



UNIVERSITAS INDONESIA

**PENGEMBANGAN MODEL PHONETIC SIMILARITY
BAHASA INDONESIA BERDASARKAN KAMUS FONETIK
BAHASA INDONESIA ZAHRA**

SKRIPSI

**DWI PUTRA HANDI PAREH
1204000289**

**FAKULTAS ILMU KOMPUTER
PROGRAM STUDI S1 REGULER ILMU KOMPUTER
DEPOK
JANUARI 2009**



UNIVERSITAS INDONESIA

**PENGEMBANGAN MODEL PHONETIC SIMILARITY
BAHASA INDONESIA BERDASARKAN KAMUS FONETIK
BAHASA INDONESIA ZAHRA**

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana

**DWI PUTRA HANDI PAREH
1204000289**

**FAKULTAS ILMU KOMPUTER
PROGRAM STUDI S1 REGULER ILMU KOMPUTER
DEPOK
JANUARI 2009**

HALAMAN PERNYATAAN ORISINALITAS

**Skripsi ini adalah hasil karya saya sendiri
dan semua sumber baik yang dikutip maupun dirujuk
telah saya nyatakan dengan benar**

Nama : Dwi Putra Handi Pareh

NPM : 1204000289

Tanda Tangan :

Tanggal : 12 Januari 2009



HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama : Dwi Putra Handi Pareh

NPM : 1204000289

Judul Skripsi : Pengembangan model phonetic similarity bahasa Indonesia berdasarkan kamus fonetik bahasa Indonesia Zahra

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Sarjana Fakultas Ilmu Komputer, Universitas Indonesia

DEWAN PENGUJI

Pembimbing : Dr. Hisar Maruli Manurung ()

Penguji : Adila Alfa Krisnadhi S.Kom., MSc. ()

Penguji : Dra. Mirna Adriani, Ph.D ()

Ditetapkan di : Depok

Tanggal : 12 Januari 2009

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena atas kasih karunia dan penyertaan-Nya akhirnya penulis dapat menyelesaikan kerja praktik dan penulisan laporan kerja praktik ini dalam jangka waktu yang telah ditentukan. Selama kerja praktik yang dilakukan penulis, penulis mendapatkan bantuan dari berbagai pihak yang sangat berarti bagi penulis. Oleh sebab itu, ucapan terima kasih hendak penulis sampaikan kepada pihak-pihak berikut:

1. Ibu dan anggota keluarga saya atas dukungan moral dan perhatian yang tidak henti-hentinya.
2. Pak Ruli sebagai dosen pembimbing tugas akhir yang sudah menyediakan waktu, tenaga, saran, dan dukungan sehingga penulis bisa menyelesaikan penulisan tugas akhir ini.
3. Amalia Zahra yang sudah menyediakan kamus fonetik bahasa Indonesia dan daftar kata yang disertai simbol fonemnya.
4. Ratna Tazkia yang sudah memberi masukan berharga mengenai pembuatan dan penyajian kuesioner.
5. Arudea Mahartianto sebagai rekan seperjuangan pengerjaan tugas akhir dan teman sepelempangan atas dukungannya selama penulis melakukan penelitian juga atas stok Gossip Girl dan Heroes yang disediakan.
6. Ranga Mahendra Jati sebagai rekan seperjuangan pengerjaan tugas akhir dan teman sepelempangan atas dukungannya selama penulis melakukan penelitian juga sebagai sesama penggemar One Piece.
7. Desmond Darma Putra atas bantuannya dalam melakukan *troubleshooting* program dan penggunaan Microsoft Word.
8. Dyta Anggraeni atas jasa-jasanya membantu penulis menggunakan Microsoft Word dan dukungan terhadap penelitian secara umum.

