

**PERBANDINGAN HASIL DETEKSI KEMIRIPAN TOPIK SKRIPSI  
DENGAN MENGGUNAKAN METODE N-GRAM DAN EKSPANSI**

**KUERI**



**SKRIPSI**

Disusun sebagai salah satu syarat menyelesaikan Program Studi Strata I  
pada Jurusan Teknik Informatika Fakultas Komunikasi dan Informatika

Universitas Muhammadiyah Surakarta

Oleh :

**DWI ISWANTO**

**NIM : L200100014**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS KOMUNIKASI DAN INFORMATIKA  
UNIVERSITAS MUHAMMADIYAH SURAKARTA**

**2014**

## HALAMAN PERSETUJUAN

Skripsi dengan judul

### PERBANDINGAN HASIL DETEKSI KEMIRIPAN TOPIK SKRIPSI DENGAN MENGGUNAKAN METODE N-GRAM DAN EKSPANSI KUERI

ini telah diperiksa, disetujui dan disahkan pada :

Hari : Sabtu .....

Tanggal : 13 Desember 2014.

Pembimbing



Husni Thamrin, S.T., M.T., Ph.D.

NIK : 706

## HALAMAN PENGESAHAN

### PERBANDINGAN HASIL DETEKSI KEMIRIPAN TOPIK SKRIPSI DENGAN MENGGUNAKAN METODE N-GRAM DAN EKSPANSI KUERI

dipersiapkan dan disusun oleh :

**DWI ISWANTO**

**NIM : L200100014**

telah dipertahankan di depan Dewan Penguji

pada tanggal 13 Desember 2014

#### Susunan Dewan Penguji

Pembimbing

Husni Thamrin, S.T., M.T., Ph.D.  
NIK : 706

Dewan Peguji I

Heru Supriyono, S.T., M.Sc.Ph.D.  
NIK : 970

Dewan Penguji II

Endah Sudarmilah, S.T., M.Eng.  
NIK : 969

Skripsi ini telah diterima sebagai salah satu persyaratan

Untuk memperoleh gelar sarjana

Tanggal 13 Desember 2014

Dekan

Fakultas Komunikasi dan Informatika

Husni Thamrin, S.T., M.T., Ph.D.  
NIK : 706

Ketua Program Studi  
Teknik Informatika



Heru Supriyono, S.T., M.Sc.Ph.D  
NIK : 970

## DAFTAR KONTRIBUSI

Bersama ini saya menyatakan bahwa skripsi ini tidak terdapat karya orang lain yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat orang lain yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Berikut saya sampaikan daftar kontribusi dalam penyusunan skripsi :

1. Saya merancang Program ini 80% sendiri dengan bantuan dari pembimbing, dan teman sebanyak 20%, buku, internet yang dilampirkan dalam daftar pustaka.
2. Saya merancang program ini menggunakan sebuah laptop dengan spesifikasi Intel® core I3® @2.13GHz, RAM 2GB, dan menggunakan aplikasi Xampp versi 1.8.2, *browser* google chrome versi 34, serta Sublime text 3.

Demikian pernyataan dan daftar kontribusi ini saya buat dengan sejurnya. Saya bertanggung jawab atas isi dan kebenaran daftar diatas.

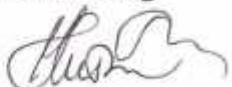
Surakarta, Desember 2014



Dwi iswanto

Mengetahui :

Pembimbing



Husni Thamrin, S.T., M.T., Ph.D.  
NIK : 706

## MOTTO DAN PERSEMBAHAN

### MOTTO

- Segala sesuatu jika dapat kita maksimalkan maka akan mendapatkan hasil yang maksimal juga.
- Memiliki impian, artinya kita punya alasan untuk terus berdoa, berjuang, dan selalu belajar dgn keras.
- Tujuan yang besar, selalu bisa dicapai dengan usaha yang besar pula. Dan usaha yang besar, selalu mengandung resiko yang tak kalah besarnya.
- Berhasil itu bukan berangan sesuatu yang besar, tapi menghargai yang kecil dan berusaha membesarkannya.
- seberapa dahsyat Cita-Cita mu, sedahsyat itu pula dinding yg kelak kan menghadang.
- Semangat yang dipupuk hari demi hari, akan menjauhi diri dari kemalasan yang diam-diam ingin menghampiri.
- Banyak Percakapan Menghasilkan Banyak Keluhan.  
Banyak Melakukan Menghasilkan Satu Percakapan,  
"Terimakasih Ya Allah, Saya Berhasil" .

