

**KLASIFIKASI USULAN JUDUL SKRIPSI  
UNTUK MENENTUKAN CALON DOSEN PEMBIMBING  
MENGUNAKAN ALGORITMA  
ADD-ONE SMOOTHING NAIVE BAYES**

Skripsi



oleh  
**EVLIN DENITA**  
**22094676**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS KRISTEN DUTA WACANA  
2013**

**KLASIFIKASI USULAN JUDUL SKRIPSI**

**UNTUK MENENTUKAN CALON DOSEN PEMBIMBING  
MENGUNAKAN ALGORITMA  
ADD-ONE SMOOTHING NAIVE BAYES**

Skripsi



Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi Informasi  
Universitas Kristen Duta Wacana  
Sebagai Salah Satu Syarat dalam Memperoleh Gelar  
Sarjana Komputer

Disusun oleh  
**EVLIN DENITA**  
**22094676**

PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS KRISTEN DUTA WACANA  
2013

## PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

### **KLASIFIKASI USULAN JUDUL SKRIPSI UNTUK MENENTUKAN CALON DOSEN PEMBIMBING MENGGUNAKAN ALGORITMA ADD-ONE SMOOTHING NAIVE BAYES**

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi keserjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar keserjanaan saya.

Yogyakarta, 27 Mei 2013



EVLINE DENITA  
22094676

## HALAMAN PERSETUJUAN

Judul Skripsi : KLASIFIKASI USULAN JUDUL SKRIPSI UNTUK  
MENENTUKAN CALON DOSEN PEMBIMBING  
MENGUNAKAN ALGORITMA ADD-ONE  
SMOOTHING NAIVE BAYES

Nama Mahasiswa : EVLINE DENITA

N I M : 22094676

Matakuliah : Skripsi (Tugas Akhir)

Kode : TIW276

Semester : Genap

Tahun Akademik : 2012/2013

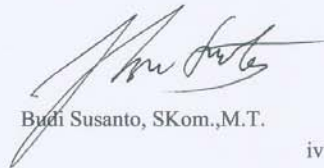
Telah diperiksa dan disetujui di  
Yogyakarta,  
Pada tanggal 27 Mei 2013

Dosen Pembimbing I



Antonius Rachmat C., SKom.,M.Cs

Dosen Pembimbing II



Budi Susanto, SKom.,M.T.

## HALAMAN PENGESAHAN

### KLASIFIKASI USULAN JUDUL SKRIPSI UNTUK MENENTUKAN CALON DOSEN PEMBIMBING MENGGUNAKAN ALGORITMA ADD-ONE SMOOTHING NAIVE BAYES

Oleh: EVLINE DENITA / 22094676

Dipertahankan di depan Dewan Penguji Skripsi  
Program Studi Teknik Informatika Fakultas Teknologi Informasi  
Universitas Kristen Duta Wacana - Yogyakarta  
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana Komputer  
pada tanggal 27 Mei 2013

Yogyakarta, 27 Mei 2013  
Mengesahkan,

Dewan Penguji:

1. Antonius Rachmat C., SKom.,M.Cs
2. Budi Susanto, SKom.,M.T.
3. Nugroho Agus Haryono, M.Si
4. Ignatia Dhian E K R, S.Kom



  
Dekan  
  
(Drs. Wimmie Handwidjojo, MIT.)

Ketua Program Studi  
  
(Nugroho Agus Haryono, M.Si)

## UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa yang telah melimpahkan rahmat dan anugerah, sehingga penulis dapat menyelesaikan dengan baik dan tepat waktu Tugas Akhir dengan judul Klasifikasi Usulan Judul Skripsi untuk Menentukan Calon Dosen Pembimbing menggunakan Algoritma *Add One Smoothing* Naive Bayes.

Penulisan laporan ini merupakan kelengkapan dan pemenuhan dari salah satu syarat dalam memperoleh gelar Sarjana Komputer. Selain itu bertujuan melatih mahasiswa untuk dapat menghasilkan suatu karya yang dapat dipertanggungjawabkan secara ilmiah, sehingga dapat bermanfaat bagi penggunanya.

