

TUGAS AKHIR

MENGKLASIFIKASI LAPORAN PRAKTEK KERJA LAPANGAN UNTUK MENCEGAH TERJADINYA PLAGIARISM MENGGUNAKAN *JACCARD MEASURE*

Disusun Untuk Melengkapi dan Memenuhi Syarat Kelulusan Program Strata 1
Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember

O L E H

**ADITYA ANGGRAINI YOGA KRESNA WIJAYANTI
NIM 1110651172**

**UNIVERSITAS MUHAMMADIYAH JEMBER
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA BIDANG
KEAHLIAN SISTEM BISNIS CERDAS**

2015



M O T T O

"Jika seseorang bepergian dengan tujuan mencari ilmu, maka Allah akan menjadikan perjalanannya seperti perjalanan menuju surga."
(*Nabi Muhammad SAW*)

Genius is one percent inspiration and ninety-nine percent perspiration.
(*Thomas Alva Edison*)

"Berotak London Berhati Masjidil Haram"
(*Dr. K.H. Musta'in Romly*)

"We must have perseverance and above all confidence in ourselves. We must believe that we are gifted for something."
(*Marie Curie*)

"Sesungguhnya sesudah kesulitan itu ada kemudahan. Maka apabila kamu telah selesai (dari suatu urusan), kerjakanlah dengan sungguh-sungguh (urusan) yang lain."
(*Q.S Al-Insyirah 6-7*)

"Kecerdasan tidak banyak berperan dalam proses penemuan. Ada suatu lompatan dalam kesadaran, sebutlah itu intuisi atau apapun namanya, solusinya muncul begitu saja dan kita tidak tahu bagaimana atau mengapa"
(*Albert Einstein*)

"Jangan terlalu mengharap pada apa yang dapat dunia berikan padamu, tetapi berikan yang terbaik pada dunia. Maka niscaya dunia pun bisa jadi tempat yang lebih indah"
(*Mario Teguh*)

"Ketika dunia ternyata jahat padamu,Maka kau harus menghadapinya. Karena tidak seorangpun yang akan menyelamatkanmu jika kau tidak berusaha."
(*Zorro*)

HALAMAN PERSETUJUAN

MENGKLASIFIKASI LAPORAN PRAKTEK KERJA LAPANGAN UNTUK MENCEGAH TERjadinya PLAGIARISM MENGGUNAKAN JACCARD MEASURE

Oleh :

ADITYA ANGGRAINI YOGA KRESNA WIJAYANTI (1110651172)

Telah diuji dan dipertahankan pada,

Hari dan Tanggal : Jum'at, 30 Januari 2015

Jam : 13.00 WIB

Tempat : Ruang 1.7 Gedung B Universitas Muhammadiyah Jember
Dewan Pengaji

Ketua

**1. Ir. Rusgianto, M.M
NIP. 13 1863 867**

Anggota

**2. Agung Nilogiri, S.T.,M.Kom
NIP. 19770330 200501 100 2**

*Jember, 30 Januari 2015
Mengetahui*

Dekan Fakultas Teknik

Ketua Program Studi Teknik Informatika

**Ir. Rusgianto, M.M
NIP. 13 1863 867**

**Agung Nilogiri, S.T., M.Kom
NIP. 19770330 200501 100 2**

HALAMAN PENGESAHAN

MENGKLASIFIKASI LAPORAN PRAKTEK KERJA LAPANGAN UNTUK MENCEGAH TERJADINYA PLAGIARISM MENGGUNAKAN JACCARD MEASURE

Oleh :

**ADITYA ANGGRAINI YOGA KRESNA WIJAYANTI
1110651172**

**Proyek Akhir Ini Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh
Gelar Sarjana Komputer (S.Kom)**
di
Universitas Muhammadiyah Jember

Disetujui oleh :

Tim Penguji

Dosen Pembimbing

- | | |
|---|---|
| 1. <u>Ir. Rusgianto, M.M</u>
NIP. 13 1863 867 | 1. <u>Ratih Avuninghemi, S.St, M.Kom</u>
 |
| 2. <u>Agung Nilogiri, S.T, M.Kom.</u>
NIP. 19770330 200501 1 002 | 2. <u>Deni Arifianto, S.Kom</u>
NPK. 11 03 588 |