## PERSEMBAHAN

- ✿ Allah SWT yang Maha Memberkati dan Maha Pemberi Nikmat yang senantiasa menjaga hati dan pikiran tetap konsisten selama proses penyusunan skripsi.
- ✿ Ibu Kandung ku tercinta yang selalu memberikan restu dan doanya.
- ✿ Bapak Ngadiman dan Ibu Sukini yang telah merawat serta mendidik ku sampai sekarang.
- ✿ Teman saya mas Nanang, mas Rozy dan seluruh teman-teman kos yang tidak bisa disebutkan satu persatu, dan Semua teman-temanku Teknik Informatika angkatan 2010.
- ✿ Seluruh partner kerja di Exabytes Indonesia yang telah mempercayai saya sampai saat ini.
- ✿ Seluruh dosen dan karyawan Teknik Informatika Universitas Muhammadiyah Surakarta.
- ✿ Rekan-rekan yang tidak dapat saya sebutkan satu persatu.

## KATA PENGANTAR

**Assalamu‘alaikum Wr. Wb.**

Syukur Alhamdulillah kita panjatkan kehadirat Allah SWT yang selalu memberikan rahmat, hidayah, kesehatan dan kekuatan sehingga penulis dapat menyelesaikan skripsi dengan judul “**PERBANDINGAN HASIL DETEKSI KEMIRIPAN TOPIK SKRIPSI DENGAN MENGGUNAKAN METODE N-GRAM DAN EKSPANSI KUERI**”.

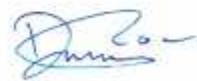
Skripsi ini disusun agar dapat memenuhi syarat dalam mendapatkan gelar sarjana pada Program Studi Teknik Informatika Fakultas Komunikasi dan Informatika Universitas Muhammadiyah Surakarta. Terselesaikannya skripsi ini tidak lepas dari dukungan dan bantuan dari berbagai pihak. Oleh karena itu dalam kesempatan ini penulis mengucapkan terimakasih kepada :

1. Allah SWT yang Maha Besar Engkau yang telah memberikan kesempatan kepada penulis untuk menyelesaikan skripsi ini.
2. Husni Thamrin, S.T, MT., Ph.D selaku Dekan Fakultas Komunikasi dan Informatika dan selaku pembimbing yang senantiasa memberikan waktu, nasehat dan bimbingannya kepada penulis sehingga dapat menyelesaikan skripsi ini.
3. Heru Supriyono , S.T, M.Eng selaku Ketua Jurusan Teknik Informatika Universitas Muhammadiyah Surakarta.
4. Kepada orang tua dan saudara-saudaraku yang selalu memberikan doa restu dan dukungan kepada penulis.

5. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah memberikan kontribusinya hingga skripsi ini selesai.

Akhir kata, penulis sangat berharap skripsi yang telah disusun ini dapat bermanfaat bagi penulis khususnya dan bagi seluruh pembaca pada umumnya untuk menambah pengetahuan kepada kita semua.

Surakarta, Desember 2014



Penulis

*Wassalamu'alaikum Wr. Wb.*

## DAFTAR ISI

|  |      |
|--|------|
| HALAMAN PERSETUJUAN .....                      | ii   |
| HALAMAN PENGESAHAN .....                       | iii  |
| HALAMAN KONTRIBUSI .....                       | iv   |
| HALAMAN MOTTO DAN PERSEMBAHAN .....            | v    |
| KATA PENGANTAR .....                           | vii  |
| DAFTAR ISI .....                               | ix   |
| DAFTAR TABEL .....                             | xii  |
| DAFTAR GAMBAR .....                            | xiii |
| ABSTRAKSI .....                                | xv   |
| BAB I PENDAHULUAN .....                        | 1    |
| A. Latar Belakang Masalah .....                | 1    |
| B. Rumusan Masalah .....                       | 2    |
| C. Batasan Masalah .....                       | 2    |
| D. Tujuan Penelitian .....                     | 3    |
| E. Manfaat Penelitian .....                    | 3    |
| F. Sistematika Penulisan .....                 | 4    |
| BAB II TINJAUAN PUSTAKA .....                  | 6    |
| A. Telaah Penelitian .....                     | 6    |
| B. Landasan Teori .....                        | 7    |
| 1. N-gram .....                                | 7    |
| 2. Ekspansi Kueri .....                        | 7    |
| 3. HTML 5 .....                                | 7    |
| 4. <i>bootstrap</i> .....                      | 8    |
| 5. <i>MySQL</i> .....                          | 8    |
| 6. PHP ( <i>Hypertext Preprocessor</i> ) ..... | 8    |

