

Pendeteksi Kesalahan Pengetikan Kata Non Baku pada Karya Tulis Menggunakan Metode N-Gram

Titi Hartina¹, Agustin²

¹STMIK Amik Riau/Teknik Informatika
e-mail: 1titihartina@gmail.com

²STMIK Amik Riau /Teknologi Informasi
e-mail: agustin@sar.ac.id

Abstrak

Karya tulis sudah menjadi salah satu sarana bagi para generasi muda untuk menuangkan idenya, ide tersebut biasa di susun membentuk sebuah karangan berupa tulisan dari hari pemikiran, pengamatan, dan tinjauan dalam berbagai bidang. Karya tulis yang dibuat haruslah memiliki kualitas yang baik sehingga pembaca akan mudah untuk menerima makna dan tujuan dari penulisan karya tulis tersebut. Akan tetapi tidak semua karya tulis yang dibuat memiliki kualitas seperti yang diharapkan. Salah satu faktor yang mempengaruhi rendahnya kualitas penulisan karya tulis yang baik adalah kesalahan dalam penggunaan kata non baku yang disengaja maupun yang tidak disengaja. Mengacu pada masalah tersebut, maka digunakan metode N-Gram untuk membantu proses penelitian yang akan dilakukan. Metode N-Gram ini digunakan untuk pembangkitan kata dan karakter yang tidak sesuai dengan penulisan kata baku dalam Bahasa Indonesia. Dengan cara ini, maka kata nonbaku yang menjadi inti permasalahan tersebut akan mudah terdeteksi. Agar dapat mengimplementasikan tujuan dari penelitian maka metode ini dituangkan kedalam sebuah aplikasi pendeteksi kesalahan pengetikan kata nonbaku pada karya tulis berbahasa indonesia menggunakan metode N-Gram. Aplikasi ini dapat meminimalisasi kekurangan dalam penulisan karya tulis.

Kata Kunci: deteksi, karya tulis, kata non baku

Abstract

The writing has become one of the means for the young generation to express their ideas, the idea is usually arranged to form an essay in the form of writing from the day of thought, observation, and review in various fields. The written work must be of good quality so that the reader will be easy to accept the meaning and purpose of writing the written work. However, not all written works have the quality as expected. One of the factors that influences the poor quality of good writing is the use of intentional or unintentional non standard words. Referring to these problems, the N-Gram method is used to assist the research process to be conducted. The N-Gram method is used to generate words and characters that are not in accordance with the standard word writing in Indonesian. In this way, the non-standard word at the heart of the problem will be easily detected. In order to implement the objectives of the research, this method is poured into a non-standard typing error detection application in Indonesian language writing using the N-Gram method. This application can minimize deficiencies in writing the paper

Keywords: detection, article, non standard word

Pendahuluan

Karya tulis ilmiah di perguruan tinggi merupakan bagian dari tuntutan formal akademik. Melalui karya ilmiah, mahasiswa mengungkapkan pemikirannya secara sistematis sesuai dengan kaidah-kaidah keilmuan. (Sarmadan, 2017). Karya tulis ilmiah memiliki tiga ciri yaitu harus memiliki makna tunggal, tepat mendefinisikan setiap

istilah agar tidak menimbulkan kerancuan atau keraguan, dan singkat berlandaskan ekonomi bahasa (Bahasa Indonesia, 2012)

Berdasarkan ciri dari karya tulis yang telah dijelaskan diatas, tidak menutup kemungkinan karya tulis yang dibuat tidak mengalami kesalahan, baik itu kesalahan yang disengaja maupun kesalahan yang tidak disengaja. Ada beberapa faktor yang

mempengaruhi kesalahan dalam penulisan karya tulis diantaranya adalah:

- (a) Ketidakteknelitian dalam pengetikan.
- (b) Kurangnya pengetahuan penulisan dalam penggunaan bahasa baku serta tanda baca.
- (c) Kebiasaan menggunakan bahasa sehari-hari yang tanpa sengaja terbawa kedalam proses pengetikan
- (d) Penulisan tidak sesuai dengan Ejaan Yang Disempurnakan (EYD).

