

**IMPLEMENTASI ALGORITMA NAÏVE BAYES
MENGUNAKAN ISEAR UNTUK KLASIFIKASI EMOSI LIRIK
LAGU BERBAHASA INGGRIS**

Skripsi



PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI
INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA 2016

IMPLEMENTASI ALGORITMA NAÏVE BAYES MENGUNAKAN ISEAR UNTUK KLASIFIKASI EMOSI LIRIK LAGU BERBAHASA INGGRIS

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
LAKSMITA WIDYA ASTUTI
71120020

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI
INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA 2016

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

IMPLEMENTASI ALGORITMA NAÏVE BAYES MENGGUNAKAN ISEAR UNTUK KLASIFIKASI EMOSI LIRIK LAGU BERBAHASA INGGRIS

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, 10 Mei 2016



LAKSMITA WIDYA ASTUTI

71120020

HALAMAN PERSETUJUAN

Judul Skripsi : IMPLEMENTASI ALGORITMA NAÏVE BAYES
MENGUNAKAN ISEAR UNTUK KLASIFIKASI EMOSI
LIRIK LAGU BERBAHASA INGGRIS

Nama Mahasiswa : LAKSMITA WIDYA ASTUTI

N I M : 71120020

Matakuliah : Skripsi (Tugas Akhir)

Kode : TIW276

Semester : Genap

Tahun Akademik : 2015/2016

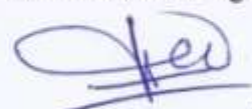
Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 10 Mei 2016

Dosen Pembimbing I



Antonius Rachmat C., S.Kom.,M.Cs.

Dosen Pembimbing II



Yuan Lukito, S.Kom., M.Cs.

HALAMAN PENGESAHAN

**IMPLEMENTASI ALGORITMA NAÏVE BAYES MENGGUNAKAN
ISEAR UNTUK KLASIFIKASI EMOSI LIRIK LAGU BERBAHASA
INGGRIS**

Oleh: LAKSMITA WIDYA ASTUTI / 71120020

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 23 Mei 2016

Yogyakarta, 6 Juni 2016
Mengesahkan,

Dewan Penguji:

1. Antonius Rachmat C., S.Kom., M.Cs.
2. Yuan Lukito, S.Kom., M.Cs.
3. Rosa Delima, S.Kom., M.Kom.
4. Willy Sudiarto Raharjo., S.Kom., M.Cs.

Dekan



(Budi Susanto, S.Kom., M.T.)

Ketua Program Studi

(Gloria Virginia, Ph.D.)

KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa atas rahmat dan berkatnya sehingga penyusunan laporan tugas akhir yang berjudul “Implementasi Algoritma Naïve Bayes Menggunakan ISEAR untuk Klasifikasi Emosi Lirik Lagu Berbahasa Inggris” dapat diselesaikan dengan baik. Pembuatan laporan ini merupakan suatu kelengkapan dan pemenuhan dari tugas akhir yang diberikan di Universitas Kristen Duta Wacana. Selain itu, pembuatan tugas akhir ini bertujuan melatih mahasiswa untuk dapat membuat suatu karya bukan hanya untuk pemenuhan tugas matakuliah yang diberikan, tetapi dapat dipertanggungjawabkan sehingga dapat bermanfaat bagi pengguna dan pengembangan selanjutnya.

Dalam menyusun laporan tugas akhir ini, tentunya penulis tidak terlepas dari dukungan dari berbagai pihak, baik kritik, saran serta bimbingannya. Dalam kesempatan ini penulis mengucapkan terimakasih yang sedalam-dalamnya kepada semua pihak yang telah mendukung sehingga penyusunan laporan tugas akhir ini dapat terselesaikan dengan baik terkhusus untuk:

