

**EVALUASI PENGUKURAN JARAK PADA ALGORITMA RABIN
KARP UNTUK MENDETEKSI PLAGIASI BILINGUAL**

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Jenjang Strata Satu (S1)
Pada Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Ponorogo



IMAS PRADHIPTA BAGASKARA

16532674

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK**

UNIVERSITAS MUHAMMADIYAH PONOROGO

2020

HALAMAN PENGESAHAN

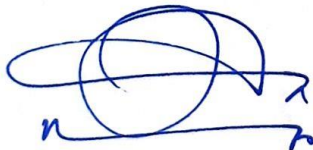
Nama : Imas Pradhipta Bagaskara
NIM : 16532674
Program Studi : Teknik Informatika
Fakultas : Teknik
Judul Skripsi : "Evaluasi Pengukuran Jarak Pada Algoritma Rabin Karp Untuk Mendeteksi Plagiasi Bilingual"

Isi dan formatnya telah disetujui dan dinyatakan memenuhi syarat untuk melengkapi persyaratan guna memperoleh Gelar Sarjana pada Program Studi Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Ponorogo.

Ponorogo, Agustus 2020

Menyetujui,

Dosen Pembimbing I,



Dra. Ida Widaningrum, M.Kom

NIK. 19660417 201101 13

Dosen Pembimbing II,



Dyah Mustikasari, S.T., M.Eng

NIK. 19871007 201609 13

Mengetahui,

Dekan Fakultas Teknik,



Dr. Ir. Aliyadi, M.M., M.Kom

NIK. 19640103 19909 12

Ketua Program Studi Teknik Informatika,



Dyah Mustikasari, S.T., M.Eng

NIK. 19871007 201609 13

PERNYATAAN ORISINALITAS SKRIPSI

Yang bertandatangan dibawah ini :

Nama : Imas Pradhipta Bagaskara

NIM : 16532674

Program Studi : Teknik Informatika

Dengan ini menyatakan bahwa Skripsi saya dengan judul : “Evaluasi Pengukuran Jarak Pada Algoritma Rabin Karp Untuk Mendeteksi Plagiasi Bilingual” bahwa berdasarkan hasil penelusuran berbagai karya ilmiah, gagasan, dan masalah ilmiah yang saya rancang / teliti di dalam Naskah Skripsi ini adalah asli dari pemikiran saya. Tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip dalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

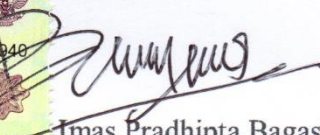
Apabila ternyata di dalam Naskah Skripsi ini dapat dibuktikan terdapat unsur- unsur plagiarisme, saya bersedia Ijazah saya dibatalkan, serta diproses sesuai dengan peraturan perundang- undangan yang berlaku.

Demikian pernyataan ini dibuat dengan sesungguhnya dan dengan sebenar- benarnya.

Ponorogo, Agustus 2020

Mahasiswa,




Imas Pradhipta Bagaskara
NIM. 16532674

HALAMAN BERITA ACARA UJIAN

Nama : Imas Pradhipta Bagaskara
NIM : 16532674
Program Studi : Teknik Informatika
Fakultas : Teknik
Judul Skripsi : "Evaluasi Pengukuran Jarak Pada Algoritma Rabin Karp Untuk Mendeteksi Plagiasi Bilingual"

Telah diuji dan dipertahankan dihadapan Dosen penguji tugas akhir jenjang Strata

Satu (S1) pada :

Hari : Rabu
Tanggal : 5 Agustus 2020
Nilai :

Dosen Penguji,

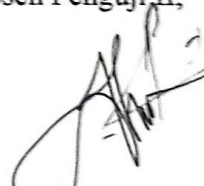
Dosen Penguji I,



Dr. Aslan Alwi, S.SI., M.Cs

NIK. 19720324 201101 13

Dosen Penguji II,



Indah Puji Astuti, S.Kom., M.Kom

NIK. 19860424 201609 13

Mengetahui,

Dekan Fakultas Teknik,



Dr. Ir. Aliyadi, M.M., M.Kom

NIK. 19640103 19909 12

Ketua Program Studi Teknik Informatika,








Dyah Mustikasari, S.T., M.Eng

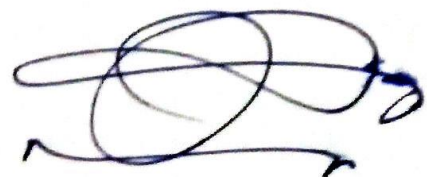
NIK. 19871007 201609 13

BERITA ACARA
BIMBINGAN SKRIPSI

13. Nama : Imas Pradhipta Bagaskara
14. NIM : 16532674
15. Program Studi : Teknik Informatika
16. Fakultas : Teknik
10. Judul Skripsi : Evaluasi Pengukuran Jarak Pada Algoritma Rabin-Karp Untuk Mendeteksi Plagiasi Bilingual
17. Dosen Pembimbing : Dra. Ida Widaningrum, M.Kom.
18. Konsultasi :