Dalam menyelesaikan pembuatan program dan laporan Tugas Akhir ini, penulis telah banyak menerima bimbingan, saran dan masukan dari berbagai pihak baik secara langsung maupun secara tidak langsung. Untuk itu, dengan segala kerendahan hati, pada kesempatan ini penulis menyampaikan ucapan terimakasih kepada :

1. Bapak Antonius Rachmat C., S.Kom.,M.Cs selaku dosen pembimbing I, atas bimbingan, petunjuk, saran dan masukan yang diberikan selama pengerjaan Tugas Akhir ini sejak awal hingga akhir.
2. Bapak Budi Susanto, S.Kom.,M.T. selaku dosen pembimbing II, atas bimbingan, petunjuk, saran dan masukan yang diberikan selama pengerjaan Tugas Akhir ini sejak awal hingga akhir.
3. Dosen-dosen Universitas Kristen Duta Wacana yang telah memberikan bimbingan, pengajaran, dan dukungan kepada penulis selama ini.
4. Orang tua dan kakak tercinta yang telah memberi dukungan, semangat, motivasi serta dorongan kepada penulis.
5. Fendy Christiawan yang telah memberi semangat, motivasi, serta telah mendukung penulis dalam menyelesaikan Tugas Akhir ini.

6. Para sahabat terkasih “ALOHA” (Diana, Rosi, Irma, Selvi, Steven, Taffy, Andree, Ivan, Rudy, William, Setiyo, Mario, Wira, Okky, Yonas, Rendy), Feli, Zhen-Zhen, Adel, Cik Vivi, Novi, Nivy, dan sahabat lainnya yang telah memberikan semangat dan dukungan untuk penulis selama mengerjakan Tugas Akhir ini serta yang telah menjadi sahabat yang tulus kepada penulis selama menempuh perkuliahan di Universitas Kristen Duta Wacana.
7. Para teman seperjuangan yang mengambil Tugas Akhir di semester ini, khususnya Ria, Fanny, Rieka, Friska, Kak Michel, Putri, dan juga teman-teman lainnya yang telah berjuang bersama, saling mendukung, dan saling mengingatkan selama pengerjaan Tugas Akhir ini.
8. Pihak lain yang tidak dapat penulis sebutkan satu per satu, sehingga Tugas Akhir ini dapat terselesaikan dengan baik.

Penulis menyadari bahwa program dan laporan Tugas Akhir ini masih jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari pembaca sekalian. Sehingga suatu saat penulis dapat memberikan karya yang lebih baik lagi.

Akhir kata penulis ingin meminta maaf apabila ada kesalahan baik dalam penyusunan laporan maupun yang pernah penulis lakukan selama proses pembuatan program Tugas Akhir ini. Dan semoga Tugas Akhir ini dapat berguna bagi kita semua.

Yogyakarta, 9 Mei 2013

Penulis

## INTISARI

Klasifikasi Usulan Judul Skripsi untuk Menentukan Calon Dosen Pembimbing  
menggunakan Algoritma *Add One Smoothing* Naive Bayes

SinTA UKDW (Sistem Informasi Tugas Akhir Universitas Kristen Duta Wacana) merupakan sistem informasi yang menyimpan dokumen-dokumen skripsi dari mahasiswa UKDW. Permasalahan yang sering timbul saat mahasiswa berencana kolokium adalah apabila mahasiswa tersebut sudah mempunyai ide judul untuk skripsinya, namun ia tidak mengerti siapa dosen pembimbing yang paling tepat untuk judul tersebut.

Melihat permasalahan tersebut, penulis membangun sebuah sistem yang dapat melakukan analisis terhadap data-data judul skripsi beserta abstraknya untuk setiap dosen pembimbing yang ada. Sehingga nantinya sistem tersebut dapat menentukan calon dosen pembimbing dengan menerapkan Algoritma *Add One Smoothing Naive Bayes*, berdasarkan inputan *user* yang berupa usulan judul skripsi dan rumusan masalah yang diajukan.

Sistem klasifikasi usulan judul skripsi menggunakan *Add One Smoothing Naive Bayes* memberikan akurasi tertinggi untuk *feature selection* 50% yaitu sebesar 92% serta memberikan akurasi terendah untuk *feature selection* 10% yaitu sebesar 66%.

