

PERANCANGAN APLIKASI PENDETEKSI TINGKAT KESAMAAN ANTAR DOKUMEN DENGAN ALGORITMA WINNOWING

Arnold Pramudita Tjiawi, Dyah E. Herwindiati, Lely Hiryanto

Teknik Informatika Universitas Tarumanagara

Jl. Letjen S.Parman No.1, Jakarta 11440 Indonesia

email : tjiawi.rush@gmail.com

Abstrak

Perancangan aplikasi pendeteksian tingkat kesamaan antar dokumen ini dimaksudkan untuk menyimpan data tesis mahasiswa dan aplikasi dapat digunakan untuk membandingkan data skripsi mahasiswa dengan data baru yang ada. Hasil dari proses deteksi adalah persentase tingkat kesamaan. Algoritma yang digunakan dalam merancang aplikasi ini adalah Algoritma Winnowing. Algoritma ini termasuk dalam salah satu metode fingerprinting dokumen. Algoritma ini menggunakan rolling hashing untuk melakukan proses hashing dan menggunakan koefisien jaccard untuk menghitung tingkat kemiripan. Dalam pengujian aplikasi ini dengan data dummy diketahui bahwa aplikasi ini sangat bergantung pada urutan masing-masing lokasi substring. Aplikasi ini juga telah diuji untuk membandingkan sejumlah data skripsi yang ada dengan persentase tingkat kesamaan di bawah 30%. Aplikasi ini tidak bisa menentukan apakah sebuah dokumen plagiat atau tidak, namun aplikasi ini bisa memberikan informasi berupa persentase tingkat kesamaan dan persentase dapat digunakan oleh pihak yang memiliki wewenang untuk menentukan.

Kata kunci—Algoritma Winnowing, plagiarisme, rolling hashing, tingkat kesamaan antar dokumen, jaccard coefficient

Abstract

The design of detection application for similarity level between documents is intended to store student thesis data and applications can be used to compare a student thesis data with existing with new data. The result of the detection process is the percentage of similarity level. The algorithm used in designing this application is winnowing algorithm. This algorithm is included in one of the document fingerprinting methods. This algorithm uses rolling hashing to do the hashing process and use jaccard coefficient to calculate the similarity level. In testing this application with dummy data it is found that this application is very dependent on the order of each substring location. This application has also been tested to compare a number of existing thesis data with a percentage of similarity rate below 30%. This application can not specify a plagiarism document or not, but this application can provide information in the form of percentage level of similarity and percentage can be used by who has the authority to determine.

Keywords—Document Fingerprinting, Jaccard Coefficient, Plagiarism, Rolling Hashing, Similarity, Winnowing algorithm

1. PENDAHULUAN

Penulisan karya ilmiah dalam lingkup pendidikan tinggi merupakan suatu hal yang wajib ada. Pada jenjang Strata satu (S1) dikenal makalah dan skripsi (tugas akhir). Dalam penulisan skripsi biasa diperlukan data ataupun informasi-informasi pendukung. Penulisan karya ilmiah tidak terlepas dari sumber-sumber pustaka yang dijadikan acuan penulisan. Sumber-sumber pustaka didapat dari buku, koran, majalan jurnal dan media tulis lainnya. Data dan informasi ini dapat dengan mudah diakses dengan adanya kemajuan teknologi informasi yang ada sekarang. Tetapi kemudahan ini memiliki efek negatif, seperti penjiplakan terhadap karya ilmiah orang lain. Hal ini berbahaya dan merugikan bagi pemilik karya ilmiah maupun bagi institusi tempat penjiplak berada.

Kegiatan penjiplakan ini biasa disebut sebagai plagiarisme. Plagiarisme ini sulit untuk di deteksi jika dengan kasat mata, karena banyaknya informasi-informasi yang biasa terdapat dalam sebuah karya ilmiah. Untuk itu diperlukan sebuah perancangan aplikasi pendeteksi tingkat kesamaan antar dokumen. Aplikasi yang dirancang ini akan menggunakan algoritma winnowing. Algoritma ini menggunakan *rolling hashing* untuk melakukan *hashing* dan menggunakan *jaccard coefficient* untuk menghitung tingkat kesamaan. Hasil yang diharapkan dari aplikasi ini adalah sebuah prosentase tingkat kesamaan antar dokumen satu dengan dokumen lainnya, sehingga prosentase ini dapat digunakan untuk mengambil keputusan oleh pemilik otoritas.

2. DASAR TEORI

2.1 Plagiarisme

Plagiarisme merupakan sebuah tindakan plagiat atau melakukan plagiat. Plagiat atau penjiplakan adalah pengambilan karangan, pendapat, dan sebagainya orang lain dan menjadikannya seolah karangan dan pendapat sendiri. Misalnya menerbitkan karya tulis orang lain yang belum sempat diterbitkan oleh penulisnya atas nama diri sendiri[1]. Metode untuk mendeteksi plagiarisme dapat dibagi menjadi tiga bagian, yaitu metode perbandingan teks lengkap, metode dokumen fingerprinting, dan metode kesamaan kata kunci[2]. Metode yang digunakan dalam perancangan ini adalah metode dokumen *fingerprinting*. Metode ini menggunakan teknik hashing, berikut adalah beberapa algoritmanya yaitu *rabin-karp*, *winnowing*, dan *manber*. Pada perancangan ini digunakan algoritma winnowing.

2.2 Algoritma Winnowing

Algoritma winnowing merupakan salah satu metode dalam dokumen fingerprinting. Input dari algoritma ini adalah dokumen teks yang diproses sehingga mendapatkan output nilai hash. Algoritma yang digunakan untuk mendapatkan nilai hash pada winnowing adalah *rolling hashing*. Berikut ini tahapan proses dalam algoritma winnowing :

1. Text Pre-processing

Dalam proses ini dokumen teks diolah agar menjadi data siap proses. Pengolahan ini meliputi spasi, jenis huruf (kapital atau normal), tanda baca dan sebagainya (*whitespace insensitivity*).

Contoh teks:

Indonesia tanah air Beta.

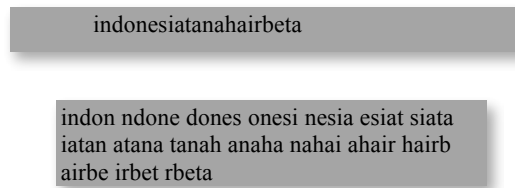
Indonesia tanah air Beta.

indonesiatanahairbeta



2. Membentuk rangkaian k-grams dari teks

Percobaan menggunakan nilai k-grams = 5, berikut ini contohnya :

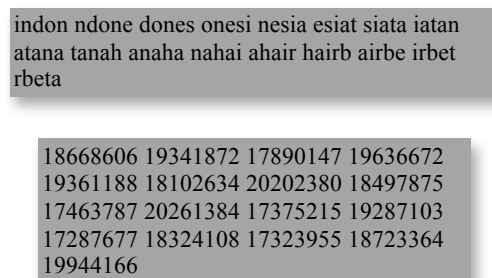


3. Melakukan rolling hashing untuk setiap k-grams

Persamaan (1) adalah perhitungan fungsi *hash* dari algoritma *winnowing* dengan *c* sebagai nilai ASCII, *b* sebagai nilai *basis* bilangan prima, dan banyak karakter *k* [3] :

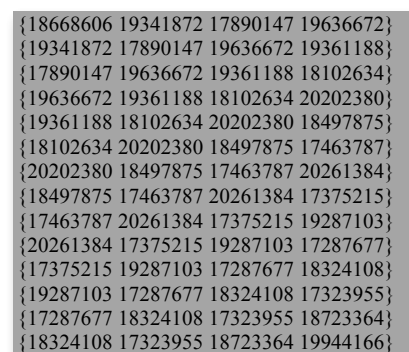
$$H_{(c_1 \dots c_k)} = c_1 * b^{(k-1)} + c_2 * b^{(k-2)} + \dots + c_{n-1} * b + c_k \quad (1)$$

Berikut ini contoh pernggunaan *rolling hashing* pada dokumen teks yang sudah melewati proses nomor 2. Dengan menggunakan nilai $b = 11$.