Jember, 30 Januari 2015
Mengetahui

Dekan Fakultas Teknik

Ketua Program Studi Teknik Informatika

Ir. Rusgianto, M.M
NIP. 13 1863 867

Agung Nilogiri, S.T, M.Kom.
NIP. 19770330 200501 1002

PERNYATAAN

Yang bertanda tangan di bawah ini :

N a m a : ADITYA ANGGRAINI YOGA KRESNA WIJAYANTI
N I M : 1110651172
INSTITUSI : Strata-1 Program Studi Teknik Informatika, Fakultas Teknik,
Universitas Muhammadiyah Jember.

Menyatakan bahwa Tugas Akhir yang berjudul **“Mengklasifikasi Laporan Praktek Kerja Lapangan Untuk Mencegah Terjadinya Plagiarism Menggunakan Jaccard Measure”** bukan merupakan Tugas Akhir orang lain baik sebagian maupun keseluruhan kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya dan apabila pernyataan ini tidak benar penulis bersedia mendapatkan sanksi dari akademik.

Jember, 30 Januari 2015

Aditya Anggraini Yoga Kresna Wijayanti
NIM. 1110651172

MENGKLASIFIKASI LAPORAN PRAKTEK KERJA LAPANGAN UNTUK MENCEGAH TERJADINYA PLAGIARISM MENGGUNAKAN JACCARD MEASURE

¹Aditya Anggraini Yoga Kresna Wijayanti (1110651172),

²Ratih Ayuninghemi S.ST, M.Kom, ³Deni Arifianto, S.Kom

Jurusan Teknik Informatika Fakultas Teknik Univertas Muhammadiyah Jember

E-mail : aditya.taur05.ict@gmail.com

ABSTRAK

Dokumen laporan praktek kerja lapangan dibuat setelah melakukan praktek kerja lapangan pada suatu instansi pemerintah maupun swasta. pada laporan masih dapat ditemukan kemiripan atau plagiasi pada isi laporan tanpa mencantumkan sumber referensi. Hal ini diperlukan agar dapat meningkatkan kualitas dokumen laporan praktek kerja lapangan. Penelitian ini membahas kemiripan beberapa dokumen laporan untuk mendeteksi kesamaan dari isi dokumen tersebut. Untuk itu diperlukan suatu mekanisme mendeteksi plagiasi dokumen. Tahapan dalam metodologi penelitian terdiri dari: pengumpulan dokumen, isi laporan praktek kerja lapangan sebagai dataset, melakukan proses *text mining* yaitu : *casefolding*, *tokenizing*, *stemming* dan *filtering*, pembobotan dengan *TF-IDF*, perhitungan dengan metode *jaccard measure* sebagai alat untuk menukan apakah dokumen termasuk dinyatakan sebagai dokumen plagiat atau tidak plagiat dengan melihat dari isi laporan praktek kerja lapangan tersebut. Sebelum di lakukan penghitungan dengan *jaccard measure* , di lakukan pra proses *TF-IDF* terlebih dahulu, yaitu tingkat kemunculan kata dari suatu kalimat dengan koleksi dokumen yaitu jumlah yang akan di hitung bobotnya, dari bobot tersebut akan dijadikan masukkan untuk *jaccard measure*. Hasilnya berupa pengkalsifikasian tersebut telah diberikan 3 nilai ambang batas yaitu 0.5, 0.6, dan 0.7 . Data uji coba diambil dari beberapa laporan praktek kerja lapangan sebanyak 50 dokumen. Hasil pengujian menunjukan bahwa metode *jaccard measure*, mampu mengklasifikasikan dokumen plagiat maupun tidak plagiat dengan nilai ambang batas yang diberikan. *threshold* 0.5 sebesar 88%, *threshold* 0.6 sebesar 87%, dan *threshold* 0.7 sebesar 88% . Dengan hasil rata-rata akurasi sebesar 88%.

Kata kunci : *text mining*, *tf – idf*, *laporan praktek kerja lapangan*, *koleksi dokumen*, *jaccard measure*.