Kata kunci : klasifikasi, judul, skripsi, dosen pembimbing, *Add One Smoothing Naive Bayes*, akurasi, *feature selection*



## DAFTAR ISI

HALAMAN SAMBUNG .....	
PERNYATAAN KEASLIAN SKRIPSI.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN.....	v
UCAPAN TERIMA KASIH .....	vi
INTISARI .....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL .....	xii
DAFTAR GAMBAR DAN GRAFIK .....	xiii
BAB 1 PENDAHULUAN .....	1
1.1 Latar Belakang Masalah .....	1
1.2 Perumusan Masalah .....	1
1.3 Batasan Masalah .....	2
1.4 Tujuan Penelitian .....	2
1.5 Metode Penelitian .....	2
1.6 Sistematika Penulisan .....	3
BAB 2 TINJAUAN PUSTAKA .....	5
2.1 Tinjauan Pustaka .....	5
2.2 Landasan Teori .....	6
2.2.1 <i>Text Mining</i> .....	6
2.2.2 Klasifikasi .....	6
2.2.2.1 <i>Text Preprocessing</i> .....	7
2.2.2.2 <i>Text Transformation</i> .....	7
2.2.2.3 Pembobotan TF-IDF .....	8
2.2.2.4 <i>Frequency Based Feature Selection</i> .....	10
2.2.2.5 Algoritma Naive Bayes .....	10
2.2.3 Evaluasi Sistem .....	12
BAB 3 ANALISIS DAN PERANCANGAN SISTEM .....	13
3.1 Spesifikasi Kebutuhan Non Fungsional dan Fungsional .....	13

3.1.1 Spesifikasi Kebutuhan Non Fungsional .....	13
3.1.1.1 Spesifikasi Batasan Sistem .....	13
3.1.1.2 Spesifikasi Kebutuhan Perangkat Keras .....	13
3.1.1.3 Spesifikasi Kebutuhan Perangkat Lunak .....	14
3.1.1.4 Spesifikasi Sistem .....	14
3.1.2 Spesifikasi Kebutuhan Fungsional .....	16
3.1.2.1 Diagram Use Case .....	16
3.2 Perancangan Proses Sistem .....	18
3.2.1 Flowchart .....	18
3.2.1.1 Flowchart Sistem .....	18
3.2.1.2 Flowchart <i>Preprocessing Dokumen</i> .....	19
3.2.1.3 Flowchart Tokenisasi .....	21
3.2.1.4 Flowchart Pembobotan TF-IDF .....	22
3.2.1.5 Flowchart Normalisasi Pembobotan TF-IDF .....	24
3.2.1.6 Flowchart Klasifikasi dengan <i>Add One Smoothing</i> Naive Bayes .....	26
3.3 Perancangan Basis Data .....	29
3.3.1 Kamus Data .....	29
3.3.2 Diagram Skema Basis Data .....	30
3.4 Rancangan Antarmuka Sistem .....	32
3.5 Rancangan Pengujian dan Evaluasi Sistem .....	35
3.5.1 Rancangan Pengujian .....	35
3.5.2 Rancangan Evaluasi Sistem .....	36
3.6 Contoh Kasus Klasifikasi .....	36
3.7 <i>Test Case</i> .....	41
<b>BAB 4 IMPLEMENTASI DAN ANALISIS SISTEM .....</b>	<b>45</b>
4.1 Implementasi Sistem .....	45
4.1.1 Antarmuka Sistem.....	45
4.1.2 Pengumpulan Dokumen.....	55
4.1.3 Pseudocode Sistem.....	56
4.1.3.1 Tahap <i>Preprocessing</i> .....	56

4.1.3.2 Tahap <i>Feature Selection</i> .....	57
4.1.3.3 Tahap Klasifikasi Dokumen.....	58
4.2 Evaluasi Sistem .....	60
4.2.1 <i>Feature Selection</i> 10%.....	61
4.2.2 <i>Feature Selection</i> 20%.....	63
4.2.3 <i>Feature Selection</i> 30%.....	66
4.2.4 <i>Feature Selection</i> 40%.....	68
4.2.5 <i>Feature Selection</i> 50%.....	70
4.2.6 Grafik Hasil Pengujian .....	72
4.2.6.1 Grafik Keakuratan Sistem .....	72
4.3 Analisis Bobot Token untuk Kategori Joko Purwadi dan Rosa Delima .....	73
BAB 5 KESIMPULAN DAN SARAN .....	75
5.1 Kesimpulan .....	75
5.2 Saran .....	75
DAFTAR PUSTAKA .....	76

©UKDW

## DAFTAR TABEL

Tabel 2.1	Frekuensi Kemunculan Token .....	8
Tabel 3.1	Kamus Data .....	29
Tabel 3.2	Hasil Tokenisasi Dokumen.....	38
Tabel 3.3	<i>Test Case</i> .....	41
Tabel 4.1	Hasil pengujian dengan FS 10% .....	62
Tabel 4.2	Hasil pengujian dengan FS 20% .....	64
Tabel 4.3	Hasil pengujian dengan FS 30% .....	66
Tabel 4.4	Hasil pengujian dengan FS 40% .....	68
Tabel 4.5	Hasil pengujian dengan FS 50% .....	70

