-kotak yang masing-masing menyimpan satu *bookmark* menuju sebuah situs. Beberapa *browser* dapat menampilkan lebih dari satu grup *speed dial* yang masing-masing ditampilkan dalam sebuah *tab*. Pengaturan *bookmark* yang ditugaskan pada setiap kotak ini dilakukan secara manual. Di satu sisi, fitur ini memberikan kebebasan pada penggunanya. Namun di sisi lain, cara ini tentunya tidak jauh beda dengan harus mencari alamat situs tersebut dan menambahkan *bookmark*-nya satu demi satu sehingga membuang waktu. Pada akhirnya, orang jadi cenderung untuk tidak menggunakannya.

Terkait dengan permasalahan tersebut, diperlukan adanya solusi berupa sebuah sistem perangkat lunak cerdas yang mempunyai potensi untuk mengatasinya. Salah satunya adalah dengan memanfaatkan *weblog* access yang dapat dipakai untuk dasar menganalisis perilaku jelajah pengguna. Analisis perilaku jelajah dilakukan dengan menggunakan kombinasi algoritma *clustering* *Self-Organizing Map* (SOM) dan *Fuzzy Inference System*. Algoritma *Self-Organizing Map* dipilih untuk proses mengklasterkan pola perilaku jelajah pengguna karena kelebihannya dalam tidak saja mengelompokkan situs-situs ke dalam klaster yang sesuai, tetapi juga dalam menampilkan representasi hubungan antar klaster secara visual. Sedangkan, *Fuzzy Inference System* dipilih untuk mengambil kesimpulan rekomendasi *link* yang sesuai dari pola perilaku jelajah pengguna yang sifatnya tidak pasti (*fuzzy*).

1. **RUMUSAN MASALAH**

Dengan mendasarkan pada potensi analisis preferensi dengan menggunakan kombinasi algoritma *Self-Organizing Map* (SOM) dan *Fuzzy Inference System*, maka rumusan masalah yang diangkat dalam Tugas Akhir ini dapat dipaparkan sebagai berikut :

* + - 1. Bagaimana melakukan praproses dan persiapan dataset *weblog* yang dipakai
      2. Bagaimana melakukan pengelompokan dataset *weblog* ke dalam klaster yang sesuai
      3. Bagaimana melakukan analisis dan rekomendasi untuk situs favorit berdasarkan hasil pengelompokan ke dalam klaster
      4. Bagaimana membuat sistem perangkat lunak yang mampu merepresentasikan hasil analisis dalam bentuk *bookmark* yang dapat diakses oleh pengguna

1. **BATASAN MASALAH**

Permasalahan yang dibahas dalam Tugas Akhir ini memiliki beberapa batasan. Dataset yang digunakan berasal dari sumber eksternal dan diinputkan secara manual ke dalam sistem perangkat lunak. Kemudian, data akses *weblog* yang digunakan terdiri atas alamat IP pengakses, waktu akses, *page request*, dan *volume request*.

1. **TUJUAN TUGAS AKHIR**

Tugas Akhir ini bertujuan mendesain dan membangun perangkat lunak sederhana yang mampu menghasilkan rekomendasi *link* situs favorit dari analisis *weblog* berupa daftar *bookmark* seperti dalam fungsi *speed dial* pada *browser* dengan menggunakan kombinasi algoritma *Self-Organizing Map* dan *Fuzzy Inference System*.

1. **MANFAAT TUGAS AKHIR**

Manfaat yang diharapkan dari Tugas Akhir ini yaitu mampu menghasilkan sebuah sistem aplikasi perangkat lunak yang memberikan kemudahan bagi pengguna untuk menjelajah internet. Kemudahan tersebut berupa pengaturan *bookmark* *link*-*link* situs favorit secara otomatis yang sesuai dengan pola perilaku jelajah pengguna, sehingga pengguna dapat menghemat waktu penjelajahan internet.

1. **RINGKASAN TUGAS AKHIR**

Pada Tugas Akhir ini, penulis akan membuat perangkat lunak yang digunakan untuk membuat rekomendasi *link* situs-situs internet favorit dari *weblog history* dengan kombinasi metode algoritma *Self-Organizing Map* (SOM) dan *Fuzzy Inference System*. Dengan metode ini, situs-situs tersebut akan dikelompokkan ke dalam grup-grup kecil berdasarkan frekuensi dan perilaku pengguna dalam menjelajahi internet dan dikembalikan sebagai kumpulan *bookmark* rekomendasi *link* situs yang siap digunakan sehingga berfungsi layaknya fitur *speed dial* pada *browser* biasa. Sistem perangkat lunak dikembangkan dalam dua tahap, yaitu tahap pembelajaran sistem dan pengujian sistem. Berikut adalah penjelasan mengenai masing-masing tahapan tersebut :

* Tahap pembelajaran sistem

Tahap pembelajaran sistem diawali dengan tahap analisis *weblog*, untuk mendapatkan pola jelajah akses internet dari pengguna yang akan dikelompokkan dalam klaster-klaster. Dalam Tugas Akhir ini, pengguna satu dan yang lainnya dibedakan dengan *IP Address* saat melakukan akses internet.