1. Bapak Antonius Rachmat C., S.Kom.,M.Cs. dan Bapak Yuan Lukito, S.Kom., M.Cs. selaku dosen pembimbing atas segala bimbingan dan masukkannya dalam proses pengerjaan tugas akhir.
2. Keluarga yang sudah memberi dukungan dan semangat.
3. Michael Abadi Santoso yang sudah membantu memberikan semangat selama pengerjaan tugas akhir.
4. Valonia Inge Santoso, Melisa Fatmasari, Johan Sutanto, Jovani Ineke Chandra, Ireene Wijaya, Susy Valentina Rahardjo, Evelin Fanny Larissa, Sienne Malista Chandra, Monika Margi Maria Dina selaku teman

seperjuangan yang sudah memberi dukungan penuh baik sebelum pelaksanaan maupun setelah pelaksanaan tugas akhir.

5. Teman-teman serta saudara-saudara yang lain yang tidak bisa disebutkan satu persatu untuk segala masukkan dan dukungan sehingga pengerjaan tugas akhir ini dapat terselesaikan dengan baik.

Penulis menyadari bahwa dalam pelaksanaan tugas akhir mulai dari pelaksanaan hingga penyusunan laporan tidak luput dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun guna bekal bagi penulis untuk lebih baik di masa yang akan datang.

Yogyakarta, 10 Mei 2016

Penulis

INTISARI

Lirik lagu merupakan suatu ungkapan perasaan seseorang terhadap sesuatu hal yang sudah dilihat, didengar maupun dialaminya. Ungkapan perasaan tersebut tidak jarang menjadi salah satu kriteria user dalam melakukan pencarian lirik lagu, sehingga pencarian tidak hanya terbatas berdasarkan genre atau judul lagu namun melalui emosi dari lirik lagu yang diungkapkan. Pemilihan lirik lagu terutama saat user melakukan sebuah pencarian melekat dengan sebuah kategori agar lebih mudah dan terstruktur. Agar dapat mencapai tujuan tersebut, diperlukan suatu sistem pengkategorian agar dapat mengklasifikasikan lirik lagu secara otomatis.

Dengan latar belakang seperti yang disebutkan diatas, penulis membangun sebuah sistem yang dapat mengklasifikasikan lirik lagu berdasarkan emosinya dalam 3 kategori emosi yaitu *sad*, *happy* dan *angry*. Klasifikasi dilakukan dengan menggunakan algoritma Naïve Bayes dengan bantuan ISEAR sebagai *feature selection*. Tahap proses yang dilakukan pada penelitian ini yaitu *preprocessing*, *feature selection* dan klasifikasi. *Preprocessing* yang digunakan pada penelitian ini adalah *tokenisasi*, *stopword* dan *stemming*. Sedangkan proses *feature selection* yang digunakan dalam penelitian ini yaitu melakukan perhitungan TF-IDF dengan bantuan ISEAR.

Hasil dari penelitian ini menunjukkan dengan menggunakan ISEAR akurasi tertinggi terdapat pada *feature set* 40% - 100% yaitu sebesar 82,2%. Perbedaan signifikan dihasilkan pada penggunaan ISEAR dengan akurasi rata-rata keseluruhan porsi *featureset* sebesar 79% sedangkan tanpa menggunakan ISEAR rata-rata akurasi sebesar 54%. Dokumen paling relevan untuk pengujian menggunakan ISEAR terdapat pada kategori *angry* dengan rata-rata *f-measure* sebesar 0.7267.

Kata kunci : ISEAR, tf-idf, Naïve Bayes, Emosi, Lirik Lagu

DAFTAR ISI

PERNYATAAN KEASLIAN SKRIPSI.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN.....	v
KATA PENGANTAR	vi
INTISARI.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiv
DAFTAR LISTING	xv
BAB 1	1
PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan Penelitian.....	3
1.5 Metode Penelitian.....	4
1.6 Sistematika Penulisan.....	4
BAB 2	6
TINJAUAN PUSTAKA DAN DASAR TEORI	6
2.1 Tinjauan Pustaka	6
2.2 Dasar Teori.....	8
2.2.1. Emosi dan Musik	8
2.2.1 Tokenisasi	10
2.1.2 Stopword.....	10
2.1.3 Stemming	11