©UKYDN

## DAFTAR GAMBAR DAN GRAFIK

Gambar 2.1	Contoh perhitungan tf-idf .....	10
Gambar 2.2	Contoh perhitungan <i>add one smoothing</i> naive bayes .....	12
Gambar 3.1	<i>Use Case Diagram</i> .....	16
Gambar 3.2	<i>Flowchart</i> Sistem .....	18
Gambar 3.3	<i>Flowchart Preprocessing</i> Dokumen.....	20
Gambar 3.4	<i>Flowchart</i> Tokenisasi .....	21
Gambar 3.5	<i>Flowchart</i> Pembobotan TF-IDF .....	23
Gambar 3.6	<i>Flowchart</i> Normalisasi Pembobotan TF-IDF .....	25
Gambar 3.7	<i>Flowchart</i> Klasifikasi dengan <i>Add One Smoothing</i> Naive Bayes .....	28
Gambar 3.8	Skema Diagram Basis Data .....	30
Gambar 3.9	Rancangan Halaman <i>Home</i> .....	32
Gambar 3.10	Rancangan Halaman <i>Home</i> Admin .....	33
Gambar 3.11	Rancangan Halaman Tambah Data Pelatihan .....	34
Gambar 3.12	Rancangan Halaman Pengujian .....	35
Gambar 4.1	Halaman <i>Home</i> .....	45
Gambar 4.2	Halaman <i>Home</i> Admin .....	46
Gambar 4.3	Halaman Tambah Data Pelatihan .....	47
Gambar 4.4	Halaman <i>Update Feature Selection</i> .....	48
Gambar 4.5	Halaman Pengujian .....	49
Gambar 4.6	Halaman <i>Case Folding</i> .....	50
Gambar 4.7	Halaman Penghilangan Tanda Baca dan Angka .....	50
Gambar 4.8	Halaman Tokenisasi .....	51
Gambar 4.9	Halaman Penghapusan <i>Stopwords</i> .....	52
Gambar 4.10	Halaman Penghitungan TF .....	53
Gambar 4.11	Halaman Hasil Klasifikasi .....	54
Gambar 4.12	Halaman Hasil Klasifikasi ( <i>Sorting</i> ) .....	55

Listing 4.1	<i>Pseudocode</i> Tokenisasi dan Penghapusan <i>Stopword</i> .....	56
Listing 4.2	<i>Pseudocode</i> Perhitungan TF-IDF .....	57
Listing 4.3	<i>Pseudocode Feature Selection</i> .....	58
Listing 4.4	<i>Pseudocode</i> Klasifikasi Dokumen .....	59
Gambar 4.13	Contoh Pengujian dengan FS 10% dengan Hasil Benar .....	60
Gambar 4.14	Contoh Pengujian dengan FS 10% dengan Hasil Salah .....	61
Grafik 4.1	Grafik Keakuratan Sistem .....	72
Grafik 4.2	Grafik Perbandingan Bobot Token .....	74

©UKDW

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

SinTA UKDW (Sistem Informasi Tugas Akhir Universitas Kristen Duta Wacana) merupakan sistem informasi yang menyimpan dokumen-dokumen skripsi dari mahasiswa UKDW. Dengan adanya sistem informasi ini, mahasiswa amat dimudahkan dalam melakukan pencarian laporan skripsi yang telah dibuat oleh mahasiswa-mahasiswa UKDW terdahulu. Permasalahan yang sering timbul saat mahasiswa berencana kolokium adalah apabila mahasiswa tersebut sudah mempunyai ide judul untuk skripsinya, namun ia tidak mengerti siapa dosen pembimbing yang paling tepat untuk judul tersebut.

Untuk mendapatkan informasi mengenai nama dosen yang paling sering membimbing mahasiswa lain yang mempunyai judul skripsi dengan topik yang hampir serupa, dapat dilakukan pengklasifikasian data-data skripsi dengan menggunakan algoritma *Add One Smoothing Naive Bayes*.

Dengan demikian penyelesaian untuk masalah ini diharapkan dapat diwujudkan dengan dibuatnya sebuah sistem yang dapat melakukan analisis terhadap data-data judul skripsi beserta abstraknya untuk setiap dosen pembimbing yang ada. Sehingga dengan adanya sistem tersebut, para mahasiswa dapat terbantu dalam memilih calon dosen pembimbing yang paling tepat untuk judul skripsi yang akan ia ajukan.