Semua kendala diatas, secara langsung maupun tidak langsung berpengaruh terhadap penurunan tingkat kualitas karya tulis yang dibuat. Kemampuan setiap manusia dalam memeriksa kembali kesalahan yang dibuat berbeda-beda dan membutuhkan tingkat ketelitian yang sangat tinggi, maka ditawarkan sebuah aplikasi yang akan mendeteksi kesalahan pengetikan khususnya pada kata nonbaku yang masih banyak digunakan penulis dalam proses pembuatan karya tulis

Aplikasi berupa perangkat lunak yang berbentuk *software* yang berisi kesatuan perintah atau program yang dibuat untuk melaksanakan sebuah pekerjaan yang diinginkan. (Afandi, 2013).

Kosakata atau kumpulan kata adalah daftar kata-kata yang dimiliki oleh seseorang dan dipakai dalam bidang ilmu pengetahuan (Supriadin, 2016).

Kata baku adalah kata-kata yang lazim digunakan dalam situasi formal atau resmi yang penulisannya sesuai dengan kaidah-kaidah yang dibakukan. Baku tidaknya sebuah kata dapat dilihat dari segi lafal, ejaan, gramatika, dan kenasionalan-nya. Kaidah standar yang dimaksud dapat berupa pedoman ejaan (EYD), tata bahasa baku, dan kamus (Setiawati, 2016).

Pembuatan aplikasi pendeteksi kesalahan kata nonbaku ini menggunakan metode *N-Gram*, dimana metode tersebut memiliki fungsi untuk mengambil potongan-potongan karakter huruf yang dimaksud dalam sebuah kata yang secara kontinu akan dibaca dari kata sumbernya hingga akhir dan menemukan kata yang tidak sesuai dengan Ejaan Yang Disempurnakan (EYD). Keuntungan *N-gram* dalam *matching* string adalah berdasarkan karakteristik *N-gram* sebagai bagian dari suatu string, sehingga kesalahan pada

sebagian string hanya akan berakibat perbedaan pada sebagian *N-gram*. Sebagai contoh jika *N-gram* dari dua string dibandingkan, kemudian kita menghitung cacah *N-gram* yang sama dari dua string tersebut maka akan didapatkan nilai similaritas atau kemiripan dua string tersebut yang bersifat resistan terhadap kesalahan tekstual. (Hamzah, 2010)

Aplikasi ini berpotensi untuk digunakan oleh editor ataupun penulis agar dapat membantu mendeteksi kesalahan pengetikan yang disengaja maupun tidak disengaja pada karya tulis Berbahasa Indonesia untuk meningkatkan kualitas.

Metode Penelitian

Penelitian ini menggunakan *Metode N-gram* sebagai algoritma untuk menghitung nilai probabilitas dari kata-kata yang dianggap tidak sesuai atau salah. Adapun tahapannya sebagai berikut:

1. Tahap Pengumpulan Data

Data yang diperoleh berasal dari wawancara yang dilakukan dengan para Mahasiswa/I yang pernah membuat sebuah karya tulis baik itu karya tulis ilmiah maupun karya tulis non ilmiah. Berikut ini penjelasan tentang data yang digunakan:

a. Jenis Data

Jenis data yang digunakan dalam penelitian ini yaitu data sekunder. Data sekunder merupakan sumber data penelitian yang diperoleh secara tidak langsung atau melalui media perantara seperti literatur, buku, serta referensi lainnya yang berhubungan dengan permasalahan pada penelitian ini.

b. Teknik Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data secara lengkap kemudian dianalisa dan didefinisikan. Teknik yang digunakan untuk pengumpulan data dalam penelitian ini adalah dengan melakukan tinjauan pustaka seperti mereview jurnal, skripsi, dan membaca buku yang berkaitan dengan penelitian yang dilakukan oleh peneliti.

2. Tahap Analisa Sistem

Dalam tahapan ini, informasi yang diperoleh pada tahap pengumpulan data kemudian dianalisa untuk memenuhi kebutuhan sistem yang diusulkan, serta menganalisa dan merancang sistem yang akan dibangun sehingga menghasilkan

sistem yang bisa dimanfaatkan secara terus menerus.

3. Tahap Perancangan Sistem

Tahapan perancangan dibagi atas perancangan sistem global dan perancangan sistem terinci. Perancangan sistem global dilakukan dengan pemodelan *Unified Modelling Language* (UML), *Unified Modelling Language* (UML) adalah suatu alat untuk memvisualisasikan dan mendokumentasikan hasil analisa dan desain yang berisi sintak dalam memodelkan sistem secara visual. Juga merupakan suatu kumpulan konvensi pemodelan yang digunakan untuk menentukan atau menggambarkan sebuah sistem *software* yang terkait dengan objek (Haviluddin, 2011). Sedangkan perancangan sistem terinci terbagi atas perancangan *output*, perancangan *input*, dan perancangan *database*.