2.1.4	Pembobotan TF-IDF	12
2.1.5	ISEAR	12
2.1.6	Klasifikasi Naïve Bayes	14
2.1.7	Pengukuran Akurasi	16
BAB 3	17
METODOLOGI PENELITIAN	17
3.1	Spesifikasi Sistem.....	17
3.1.1	Spesifikasi Kebutuhan Non Fungsional	17
3.1.2	Spesifikasi Kebutuhan Fungsional	19
3.2	Perancangan Proses Sistem	20
3.2.1	Flowchart	21
3.2.2	Perancangan Basis Data	24
3.3	Perancangan Antar Muka Sistem	25
3.3.1	Rancangan Antar Muka Login.....	25
3.3.2	Rancangan Antar Muka <i>Testing Data</i>	26
3.3.3	Rancangan Antar Muka Lihat Hasil.....	26
3.3.4	Rancangan Antar Muka Manajemen Korpus <i>Training Data</i>	27
3.3.5	Rancangan Antar Muka Manajemen Korpus <i>Add Data</i>	27
3.3.6	Rancangan Antar Muka Manajemen Korpus <i>Generate TF-IDF</i>	28
3.3.7	Rancangan Antar Muka Tambah Data ISEAR	29
3.3.8	Rancangan Antar Muka Pengujian (<i>Testing</i> dan Akurasi).....	29
3.3.9	Rancangan Antar Muka Manajemen <i>User</i> bagian <i>User Data</i>	30
3.3.10	Rancangan Antar Muka Manajemen <i>User</i> bagian <i>Add User Data</i>	31
3.4	Perancangan Pengujian dan Evaluasi Sistem	31
3.4.1	Rancangan Pengujian dan Evaluasi Sistem	31
BAB 4	33
HASIL DAN PEMBAHASAN	33
4.1	Implementasi Sistem	33

4.1.1	Antar Muka Sistem	33
4.2	Analisis Sistem	48
4.2.1	Evaluasi Sistem	48
	73.86%	49
4.3	Pembahasan	50
BAB 5	57
KESIMPULAN DAN SARAN.....		57
5.1	Kesimpulan.....	57
5.2	Saran.....	57
DAFTAR PUSTAKA		59
LAMPIRAN.....		

©UKYDWN

DAFTAR GAMBAR

<i>Gambar 2. 1</i> Visualisasi Model Emosi (Jee, Cheong, Kim, Kwon, & Kobayashi, 2009)	9
<i>Gambar 2. 2</i> Algoritma porter (Tala, 1999)	11
<i>Gambar 3. 1</i> Usecase Diagram Sistem Klasifikasi Lirik Lagu	19
<i>Gambar 3. 2</i> Flow Chart Sistem.....	21
<i>Gambar 3. 3</i> Flow Chart Preprocessing	22
<i>Gambar 3. 4</i> Flow Chart Feature Selection	23
<i>Gambar 3. 5</i> Flow Chart Algoritma Naïve Bayes	24
<i>Gambar 3. 6</i> Rancangan Basis Data	24
<i>Gambar 3. 7</i> Rancangan Antar Muka Cari Emosi.....	25
<i>Gambar 3. 8</i> Rancangan Antar Muka <i>Testing</i> Data.....	26
<i>Gambar 3. 9</i> Rancangan Antar Muka Lihat Hasil.....	26
<i>Gambar 3. 10</i> Rancangan Antar Muka Manajemen Korpus <i>Training</i> Data.....	27
<i>Gambar 3. 11</i> Rancangan Antar Muka <i>Add Training Data</i>	27
<i>Gambar 3. 12</i> Rancangan Antar Muka <i>Preprocessing Training Data</i>	28
<i>Gambar 3. 13</i> Rancangan Antar Muka <i>Generate TF-IDF</i>	28
<i>Gambar 3. 14</i> Rancangan Antar Muka <i>ISEAR</i> Data.....	29
<i>Gambar 3. 15</i> Rancangan Antar Muka <i>Testing Data</i>	29
<i>Gambar 3. 16</i> Rancangan antar muka Pengujian dan Hasil Probabilitas.....	30
<i>Gambar 3. 17</i> Rancangan Antar Muka <i>User Data</i>	30
<i>Gambar 3. 18</i> Rancangan Antar Muka <i>Add User Data</i>	31
<i>Gambar 4. 1</i> Halaman <i>User Testing</i>	34
<i>Gambar 4. 2</i> Halaman <i>Process Data Testing User</i>	34
<i>Gambar 4. 3</i> Halaman <i>User Login</i>	36
<i>Gambar 4. 4</i> Halaman Manajemen Korpus Lihat Data <i>Training</i>	37
<i>Gambar 4. 5</i> Halaman Manajemen Korpus <i>Add Data Training</i>	37
<i>Gambar 4. 6</i> Halaman Manajemen Korpus <i>Preprocessing Data Training</i>	38