4. Memilih nilai *fingerprint* dari hasil *hashing* dengan pembagian hasil *hash* berdasarkan satu nilai *window* merupakan hasil dari perhitungan $w = t-k+1$ [3].

Misalkan nilai $t = 8$, sehingga $w = 8-5+1 = 4$. Dalam satu ukuran *window* dapat terisi dengan 4 nilai *hash*.



Setelah *window* terbentuk sebanyak 14 buah, maka dicarilah nilai *fingerprint* dengan melihat nilai terkecil dari setiap *window*. Maka didapati nilai *fingerprint* : 17890147 18102634 17463787 17375215 17287677 17323955. Dari 14 *window* hanya terdapat 6 *fingerprint*, hal ini disebabkan karena ada beberapa *window* yang memiliki nilai terendah yang sama sehingga tidak perlu dituliskan kembali.

5. Menghitung tingkat kemiripan

Untuk menghitung tingkat kemiripan diperlukan minimal dua buah dokumen teks untuk dibandingkan. Carany dengan mengikuti dari proses ke-1 hingga mendapatkan *fingerprint* dari setiap dokumen teks, lalu akan digunakan *jaccard coefficient* untuk menghitung kesamaannya.

$$\text{Similarity}(d_i, d_j) = \frac{|w(d_i) \cap w(d_j)|}{|w(d_i) \cup w(d_j)|} \times 100\% \quad (2)$$

Misalkan file teks D_1 dan D_2 dengan *fingerprint* $D_1 = 17890147, 18102634, 17463787, 17375215, 17287677, 17323955$ dan *fingerprint* $D_2 = 17890147, 18102524, 17318708$ maka sesuai dengan persamaan (2) diperoleh nilai kesamaan sebesar 12.5%

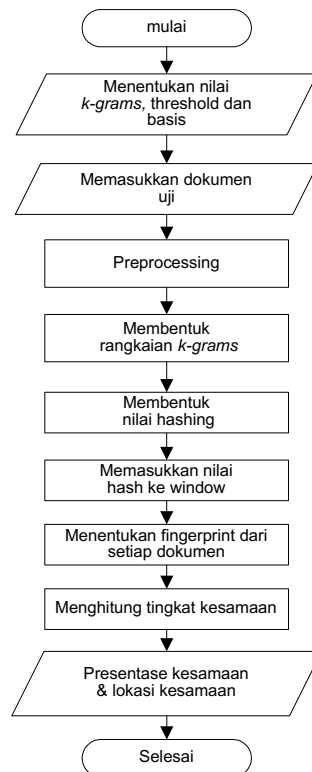
$$\text{Similarity}(d_i, d_j) = \frac{|1|}{|8|} \times 100\% = 12.5\%$$

2.3 Penentuan Nilai Parameter

Penentuan nilai parameter untuk pengujian sudah dilakukan berdasarkan pengujian yang sudah pernah dilakukan oleh Milani Winangga, Drs. Marji, M.T., dan Drs. Achmad Ridok, M.Kom. Kinerja terbaik diperoleh pada konfigurasi parameter nilai kgram=5, threshold=25, dan basis=11 dengan hasil tingkat prosentase error yang berbanding lurus dengan prosentase perubahan dokumen [4].

2.4 Sistem yang Dibuat

Sistem aplikasi pendeteksi tingkat kesamaan antar dokumen dengan algoritma winnowing ini memiliki alur yang dapat dilihat pada Gambar 1. Pada proses pendeteksian diperlukan untuk menentukan nilai *k-grams*, *guarantee threshold (threshold)*, dan *basis*. Setelah itu user harus memasukkan dokumen uji dengan format ekstensi .docx serta memilih dokumen pembanding yang sudah ada datanya dalam aplikasi. Selanjutnya proses akan menjalankan proses dengan algoritma *winnowing*. Jika aplikasi sudah menjalankan urutan proses dalam algoritma *winnowing* maka akan didapati hasil prosentase tingkat kesamaan serta menampilkan kata yang sama. Prosentase ini dapat dilihat dan digunakan sebagai data untuk menentukan apakah dokumen uji tersebut plagiat atau tidak. Penentuan ini ditentukan oleh pihak yang memiliki otoritas dalam menentukan dan aplikasi hanya menampilkan prosentase tingkat kesamaan.



Gambar 1 Alur kerja sistem yang di buat

3. HASIL DAN PEMBAHASAN

Pengujian dilakukan terhadap aplikasi pendeteksi tingkat kesamaan antar dokumen dengan algoritma *winnowing* ini. Pengujian dibagi menjadi 2 yaitu pengujian terhadap modul dan pengujian terhadap data. Pengujian terhadap modul ini dilakukan untuk memastikan bahwa setiap modul dapat berjalan dan berfungsi dengan baik. Sedangkan untuk pengujian terhadap data dilakukan untuk memastikan bahwa penerapan algoritma *winnowing* pada aplikasi ini sudah berjalan sesuai dengan yang diharapkan.

Hasil pengujian terhadap modul menghasilkan bahwa setiap modul dapat berjalan dengan baik. Berikut ini daftar modul yang ada dalam aplikasi :

- Modul Login
- Modul Home
- Modul Mahasiswa
- Modul Dosen
- Modul Archieve
- Modul Deteksi
- Modul Hasil
- Modul Unggah Skripsi
- Modul Log Out

Pengujian terhadap data dilakukan dengan 2 tahap, yaitu tahap pertama dengan data dummy dan tahap kedua dengan data skripsi asli. Pengujian dengan data dummy ini untuk memastikan bahwa algoritma dalam aplikasi berjalan sesuai yang diharapkan dan untuk mencoba kinerja aplikasi. Sedangkan untuk pengujian terhadap data skripsi asli untuk memastikan bahwa aplikasi dapat berjalan sesuai dengan yang diharapkan.

Untuk pengujian menggunakan data dummy telah dilakukan dengan sebuah dokumen teks yang berisi sebuah paragraf yang diatur komposisinya sebagai berikut :

- Data dummy 100% sama persis
- Data dummy dipotong 20% kata, sehingga tersisa 80% kata
- Data dummy dipotong 40% kata, sehingga tersisa 60% kata
- Data dummy dipotong 60% kata, sehingga tersisa 40% kata
- Data dummy dipotong 80% kata, sehingga tersisa 20% kata
- Data dummy diacak 20% kata, pengacakan dilakukan secara manual
- Data dummy diacak 60% kata, pengacakan dilakukan secara manual

Hasil dari pengujian data dummy dapat dilihat pada Gambar 2.