Tahap awal dari analisis *weblog* yaitu melakukan persiapan dataset weblog secara manual untuk mendapatkan dataset yang relevan dan mudah dianalisis. Selanjutnya data yang memiliki format teks direpresentasikan ke dalam format numerik agar dapat diproses oleh *Self-Organizing Map* (SOM). Setelah mendapatkan klaster pola akses jelajah situs dari SOM dari analisis weblog, tahapan berikutnya adalah proses inferensi atau pengambilan kesimpulan untuk prediksi rekomendasi *link* favorit dengan menggunakan *Fuzzy Inference System*. Hasil prediksi ini kemudian akan direpresentasikan dalam bentuk *bookmark* pada pengguna. Skema rancangan arsitektur sistem perangkat lunak rekomendasi *link* favorit ini adalah sebagai berikut :



Gambar Arsitektur Sistem Perangkat Lunak Rekomendasi Link Favorit

Agar lebih memahami cara kerja sistem, berikut ini akan coba dijelaskan metode-metode yang akan dipakai :

* Self-Organizing Map (SOM)

Algoritma *Self-Organizing Map* atau biasa disingkat dengan SOM pertama kali diperkenalkan oleh Teuvo Kohonen pada tahun 1982 [3]. Algoritma SOM ini merupakan bentuk pemodelan jaringan saraf yang mana antara sel-sel saraf yang bertetangga (*neighbourhood*) akan saling berkompetisi dalam aktifitas mereka dalam sebuah jaringan lateral yang bersifat mutual [2]. Konsep dari algoritma ini adalah memetakan inputan berdimensi banyak ke dalam sebuah peta dua dimensi yang tidak hanya merepresentasikan klaster-klaster secara visual, tetapi juga menampilkan keterkaitan antar klaster-klaster tersebut.



Gambar Bentuk peta dua dimensi representasi SOM

Langkah-langkah dari algoritma jaringan SOM ini dapat dipaparkan sebagai berikut :

* 1. Inisialisasikan :
* Nilai bobot *wij* secara acak dengan nilai yang kecil
* Besar ukuran *neighbourhood* *Nm*(0) dengan nilai yang cukup besar tetapi lebih kecil dari jumlah *node* yang pada satu dimensi peta
* Parameter fungsi *α*(t) dan *σ2*(t) antara 0 sampai 1
  1. Masukkan *input* *x* ke dalam *layer input* dan hitung *distance* (*d*) dari *input* ini ke bobot *w* dari setiap *node* *j*

dj = || *x* – *wj* || =

* 1. Pilih *node* dengan *distance* terkecil sebagai pemenang *m* (disebut dengan *best matching*)
  2. Lakukan *update* pada *node* pemenang *m* dan *node* tetangganya dengan rumus :



Dimana c = *α*(t)exp( - || ri – rm || / *σ2*(t)) untuk semua *node* *j* dalam *Nm*(t) dan ri - rm adalah jumlah *node* antara *node* *i* dan *node* pemenang *m*.

* 1. Lanjutkan dari langkah kedua untuk *Ω* epochs, tambahkan 1 pada *t*, dan kurangi ukuran *neighbourhood*, *α*(t) dan *σ2*(t)



Ulangi hingga bobot tidak berubah.

* 1. Ulangi langkah-langkah di atas hingga *t*=jumlah masukan.

Bentuk dataset yang digunakan oleh algoritma ini adalah numerik [4]. Untuk itu, dataset weblog yang digunakan dalam Tugas Akhir ini harus diubah ke dalam bentuk matriks yang bisa diproses oleh SOM. Representasi matriks dari dataset *weblog* dibuat dalam bentuk grup yang berisi transaksi yang dilakukan pengguna [1][2]. Maksud dari transaksi ini adalah pola jelajah yang dilakukan oleh satu pengguna dalam satu satuan waktu.

Diasumsikan terdapat sebuah set berisi *n* *link* yang berbeda yang muncul dalam *weblog* :

U = {*link1, link2, …, linkn*}

Dan sebuah set yang terdiri dari *m* transaksi pengguna :

T = { t1, t2, …, tm}

Representasikan transaksi dalam vektor bit

dimana akan bernilai 1 jika *link* yang dimaksud muncul dalam transaksi dan bernilai 0 jika tidak. Contoh representasi transaksi dalam bentuk vektor bit adalah sebagai berikut :

Table Representasi transaksi yang dilakukan pengguna dalam bentuk vektor bit

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Link1 | Link2 | … | LinkN |
| Trans1 | 1 | 0 | … | 1 |
| Trans2 | 0 | 0 | … | 1 |
| … | … | … | … | … |
| TransM | 1 | 1 | … | 0 |

Selanjutnya untuk mengurangi panjang dimensinya, transaksi-transaksi tersebut dikelompokkan lagi dalam grup pengguna dengan awalan IP tertentu sebagai matriks yang akan diinputkan ke dalam SOM.