<i>Gambar 4. 7</i> Halaman Manajemen Korpus <i>Preprocessing</i> Data Training	38
<i>Gambar 4. 8</i> Halaman Manajemen Korpus <i>Preprocessing</i> Data Training	39
<i>Gambar 4. 9</i> Halaman Manajemen Korpus <i>Preprocessing</i> Data Training	39
<i>Gambar 4. 10</i> Halaman Manajemen Korpus <i>Preprocessing</i> Data Training	40
<i>Gambar 4. 11</i> Halaman Manajemen Korpus <i>Add Data Training Multiple Input</i>	40
<i>Gambar 4. 12</i> Struktur Data Teks <i>Training</i>	41
<i>Gambar 4. 13</i> Halaman Manajemen Korpus <i>Generate TF-IDF</i>	42
<i>Gambar 4. 14</i> Halaman Manajemen Korpus <i>ISEAR Data</i>	43
<i>Gambar 4. 15</i> Struktur data <i>testing</i> Halaman Admin	44
<i>Gambar 4. 16</i> Halaman <i>Testing</i> dan <i>Evaluation</i>	44
<i>Gambar 4. 17</i> Halaman Lihat Hasil Uji	45
<i>Gambar 4. 18</i> Halaman Lihat Akurasi Berdasarkan <i>Featureset</i> dan <i>ISEAR</i>	45
<i>Gambar 4. 19</i> <i>Confusion Matrix</i> Hasil Evaluasi	46
<i>Gambar 4. 20</i> Halaman Manajemen Data Pengguna	46
<i>Gambar 4. 21</i> Halaman Tambah Pengguna.....	47
<i>Gambar 4. 22</i> Grafik Akurasi menggunakan <i>ISEAR</i>	52
<i>Gambar 4. 23</i> Grafik Pengujian tidak Menggunakan <i>ISEAR</i>	52
<i>Gambar 4. 24</i> Hasil Klasifikasi Per-Kalimat	54
<i>Gambar 4. 25</i> <i>Testing</i> Lirik Lagu Secara Utuh	54
<i>Gambar 4. 26</i> Contoh Hasil Pengujian tanpa Logaritma	56
<i>Gambar 4. 27</i> Hasil Pengujian Perhitungan Probabilitas dengan Modifikasi Naive Bayes	56

DAFTAR TABEL

Tabel 2. 1	8
Tabel 2. 2	13
Tabel 2. 3	15
Tabel 3. 1	32
Tabel 4. 1	43
Tabel 4. 2	49
Tabel 4. 3	50
Tabel 4. 4	51
Tabel 4. 5	53

©UKYDWN

DAFTAR LISTING

Listing 4. 1 Pseudocode Tokenisasi.....	35
Listing 4. 2 Pseudocode <i>Stopword</i>	35
Listing 4. 3 Pseudocode Perhitungan TF-IDF.....	42