CLASSIFYING FIELD WORK PRACTICE REPORT TO PREVENT PLAGIARISM USE JACCARD MEASURE

¹*Aditya Anggraini Yoga Kresna Wijayanti (1110651172),*

²*Ratih Ayuninghemi S.ST, M.Kom,* ³*Deni Arifianto, S.Kom*

Department of Informatics Faculty of Engineering

University of Muhammadiyah Jember

E-mail : aditya.taur05.ict@gmail.com

ABSTRACT

Document field practice reports made after doing field work practice in a government agency or private. the report can still be found similarities or plagiarism on the contents of the report without including the source reference. This is necessary in order to improve the quality of the report documents the practice of field work. This research report discusses the similarity of some of the documents to detect the similarity from the contents of the document. For that required an mechanism to detect plagiarism document. Stages in research methodology consists of: collection of documents, the contents of the report as a dataset field practice, perform text mining process that is: casefolding, tokenizing, stemming and filtering, with TF-IDF the weighting, Jaccard measure calculation method as a tool for mentukan whether documents including expressed as a document plagiarism or not plagiarism by looking at the contents of the report of the field practice. Before doing the calculation with the Jaccard measure, in doing the pre-process of TF-IDF advance, that is the level of occurrence of the word of a sentence with a collection of documents that amount to be calculated weight, of the weight will be used as fill for the Jaccard measure. The result is pengkalsifikasian has been awarded three threshold values of 0.5, 0.6, and 0.7. Trial data taken from some field practice reports as many as 50 documents. Test results showed that the method Jaccard measure, able to classify documents plagiarism or not plagiarism with a given threshold value. The threshold 0.5 of 88%, the threshold 0.6 of 87%, and the threshold 0.7 of 88%. With an average accuracy result of 88%.

Keywords: *text mining, tf - idf, field practice reports, collections of documents, Jaccard measure.*

UNGKAPAN TERIMA KASIH

Bismillahirrahmanirrahiim.

Alhamdulillah, segala puji bagi Allah yang senantiasa menetapkan nikmat terbesar dalam hidup ini kepada saya berupa nikmat iman dan Islam Dengan rahmat-Nya penulis diberikan kemudahan dalam menyelesaikan studi dan di kampus Universitas Muhammadiyah Jember.

Atas segala upaya, bimbingan, dan arahan dari semua pihak, tanpa mengurangi rasa hormat penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Ir. Rusgianto M.M selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember dan selaku dosen pengaji 1 yang memberikan saran dan kritik yang membangun dalam penelitian ini,
2. Bapak Agung Nilogiri, S.T., M.Kom selaku Kepala Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember dan selaku dosen pengaji 2 yang memberikan saran dan kritik yang membangun dalam penelitian ini,,
3. Ibu Ratih Ayuninghemi, S.St., M.Kom selaku dosen pembimbing 1 yang juga telah meluangkan waktunya untuk membimbing dalam penelitian dan penulisan laporan ini,
4. Bapak Deni Arifianto, S.Kom selaku dosen pembimbing 2 yang juga telah meluangkan waktunya untuk membimbing dalam penelitian dan penulisan laporan ini,
5. Para dosen dan karyawan Fakultas Teknik Informatika, terima kasih atas semua ilmu yang telah diberikan,
6. Semua teman-teman yang membantu, menyemangati dan memotivasi saya untuk terus maju sehingga saya dapat menyelesaikan studi ini dengan tepat waktu.
7. Semua orang yang telah bersedia membagi ilmunya kepada saya. Semoga Allah meninggikan derajat mereka dengan ilmu.

8. Dan seluruh pihak yang namanya tidak mungkin saya sebutkan satu per satu di sini. Terimaskasih telah membantu saya dalam menyelesaikan studi dan proyek akhir ini.

Dan untuk semuanya yang tersebut di atas, semoga Allah menunjukkan kita pada jalan yang benar, menghimpunkan kita dengan orang-orang yang beriman dan berilmu, diampuni dosa-dosa kita dan senantiasa tercurah rahmat, taufiq dan hidayahNya kepada kita dan keluarga kita sekalian. Aamiin ...