#	FILE 1	FILE 2 (ARCHIVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
38.	dummy 100.docx	dummy 100.docx	5	11	21	100
39.	dummy 80 (potong 20).docx	dummy 100.docx	5	11	21	76.92
40.	dummy 60 (potong 40).docx	dummy 100.docx	5	11	21	61.54
41.	dummy 40 (potong 60).docx	dummy 100.docx	5	11	21	38.46
42.	dummy 20 (potong 80).docx	dummy 100.docx	5	11	21	15.38
43.	dummy acak 20.docx	dummy 100.docx	5	11	21	53.13
44.	dummy acak 60.docx	dummy 100.docx	5	11	21	23.68

Gambar 2 Hasil pengujian data dummy

Pengujian terhadap data skripsi asli menggunakan laporan skripsi dari 18 orang mahasiswa. Laporan tersebut dipecah menjadi 3 bagian yaitu BAB I, BAB III, dan BAB IV. Pengujian dilakukan dengan mengambil data BAB I seorang mahasiswa dan dibandingkan

dengan 18 data lainnya. Berikut ini hasil pengujian BAB I bisa dilihat di Gambar 3, pengujian pada BAB III dapat dilihat di Gambar 4, dan pengujian pada BAB IV dapat dilihat di Gambar 5

#	FILE 1	FILE 2 (ARCHIVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
8.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Ayu Windy Astuti).docx	5	11	21	22.56
9.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Elizabeth Erlsha).docx	5	11	21	14.81
10.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Farencio).docx	5	11	21	19.62
11.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Ferry Dharmawan).docx	5	11	21	13.41
12.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Fransisca Regina).docx	5	11	21	17.1
13.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Irawati Djajadi).docx	5	11	21	19.86
14.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Jacklin Shintia Thio).docx	5	11	21	21.32
15.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Josselyn Shintia Thio).docx	5	11	21	20.84
16.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Mariana).docx	5	11	21	21.8
17.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Mishelle Tirta Winarta).docx	5	11	21	18.53
18.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	18.07
19.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Renaldo Ali).docx	5	11	21	15.66
20.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Rionaldy Triasaputra).docx	5	11	21	13.44
21.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Rosalinda).docx	5	11	21	17.28
22.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Paryanka Victorius Abadi).docx	5	11	21	100
23.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	26.64
24.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Vina Tendeen).docx	5	11	21	18.12
25.	BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Yunita).docx	5	11	21	19.07

Gambar 3 Pengujian pada BAB I

#	FILE 1	FILE 2 (ARCHIVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
26.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Ayu Windy Astuti).docx	5	11	21	26.3
27.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Elizabeth Erlsha).docx	5	11	21	18.08
28.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Farencio).docx	5	11	21	19.6
29.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Ferry Dharmawan).docx	5	11	21	20.15
30.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Fransisca Regina).docx	5	11	21	20.55
31.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Irawati Djajadi).docx	5	11	21	18.54
32.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Mariana).docx	5	11	21	21.04
33.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Mishelle Tirta Winarta).docx	5	11	21	19.48
34.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	18.66
35.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Renaldo Ali).docx	5	11	21	15.63
36.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Rionaldy Triasaputra).docx	5	11	21	15.74
37.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Rosalinda).docx	5	11	21	18
38.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Paryanka Victorius Abadi).docx	5	11	21	100
39.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Jacklin Shintia Thio).docx	5	11	21	21.91
40.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Josselyn Shintia Thio).docx	5	11	21	20.52
41.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	29.11
42.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Vina Tendeen).docx	5	11	21	20.84
43.	BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Yunita).docx	5	11	21	20.09

Gambar 4 Pengujian pada BAB III

#	FILE 1	FILE 2 (ARCHIVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
44.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Ayu Windy Astuti).docx	5	11	21	19.52
45.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Elizabeth Erlsha).docx	5	11	21	10.83
46.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Farencio).docx	5	11	21	19.9
47.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Ferry Dharmawan).docx	5	11	21	12.97
48.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Fransisca Regina).docx	5	11	21	16.13
49.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Irawati Djajadi).docx	5	11	21	14.97
50.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Jacklin Shintia Thio).docx	5	11	21	22.23
51.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Josselyn Shintia Thio).docx	5	11	21	19.42
52.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Mariana).docx	5	11	21	23.08
53.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	14.91
54.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	12.89
55.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Renaldo Ali).docx	5	11	21	13.29
56.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Rionaldy Triasaputra).docx	5	11	21	10.06
57.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Rosalinda).docx	5	11	21	13.25
58.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Paryanka Victorius Abadi).docx	5	11	21	100
59.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	25.35
60.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Vina Tendeen).docx	5	11	21	15.73
61.	BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Yunita).docx	5	11	21	15.15

Gambar 5 Pengujian pada BAB IV

Pengujian selanjutnya adalah dengan membandingkan beberapa skripsi mahasiswa yang memiliki topik yang mirip, berikut ini daftarnya :

- a. Pengujian antara Skripsi Farenco dengan Paryanka Victorius Abadi dapat dilihat pada Gambar 6.

FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTAS
BAB 1 (Paryanka Victorius Abadi).docx	BAB 1 (Farenco).docx	5	11	21	19.62
BAB 3 (Paryanka Victorius Abadi).docx	BAB 3 (Farenco).docx	5	11	21	19.6
BAB 4 (Paryanka Victorius Abadi).docx	BAB 4 (Farenco).docx	5	11	21	19.9

Gambar 12 Pengujian Skripsi Farenco dengan Paryanka

- b. Pengujian antara Skripsi Fransisca Regina dengan Vina Tendean dapat dilihat pada Gambar 7.

FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
BAB 1 (Vina Tendean).docx	BAB 1 (Fransisca Regina).docx	5	11	21	27.97
BAB 3 (Vina Tendean).docx	BAB 3 (Fransisca Regina).docx	5	11	21	36.81
BAB 4 (Vina Tendean).docx	BAB 4 (Fransisca Regina).docx	5	11	21	22.37

Gambar 7 Pengujian Skripsi Fransisca Regina dengan Vina Tendean

- c. Pengujian antara Skripsi Ferry Dharmawan dengan Renaldo Ali dapat dilihat pada Gambar 8.

FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTASE
BAB 1 (Ferry Dharmawan).docx	BAB 1 (Renaldo Ali).docx	5	11	21	30.36
BAB 3 (Ferry Dharmawan).docx	BAB 3 (Renaldo Ali).docx	5	11	21	20.6
BAB 4 (Ferry Dharmawan).docx	BAB 4 (Renaldo Ali).docx	5	11	21	27.41

Gambar 8 Pengujian Skripsi Ferry Dharmawan dengan Renaldo Ali

- d. Pengujian antara Skripsi Ayu Windy, Jacklin Shintia Thio, Josselyn Shintia Thio, Mishelle Tirta, Nadia Yanitra, Sri Whisnu, dan Mariana dapat dilihat pada Gambar 9.

FILE 1	FILE 2 (ARCHIEVE)	NILAI K	NILAI B	NILAI W	PERSENTASI
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Mariana).docx	5	11	21	26.66
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Jacklin Shintia Thio).docx	5	11	21	25.49
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Josselyn Shintia Thio).docx	5	11	21	25.98
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Mishelle Tirta Winarta).docx	5	11	21	20.57
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	20.56
BAB 1 (Ayu Windy Astuti).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	21.82
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Josselyn Shintia Thio).docx	5	11	21	29.03
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Mariana).docx	5	11	21	27.13
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Mishelle Tirta Winarta).docx	5	11	21	27.21
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	24.31
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	42.57
BAB 3 (Ayu Windy Astuti).docx	BAB 3 (Jacklin Shintia Thio).docx	5	11	21	26.35
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Jacklin Shintia Thio).docx	5	11	21	22.64
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Josselyn Shintia Thio).docx	5	11	21	28.17
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Mariana).docx	5	11	21	22.13
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	15.07
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	15.4
BAB 4 (Ayu Windy Astuti).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	22.94

BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Josselyn Shintia Thio).docx	5	11	21	21.76
BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Mariana).docx	5	11	21	24.53
BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Mishelle Tirta Winarta).docx	5	11	21	18.22
BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	20.02
BAB 1 (Jacklin Shintia Thio).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	28
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Josselyn Shintia Thio).docx	5	11	21	26.64
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Mariana).docx	5	11	21	22.42
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Mishelle Tirta Winarta).docx	5	11	21	22.65
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	20.87
BAB 3 (Jacklin Shintia Thio).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	28.59
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Josselyn Shintia Thio).docx	5	11	21	26.26
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Mariana).docx	5	11	21	24.83
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	15
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	14.3
BAB 4 (Jacklin Shintia Thio).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	26.55
BAB 1 (Josselyn Shintia Thio).docx	BAB 1 (Mariana).docx	5	11	21	23.23
BAB 1 (Josselyn Shintia Thio).docx	BAB 1 (Mishelle Tirta Winarta).docx	5	11	21	22.08
FILE 1	FILE 2 (ARCHIVE)	NILAI K	NILAI B	NILAI W	PERSE
BAB 1 (Josselyn Shintia Thio).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	20.9
BAB 1 (Josselyn Shintia Thio).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	20.43
BAB 3 (Josselyn Shintia Thio).docx	BAB 3 (Mariana).docx	5	11	21	25.58
BAB 3 (Josselyn Shintia Thio).docx	BAB 3 (Mishelle Tirta Winarta).docx	5	11	21	28.64
BAB 3 (Josselyn Shintia Thio).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	24.16
BAB 3 (Josselyn Shintia Thio).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	29.21
BAB 4 (Josselyn Shintia Thio).docx	BAB 4 (Mariana).docx	5	11	21	22.65
BAB 4 (Josselyn Shintia Thio).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	14.35
BAB 4 (Josselyn Shintia Thio).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	16.08
BAB 4 (Josselyn Shintia Thio).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	22.57
BAB 1 (Mariana).docx	BAB 1 (Mishelle Tirta Winarta).docx	5	11	21	23.61
BAB 1 (Mariana).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	18.19
BAB 1 (Mariana).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	25.18
BAB 3 (Mariana).docx	BAB 3 (Mishelle Tirta Winarta).docx	5	11	21	21.22
BAB 3 (Mariana).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	19.99
BAB 3 (Mariana).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	25.25
BAB 4 (Mariana).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	16.67
BAB 3 (Mariana).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	19.99
BAB 3 (Mariana).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	25.25
BAB 4 (Mariana).docx	BAB 4 (Mishelle Tirta Winarta).docx	5	11	21	16.67
BAB 4 (Mariana).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	19.21
BAB 4 (Mariana).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	25
BAB 1 (Mishelle Tirta Winarta).docx	BAB 1 (Nadia Yanitra).docx	5	11	21	28.04
BAB 1 (Mishelle Tirta Winarta).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	19.63
BAB 3 (Mishelle Tirta Winarta).docx	BAB 3 (Nadia Yanitra).docx	5	11	21	43.36
BAB 3 (Mishelle Tirta Winarta).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	24.07
BAB 4 (Mishelle Tirta Winarta).docx	BAB 4 (Nadia Yanitra).docx	5	11	21	25.44
BAB 4 (Mishelle Tirta Winarta).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	14.65
BAB 1 (Nadia Yanitra).docx	BAB 1 (Sri Whisnu A.W).docx	5	11	21	19.97
BAB 3 (Nadia Yanitra).docx	BAB 3 (Sri Whisnu A.W).docx	5	11	21	23.19
BAB 4 (Nadia Yanitra).docx	BAB 4 (Sri Whisnu A.W).docx	5	11	21	15.33

Gambar 9 Pengujian antara skripsi Ayu Windy, Jacklin Shintia Thio, Josselyn Shintia Thio, Mishelle Tirta, Nadia Yanitra, Sri Whisnu, dan Mariana

4. KESIMPULAN

Kesimpulan yang dapat diperoleh berdasarkan pembuatan dan pengujian dari aplikasi pendeteksi tingkat kesamaan antar dokumen dengan algoritma winnowing ini adalah sebagai berikut :

1. Aplikasi pendeteksi tingkat kesamaan antar dokumen ini dapat dijalankan dan mengoperasikan algoritma winnowing dengan baik serta dapat memberikan hasil yang dapat di tampilkan prosesnya.
2. Berdasarkan hasil pengujian diketahui bahwa susunan kata antar dokumen sangat menentukan nilai prosentase kesamaan. Jika urutan kata antar dokumen yang sama dirubah, maka hasilnya persentase kesamaannya akan turun.
3. Aplikasi dapat melakukan perbandingan dokumen uji dengan beberapa dokumen pembanding dan dapat ditampilkan secara bersamaan.

Saran untuk yang ingin mengembangkan aplikasi pendeteksi tingkat kesamaan antar dokumen dengan algoritma winnowing ini adalah sebagai berikut :

1. Dapat mengirimkan data hasil deteksi melalui email ketika proses deteksi sudah selesai
2. Menggunakan tambahan metode dokumen *fingerprinting* lainnya agar dapat digunakan sebagai tambahan informasi ukuran kesamaan antar dokumen.

UCAPAN TERIMA KASIH

Terima kasih saya ucapkan kepada dosen-dosen Fakultas Teknologi Informasi Universitas Tarumanagara yang sudah mau membimbing penulis dalam perancangan aplikasi ini. Terima kasih juga kepada seluruh pendahulu saya yang sudah digunakan data skripsinya untuk data pembanding dalam proses perancangan aplikasi ini.

DAFTAR PUSTAKA

- [1] Depdikbud, 1997, *Kamus Besar Bahasa Indonesia*. Jakarta.
- [2] Stein, Benno & Eissen, Sven Meyer Zu. 2006. Fingerprint-based Similarity Search and its Applications. Bauhaus-Universitat Weimar.
- [3] Schleimer, Saul. Wilkerson, Daniel S., dan Aiken, Alex. 2003. *Winnowing : Local Algorithms for document Fingerprinting*. San Diego.
- [4] Milani W, Marji, Achmad Ridok. 2014. *Deteksi Plagiarisme Pada Dokumen Teks Bahasa Indonesia Menggunakan Algoritma Winnowing Dengan Stemming*. Malang : Universitas Brawijaya.

SISTEM PENDETEKSI PERUBAHAN LAHAN HIJAU DI JABODETABEK

Andree Phanderson¹, Dyah Erny Herwindiati², Bagus Mulyawan³

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara,
Jln. Letjen S. Parman No. 1, Jakarta, 11440, Indonesia

E-mail:¹andreep.ti@stu.untar.ac.id, ²dyahh@fti.untar.ac.id, ³bagus@fti.untar.ac.id

Abstrak

Perubahan lahan hijau di Jabodetabek telah menjadi hal yang sangat penting untuk diperhatikan oleh pemerintah Indonesia. Klasifikasi lahan hijau ini menggunakan citra Band 1, 2, 3, 4 dan 5 dari Landsat 8. Sebelum dilakukan klasifikasi, citra yang didapat dari satelit Landsat 8 akan dikoreksi terlebih dahulu menggunakan algoritma koreksi radiometri yaitu algoritma mini-max. Setelah dilakukan proses koreksi radiometri, klasifikasi lahan hijau akan dilakukan dengan metode NDVI (Normalized Difference Vegetation Index) dan SAVI (Soil Adjusted Vegetation Index). Objek atau area tersebut akan diklasifikasikan dengan lahan hijau apabila hasil dari nilai NDVI lebih besar atau sama dengan 0.3. Setelah itu hasil klasifikasi dibagi menjadi 2 (dua) kategori, yaitu Y_1 dan Y_2 dengan metode dummy dependent variable. Dummy Dependent Variable ini akan digunakan dalam metode regresi linier untuk melakukan klasifikasi area yang telah dipilih di Jabodetabek. Untuk membuktikan apakah hasil dari sistem dapat dipercaya atau tidak, hasil klasifikasi akan dibandingkan dengan gambar asli yang terlihat di Google Earth. Penurunan jumlah lahan hijau terbesar, terdapat pada wilayah Bogor, bulan Mei 2015, yaitu sebesar 325.7368 Km².

Kata kunci—Dummy Dependent Variable, Koreksi Radiometri, NDVI, Regresi Linier, SAVI.