©UKDW

INTISARI

Lirik lagu merupakan suatu ungkapan perasaan seseorang terhadap sesuatu hal yang sudah dilihat, didengar maupun dialaminya. Ungkapan perasaan tersebut tidak jarang menjadi salah satu kriteria user dalam melakukan pencarian lirik lagu, sehingga pencarian tidak hanya terbatas berdasarkan genre atau judul lagu namun melalui emosi dari lirik lagu yang diungkapkan. Pemilihan lirik lagu terutama saat user melakukan sebuah pencarian melekat dengan sebuah kategori agar lebih mudah dan terstruktur. Agar dapat mencapai tujuan tersebut, diperlukan suatu sistem pengkategorian agar dapat mengklasifikasikan lirik lagu secara otomatis.

Dengan latar belakang seperti yang disebutkan diatas, penulis membangun sebuah sistem yang dapat mengklasifikasikan lirik lagu berdasarkan emosinya dalam 3 kategori emosi yaitu *sad*, *happy* dan *angry*. Klasifikasi dilakukan dengan menggunakan algoritma Naïve Bayes dengan bantuan ISEAR sebagai *feature selection*. Tahap proses yang dilakukan pada penelitian ini yaitu *preprocessing*, *feature selection* dan klasifikasi. *Preprocessing* yang digunakan pada penelitian ini adalah *tokenisasi*, *stopword* dan *stemming*. Sedangkan proses *feature selection* yang digunakan dalam penelitian ini yaitu melakukan perhitungan TF-IDF dengan bantuan ISEAR.

Hasil dari penelitian ini menunjukkan dengan menggunakan ISEAR akurasi tertinggi terdapat pada *feature set* 40% - 100% yaitu sebesar 82,2%. Perbedaan signifikan dihasilkan pada penggunaan ISEAR dengan akurasi rata-rata keseluruhan porsi *featureset* sebesar 79% sedangkan tanpa menggunakan ISEAR rata-rata akurasi sebesar 54%. Dokumen paling relevan untuk pengujian menggunakan ISEAR terdapat pada kategori *angry* dengan rata-rata *f-measure* sebesar 0.7267.

Kata kunci : ISEAR, tf-idf, Naïve Bayes, Emosi, Lirik Lagu

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Musik memiliki peran yang penting bagi kehidupan manusia. Selain sebagai media hiburan, musik juga merupakan media bantu diri untuk menangani perasaan emosi yang sedang di hadapi (Lerik & Prawitasari, 2005). Elemen musik selain dari irama, melodi dan harmoni juga tidak terlepas dengan makna dari kata-kata yang terkandung didalamnya. Maksud dari setiap kata di dalam lirik lagu juga memiliki arti penting bagi kebanyakan orang yaitu menyembuhkan perasaan dari kesedihan, memotivasi, melampiaskan kemarahan, melampiaskan kebahagiaan dan lain sebagainya.

Banyak orang salah mengartikan bahwa lagu-lagu yang memiliki irama bersemangat seperti *rock*, dangdut, dan *country* menggambarkan lagu yang ceria dan penuh kegembiraan, sedangkan lagu *mellow* seperti *jazz*, *pop*, dan *ballad* menggambarkan makna lagu yang sedih karena memiliki irama yang lambat. Setiap bait pada lirik lagu umumnya mengandung maksud melalui gaya bahasa yang digunakan oleh pengarang. Pengarang bebas menyampaikan emosi dengan *genre* yang bervariasi sesuai dengan keinginannya. Pengungkapan emosi tersebut tidak jarang menjadi salah satu kebutuhan umum para penikmat musik, sehingga pemilihan kategori dalam pencarian lagu tidak hanya terbatas berdasarkan *genre*, namun juga melalui emosi dari para pencipta lagu.