Table Total angka 1 pada setiap kolom menjadi vektor aktifitas yang menjadi representasi grup transaksi

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Link1 | Link2 | … | LinkN |
| Trans1 | 1 | 0 | … | 1 |
| Trans2 | 0 | 0 | … | 1 |
| … | … | … | … | … |
| TransM | 1 | 1 | … | 0 |
| Total Trans | 100 | 1 | … | 12 |
| **GRUP 1** |  |  |  |  |

Table Representasi Link sebagai vektor dari grup transaksi

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Grup Transaksi |  |  |  |
|  | 1 | 2 | … | J |
| Link1 | 100 | 0 | … | 32 |
| Link2 | 1 | 0 | … | 1 |
| … | … | … | … | … |
| LinkN | 1 | 1 | … | 0 |

* Fuzzy Inference System

*Fuzzy* *Inference System* atau biasa disebut FIS adalah proses menyusun hasil pemetaan masukan yang diberikan menjadi sebuah keluaran dengan menggunakan *Fuzzy Logic* [6]. Pemetaan tersebut kemudian menjadi basis pengambilan keputusan atau pola yang tak terlihat. FIS merupakan sebuah *framework* komputasi populer yang didasarkan pada konsep *fuzzy set*, *fuzzy if-then rules*, dan *fuzzy reasoning* [4]. *Fuzzy set* sendiri adalah sebuah set tanpa batasan yang jelas/pasti. Struktur dasar dari sebuah FIS terdiri atas tiga komponen konseptual :

*Rule base*, yang berisipilihan-pilihan dari *fuzzy rules*

*Database*, yang mendefinisikan fungsi keanggotaan yang digunakan pada *fuzzy rules*

Mekanisme *reasoning*, yang melakukan proses inferensi pada *rules* dan fakta yang diberikan untuk mendapatkan kesimpulan.

Langkah-langkah untuk *Fuzzy* *Inference System* secara umum dapat dipaparkan sebagai berikut :

1. Ubah masukan ke dalam bentuk *fuzzy sets*
2. Aplikasikan operator *fuzzy* pada masukan (AND, OR)
3. Aplikasikan metode implikasi (*if-then*)
4. Agregasikan semua keluaran
5. Lakukan proses *defuzzify*

Dalam Tugas Akhir ini, FIS mendapatkan masukan berupa hasil cluster dari SOM. Dari masukan ini, setiap *rule* dalam FIS merepresentasikan sebuah *fuzzy relation* antara setiap transaksi t = {t1, t2, .., tm} yang dilakukan pengguna dan *link* yang akan direkomendasikan [7] dalam bentuk sebagai berikut :

**IF** (t1 adalah *A*1k) **AND** … **AND** (tm adalah *A*jk)

**THEN** (relevansi dari link1 adalah *b*1k) **AND** …. **AND** (relevansi dari linkN adalah *b*nk)

untuk k = {1,…,K} dimana K adalah jumlah *rule*, *Ajk* (*j*=1,…, *n* ) adalah *fuzzy* sets dengan fungsi keanggotaan Gaussian yang didefinisikan dengan variabel masukan *t* dan *j*, dan *bjk* adalah *fuzzy singleton* yang merepresentasikan jumlah rekomendasi (derajat relevansi) dari *link*  ke-*j*. Derajat relevansi *link* dihitung dengan fungsi:



dimana adalah derajat relevansi, s0 adalah transaksi yang dilakukan pengguna, dan µk adalah derajat kecocokan transaksi s0 dengan *rule* ke-k.

Hasil akhir dari proses yang terjadi dalam FIS ini akan menghasilkan sebuah bentuk implikasi yang dapat dibaca sebagai berikut :

**IF** (derajat *link1* adalah rendah) **AND**..**AND** (derajat *linkN* adalah tinggi)

**THEN** (rekomendasi *link1* dengan relevansi 0.3) **AND**..**AND** (rekomendasi *linkN* dengan relevansi 0.8)

* Tahap pengujian sistem

Tahap berikutnya adalah tahap pengujian sistem. *Weblog* yang akan diuji diinputkan ke sistem. Kemudian dilakukan pemrosesan seperti pada tahap pembelajaran sistem dengan tujuan untuk mendapatkan klasterisasi pola jelajah akses internet. Dan tahap terakhir adalah menampilkan hasil prediksi rekomendasi *link* situs favorit yang sesuai dengan pengguna dari analisis *weblog* berupa *bookmark* yang dapat disimpan ke dalam sistem.