Abstract

Change of green area in Jakarta, Bogor, Depok, Tangerang and Bekasi (Jabodetabek) has been something very important. Classification of green area aims to do classification in Jabodetabek using Landsat 8 satellite images, band 1, 2, 3, 4 and 5. Before classification was done, the satellite images will be corrected using Radiometric Correction method called Mini-max algorithm. After doing radiometric correction, the classification will use the NDVI (Normalized Difference Vegetation Index) and SAVI (Soil Adjusted Vegetation Index) method. The selected area will be classified as green when NDVI values similar or has more than 0.3. After perform two categories, Y_1 and Y_2 are selected by NDVI values using dummy dependent variable. Linear regression method use that dummy dependent variable to classify the selected area in Jabodetabek. To see how can we trust the result, the classified area will be compared with the appearance of selected area in Google Earth. The highest degradation of green area is in Bogor, May 2015, 325.7368Km².

Keywords—Dummy Dependent Variable, NDVI, Linear Regression, Radiometric Correction, SAVI.

1. PENDAHULUAN

Penggunaan lahan hijau di Jabodetabek telah menjadi hal yang sangat penting untuk diperhatikan oleh pemerintah. Seiring berkembangnya jumlah penduduk yang bermukim atau bermigrasi dari kota lain ke Jabodetabek, maka ketersediaan lahan hijau pun semakin berkurang. Dengan berkurangnya ketersediaan lahan hijau, maka banjir akan semakin mudah terjadi karena daya serap air semakin berkurang.

Tujuan penelitian ini adalah untuk mengetahui jumlah perubahan lahan hijau yang terjadi di Jabodetabek dari tahun ke tahun secara periodic yaitu 6 (enam) bulan sekali. Pengklasifikasian lahan hijau ini menggunakan algoritma *mini-max* untuk melakukan koreksi radiometri terhadap citra, metode NDVI (*Normalized Difference Vegetation Index*), SAVI (*Soil Adjusted Vegetation Index*) dan Regresi Linier untuk mengklasifikasikan lahan hijau.

2. METODE PENELITIAN

2.1 Minimum-Maximum Normalization

Metode ini mengasumsikan gambar subjek harus memiliki piksel minimum dan piksel maksimum yang sama seperti gambar referensinya pada setiap *band*-nya. Koefisien normalisasi untuk metode minimum dan maksimum ini adalah [1]:

$$a_k = \frac{y_{kmax} - y_{kmin}}{x_{kmax} - x_{kmin}} \quad (1)$$

$$b_k = y_{kmin} - a_k x_{kmin} \quad (2)$$

$$y_k'[i,j] = a_k x_k \text{piksel} + b_k \quad (3)$$

Keterangan:

k = indeks ke pada citra

a_k, b_k = Konstanta normalisasi untuk *band* ke k

Y_{kmax} = Nilai maksimum piksel citra referensi

Y_{kmin} = Nilai minimum piksel citra referensi

X_{kmax} = Nilai maksimum piksel citra yang dikoreksi

X_{kmin} = Nilai minimum piksel citra yang dikoreksi

$y_k'[i,j]$ = Nilai piksel pada baris ke i dan kolom ke j citra yang dikoreksi

2.2 Normalized Difference Vegetation Index(NDVI)

Normalized Difference Vegetation Index (NDVI) adalah metode untuk menghitung kuantitas vegetasi dengan membandingkan perbedaan antara *near-infrared* (refleksi dari vegetasi) dan *red light* (yang diserap oleh vegetasi). NDVI selalu berada pada jarak antara -1 sampai 1.[2]

Formula untuk menghitung NDVI adalah sebagai berikut:

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \quad (4)$$

Keterangan:

NDVI = Nilai vegetasi dari suatu area.

NIR = Citra *band* 5 pada satelit Landsat 8.

Red = Citra *band* 4 pada satelit Landsat 8.

Pada persamaan NDVI lebih besar atau sama dengan 0.3 maka nilai objek atau area tersebut diklasifikasikan sebagai objek atau area hijau. Sebaliknya, jika persamaan NDVI lebih kecil dari 0.3, maka objek atau area tersebut diklasifikasikan sebagai objek atau area yang bukan hijau.

2.3 Soil Adjusted Vegetation Index(SAVI)

Di daerah dimana tutupan vegetatif rendah, yaitu < 40 persen dan permukaan tanah terpapar, pantulan cahaya dari spektrum merah dan *near-infrared* dapat mempengaruhi nilai indeks vegetasi. Permasalahan ini lah yang menjadi hal utama ketika membandingkan nilai vegetasi pada permukaan tanah yang berbeda. Indeks vegetasi ini dikembangkan sebagai modifikasi dari *Normalized Difference Vegetation Index* untuk mengoreksi kecerahan tanah saat tutupan vegetatif rendah.[3]

Formula untuk menghitung SAVI adalah:

$$SAVI = \frac{(NIR-Red)}{(NIR+Red+L)} * (1 + L) \quad (5)$$

Keterangan:

SAVI = Nilai vegetasi dari suatu area.

NIR = Citra *band* 5 pada satelit Landsat 8.

Red = Citra *band* 4 pada satelit Landsat 8.

L = Faktor kalibrasi tanah adalah 0.5 [4]

2.4 Regresi Linier

Regresi linier sederhana adalah metode statistik yang berfungsi untuk menguji sejauh mana hubungan sebab akibat antara variabel faktor penyebab terhadap variabel faktor akibatnya. Faktor penyebab pada umumnya dilambangkan dengan X sedangkan faktor akibat umumnya dilambangkan dengan Y [5].

Formula untuk menghitung regresi linier sederhana satu variabel adalah sebagai berikut:

$$Y_1 = b_0 + b_1X \quad (6)$$

$$Y_2 = b_0 + b_1X \quad (7)$$

Dengan

$$X = X_0, X_1 \quad (8)$$

$$B_a = (X^T X)^{-1} * (X^T * Y_1) \quad (9)$$

$$B_b = (X^T X)^{-1} * (X^T * Y_2) \quad (10)$$

Keterangan:

b_0 = Nilai titik potong garis regresi

b_1 = Nilai koefisien regresi berkaitan dengan faktor penyebab

X = Matriks yang berisi X_0 dan X_1

Y_1, Y_2 = Vektor nilai dengan *dummy dependent variable*.

X^T = *Transpose* matriks X

$(X)^{-1}$ = *Invers* dari matriks X

Formula untuk menghitung regresi linier sederhana dua variabel adalah sebagai berikut:

$$Y_1 = b_0 + b_1X_1 + b_2X_2 \quad (11)$$

$$Y_2 = b_0 + b_1X_1 + b_2X_2 \quad (12)$$

Dengan

$$X = X_0, X_1, X_2 \quad (13)$$

$$B_a = (X^T X)^{-1} * (X^T * Y_1) \quad (14)$$

$$B_b = (X^T X)^{-1} * (X^T * Y_2) \quad (15)$$

Keterangan:

b_0 = Nilai titik potong garis regresi

b_1 = Nilai koefisien regresi berkaitan dengan faktor penyebab pertama.

b_2 = Nilai koefisien regresi berkaitan dengan faktor penyebab kedua.

X = Matriks yang berisi X_0 dan X_1 dan X_2 .

Y_1, Y_2 = Vektor nilai dengan *dummy dependent variable*.

X^T = Transpose matriks X .

$(X)^{-1}$ = Invers dari matriks X .