Jember, 30 Januari 2015

Penulis

PERSEMBAHAN

Skripsi ini dengan hormat akan dipersembahkan kepada :

1. Ayahanda dan Ibunda tercinta yang telah memberikan dukungan baik moral maupun moril serta kasih sayangnya selamanya ini
2. Kakanda Ajad Suderajat yang selalu mensuport penulis dalam mengerjakan tugas akhir ini serta selalu ada untuk membantu dan menghibur penulis
3. Yayasan dan Para Alumni Pondok Pesantren Darul Ulum yang telah memberikan semangat untuk menjadi yang terbaik dalam mencari ilmu.
4. Ibu Ratih Ayuninghemi, S.St. M.Kom dan Bapak Deni Arifianto, S.Kom yang sudah meluangkan waktu untuk membantu dalam menyelesaikan tugas akhir ini,
5. Seluruh Keluarga Yogyakarta-Ponorogo yang selalu memberikan pandangan positif untuk kemasa yang akan datang,
6. Saudara Imam Syahroni, S.Kom yang sudah membagikan ilmunya dalam penggerjaan program,
7. Kakak tingkat angkatan 2008-2010 yang memberi banyak inspirasi dan pelajaran hidup,
8. Sahabat dan teman-temanku angkatan 2011 yang memberi banyak dukungan dalam menyelesaikan tugas akhir ini,
9. Almamaterku tercinta, Universitas Muhammadiyah Jember.

Jember, 30 Januari 2015
Penulis

KATA PENGANTAR



Puji syukur kehadirat Allah SWT yang Maha Pengasih lagi Maha Penyayang, Yang hanya kepadaNya-lah segala sesuatu bergantung. Alhamdulillah tak lupa senantiasa saya panjatkan karena hanya dengan ridho, kemurahan dan kekuasaanNya-lah proyek akhir yang berjudul:

“ MENGKLASIFIKASI LAPORAN PRAKTEK KERJA LAPANGAN UNTUK MENCEGAH TERJADINYA PLAGIARISM MENGGUNAKAN JACCARD MEASURE”

dapat diselesaikan dengan segala kelebihan dan tak lepas dari kekurangan yang terdapat di dalamnya.

Shalawat serta salam semoga senantiasa tercurah kepada baginda Rasulullah Muhammad SAW, keluraga beliau dan para sahabat hingga pengikutnya hingga akhir zaman, orang-orang yang senantiasa istiqomah menegakkan kebenaran dan menebar kebaikan di bumi Allah SWT.

Proyek akhir ini menjelaskan tentang bagaimana mengklasifikasian dokumen menjadi dokumen yang plagiat atau tidak menggunakan metode *jaccard measure*.

Dengan segala kerendahan hati, penulis memohon maaf jika ternyata di kemudian hari diketahui bahwa hasil dari proyek akhir ini masih jauh dari kesempurnaan. Semoga hasil dari proyek akhir ini dapat mempermudah dalam proses penentuan kualitas tanaman tembakau sehingga lebih akurat. Dan lebih dari itu semoga bermanfaat bagi setiap insan yang mempergunakannya untuk kebaikan di jalan Allah SWT.

Jember, 30 Januari 2015
Penulis

DAFTAR ISI

LEMBAR SAMPUL	i
LEMBAR LOGO	ii
MOTTO	iii
LEMBAR PERSETUJUAN	iv
LEMBAR PENGESAHAN	v
LEMBAR PERNYATAAN	vi
ABSTRAK	vii
ABSTRACT	viii
LEMBAR UCAPAN TERIMAKASIH	ix
LEMBAR PERSEMBAHAN	xi
KATA PENGANTAR.....	xii
DAFTAR ISI.....	xiii
DAFTAR TABEL	xv
DAFTAR GAMBAR.....	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan	3
1.5 Manfaat	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Praktek Kerja Lapangan.....	4
2.2 Laporan	4
2.3 Plagiasi	5
2.4 Klasifikasi	6
2.5 <i>Text Mining</i>	7
2.6 Tf/IDF	15