9. Seluruh rekan-rekan mahasiswa Lab IR seperti Arawinda, Charles Christian, Teddy, Aurora Marsye, Femphy Pisceldo, Sadar Baskoro dll atas kebersamaannya sebagai anggota Lab IR dan teman seperjuangan dalam mengerjakan tugas akhir.
10. Ibu Mirna dan Pak Adila sebagai rekan Lab IR atas sesi pertemuan mingguan yang menarik.
11. Teman-teman yang sering berada di kantin, Bama, Xoxpe, Wanto, Widut, dan lain-lain atas kesediaannya mengobati kejenuhan selama rutinitas penelitian.
12. Tekfuk Sulianto atas sesi tutorial Dota dan pendapatnya mengenai kuesioner.
13. Rekan-rekan lab Ristek atas kebersamaannya dengan penulis pada pukul 18.00 hingga 21.00.
14. Para subjek kuesioner yang telah meluangkan waktu dan tenaganya untuk membantu penelitian ini.
15. Kepada semua pihak yang terkait yang namanya belum disebutkan disini.

Penulis sangat sadar bahwa dalam melakukan penelitian ini banyak kekurangan dan kesalahan yang telah penulis lakukan. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan untuk perbaikan ke depannya. Semoga penelitian ini dapat membawa manfaat bagi semua pihak yang berkepentingan.

Depok, Desember 2008

Dwi Putra H. P

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS
AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademis Universitas Indonesia, saya yang bertanda tangan di bawah ini:

Nama : Dwi Putra Handi Parih

NPM : 1204000289

Program Studi : Sarjana

Fakultas : Ilmu Komputer

Jenis karya : Skripsi

demikian demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul:

Pengembangan model phonetic similarity bahasa Indonesia berdasarkan kamus fonetik bahasa Indonesia Zahra

berserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok

Pada tanggal : 12 Januari 2009

Yang menyatakan

(Dwi Putra Handi Parih)

DAFTAR ISI

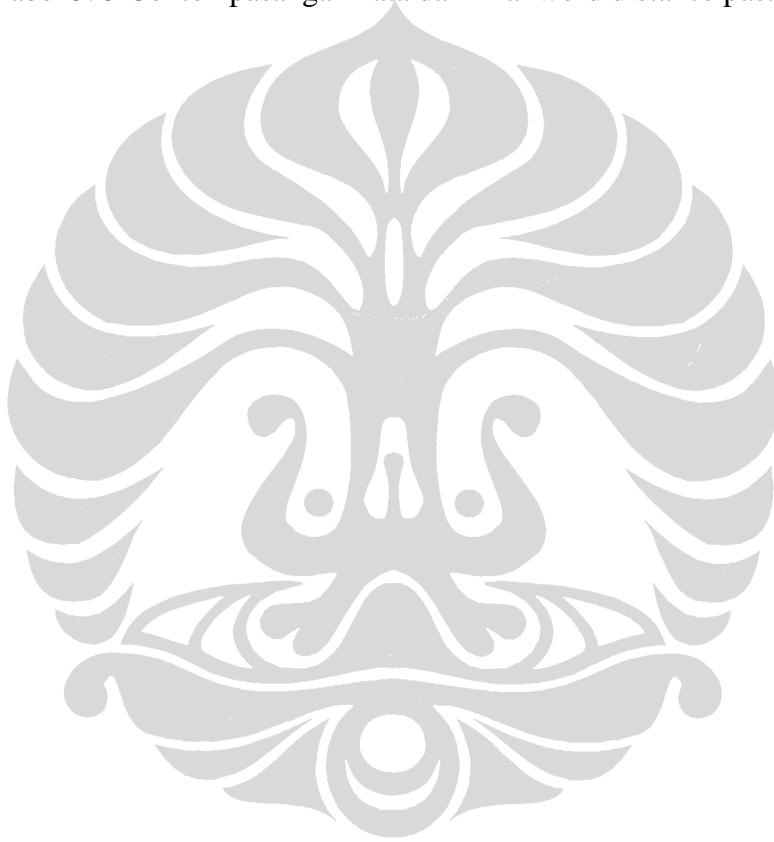
HALAMAN PERNYATAAN ORISINALITAS.....	iii
HALAMAN PENGESAHAN.....	iv
KATA PENGANTAR	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	vii
ABSTRAK	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
BAB 1 PENDAHULUAN	1
1.1. Latar belakang	1
1.2. Permasalahan.....	1
1.3. Tujuan	2
1.4. Ruang Lingkup.....	2
1.5. Metodologi penelitian	2
1.6. Sistematika Penulisan	3
BAB 2 LANDASAN TEORI.....	4
2.1. Phonetic similarity	4
2.1.1 Fonologi	4
2.1.2. Fonem.....	4
2.1.3. Penerapan phonetic similarity	4
2.1.4. Model phonetic similarity Stephen Rooney	5
2.1.5. Model phonetic similarity STANDUP.....	6
2.2. Kamus fonetik Amalia Zahra.....	8
2.3. Levenshtein Distance	9
BAB 3 PERANCANGAN	9
3.1. Model phonetic similarity	9
3.1.1. Nilai kemiripan dan nilai jarak kemiripan	10
3.1.2. Rentang nilai jarak kemiripan	10
3.1.3. Nilai <i>default</i> pada matriks substitusi.....	10
3.2. Perancangan kuesioner.....	11
3.2.1. Metodologi kuesioner	11
3.2.2. Penyusunan kuesioner	13
3.3. Ekstraksi nilai kuesioner	24
3.4. Penghitungan word distance	26
3.3.1. Post-normalized word distance	26
3.4. Rangkuman	27
BAB 4 IMPLEMENTASI.....	28
4.1. Implementasi word distance.....	28
4.1.1. Implementasi Levenshtein Distance	28
4.1.2. Modifikasi umum	30
4.1.3. Modifikasi nilai operasi Levenshtein distance.....	32
4.1.4. Modifikasi post-normalized word distance	33
4.1.5. Modifikasi masukan dan keluaran	34
4.1.6. Nilai <i>threshold</i>	36
4.1.7. Aplikasi word distance	37