### 1.2 Perumusan Masalah

Berdasarkan latar belakang masalah, akan diuji seberapa tepat algoritma *Add One Smoothing Naive Bayes* digunakan dalam memberikan pertimbangan calon dosen pembimbing untuk suatu judul skripsi yang diajukan?

### 1.3 Batasan Masalah

Adapun cakupan dalam penelitian ini antara lain :

- a. Data pelatihan yang digunakan sebagai proses analisis adalah data judul skripsi, abstrak, beserta dosen pembimbing yang diambil dari website [sinta.ukdw.ac.id](http://sinta.ukdw.ac.id).
- b. Sampel data yang digunakan sebagai proses analisis adalah sejumlah 150 data judul skripsi jurusan Teknik Informatika di UKDW yang meliputi 10 dosen pembimbing dengan masing-masing 15 sampel judul skripsi beserta abstraknya.
- c. Inputan berupa judul skripsi beserta rumusan masalahnya. Outputnya adalah nama dosen pembimbing yang menurut sistem paling tepat.
- d. Tidak dilakukan *stemming* pada tahap *preprocessing* dokumen.
- e. Sistem melakukan pembobotan dengan algoritma TF-IDF serta pengklasifikasian dengan algoritma *Add One Smoothing Naive Bayes*.

### 1.4 Tujuan Penelitian

Tujuan dari penelitian Tugas Akhir ini adalah :

- a. Membuat program bantu klasifikasi usulan judul skripsi untuk menentukan calon dosen pembimbing dengan menggunakan algoritma *Add One Smoothing Naive Bayes*.
- b. Program ini nantinya bisa digunakan oleh mahasiswa saat akan mengajukan judul skripsi di mana program dapat melakukan klasifikasi terhadap judul skripsi yang diajukan.

### 1.5 Metode Penelitian

Metodologi yang digunakan dalam penelitian ini adalah :

#### 1. Studi Pustaka

Studi Pustaka dilakukan dengan mempelajari teori-teori melalui buku, artikel, jurnal dan bahan lain yang mendukung yang



berhubungan dengan *text mining* dan Algoritma *Add One Smoothing Naive Bayes*, serta metode-metode pendukung lainnya yang dibutuhkan.

## 2. Perancangan sistem

Tahap ini berisi perancangan basis data dan perancangan antarmuka untuk sistem yang akan dibangun.

## 3. Pembangunan sistem

Tahap ini merupakan tahap pembuatan program sesuai dengan rancangan yang telah dibuat.

## 4. Implementasi dan testing

Pada tahap ini dilakukan pengujian terhadap program, dengan memasukkan inputan berupa judul skripsi yang akan diajukan, kemudian judul tersebut akan diklasifikasikan termasuk judul skripsi yang sering dibimbing oleh dosen siapa.

## 5. Analisis hasil percobaan dan evaluasi

Setelah dilakukan pengujian terhadap program, tahap selanjutnya adalah menganalisis keefektifan program dalam melakukan klasifikasi terhadap judul-judul skripsi berdasarkan dosen pembimbingnya.

### 1.6 Sistematika Penulisan

Penulisan laporan tugas akhir ini dibagi menjadi lima bab yaitu:

Bab 1, Pendahuluan, yang memberikan gambaran umum mengenai apa yang diteliti dalam tugas akhir ini. Pendahuluan berisi latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

Bab 2, Tinjauan Pustaka, yang terdiri dari dua bagian utama, yakni tinjauan pustaka dan landasan teori. Tinjauan pustaka menguraikan berbagai teori mengenai klasifikasi dengan algoritma *Add One Smoothing Naive Bayes* yang didapatkan dari berbagai sumber pustaka yang digunakan untuk penyusunan tugas

akhir. Landasan teori memuat penjelasan tentang konsep dan prinsip utama yang diperlukan untuk memecahkan masalah dalam penelitian. Hanya penjelasan yang berhubungan dengan penelitian yang dilakukan yang akan dicantumkan di sini.

Bab 3, Analisis dan Perancangan Sistem, yang mencakup perancangan sistem yang akan dibuat, yakni mengenai kebutuhan *hardware* dan *software*, spesifikasi sistem, arsitektur sistem, *use case diagram*, algoritma yang digunakan, skema *database*, dan rancangan antarmuka sistem.