Pemilihan lirik lagu terutama saat *user* melakukan sebuah pencarian melekat dengan sebuah kategori agar lebih mudah dan terstruktur. Agar dapat mencapai tujuan tersebut, diperlukan mesin pengkategori agar lirik lagu dapat diklasifikasikan

secara otomatis. Mesin kategori lirik lagu berdasarkan emosinya dapat dibuat dengan menggunakan banyak algoritma. Algoritma tersebut contohnya yaitu *K-Nearest Neighbour*, *Naïve Bayes*, *Decision Tree*, dan *SVN*. Pada penelitian ini, penulis menggunakan algoritma *Naïve Bayes* untuk mengklasifikasikan lirik lagu berdasarkan emosinya. Algoritma ini dipilih karena memiliki kinerja yang cukup tinggi untuk klasifikasi dokumen teks (Hamzah, 2012).

Tidak cukup menggunakan algoritma klasifikasi supaya lirik lagu dikategorikan tepat sesuai emosinya, namun penggunaan *feature set* ISEAR juga akan di terapkan di dalam penelitian ini. *Feature set* ISEAR dapat di unduh dari <http://www.affective-sciences.org/researchmaterial>. Tujuan dari penggunaan *feature set* ISEAR ini di maksudkan agar hasil klasifikasi yang dihasilkan semakin baik. Pengukuran akurasi dilakukan dengan membandingkan perhitungan akurasi sistem, *precision*, *recall* dan *f-meassure* baik menggunakan *feature set* ISEAR maupun tidak menggunakan data ISEAR. Porsi *feature set* yang digunakan beranekaragam yaitu 10%, - 100% dengan range 10 dari total *term* terbaik yang sudah melalui proses *feature selection*.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, rumusan masalah yang dapat diperoleh dari penelitian yang akan dilakukan adalah bagaimana implementasi serta akurasi dari algoritma *Naïve Bayes* menggunakan bantuan dataset ISEAR (*International Survey on Emotion Antecedents and Reaction*) sebagai *feature selection* dalam proses klasifikasi.

1.3 Batasan Masalah

Batasan-batasan yang diterapkan pada penelitian yang akan dilakukan adalah sebagai berikut:

1. Data yang digunakan dalam penelitian merupakan sekumpulan korpus teks lirik lagu baik yang sudah di kategorikan maupun yang belum dikategorikan.
2. Menggunakan dataset ISEAR (*International Survey on Emotion Antecedents and Reaction*) sebagai *feature selection* saat melakukan pembobotan tf-idf. Dataset tersebut di ambil dari:
<http://www.affective-sciences.org/researchmaterial>
3. Sumber alamat data uji dan data *training* yang meliputi daftar lagu beserta kategorinya diambil dari:
 - a. <http://www.pastemagazine.com/>
 - b. <https://www.upvenue.com/>
 - c. <http://www.maxtv.com.au/>
 - d. <http://www.ranker.com/>
4. Sumber data lirik diambil dari link berikut:
 - a. <http://www.metrolyrics.com>
 - b. <http://www.azlyrics.com>
5. Sumber daftar stopword :
<http://jmlr.csail.mit.edu/papers/volume5/lewis04a/a11-smart-stop-list/english.stop>.
6. Penelitian ini menggunakan 3 kelas emosi yaitu *sad*, *happy* dan *angry*.
7. Inputan berupa data lirik lagu dan output berupa jenis emosi yang sesuai dengan teks lirik lagu.
8. Sistem tidak mementingkan arti dari suatu bentuk frase.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah mengimplementasikan dan mengukur tingkat akurasi dari sistem baik *precision*, *recall* dan *f-measure* dari algoritma Naïve Bayes

menggunakan bantuan ISEAR (*International Survey on Emotion Antecedents and Reaction*) sebagai *featureset* pada proses *feature selection*.

1.5 Metode Penelitian

Metode yang akan digunakan sebagai langkah-langkah untuk pemecahan masalah dalam penelitian adalah sebagai berikut:

1. Studi Literatur

Studi literatur dilakukan dengan cara mempelajari teori-teori yang diperlukan dalam penelitian melalui jurnal-jurnal, video dan buku elektronik yang berkaitan dengan algoritma klasifikasi Naïve Bayes dan *feature set* ISEAR untuk *feature selection*.