3. HASIL DAN PEMBAHASAN

Hasil percobaan klasifikasi lahan hijau yang dilakukan dengan metode regresi linier pada beberapa sampel di kota Jakarta, Bogor, Depok, Tangerang atau Bekasi dapat dilihat pada gambar dibawah:

Tabel 1 Model Regresi Satu Variabel

	B0	B1
Y1	-0.415872	3.0618505
Y2	1.415872	-3.0618505

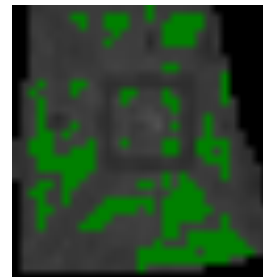
Tabel 2 Model Regresi Dua Variabel

	B0	B1	B2
Y1	-0.328472	-712.793704	479.396811
Y2	1.3284727	712.793700	-479.396808

Model regresi tersebut digunakan untuk mengklasifikasikan lahan hijau karena telah diuji dan mendapatkan keakuratan diatas 90%. Untuk model regresi satu variabel, rata-rata keakuratan yang dihasilkan mencapai 96% dan untuk model regresi dua variabel, rata-rata keakuratan yang dihasilkan mencapai 90%.



(a)



(b)

Gambar 1(a) Deteksi Lahan Hijau pada Monas dengan Model Regresi Satu Variabel (b)Deteksi Lahan Hijau pada Monas dengan Model Regresi Dua Variabel



(a)



(b)

Gambar 2(a) Deteksi Lahan Hijau pada Kebun Raya Bogor dengan Model Regresi Satu Variabel
(b) Deteksi Lahan Hijau pada Kebun Raya Bogor dengan Model Regresi Dua Variabel



(a)



(b)

Gambar 3(a) Deteksi Lahan Hijau pada Situ Rawabadak dengan Model Regresi Satu Variabel (b)
Deteksi Lahan Hijau pada Situ Rawabadak dengan Model Regresi Dua Variabel

Keterangan untuk Gambar 1 sampai Gambar 3:

Warna Hijau: Lahan hijau

Selain warna hijau: Bukan lahan hijau



(a)



(b)

Gambar 4(a) Deteksi Lahan Tidak Serap Air pada Jembatan Merah dengan Model Regresi Satu Variabel
(b) Deteksi Lahan Tidak Serap Air pada Jembatan Merah dengan Model Regresi Dua Variabel



(a)



(b)

Gambar 5(a) Deteksi Lahan Tidak Serap Air pada Margonda dengan Model Regresi Satu Variabel
(b) Deteksi Lahan Tidak Serap Air pada Margonda dengan Model Regresi Dua Variabel



Gambar 6(a) Deteksi Lahan Tidak Serap Air pada Sukaresmi dengan Model Regresi Satu Variabel
(b) Deteksi Lahan Tidak Serap Air pada Sukaresmi dengan Model Regresi Dua Variabel

Keterangan untuk Gambar 4 sampai Gambar 6:

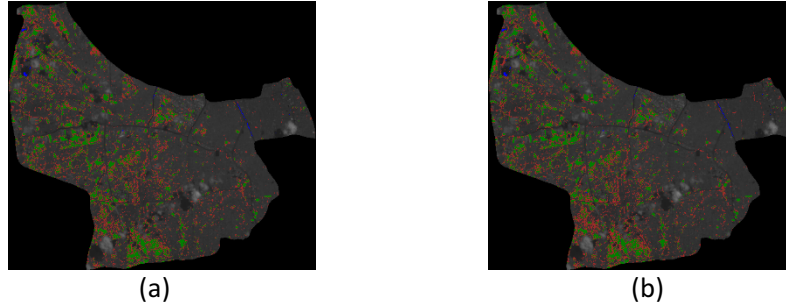
Warna merah: Lahan tidak Serap Air

Selain warna merah: Lahan serap air

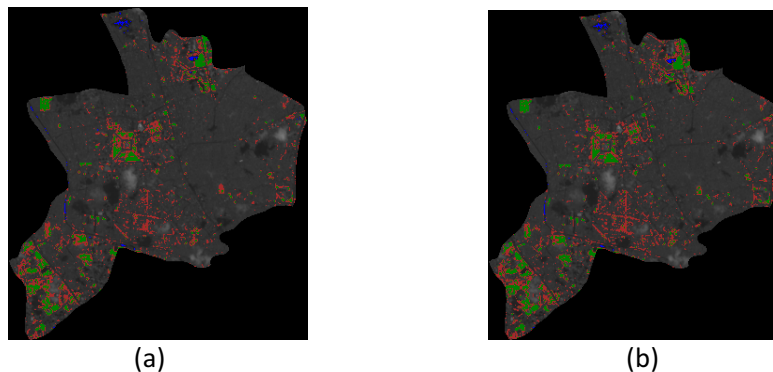
Tabel 3 Hasil Evaluasi Deteksi Lahan Hijau dan Lahan Tidak Serap Air

Jenis Lahan	Sampel	Regresi	
		Satu Variabel	Dua Variabel
Hijau	Monas	96.9483 %	94.7183 %
Hijau	Kebun Raya Bogor	97.159 %	89.0422 %
Hijau	Situ Rawabadak	100 %	88.8888 %
Tidak Serap Air	Jembatan Merah	83.1369 %	98.9436 %
Tidak Serap Air	Margo City	68.5393 %	64.0449 %
Tidak Serap Air	Sukaresmi, Cikarang Selatan	72.1017 %	61.4043 %

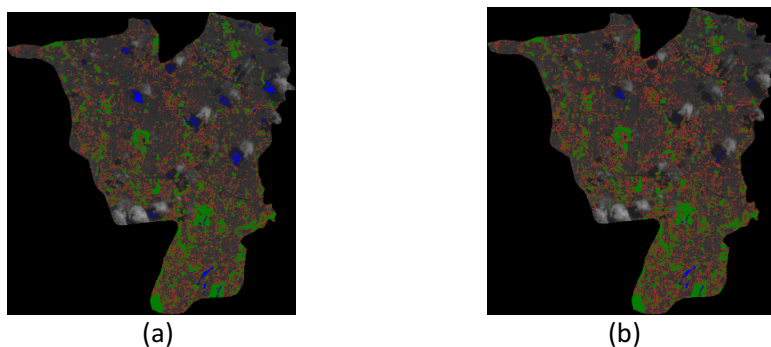
Selain itu, percobaan juga dilakukan pada wilayah Jakarta dan Bogor. Hasil percobaan untuk semester ganjil tahun 2014 dapat dilihat pada gambar di bawah:



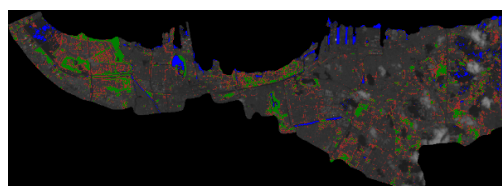
Gambar 7(a) Deteksi Lahan Jakarta Barat Tahun 2014 dengan Model Regresi Satu Variabel (b) Deteksi Lahan Jakarta Barat Tahun 2014 dengan Model Regresi Dua Variabel



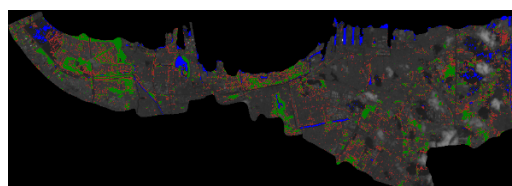
Gambar 8 (a) Deteksi Lahan Jakarta Pusat Tahun 2014 dengan Model Regresi Satu Variabel (b) Deteksi Lahan Jakarta Pusat Tahun 2014 dengan Model Regresi Dua Variabel



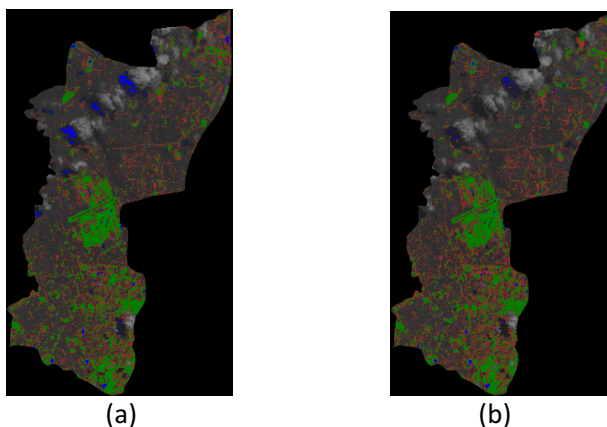
Gambar 9 Deteksi Lahan Jakarta Selatan Tahun 2014 dengan Model Regresi Satu Variabel (b) Deteksi Lahan Jakarta Selatan Tahun 2014 dengan Model Regresi Dua Variabel