1. **PERENCANAAN TUGAS AKHIR**

Perencanaan Tugas Akhir ini terdiri dari beberapa tahap, yaitu :

1. Penyusunan Proposal Tugas Akhir

Tahap awal untuk memulai pengerjaan Tugas Akhir adalah penyusunan Proposal Tugas Akhir. Pada proposal ini, penulis mengajukan gagasan pembuatan perangkat lunak rekomendasi *link* favorit dengan kombinasi metode SOM dan Fuzzy dan implementasi *Desktop* *Application*.

1. Studi Literatur

Pada tahap ini dilakukan pencarian, pengumpulan, penyaringan, pembelajaran dan pemahaman literatur yang berhubungan dengan metode yang digunakan, khususnya metode yang berkaitan dengan *clustering* dan *inference*. Algoritma tersebut di antaranya :

* + - * 1. Algoritma *Self-Organizing Map* (SOM)
        2. *Fuzzy Inference System*

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 membangun sistem perangkat lunak rekomendasi *link* situs favorit.

1. Pengujian dan Evaluasi

Pada tahap ini dilakukan uji coba terhadap sistem yang telah dibuat, dengan cara menilai apakah sistem perangkat lunak sudah dapat menampilkan preferensi situs favorit sesuai penggunaan dan menyimpan hasil analisis tersebut.

1. Penyusunan Buku Tugas Akhir

Tahap terakhir merupakan penyusunan laporan yang memuat dokumentasi mengenai pembuatan serta hasil dari implementasi perancangan dan *prototyping* yang telah dibuat. Secara garis besar, buku laporan Tugas Akhir ini terdiri atas beberapa bagian yaitu:

1. Pendahuluan
   1. Latar Belakang
   2. Permasalahan
   3. Batasan Tugas Akhir
   4. Tujuan
   5. Metodologi
   6. Sistematika Penulisan
2. Tinjauan Pustaka
3. Desain dan Implementasi
4. Uji Coba dan Evaluasi
5. Kesimpulan dan Saran
6. Daftar Pustaka
7. **JADWAL KEGIATAN TUGAS AKHIR**

Tugas Akhir ini diharapkan bisa dikerjakan menurut jadwal sebagai berikut:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No** | **Kegiatan** | **Bulan** | | | | | | | | | | | |
| **September** | | **Oktober** | | **November** | | **Desember** | | **Januari** | | **Februari** | |
| 1 | Penyusunan Proposal Tugas Akhir |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | Studi Literatur |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | Implementasi |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | Pengujian dan Evaluasi |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | Penyusunan Buku Tugas Akhir |  |  |  |  |  |  |  |  |  |  |  |  |

1. **DAFTAR PUSTAKA**

[1] Qi, D., Li, Chung-Chih. 2008. *Self-Organizing Map based Web Pages Clustering using Web Logs*. [pdf]. (<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.6437&rep=rep1&type=pdf>, diakses tanggal 25 Oktober 2010)

[2] Smith, K.A., Ng, A. 2003. Web Page Clustering Using A Self-Organizing Map of User Navigation Patterns. Decision Support Systems 35, 2003. pp. 245-256.

[3] Kohonen, Teuvo. *The Self-Organizing Map (Kohonen)*. Proceedings of the IEEE, Vol 78, No.9, 1990. pp. 1464-1480.

[4] X. Wang et al. *Intelligent Web Traffic Mining and Analysis*. Journal of Network and Computer Applications 28, 2005. pp. 147-165.

[5] Arnrich, Bert. 2009. *Clustering and Self Organizing Maps.* [pdf]. (<http://www.ife.ee.ethz.ch/education/WS1_HS2009_Le11.pdf>, diakses tanggal 25 Oktober 2010)

[6] Anonim. \_\_\_\_. *Fuzzy Inference Systems :: Tutorial (Fuzzy Logic Toolbox)*. [online]. (<http://www.mathworks.com/help/toolbox/fuzzy/fp351dup8.html>, diakses tanggal 25 Oktober 2010)

[7] Castellano, G., Fanelli, A. M., & Torsello, M. A. 2006. *Dynamic Link Suggestion by A Neuro-Fuzzy Web Recommendation System.* Proceedings of the 7th Asia Pacific Industrial Engineering and Management Systems Conference 2006.

**LEMBAR PENGESAHAN**

###### **Surabaya, 29 Oktober 2010**

Menyetujui,

Pembimbing I

Diana Purwitasari, S.Kom, M.Sc.

NIP : 19780410 200312 2001

Pembimbing II

Umi Laili Yuhana, S.Kom, M.Sc.

NIP : 19790626 200501 2002