| NO | TANGGAL | URAIAN | TANDA TANGAN |
|----|---------------|--|---|
| 1. | 14 April 2020 | pembahasan dataset dan penentuan dataset |  |
| 2. | 5 Juni 2020 | Demo Aplikasi |  |
| 3. | 22 Juli 2020 | Revisi use case, Sequence Diagram |  |
| 4. | 29 Juli 2020 | Revisi penulisan nomor 2 Abstrak |  |
| 5. | 30 Juli 2020 | Ace Signy |  |






Ponorogo, 30 Juli 2020
Pembimbing,




(Dra. Ida Widaningrum, M.Kom)
NIK. 19660417 201101 13

**BERITA ACARA
BIMBINGAN SKRIPSI**

7. Nama : Imas Pradhipta Bagaskara
8. NIM : 16532674
9. Program Studi : Teknik Informatika
10. Fakultas : Teknik
9. Judul Skripsi : Evaluasi Pengukuran Jarak Pada Algoritma Rabin-Karp Untuk Mendeteksi Plagiasi Bilingual
11. Dosen Pembimbing : Dyah Mustikasari, S.T. M.Eng.
12. Konsultasi :

| NO | TANGGAL | URAIAN | TANDA TANGAN |
|----|---------------|--|---|
| 1 | 14 April 2020 | Pembahasan dataset dan penentuan dataset |  |
| 2 | 5 Juni 2020 | Demo aplikasi sbg media evaluasi |  |
| 3 | 22 Juli 2020 | Revisi usecase, Sequence diagram |  |
| 4 | 29 Juli 2020 | Revisi penulisan nomor dan abstrak |  |
| 5 | 30 Juli 2020 | Aae siday |  |

Ponorogo, Juli 2020
Pembimbing,


(Dyah Mustikasari, S.T. M.Eng)
NIK. 19871007 201609 13



UNIVERSITAS MUHAMMADIYAH PONOROGO
UNIT PELAKSANA TEKNIS PERPUSTAKAAN
Jalan Budi Utomo 10 Ponorogo 63471 Jawa Timur Indonesia
Telp (0352) 481124, 487662 Fax (0352) 461796, *Website: library.umpo.ac.id*
TERAKREDITASI A
(SK Nomor 00012/ LAP.PT/ I.2017)

SURAT KETERANGAN
HASIL PEMERIKSAAN PLAGIASI SKRIPSI MAHASISWA
UNIVERSITAS MUHAMMADIYAH PONOROGO

Dengan ini kami nyatakan bahwa skripsi dengan rincian sebagai berikut:

Nama : Imas Pradhipta Bagaskara

NIM : 16532674

Prodi : Informatika

Judul : Evaluasi Pengukuran Jarak Pada Algoritma Rabin Karp Untuk Mendeteksi Plagiasi

Bilingual

Dosen pembimbing :

1. Dra. Ida Widaningrum, M.Kom. email : iwidaningrum.as@gmail.com
2. Dyah Mustikasari, S.T. M.Eng. email : dyah.mustikasari@gmail.com

Telah dilakukan check plagiasi di UPT. Perpustakaan Universitas Muhammadiyah Ponorogo dengan prosentase plagiasi sebesar 16 %

Demikian keterangan ini dibuat untuk digunakan sebagaimana mestinya.

Ponorogo, 04 Agustus 2020

Pemeriksa



(Mohamad Ulil Albab, SIP)

NIK.1989092720150322

Nb: Dosen pembimbing dimohon untuk mengecek kembali keaslian soft file karya ilmiah yang telah dicek di perpustakaan.