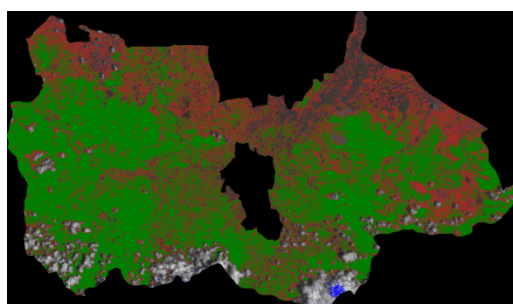
Gambar 10 Deteksi Lahan Jakarta Utara Tahun 2014 dengan Model Regresi Satu Variabel



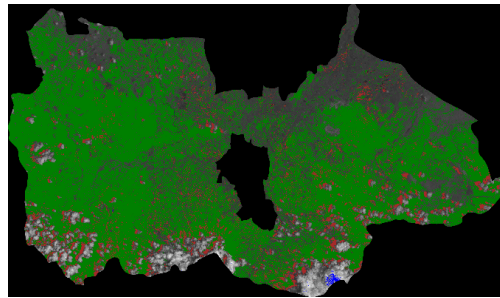
Gambar 11 Deteksi Lahan Jakarta Utara Tahun 2014 dengan Model Regresi Dua Variabel



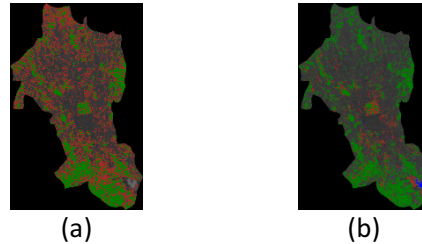
Gambar 12 (a) Deteksi Lahan Jakarta Timur Tahun 2014 dengan Model Regresi Satu Variabel (b) Deteksi Lahan Jakarta Timur Tahun 2014 dengan Model Regresi Dua Variabel



Gambar 13 Deteksi Lahan Kabupaten Bogor Tahun 2014 dengan Model Regresi Satu Variabel



Gambar 14 Deteksi Lahan Kabupaten Bogor Tahun 2014 dengan Model Regresi Dua Variabel



Gambar 15 (a) Deteksi Lahan Kota Bogor Tahun 2014 dengan Model Regresi Satu Variabel (b) Deteksi Lahan Kota Bogor Tahun 2014 dengan Model Regresi Dua Variabel

Keterangan untuk Gambar 7 sampai Gambar 15:

Warna Biru: Lahan air

Warna Coklat: Lahan kosong

Warna Hijau: Lahan hijau

Selain warna biru, warna coklat dan warna hijau: Lahan tidak serap air

Setelah berhasil melakukan percobaan pada setiap wilayah di Jakarta dan Bogor, maka selanjutnya dilakukan evaluasi lahan hijau di Jabodetabek secara periodik. Hasil evaluasi lahan hijau secara periodik untuk tahun 2014 hingga 2016 dapat dilihat pada Tabel 4 dan Tabel 5.

Tabel 4 Hasil Evaluasi Deteksi Lahan Hijau di Jabodetabek dengan Model Regresi Satu Variabel

Tahun	Kota	Luas Lahan Hijau (NDVI)	Luas Lahan Hijau (Regresi)	Ketepatan (%)
22 April 2014	Jakarta	58.2552 Km ²	58.2714 Km ²	99.9975 %
31 Oktober 2014	Jakarta	15.5502 Km ²	15.6186 Km ²	99.9896 %
27 Mei 2015	Jakarta	23.9724 Km ²	23.9733 Km ²	99.9998 %
31 Agustus 2015	Jakarta	38.3499 Km ²	41.8941 Km ²	99.9998 %
13 Mei 2016	Jakarta	57.6999 Km ²	57.7044 Km ²	99.9993 %
17 Agustus 2016	Jakarta	53.0099 Km ²	63.0225 Km ²	99.9980 %
22 April 2014	Bogor	2121.6402 Km ²	2122.0263 Km ²	99.9874 %
31 Oktober 2014	Bogor	1495.5390 Km ²	1496.8575 Km ²	99.9571 %
27 Mei 2015	Bogor	1168.1523 Km ²	1169.7822 Km ²	99.9470 %
31 Agustus 2015	Bogor	1773.1476 Km ²	1773.279 Km ²	99.9957 %
13 Mei 2016	Bogor	2088.3996 Km ²	2088.6291 Km ²	99.9925 %
17 Agustus 2016	Bogor	2140.7742 Km ²	2141.0973 Km ²	99.9895 %
22 April 2014	Depok	69.4134 Km ²	69.4143 Km ²	99.9995 %
31 Oktober 2014	Depok	11.0457 Km ²	11.0493 Km ²	99.9981 %
27 Mei 2015	Depok	31.1472 Km ²	31.1544 Km ²	99.9963 %
31 Agustus 2015	Depok	38.5245 Km ²	38.5263 Km ²	99.9990 %
13 Mei 2016	Depok	54.7893 Km ²	54.7893 Km ²	100 %
17 Agustus 2016	Depok	61.9362 Km ²	61.9398 Km ²	99.9981 %
22 April 2014	Tangerang	377.4744 Km ²	377.676 Km ²	99.9851 %
31 Oktober 2014	Tangerang	137.9448 Km ²	137.9448 Km ²	100 %
27 Mei 2015	Tangerang	186.8751 Km ²	186.9606 Km ²	99.9930 %
31 Agustus 2015	Tangerang	112.0338 Km ²	112.095 Km ²	99.9954 %

13 Mei 2016	Tangerang	249.8463 Km ²	249.9408 Km ²	99.9930 %
17 Agustus 2016	Tangerang	512.2323 Km ²	512.2773 Km ²	99.9966 %
22 April 2014	Bekasi	326.0673 Km ²	326.0862 Km ²	99.9987 %
31 Oktober 2014	Bekasi	101.4489 Km ²	101.5902 Km ²	99.9905 %
27 Mei 2015	Bekasi	309.2859 Km ²	309.4776 Km ²	99.9871 %
31 Agustus 2015	Bekasi	126.7236 Km ²	126.7983 Km ²	99.9950 %
13 Mei 2016	Bekasi	610.0254 Km ²	610.0848 Km ²	99.9960 %
17 Agustus 2016	Bekasi	299.4381 Km ²	299.6181 Km ²	99.9879 %

Tabel 5 Hasil evaluasi deteksi lahan hijau di Jabodetabek dengan model regresi dua variabel