4. Tahap Implementasi Sistem

Pada tahap ini hasil perancangan diterjemahkan kedalam kode-kode menggunakan bahasa pemrograman yang sudah ditentukan, kemudian aplikasi diuji dan siap diimplementasikan kedalam sistem pengguna atau dengan kata lain sudah siap diterapkan. Pengkodean menggunakan bahasa pemrograman berbasis web serta *MySQL* sebagai *database*.

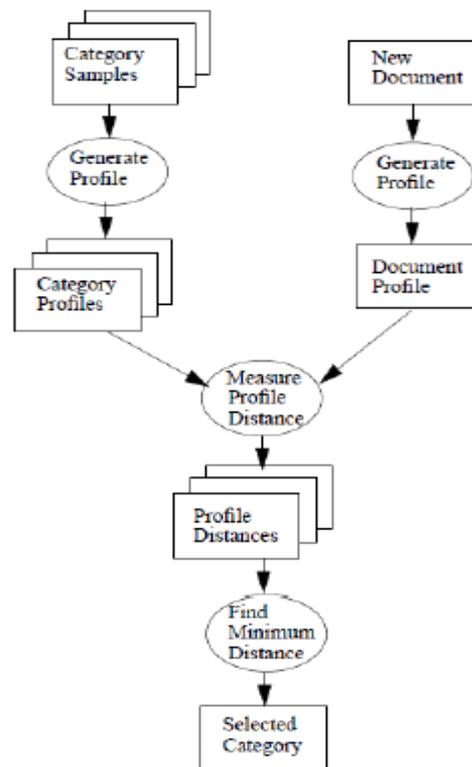
5. Tahap Pengujian Sistem

Dalam tahap ini akan dilakukan tes atau uji coba program setelah sistem dikerjakan.

Teknik *n-gram* melibatkan 2 (dua) langkah, yaitu membagi string menjadi *overlapping n-gram* (suatu set substring dengan panjang *n*) dan melakukan pengecekan untuk mendapatkan substring yang memiliki struktur yang sama. (Lisangan, Atma, & Makassar, 2015)

Gambar 1. menggambarkan proses umum yang dilakukan dalam mendeteksi suatu bahasa menggunakan *N-gram*, dimulai dengan seperangkat teks yang terdiri dari berbagai macam kategori masing-masing model bahasa, yang masing-masing kategori memiliki ribuan cacah kata. Dari seperangkat teks tersebut akan dihitung kemunculan tiap *n-gram* untuk menghasilkan satu set profil frekuensi *n-gram* untuk mewakili masing-masing kategori bahasa. Ketika dokumen baru akan dideteksi bahasanya maka akan

dihitung kemunculan tiap *n-gram* yang ada pada dokumen baru tersebut untuk menghasilkan satu set profil frekuensi *n-gram* untuk dokumen baru tersebut. Kemudian sistem akan membandingkan profil frekuensi *n-gram* dokumen baru dengan profil frekuensi *n-gram* masing-masing kategori bahasa dengan menggunakan ukuran jarak. Sistem akan mengklasifikasikan dokumen kedalam kategori bahasa tertentu yang ukuran jarak profilnya paling kecil. (Zaman & Hariyanti, 2015)



Sumber : Zaman(2015)

Gambar 1. Proses Penggunaan *N-Gram*

Menurut (Hamzah, 2010) langkah-langkah metode *n-gram* adalah sebagai berikut:

1. Hitunglah frekuensi tiap *uni-gram* (a, b, c,,) dalam koleksi dokumen bahasa Indonesia yang dipilih sebagai *training set*.
2. Tentukan probabilitas kemunculan untuk seluruh *uni-gram*: a, b,, z
3. Tentukan *profile* dokumen bahasa Indonesia dengan menetapkan nilai probabilitas tiap *uni-gram*.