|   |    |
|---|----|
| BAB III METODE PENELITIAN .....                                     | 9  |
| A. Waktu dan Tempat Penelitian .....                                | 9  |
| B. Peralatan Utama dan Pendukung .....                              | 9  |
| 1. Perangkat Keras ( <i>Hardware</i> ) .....                        | 9  |
| 2. Perangkat Lunak ( <i>Software</i> ) .....                        | 9  |
| C. Alur Penelitian .....  | 10 |
| 1. Diagram Alir ( <i>Flowchart</i> ) Penelitian .....               | 10 |
| 2. Diagram Alir ( <i>Flowchart</i> ) N-gram.....                    | 12 |
| 3. Diagram Alir ( <i>Flowchart</i> ) Ekspansi Kueri .....           | 17 |
| D. Perancangan Sistem dengan DFD ( <i>Data Flow Diagram</i> ) ..... | 21 |
| 1. DFD Level 0 .....  | 21 |
| 2. DFD Level 1 .....  | 22 |
| E. Rancangan <i>Database</i> dan Tabel .....                        | 23 |
| 1. Relasi Antar Tabel .....   | 23 |
| 2. Tabel Instan Chart (TIC) .....                                   | 24 |
| F. Rancangan Desain Tampilan aplikasi .....                         | 31 |
| 1. Tampilan Homepage pencarian .....                                | 31 |
| 2. Halaman hasil pencarian .....                                    | 33 |
| 3. Halaman Admin .....  | 35 |
| G. Membangun Sistem .....   | 37 |
| <br>BAB IV HASIL DAN PEMBAHASAN .....                               | 38 |
| A. Hasil Penelitian .....   | 38 |
| 1. Hasil Tampilan .....   | 38 |
| 2. Hasil Pengujian .....  | 49 |
| 3. Hasil Pengujian jumlah hasil temu.....                           | 52 |
| 4. Hasil Pengujian jumlah kemiripan .....                           | 55 |
| 5. Menentukan Batas skor minimal.....                               | 58 |
| B. Analisa dan Pembahasan .....                                     | 61 |
| 1. Kelebihan metode N-gram .....                                    | 61 |

|   |           |
|---|-----------|
| 2. Kekurangan metode N-gram .....         | 62        |
| 3. Kelebihan metode ekspansi kueri.....   | 63        |
| 4. Kekurangan metode ekspansi kueri ..... | 64        |
| <b>BAB V PENUTUP .....</b>                | <b>65</b> |
| A. Kesimpulan .....                       | 65        |
| B. Saran .....                            | 66        |

## **DAFTAR PUSTAKA**

## **LAMPIRAN**

## **DAFTAR TABEL**

|   |    |
|---|----|
| Tabel 3.1 Tabel skripsi .....                                     | 24 |
| Tabel 3.2 Tabel Disclipine .....                                  | 25 |
| Tabel 3.3 Tabel Lex_class.....                                    | 25 |
| Tabel 3.4 Tabel <i>phrase</i> .....                               | 26 |
| Tabel 3.5 Tabel Ref_source .....                                  | 28 |
| Tabel 3.6 Tabel Relation .....                                    | 29 |
| Tabel 3.7 Tabel Rel_type .....                                    | 29 |
| Table 3.8 Table User .....  | 30 |
| Tabel 4.1 <i>Black Box</i> Halaman Depan .....                    | 50 |
| Tabel 4.2 <i>Black Box</i> Halaman Admin .....                    | 50 |
| Table 4.3 pengujian jumlah hasil temu N-gram .....                | 53 |
| Table 4.4 pengujian jumlah hasil temu ekspansi kueri .....        | 53 |
| Table 4.5 Data hasil uji kemiripan N-gram.....                    | 55 |
| Table 4.6 Data hasil uji kemiripan ekspansi kueri .....           | 56 |
| Table 4.7 Data batas skor minimal tidak mirip N-gram.....         | 58 |
| Table 4.8 Data batas skor minimal tidak mirip ekspansi kueri..... | 58 |