UNIVERSITAS MUHAMMADIYAH PONOROGO
UNIT PELAKSANA TEKNIS PERPUSTAKAAN
Jalan Budi Utomo No. 10 Ponorogo 63471 Jawa Timur Indonesia
Telp. (0352) 481124, Fax (0352) 461796, e-mail : lib@umpo.ac.id
website : www.library.umpo.ac.id

SURAT KETERANGAN
HASIL PEMERIKSAAN ANTI PLAGIASI ARTIKEL ILMIAH MAHASISWA
UNIVERSITAS MUHAMMADIYAH PONOROGO

Telah di periksa, artikel ilmiah dengan rincian sebagai berikut :

Nama : Imas Pradhipta Bagaskara
Judul : Evaluasi Pengukuran Jarak Pada Algoritma Rabin Karp Untuk Mendeteksi Plagiasi Bilingual
Fakultas / Prodi : Fakultas Teknik / Teknik Informatika

Dosen pembimbing :

1. Dra. Ida Widaningrum, M.Kom. email : iwidaningrum.as@gmail.com
2. Dyah Mustikasari, S.T. M.Eng. email : dyah.mustikasari@gmail.com

Tingkat kesamaan pada artikel sebesar 9%

Menggunakan aplikasi anti-plagiasi *Turnitin*. Demikian, atas perhatiannya di ucapkan terima kasih.

Ponorogo, 16 September 2020

Pemeriksa



(Mohamad Ulil Albab, SIP)

NIK.1989092720150322

Nb: Dosen pembimbing dimohon untuk mengecek kembali keaslian soft file karya ilmiah yang telah dicek di perpustakaan.

HALAMAN MOTTO

“Seorang professional pastinya tidak alay dan yang pasti suka bercanda”



HALAMAN PERSEMBAHAN

Dengan nama Allah yang maha pengasih lagi maha penyayang, tuhan seru sekalian alam. Sembah sujud serta syukur atas rahmat dan ridho-Nya yang telah membekali hambanya akal dan pikiran sehingga segala kekuatan, ketabahan, dan ilmu yang bermanfaat akhirnya skripsi yang masih jauh dari kata sempurna ini dapat terselesaikan. Sholawat serta salam tetap tercurahan kepada junjungan nabi besar Muhammad SAW.

Kupersembahkan karya ini kepada orang yang sangat kucintai.

Bapak dan ibuku Tercinta

Demi senyum bahagia dari kedua orang tuaku Tanpa mengurangi rasa hormatku, semoga karya ini menjadi langkah awal untuk mewujudkan keinginan kedua orang tuaku berikutnya. Terima kasih yang sebesar-besarnya kepada mas sematawayangku. Semoga semangat dan motivasi yang telah kalian berikan akan menjadikan energi positif untuk meraih cita-citaku berikutnya. Sekali lagi, Terimakasih....

Dosen Pembimbing Tugas Akhir

Ibu Ida Widaningrum dan Bu Dyah Mustikasari selaku dosen pembimbing skripsi saya, terima kasih banyak sudah membantu selama ini, sudah dinasehati, sudah diajari, dan mengarahkan saya sampai skripsi ini selesai.

KATA PENGANTAR

Puji Syukur marilah kita haturkan kepada Tuhan Yang Maha Esa, Allah SWT, Tuhan Semesta Alam, yang mana atas semua berkat, rahmat, taufik, serta hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Evaluasi Pengukuran Jarak Pada Algoritma Rabin Karp Untuk Mendeteksi Plagiasi Bilingual”.

Pada penyusunan Skripsi ini, telah banyak melibatkan beberapa pihak yang sangat membantu dan mendukung dalam berbagai hal. Maka dari itu, sudah penulis menyampaikan banyak terimakasih kepada :

1. Ir. Aliyadi M.M.,M.Kom., selaku Dekan Fakultas Teknik Universitas Muhammadiyah Ponorogo.
2. Dyah Mustikasari, ST, M.Eng., selaku Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Ponorogo. Sekaligus Dosen Pembimbing II
3. Ibu Dra. Ida Widaningrum, M.Kom., Selaku Dosen Pembimbing I
4. Bapak dan Ibu Dosen Program Studi Teknik Informatika Universitas Muhammadiyah Ponorogo yang telah memberikan ilmunya selama perkuliahan.
5. Orang Tua tercinta dan keluarga yang telah memberikan doa dan dukungan kepada penulis
6. Teman-teman Teknik Informatika Angkatan 2016 khususnya kelas D Universitas Muhammadiyah Ponorogo.
7. Dan juga semua pihak yang tidak bisa disebutkan satu-persatu terimakasih atas segala dukungannya.