2. Perancangan Desain Sistem

Desain yang dirancang dalam penelitian ini adalah desain aplikasi berbasis web menggunakan *framework CodeIgniter*.

3. Implementasi Sistem

Sistem akan dilakukan dengan menginputkan data lirik lagu berbahasa Inggris kemudian di-*preprocessing* menggunakan tokenisasi, *stopword*, dan *stemming*. Langkah selanjutnya proses klasifikasi Naïve Bayes dengan dan tanpa menggunakan bantuan *feature set* ISEAR untuk proses *feature selection* ditahap pembobotan. Hasil klasifikasi yang didapatkan diuji dan dievaluasi baik melalui akurasi sistem, *precision*, *recall* dan *f-measure*.

4. Pengujian dan Evaluasi

Evaluasi dilakukan dengan mengukur tingkat *precision*, *recall* dan *f-measure* dari hasil klasifikasi. Porsi *feature set* yang digunakan adalah 10% - 100% dengan range 10 dari total *term* terbaik.

1.6 Sistematika Penulisan

Penyusunan laporan tugas akhir ini secara garis besar dikelompokkan menjadi lima bab sebagai berikut:

Bab 1, Pendahuluan. Bab ini memberikan gambaran umum mengenai penelitian yang akan dilakukan mencakup latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

Bab 2, Tinjauan Pustaka. Bab ini terdiri dari tinjauan pustaka dan landasan teori yang digunakan untuk referensi saat melakukan penelitian. Pada bab ini akan diuraikan berbagai konsep, teori maupun rumus-rumus yang mendukung proses penelitian.

Bab 3, Perancangan Sistem. Bab ini membahas rancangan sistem yang dibangun berupa spesifikasi sistem, rancangan diagram sistem, rancangan antar muka sistem dan tahapan-tahapan yang berkaitan dengan proses perancangan dan pembuatan sistem.

Bab 4, Implementasi dan Analisis sistem. Bab ini akan diuraikan mengenai hasil dari implementasi dari metode yang digunakan dan analisis berupa penjelasan secara teoritis.

Bab 5, Kesimpulan dan Saran. Pada bab ini, akan dibahas mengenai hasil analisis dari penelitian yang dilakukan serta saran yang mendukung supaya penelitian tersebut dapat lebih baik dari penelitian yang dilakukan. Harapannya dapat memperbaiki kinerja sistem pada riset mendatang.

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

1. Kelemahan yang dimiliki oleh Algoritma klasifikasi Naïve Bayes pada penelitian ini tidak dapat mengenali emosi dengan cara mencari hubungan semantik emosi makna antar kata namun didasarkan pada probabilitas statistik *term* antar kata.
2. Rumus modifikasi Naïve Bayes dapat meningkatkan akurasi karena dapat *handle* variable pemrograman yang memiliki keterbatasan pada limit angka tertentu.
3. Proses *testing* sering mengalami kesalahan klasifikasi karena 1 lirik lagu memiliki kemungkinan lebih dari 1 emosi yang berbeda.
4. Implementasi Algoritma Naïve Bayes dengan menggunakan ISEAR memberikan akurasi tertinggi pada *feature set* 40% - 100% sebesar 82,2%. Penggunaan *featureset* 40% dari 10 range porsi *featureset* yang ada dengan menggunakan ISEAR sudah cukup untuk menghasilkan nilai akurasi paling maksimal.
5. Berdasarkan hasil pengujian, penggunaan ISEAR dan peningkatan porsi *featureset* menghasilkan akurasi yang cukup tinggi. Perbedaan signifikan dihasilkan pada penggunaan ISEAR, akurasi rata-rata keseluruhan porsi *featureset* sebesar 79% sedangkan tanpa menggunakan ISEAR rata-rata akurasi sebesar 54%. Dokumen paling relevan terdapat pada kategori *angry* dengan rata-rata *precision* sebesar 0.5709, rata-rata *recall* sebesar 0.9071 dan rata-rata *F-Measure* sebesar 0.7267.