2.7 <i>Jaccard Measure</i>	19
BAB III METODE PENELITIAN	21
3.1 Tahapan Penelitian	21
3.2 Desain Sistem	22
3.2.1 <i>Scanning</i>	23
3.2.2 <i>CASEFOLDING</i>	23
3.2.3 <i>TOKENIZING</i>	24
3.2.4 <i>STEMMING</i>	25
3.2.5 <i>Filtering</i> dan Pembobotan TF/IDF	26
BAB IV HASIL DAN PEMBAHASAN.....	31
4.1 Data.....	31
4.2 Rekayasa Uji Coba	32
4.2.1 Rekayasa Uji Coba Data <i>Training</i>	32
4.2.2 Rekayasa Uji Coba Data <i>Testing</i>	32
4.2.3 Rekayasa Uji Coba Perhitungan TF/IDF dan <i>Jaccard Measure</i>	32
4.3 Analisa Hasil	32
4.3.1 Hasil Pengujian	32
4.3.2 <i>Accuracy</i>	33
4.3.3 <i>Precision</i>	36
4.3.4 <i>Recall</i>	39
4.4 Analisa Perbandingan Hasil Pengujian Dengan <i>Threshold</i> 0.5, 0.6, dan 0.7	42
BAB V KESIMPULAN DAN SARAN.....	44
3.1 Kesimpulan.....	44
3.2 Saran	44
DAFTAR PUSTAKA	46
LAMPIRAN.....	48

DAFTAR TABEL

Tabel 2.1	Hapus Awalan Jika Ditemukan	10
Tabel 2.2	Menentukan Tipe Awalan Untuk Kata Yang Diawali Dengan “te-“	11
Tabel 2.3	Jenis Awalan Berdasarkan Tipe Awalannya	11
Tabel 2.4	Aturan Untuk <i>Inflectional Particle</i>	12
Tabel 2.5	Aturan Untuk <i>Inflectional Possesive Pronoun</i>	13
Tabel 2.6	Aturan Untuk <i>First Order Derivational Prefix</i>	13
Tabel 2.7	Aturan Untuk <i>Second Order Derivational Prefix</i>	14
Tabel 2.8	Aturan Untuk <i>Derivational Suffix</i>	14
Tabel 2.9	<i>Term Frequency</i>	17
Tabel 2.10	Menghitung W (Pembobotan)	18
Tabel 3.1	<i>Tokenizing</i>	24
Tabel 3.2	<i>Stemming</i>	25
Tabel 3.3	<i>Filtering</i>	26
Tabel 3.4	Pembobotan	27
Tabel 3.5	Jumlah Bobot	29
Tabel 4.1	Perhitungan <i>Accuracy</i> dengan nilai <i>threshold</i> 0.5	34
Tabel 4.2	Perhitungan <i>Accuracy</i> dengan nilai <i>threshold</i> 0.6	34
Tabel 4.3	Perhitungan <i>Accuracy</i> dengan nilai <i>threshold</i> 0.7	35
Tabel 4.4	Hasil pengujian data plagiat dalam <i>Accuracy</i> dengan menggunakan seluruh <i>threshold</i>	36
Tabel 4.5	Perhitungan <i>Presicion</i> dengan nilai <i>threshold</i> 0.5.....	37
Tabel 4.6	Perhitungan Presicion dengan nilai threshold 0.6	37
Tabel 4.7	Perhitungan <i>Presicion</i> dengan nilai <i>threshold</i> 0.7.....	38
Tabel 4.8	Hasil pengujian data plagiat dalam <i>Presicion</i> dengan menggunakan seluruh <i>threshold</i>	39
Tabel 4.9	Perhitungan <i>Recall</i> dengan nilai <i>threshold</i> 0.5	40
Tabel 4.10	Perhitungan <i>Recall</i> dengan nilai <i>threshold</i> 0.6.....	40
Tabel 4.11	Perhitungan <i>Recall</i> dengan nilai <i>threshold</i> 0.7	41

Tabel 4.12 Hasil pengujian data plagiat dalam <i>Recall</i> dengan menggunakan seluruh <i>threshold</i>	42
Tabel 4.13 Analisa hasil pengujian seluruh <i>threshold</i>	43

DAFTAR GAMBAR

Gambar 2.1 Kalsifikasi Sebagai Pemetaan	7
Gambar 2.2 <i>Text Mining</i>	8
Gambar 2.3 <i>Stemming</i>	14
Gambar 2.4 Contoh Penerapan Model Ruang Vektor Dokumen.....	19
Gambar 3.1 Desain Sistem.....	22

DAFTAR LAMPIRAN

LEMBAR LAMPIRAN	47
LAMPIRAN 1	48
LAMPIRAN 2	65
LAMPIRAN 3	92