Tahun	Kota	Luas Lahan Hijau (NDVI)	Luas Lahan Hijau (Regresi)	Ketepatan (%)
22 April 2014	Jakarta	58.2552 Km ²	46.0764 Km ²	97.6640 %
31 Oktober 2014	Jakarta	15.5502 Km ²	20.6694 Km ²	99.1744 %
27 Mei 2015	Jakarta	23.9724 Km ²	13.5045 Km ²	98.3587 %
31 Agustus 2015	Jakarta	38.3499 Km ²	15.39 Km ²	96.5162 %
13 Mei 2016	Jakarta	57.6999 Km ²	40.2444 Km ²	97.2359 %
17 Agustus 2016	Jakarta	53.0099 Km ²	45.7695 Km ²	97.0948 %
22 April 2014	Bogor	2121.6402 Km ²	1846.287 Km ²	88.6621 %
31 Oktober 2014	Bogor	1495.5390 Km ²	1564.2 Km ²	93.9242 %
27 Mei 2015	Bogor	1168.1523 Km ²	1021.9464 Km ²	86.7516 %
31 Agustus 2015	Bogor	1773.1476 Km ²	1434.4614 Km ²	88.0942 %
13 Mei 2016	Bogor	2088.3996 Km ²	1677.9114 Km ²	85.3548 %
17 Agustus 2016	Bogor	2140.7742 Km ²	1756.8315 Km ²	86.2485 %
22 April 2014	Depok	69.4134 Km ²	54.5679 Km ²	92.1821 %
31 Oktober 2014	Depok	11.0457 Km ²	11.9637 Km ²	98.9567 %
27 Mei 2015	Depok	31.1472 Km ²	24.912 Km ²	95.4060 %
31 Agustus 2015	Depok	38.5245 Km ²	28.5156 Km ²	94.4658 %
13 Mei 2016	Depok	54.7893 Km ²	35.6355 Km ²	90.1932 %
17 Agustus 2016	Depok	61.9362 Km ²	44.9226 Km ²	91.1017 %
22 April 2014	Tangerang	377.4744 Km ²	354.897 Km ²	95.5959 %
31 Oktober 2014	Tangerang	137.9448 Km ²	129.6720 Km ²	96.4965 %
27 Mei 2015	Tangerang	186.8751 Km ²	177.3171 Km ²	96.9183 %
31 Agustus 2015	Tangerang	112.0338 Km ²	105.075 Km ²	97.9383 %
13 Mei 2016	Tangerang	249.8463 Km ²	217.1088 Km ²	96.1159 %
17 Agustus 2016	Tangerang	512.2323 Km ²	417.0627 Km ²	92.2250 %
22 April 2014	Bekasi	326.0673 Km ²	254.1087 Km ²	94.2107 %
31 Oktober 2014	Bekasi	101.4489 Km ²	116.9775 Km ²	97.7353 %
27 Mei 2015	Bekasi	309.2859 Km ²	280.9098 Km ²	95.5107 %
31 Agustus 2015	Bekasi	126.7236 Km ²	109.899 Km ²	97.7678 %
13 Mei 2016	Bekasi	610.0254 Km ²	530.5815 Km ²	93.6812 %
17 Agustus 2016	Bekasi	299.4381 Km ²	266.8455 Km ²	96.6015 %

Analisis berdasarkan hasil evaluasi deteksi lahan hijau diatas yaitu:

1. Hasil deteksi lahan dilakukan dengan menggunakan data citra Landsat 8. Deteksi lahan hijau dapat dikatakan berhasil karena rata-rata hasil keakuratan pengujian mencapai 90 persen.
2. Jumlah lahan hijau pada wilayah Jakarta, Bogor, Depok, Tangerang dan Bekasi berkurang secara signifikan pada semester genap tahun 2014 dan semester ganjil tahun 2015, dikarenakan terjadi kemarau panjang selama dua periode tersebut. [6]
3. Rata-rata pada proses deteksi menggunakan model regresi satu variabel mendapatkan hasil yang lebih tepat. Hal ini dikarenakan banyaknya variabel yang digunakan dapat mengurangi keakuratan dari suatu model regresi. [7]
4. Hasil yang didapat dari model regresi dua variabel kurang akurat jika dibandingkan dengan hasil yang didapat dari regresi satu variabel. Hal ini disebabkan karena adanya faktor kalibrasi tanah yang digunakan pada rumus *Soil Adjusted Vegetation Index* (SAVI) serta rumus yang memiliki kemiripan dengan rumus *Normalized Difference Vegetation Index* (NDVI) yaitu sama-sama menggunakan Band 5 dan Band 4 pada citra Landsat 8 untuk mendapatkan jenis klasifikasi lahan.

4. KESIMPULAN

Berdasarkan percobaan dan pengujian yang telah dilakukan dengan metode regresi linier, maka dapat disimpulkan bahwa:

1. Metode regresi linier dapat digunakan untuk mengklasifikasikan jenis lahan dari citra satelit Landsat 8.
2. Pada wilayah Jakarta tahun 2014, luas lahan hijau di Jakarta mencapai 58 Km², tahun 2015 mencapai 41 Km² dan tahun 2016 mencapai 57 Km².
3. Pada wilayah Bogor tahun 2014, luas lahan hijau di Bogor mencapai 2121 Km², tahun 2015 mencapai 1773 Km² dan tahun 2016 mencapai 2141 Km².
4. Pada wilayah Depok tahun 2014, luas lahan hijau di Depok mencapai 69 Km², tahun 2015 mencapai 38 Km² dan tahun 2016 mencapai 61 Km².
5. Pada wilayah Tangerang tahun 2015, luas lahan hijau di Tangerang mencapai 377 Km², tahun 2015 mencapai 186 Km² dan tahun 2016 mencapai 512 Km².
6. Pada wilayah Bekasi tahun 2014, luas lahan hijau di Bekasi mencapai 326 Km², tahun 2015 mencapai 309 Km² dan tahun 2016 mencapai 610 Km².
7. Hasil evaluasi pada wilayah Jakarta tahun 2014 hingga 2016 menunjukkan bahwa wilayah Jakarta memiliki sangat sedikit lahan hijau jika dibandingkan dengan wilayah lainnya yaitu Bogor, Depok, Tangerang dan Bekasi. Luas lahan hijau pada wilayah Jakarta tahun 2014 hingga 2016 hanya mencapai 58 Km², sedangkan Bogor mencapai 2141 Km², Depok mencapai 69 Km², Tangerang mencapai 512 Km², dan Bekasi mencapai 610 Km².
8. Hasil yang didapat dari model regresi dua variabel kurang akurat jika dibandingkan dengan hasil yang didapat dari regresi satu variabel. Hal ini disebabkan karena adanya faktor kalibrasi tanah yang digunakan pada rumus *Soil Adjusted Vegetation Index* (SAVI) serta rumus yang memiliki kemiripan dengan rumus *Normalized Difference Vegetation Index* (NDVI) yaitu sama-sama menggunakan Band 5 dan Band 4 pada citra Landsat 8 untuk mendapatkan jenis klasifikasi lahan.

DAFTAR PUSTAKA

- [1] Gore, Semma Biday. 2010. Radiometric Correction of Multitemporal Satellite Imagery, *Journal of Computer Science*, Vol. 9. h.1028, 22 Desember 2017.
- [2] GisGeoraphy. 2017, What is NDVI (Normalized Difference Vegetation Index, <http://gisgeography.com/ndvi-normalized-difference-vegetation-index/>, 22 Desember 2017.
- [3] Gtucker, 2011. Soil Adjusted Vegetation Index. http://wiki.landscapetoolbox.org/doku.php/remote_sensing_methods:soil-adjusted_vegetation_index, 22 Desember 2017.
- [4] Jensen, J.R. 2000, *Remote Sensing of The Environmental Earth Resource Perspective*, Prentice Hall, 22 Desember 2017.
- [5] Dickson Kho. 2015. Analisis Regresi Linear Sederhana (*Simpel Linear Regression*). <http://teknikelektronika.com/analisis-regresi-linear-sederhana-simple-linear-regression/>, 22 Desember 2017.
- [6] J, Pujiono S. 2015. Kemarau 2015 Jadi Musim Terpanas dalam 40 Abad, <https://www.google.com/amp/s/beritagar.id/artiket-amp/sains-teknokemarau-2015-jadi-musim-terpanas-dalam-40-abad/>, 22 Desember 2017.
- [7] Thorac, J. Dis. 2014. Too Much Covariates in Multivariable Model May Cause the Problem of Overfitting, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4178069/>, 22 Desember 2017.