4.2. Implementasi ekstraksi kuesioner	37
4.2.1. Format masukan program	37
4.2.2. Implementasi yang dilakukan	38
4.3. Rangkuman	39
BAB 5 EKSPERIMEN	41
5.1. Pelaksanaan kuesioner	41
5.1.1. Hasil kuesioner.....	41
5.1.2. Perbandingan dengan model phonetic similarity STANDUP.....	43
5.2. Eksperimen word distance	46
5.2.1. Hasil eksperimen word distance.....	47
BAB 6 PENUTUP	50
6.1. Kesimpulan	50
6.2. Saran.....	51
DAFTAR PUSTAKA	52



DAFTAR TABEL

Tabel 2. 1 Atribut level 2 untuk fonem vokal	7
Tabel 2. 2 Atribut level 2 untuk fonem konsonan.....	7
Tabel 2. 3 Kamus fonetik Amalia Zahra	8
Tabel 2. 4 Contoh matriks hasil pengerjaan Levenshtein distance	10
Tabel 5. 1 Matriks substitusi hasil ekstraksi kuesioner untuk fonem konsonan ...	42
Tabel 5. 2 substitusi hasil ekstraksi kuesioner untuk fonem vokal	43
Tabel 5. 3 Pemetaan fonem STANDUP dan fonem Amalia Zahra	44
Tabel 5. 4 Nilai jarak kemiripan untuk fonem konsonan pada model phonetic similarity STANDUP	45
Tabel 5. 5 Contoh pasangan kata dan nilai word distance pasangan kata tersebut	48



DAFTAR GAMBAR

Gambar 2. 1 <i>Pseudo-code</i> Levenshtein distance.....	10
Gambar 4. 1 Source code Levenshtein distance [MEP].....	29
Gambar 4. 2 Implementasi program setelah modifikasi umum	35
Gambar 4. 3 Tambahan pada program untuk modifikasi masukan	36
Gambar 4. 4 Source code program ekstraksi kuesioner.....	39
Gambar 4. 5 Terusan source code program ekstraksi kuesioner.....	40
Gambar 5. 1 Perbandingan standar deviasi data kuesioner dengan selisih nilai data kuesioner dan model STANDUP	46