Bab 4, Implementasi dan Analisis Sistem, yang memuat hasil implementasi dan pembahasan mengenai pengujian sistem yang dibuat berdasarkan bab 3, beserta hasil dari sistem yang dijalankan dan analisis dari sistem yang dibuat.

Bab 5, Kesimpulan dan Saran, berisi kesimpulan dari hasil penelitian yang didapatkan dan saran untuk memberikan analisis dan pengembangan yang lebih baik lagi pada penelitian ke depannya dalam topik yang serupa.

## BAB 5

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil implementasi dan analisis sistem, maka diperoleh kesimpulan sebagai berikut :

1. Sistem klasifikasi *Add One Smoothing Naive Bayes* memberikan akurasi tertinggi untuk *feature selection* 50% yaitu sebesar 92%.
2. Sistem klasifikasi *Add One Smoothing Naive Bayes* memberikan akurasi terendah untuk *feature selection* 10% yaitu sebesar 66%.
3. *Feature selection* mempengaruhi keakuratan sistem dalam memberikan usulan dosen pembimbing. Dalam kasus ini dan dalam rentang *feature selection* 10% hingga 50%, semakin tinggi nilai *feature selection*, sistem semakin akurat.
4. Dari hasil perbandingan untuk token-token yang sama dalam dokumen pelatihan kategori Joko Purwadi dan kategori Rosa Delima dapat dilihat bahwa kategori Joko Purwadi lebih banyak membimbing Tugas Akhir yang berupa game atau permainan. Sedangkan untuk kategori Rosa Delima lebih banyak membimbing Tugas Akhir yang berupa sistem pakar khususnya sistem pakar untuk mendeteksi penyakit.

#### 5.2 Saran

Saran yang diajukan penulis untuk perbaikan dan pengembangan sistem adalah sebagai berikut :

1. Jumlah dokumen pelatihan dapat ditambahkan sehingga sistem semakin akurat dalam memberikan usulan dosen pembimbing.
2. Dapat ditambahkan proses *stemming* dan pemerhatian frase dalam bahasa Indonesia untuk lebih meningkatkan akurasi sistem.

3. Diperlukan penggunaan *stored procedure* untuk melakukan perhitungan sehingga dapat mempercepat waktu yang dibutuhkan untuk *preprocessing* data dan mengurangi penggunaan memori.

©UKDW

## DAFTAR PUSTAKA

- Darujati, C. & Gumelar Agustinus B. (2012). *PEMANFAATAN TEKNIK SUPERVISED UNTUK KLASIFIKASI TEKS BAHASA INDONESIA*. Diakses tanggal 27 Januari 2013 dari [http://ejournal.narotama.ac.id/files/5-Link% 20Vol%2016%20February%202012.pdf](http://ejournal.narotama.ac.id/files/5-Link%20Vol%2016%20February%202012.pdf)
- Feldman, R. & Sanger, J. (2007). *The Text Mining Handbook*. New York: Cambridge University Press.
- Grossman, David A., Frieder, O. (2004). *Information Retrieval Algorithms and Heuristics*. New York : Springer.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining : Concepts and Techniques*. Waltham : Morgan Kauffman
- Indranandita, A., Susanto, B., & Rachmat, Antonius C. (2008). *SISTEM KLASIFIKASI DAN PENCARIAN JURNAL DENGAN MENGGUNAKAN METODE NAIVE BAYES DAN VECTOR SPACE MODEL*. Diakses tanggal 26 Januari 2013 dari <http://ti.ukdw.ac.id/ojs/index.php/informatika/article/view/48/19>
- Intan, R. & Defeng, A. *Hard: Subject-Based Search Engine Menggunakan Tf-Idf dan Jaccard's Coefficient*. Diakses tanggal 27 Januari 2013 dari <http://puslit.petra.ac.id/files/published/journals/IND/IND060801/IND06080106.pdf>.
- Manning, C.D., Raghavan, P., & Sch tze, H. (2008). *An Introduction to Information Retrieval*. New York : Cambridge University Press
- Nuraini, A., Susanto,B., & Proboyekti,U. (2007). *IMPLEMENTASI NAIVE BAYES CLASSIFIER PADA PROGRAM BANTU PENENTUAN BUKU REFERENSI MATAKULIAH*. Diakses tanggal 26 Januari 2013 dari <http://ti.ukdw.ac.id/ojs/index.php/informatika/article/view/60/12>

Weiss, Sholom M., et all. (2005). *Text Mining : Predictive Methods for Analyzing Unstructured Information*. New York : Springer.

©UKDW