Semoga Allah SWT. Memberikan balasan yang jauh sempurna. Penulis menyadari bahwa skripsi ini masih banyak mengandung kekurangan. Oleh karena itu, penulis sangat mengharapkan saran dan kritik yang membangun. Semoga skripsi ini bermanfaat bagi kita semua. Aamiin,

EVALUASI PENGUKURAN JARAK PADA ALGORITMA RABIN KARP UNTUK MENDETEKSI PLAGIASI BILINGUAL

Imas Pradhipta Bagaskara ¹⁾, Ida Widaningrum ²⁾, Dyah Mustikasari ³⁾

^{1,2,3)} Program Studi Teknik Informatika, Fakultas Teknik,

Universitas Muhammadiyah Ponorogo

e-mail: pradhiptabagaskara@gmail.com ¹⁾, iwidaningrum.as@gmail.com ²⁾,
dyah.mustikasari@gmail.com ³⁾

ABSTRAK

Pembandingan teks karya ilmiah bisa dengan cara menghitung kata atau kalimat yang sama dan ada pada karya ilmiah lainnya. Namun untuk cara ini membutuhkan pemrosesan yang panjang mengingat semua katanya akan dibandingkan serta dipengaruhi banyaknya teks karya ilmiah yang dibandingkan. Cara lain untuk membandingkan teks karya ilmiah adalah dengan menjadikannya sebagai *fingerprint*. Metode *Fingerprinting* Algoritma Rabinkarp merupakan salah satu metode yang mengubah dokumen menjadi kumpulan hash yang mewakili teks karya ilmiah dan mempersingkat penulisan karena tiap kata atau kalimat diperingskas dengan hashing. Penghitungan jarak kesamaan menggunakan beberapa metode pengukuran. Metode tersebut adalah penghitungan jarak yang akan diterapkan yaitu Cosine, Jaccard, Dice, Euclidean, Manhattan, Minkowski, Mahalanobis dan Weighted. Tiap rumus berbeda hasil pengukuran satu dan lainnya yang merupakan karakteristik dari masing-masing metode pengukuran. Dari hasil pengujian menunjukkan bahwa akurasi pengukuran terhadap nilai aproksimasi yang ditentukan menunjukkan metode pengukuran Cosine Distance merupakan metode yang lebih baik dibanding ke 7 metode pengukuran lainnya pada penelitian ini. Ini dibuktikan dengan rerata selisih 0.585 dan cukup stabil dengan standar deviasi 1.00. Pada metode Dice Distance memiliki selisih rerata 0.691 dan standar deviasi 1.00. Metode Jaccard Distance memiliki akurasi yang kurang baik ini dibuktikan dengan selisih reratanya 1.00 dan standar deviasi 1.00. Pada metode Manhattan memiliki akurasi yang cukup baik ini dibuktikan dengan selisih reratanya 0.052 dan standar deviasi 1.00. Pada metode Minkowski memiliki akurasi yang buruk karena melebihi standar deviasi yaitu selisih rerata 1.13 dan deviasi 1.00. Metode Euclidean juga memiliki akurasi yang kurang baik bisa dibuktikan dengan selisih rerata 1.32 dan deviasi 1.00. Metode Mahalanobis memiliki akurasi cukup stabil tapi kurang baik dalam akurasi ini dibuktikan dengan selisih rerata 0.86 dan deviasi 1.00. Weighted memiliki akurasi yang cukup baik dengan selisih rerata 0.03 dan deviasi 1.00.

Kata Kunci: Algoritma Rabin-Karp, *Cosine Distance*, *Jaccard Distance*, *Dice Distance*, *Euclidean Distance*, *Manhattan Distance*, *Minkowski Distance*, *Mahalanobis Distance*, *Weighted Distance*

DAFTAR ISI

| | |
|--|------|
| HALAMAN JUDUL | i |
| HALAMAN PENGESAHAN | ii |
| PERNYATAAN ORISINALITAS SKRIPSI | iii |
| HALAMAN BERITA ACARA UJIAN | iv |
| BERITA ACARA BIMBINGAN | v |
| SURAT KETERANGAN HASIL PLAGIASI SKRIPSI | vii |
| SURAT KETERANGAN HASIL PLAGIASI ARTIKEL | viii |
| HALAMAN MOTTO | ix |
| HALAMAN PERSEMBAHAN | x |
| KATA PENGANTAR | xi |
| DAFTAR ISI | xiii |
| DAFTAR GAMBAR | xv |
| BAB I PENDAHULUAN | 1 |
| 1.1. Latar Belakang | 1 |
| 1.2. Rumusan Masalah | 3 |
| 1.3. Tujuan | 3 |
| 1.4. Manfaat | 3 |
| 1.5. Batasan Masalah | 4 |
| BAB II TINJAUAN PUSTAKA | 5 |
| 2.1. Penelitian Terdahulu | 5 |
| 2.2. Plagiarisme | 7 |
| 2.3. Algoritma Rabin Karp | 7 |
| 2.4. Cosine | 8 |
| 2.5. Jaccard | 8 |
| 2.6. Koefisien Dice | 9 |
| 2.7. Euclidean | 9 |
| 2.8. Manhattan | 10 |
| 2.9. Minkowski | 10 |
| 2.10. Mahalanobis | 11 |
| 2.11. Weighted Distance | 11 |
| 2.12. Bilingual | 12 |
| 2.13. Python | 12 |