Persamaan untuk menghitung statistik *Uni-gram*:

$$\text{simQ} = \sum_{i=a}^z \frac{(f_i - f_{Ri})^2}{f_{Ri}}$$

f_i = probabilitas *uni-gram* ke- i dari dokumen yang terdeteksi
 f_{Ri} = probabilitas *uni-gram* ke- i dari dokumen *training set*

sehingga dapat disimpulkan apakah kata kata tersebut akan dideteksi atau tidak
 Kalimat yang akan di uji adalah : “Program studi teknik informatika mempelajari dan menerapkan prinsip-prinsip ilmu komputer dan analisa matematis untuk desain pengembangan pengujian dan evaluasi perangkat lunak siiiistem operasi dan kerja komputer”

Hasil dan Pembahasan

1. Metode N-Gram

Metode *n-gram* berfungsi untuk mencari persentasi bobot perbandingan,

Langkah-langkah mencari nilai menggunakan *n-gram* :

1. Tentukan *bigram* dari kata yang akan dideteksi

Tabel 1. *Bigram* dari kata yang akan dideteksi

Kata A	Bigram											
program	p	pr	ro	og	gr	ra	am	m				
studi	s	st	tu	ud	di	i						
teknik	t	te	ek	kn	ni	ik	k					
informatika	i	in	nf	fo	or	rm	ma	ti	ik	Ka	a	
mempelajari	m	me	em	mp	pe	el	la	aj	ja	Ar	ri	i
dan	d	da	an	n								
menerapkan	m	me	en	ne	er	ra	ap	pk	ka	an	n	
prinsip	p	pr	ri	in	ns	si	ip	p				
ilmu	i	il	lm	mu	u							
komputer	k	ko	om	mp	pu	ut	te	er	r			
analisa	a	an	na	al	li	is	sa	a				
pengembangan	p	pe	en	ng	ge	em	mb	ba	an	ng	ga	an
pengujian	p	pe	en	ng	gu	uj	ji	ia	an	n		
Evaluasi	e	ev	va	al	lu	ua	as	si	i			
perangkat	p	pe	er	ra	an	ng	gk	ka	at	t		
Lunak	l	lu	un	na	ak	k						
Siiistem	s	si	ii	ii	is	st	te	em	m			
Operasi	o	op	pe	er	ra	as	si	i				
kerja	k	ke	er	rj	ja	a						

2. Tentukan *bigram* dari kata pada data kamus

Tabel 2. *Bigram* dari data kamus

Kata B	Bigram											
algoritma	a	al	lg	go	or	ri	it	tm	ma	a		
aplikasi	a	ap	pl	li	ik	ka	as	si	i			
backspace	b	ba	ac	ck	ks	sp	pa	ac	ce	e		
back	b	ba	ac	ck	k							
binari	b	bi	in	na	ar	ri	i					
bit	b	bi	it	t								
brainware	b	br	ra	ai	in	nw	wa	ar	re	E		
byte	b	by	yt	te	e							
link	l	li	in	nk	k							
....
....
web	w	we	eb	b								
windows	w	wi	in	nd	do	ow	ws	s				
word	w	wo	or	rd	d							

3. Tentukan probabilitas kemunculan dari data kamus :

$$P (W_i | W_{i-1}) = \frac{Count (W_{i-1}, W_i)}{Count (W_{i-1})}$$

1. Program

- a. $P (W_p | < s >) = \frac{Count (<s>, P)}{Count (<s>)} = \frac{11}{119} = 0.09$
- b. $P (W_{pr} | p) = \frac{Count (p, pr)}{Count (p)} = \frac{5}{16} = 0.3$
- c. $P (W_{ro} | pr) = \frac{Count (pr, ro)}{Count (pr)} = \frac{2}{4} = 0.5$
- d. $P (W_{og} | ro) = \frac{Count (ro, og)}{Count (ro)} = \frac{2}{5} = 0.4$
- e. $P (W_{gr} | og) = \frac{Count (og, gr)}{Count (og)} = \frac{2}{6} = 0.3$
- f. $P (W_{ra} | gr) = \frac{Count (gr, ra)}{Count (gr)} = \frac{3}{3} = 1$
- g. $P (W_{am} | ra) = \frac{Count (ra, am)}{Count (ra)} = \frac{5}{10} = 0.5$
- h. $P (W_m | am) = \frac{Count (am, m)}{Count (am)} = \frac{5}{7} = 0.71$
- i. $P (W_{<s>} | m) = \frac{Count (m, <s>)}{Count (m)} = \frac{8}{18} = 0.4$

Jadi jumlah probabilitas dari kata Program adalah = 0.09 + 0.3 + 0.5 + 0.4 + 0.3 + 1 + 0.5 + 0.71 + 0.4 = 4.2