## **DAFTAR GAMBAR**

|   |    |
|---|----|
| Gambar 1 Diagram Alir ( <i>Flowchart</i> ) Penelitian ..... | 10 |
| Gambar 2.1 Flowchart N-gram (bagian 1) .....                | 13 |
| Gambar 2.1 Flowchart N-gram (bagian 2).....                 | 14 |
| Gambar 3.1 Flowchart ekspansi kueri (bagian 1) .....        | 18 |
| Gambar 3.1 Flowchart ekspansi kueri (bagian 2) .....        | 19 |
| Gambar 4.1 DFD Level 0 .....                                | 22 |
| Gambar 4.2 DFD Level 1 .....                                | 23 |
| Gambar 5 Entity Relationship .....                          | 24 |
| Gambar 6.1 Mockup Homepage Pencarian .....                  | 33 |
| Gambar 6.2 Mockup Halaman Hasil Pencarian .....             | 35 |
| Gambar 6.3 Prototype Halaman Admin.....                     | 39 |
| Gambar 7.1 Halaman Home .....                               | 39 |
| Gambar 7.2 Halaman N-gram .....                             | 40 |
| Gambar 7.3 Halaman Ekspansi kueri .....                     | 41 |
| Gambar 7.4 Halaman Hasil pencarian .....                    | 42 |
| Gambar 7.5 Halaman preview skripsi .....                    | 42 |
| Gambar 7.6 Tampilan pagination .....                        | 43 |
| Gambar 8.1 Tampilan Login Admin .....                       | 44 |
| Gambar 8.2 Tampilan logout Admin .....                      | 44 |
| Gambar 8.3 Tampilan Dashbord dmin.....                      | 45 |

|  |    |
|--|----|
| Gambar 8.4 Tampilan Action.....                                  | 45 |
| Gambar 8.5 Tampilan Edit Data .....                              | 46 |
| Gambar 8.6 Tampilan Hapus Data.....                              | 47 |
| Gambar 8.7 Tampilan button Add skripsi.....                      | 47 |
| Gambar 8.8 Tampilan Input Data .....                             | 48 |
| Gambar 9.1 Grafik Pengujian Jumlah hasil temu .....              | 54 |
| Gambar 9.2 Grafik Pengujian Jumlah kemiripan N-gram .....        | 57 |
| Gambar 9.3 Grafik Pengujian Jumlah kemiripan ekspansi kueri..... | 57 |

## ABSTRAKSI

Perbandingan deteksi kemiripan topik skripsi antara metode N-gram dan ekspansi kueri pada penelitian ini diartikan sebagai aplikasi pendeteksi skripsi berbasis web dengan menerapkan metode pencarian seperti metode N-gram dan ekspansi kueri lalu melakukan perbandingan di antara ke dua metode tersebut.

Saat ini jumlah topik skripsi yang dibuat semakin banyak. Topik skripsi satu dengan yang lainnya, bisa jadi akan serupa baik dari segi topik dan maksud yang akan di sampaikan, dilakukan upaya penyusunan program yang dapat melakukan sebuah proses deteksi kemiripan antar topik-topik skripsi tersebut. Jika dilakukan proses pencarian secara manual, hal tersebut sangat sulit dan membutuhkan waktu yang lama.

Tujuan dari penelitian ini, membangun sistem deteksi kemiripan topik skripsi dengan menerapkan metode pencarian yang mampu mendeteksi kemiripan dokumen skripsi, seperti N-gram dan ekspansi kueri.

Beberapa *software* yang digunakan, yaitu sistem operasi Windows 8, XAMPP versi 1.8.2 yang mendukung web server Apache, database MySQL dan Bahasa pemrograman PHP, Sublime text 3 sebagai *code editor*, *browser* Google Chrome 34.0 yang telah mendukung HTML 5.

Hasil ahir dari penelitian yang dilakukan telah terbukti bahwa metode N-gram dan ekspansi kueri mampu memberikan hasil deteksi kemiripan dan memberikan hasil berbeda di masing-masing metode yang di terapkan. Pemberian bobot di masing-masing metode akan sangat berpengaruh pada hasil pencarian. Penentuan batas skor minimal menentukan ukuran kemiripan dari setiap metode, ekspansi kueri memiliki hasil lebih baik di banding N-gram dalam deteksi kemiripan.

**Kata kunci :** metode N-gram, metode ekspansi kueri, deteksi kemiripan, topik skripsi, hasil temu. data mirip.