| | |
|---|-----------|
| BAB III METODE PENELITIAN | 13 |
| 3.1. Desain Penelitian..... | 13 |
| 3.2. Alat dan Bahan Penelitian | 15 |
| 3.2.1. Alat Penelitian..... | 15 |
| 3.2.2. Bahan Penelitian | 15 |
| 3.3. Metode Penelitian..... | 15 |
| 3.3.1. Metode Pengumpulan Data | 16 |
| 3.3.2. Metode Pengembangan Perangkat Lunak..... | 16 |
| 3.4. Model Flowcart System | 19 |
| 3.5. Skenario Eksperimen..... | 22 |
| BAB IV PEMBAHASAN..... | 25 |
| 4.1. Analisis Sistem | 25 |
| 4.1.1. Penggambaran Sistem | 25 |
| 4.1.2. Use Case Diagram | 26 |
| 4.1.3. Sequence Diagram Kelola Dataset | 27 |
| 4.2. Hasil Perancangan | 28 |
| 4.3. Pembahasan Alur Sistem..... | 41 |
| 4.3.1. Extract text..... | 42 |
| 4.3.2. Translating | 42 |
| 4.3.3. Rabin-Karp Fingerprinting..... | 42 |
| 4.4. Hasil Uji Data..... | 45 |
| BAB V PENUTUP..... | 47 |
| 5.1. Kesimpulan | 47 |
| 5.2. Saran | 49 |
| LAMPIRAN | 52 |

DAFTAR GAMBAR

| | |
|---|----|
| Gambar 3. 1 Desain Penelitian | 13 |
| Gambar 3. 2 Model Spiral (Sommerville 2011) | 16 |
| Gambar 3. 3 Desain Sistem | 19 |
| Gambar 3. 4 Skenario Eksperimen | 22 |
| Gambar 4. 1 Use Case Diagram..... | 26 |
| Gambar 4. 2 Sequence Diagram Kelola Dataset..... | 27 |
| Gambar 4. 3 Halaman Login | 29 |
| Gambar 4. 4 Menu Navigasi..... | 30 |
| Gambar 4. 5 Halaman Check How It Works..... | 31 |
| Gambar 4. 6 Percobaan Check How It Works..... | 31 |
| Gambar 4. 7 Halaman Hasil Check How It Works..... | 32 |
| Gambar 4. 8 Halaman All..... | 33 |
| Gambar 4. 9 Halaman Result..... | 34 |
| Gambar 4. 10 Halaman Upload | 35 |
| Gambar 4. 11 Halaman All Datasets..... | 36 |
| Gambar 4. 12 Halaman Show Dataset | 36 |
| Gambar 4. 13 Halaman Edit Dataset..... | 37 |
| Gambar 4. 14 Halaman Delete Dataset | 37 |
| Gambar 4. 15 Halaman Upload Dataset..... | 38 |
| Gambar 4. 16 Halaman Upload Dataset Batch..... | 39 |
| Gambar 4. 17 Halaman Percobaan..... | 40 |
| Gambar 4. 18 Halaman Hasil Plagiasi | 40 |
| Gambar 4. 19 Halaman Pengukuran | 41 |
| Gambar 4. 20 Halaman Pengukuran Lanjutan..... | 41 |
| Gambar 4. 21 Sanatize | 43 |
| Gambar 4. 22 N-Gram dengan panjang 7 kata | 43 |
| Gambar 4. 23 Hasil Rolling Hash | 44 |
| Gambar 4. 24 Hasil Fingerprint | 44 |
| Gambar 4. 25 Percobaan Uji Data 1 | 45 |
| Gambar 4. 26 Percobaan Uji Data 2 | 45 |