5.2 Saran

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

1. Memodifikasi proses klasifikasi dengan melakukan uji hipotesis dengan melakukan perhitungan statistik sehingga sistem dapat mengenali lebih dari 1 kategori emosi agar akurasi meningkat.
2. Memodifikasi algoritma *preprocessing* yang lebih baik (contoh: kata “tidak” pada kata “tidak senang” tidak menjadi sebuah stoplist tetapi menjadi sebuah frasa sehingga memiliki makna kata).
3. Menambah dokumen *training* yang baik sehingga menghasilkan lebih banyak pencari dokumen sesuai kategori yang diharapkan.

©UKDW

DAFTAR PUSTAKA

- Agusta, L. (2009). Perbandingan Algoritma Stemming Porter dengan Algoritma Nazief & Adriani untuk Stemming Dokumen Teks Bahasa Indonesia. *Konferensi Nasional Sistem dan Informatika*, 196-201.
- Aldi, S., Lailil, M., & Indriati. (2013). Klasifikasi Berita Bahasa Inggris Menggunakan Algoritma K-Nearest Neighbour (KNN) berbasis Ontologi.
- Bata, J. V. (2016). *Deteksi Jenis Emosi Dari Teks Bahasa Indonesia Menggunakan Keyword-spotting Dan Naive Bayes*. Retrieved from Postgraduated Thesis: <http://e-journal.uajy.ac.id/8903/>
- Destuardi, I., & Sumpeno, S. (2009). Klasifikasi Emosi Untuk Teks Bahasa Indonesia Menggunakan Metode Naive Bayes. *Seminar Nasional Pascasarjana IX*.
- Hamzah, A. (2012). Klasifikasi Teks dengan Naive Bayes (NBC) untuk Pengelompokan Teks Berita dan Abstract Akademis. *Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) Periode III*, B-220 - B-277.
- Hapsari, R. K., & Santoso, Y. J. (2015). Stemming Artikel Berbahasa Indonesia dengan Pendekatan Confix-Stripping. *Prosiding Seminar Nasional Manajemen Teknologi XXI*.
- International Survey on Emotion Antecedents and Reaction. (2015). *Affective Science*. Retrieved October Friday, 2015, from Swiss International Center of Competence in Research: <http://www.affective-sciences.org/researchmaterial>
- Jee, E.-S., Cheong, Y.-J., Kim, C. H., Kwon, D.-S., & Kobayashi, H. (2009). *Sound Production for the Emotional Expression of Socially Interactive Robots, Advances in Human-Robot Interaction*. Retrieved from

<http://www.intechopen.com/books/advances-in-human-robot-interaction/sound-production-for-the-emotional-expression-of-socially-interactive-robots>

Lerik, M. D., & Prawitasari, J. E. (2005). Pengaruh Terapi Musik terhadap Depresi. *SOSIOSAINS*, 211.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Oatley, K., & Jenkins, J. M. (1996). *Understanding Emotions*. Blackwell.

Saadah, M. N., Atmagi, R. W., Rahayu, D. S., & Arifin, A. Z. (2011). Sistem Temu Kembali Dokumen Teks dengan Pembobotan. *Juti*, 18.

Sakti, S. (2010). *Eksperimen Mengenai Emosi Oleh Musik Bahagia Dan Sedih Dengan Volume Tertentu*. Retrieved from Undergraduated Thesis: www.library.usd.ac.id

Scherer, K. R., & Wallbott, H. (1986). How universal and specific is emotional experience? . *Social Science Information*, 763-795.

Song, Y., & Dixon, S. (2015). *How Well Can A Music Emotion Recognition System Predict The Emotional Responses Of Participants?* Retrieved from <https://www.eecs.qmul.ac.uk/~simond/pub/2015/SongDixon-SMC2015-EmotionPrediction.pdf>

Tala, F. Z. (1999). *A Study of Stemming Effects on Information*. Retrieved from <https://www.ilic.uva.nl/Research/Publications/Reports/MoL-2003-02.text.pdf>