2. Studi

- a. $P (W_A | < s >) = \frac{Count (<s>, A)}{Count (<s>)} = \frac{7}{119} = 0.058$
- b. $P (W_{Ad} | A) = \frac{Count (A, Ad)}{Count (A)} = \frac{1}{15} = 0.066$
- c. $P (W_{dm} | Ad) = \frac{Count (Ad, dm)}{Count (dm)} = \frac{1}{1} = 0.3$
- d. $P (W_{mi} | dm) = \frac{Count (dm, mi)}{Count (dm)} = \frac{1}{1} = 1$
- e. $P (W_{in} | mi) = \frac{Count (mi, in)}{Count (mi)} = \frac{1}{3} = 0.3$
- f. $P (W_n | in) = \frac{Count (in, n)}{Count (in)} = \frac{1}{12} = 0.083$
- g. $P (W_{<s>} | n) = \frac{Count (n, <s>)}{Count (n)} = \frac{6}{8} = 0.75$

Jadi jumlah probabilitas dari kata Studi adalah = 0.058 + 0.066 + 0.3 + 1 + 0.3 + 0.083 + 0.75 = 2.55

Maka didapat nilai probabilitas dari keseluruhan data sebagai berikut:

Tabel 3. Hasil probabilitas keseluruhan data

No.	Kata B	Probabilitas (P)
1	admin	2.55
2	ai	1.19
3	algoritma	4.29
4	alu	1.17
5	analog	1.89
6	aplikasi	3.27
7	backspace	4.7
8	back	2.7
9	binari	2
10	bit	1.75
11	blogger	3.59
12	bluetooth	3.11
13	brainware	4.86
81	rom	0.87
82	script	3.69
83	signal	2.49
....
...
....
....
94	virus	3.98
95	web	1.95
96	website	4.32
97	whatsapp	4.69
98	wifi	3.03
99	windows	3.76
100	word	2.11

4. Hitunglah menggunakan rumus *n-gram* untuk mencari nilai kemiripan dari kata yang dideteksi.

Pada kalimat diatas terdapat kesalahan pengetikan pada kata "sistem" menjadi "siiiistem", maka *n-gram* akan melakukan perhitungan sebagai berikut:

$$\text{siiiistem} : \text{SimQ} = \frac{(f_i - f_{Ri})^2}{f_{Ri}} = \frac{(3.32 - 2.97)^2}{2.97} = 0.04$$

Dari perhitungan di atas dapat disimpulkan bahwa semakin kecil nilai desimal yang didapat maka semakin besar kemiripan kata dengan data kamus sehingga akan terdeteksi oleh sistem sebagai kata yang harus di ubah.

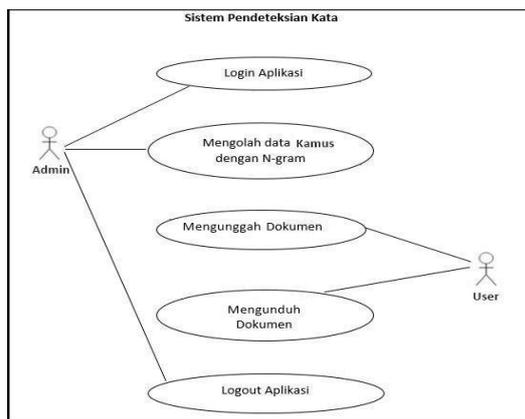
Maka kalimat yang belum di proses akan dideteksi sebagai berikut:

Kalimat sebelum dideteksi	Program studi teknik informatika mempelajari dan menerapkan prinsip-prinsip ilmu komputer dan analisa matematis
---------------------------	---

Kalimat setelah diproses untuk desain pengembangan pengujian dan evaluasi perangkat lunak sistem operasi dan kerja komputer Program studi teknik informatika mempelajari dan menerapkan prinsip-prinsip ilmu komputer dan analisa matematis untuk desain pengembangan pengujian dan evaluasi perangkat lunak <x | sistem> sistem operasi dan kerja komputer

3. Perancangan Sistem

Perancangan perangkat lunak digunakan *Unified Modelling Language* (UML) agar dapat menjelaskan alur kerja sistem yang dikembangkan atau gambaran secara umum mengenai sistem yang dikembangkan. *Use case* mendeskripsikan sebuah interaksi antara satu atau lebih aktor dengan sistem informasi yang akan dibuat (Rosa dna Shalahuddin,2013). Berikut *Use case diagram* yang digunakan dalam perancangan sistem:

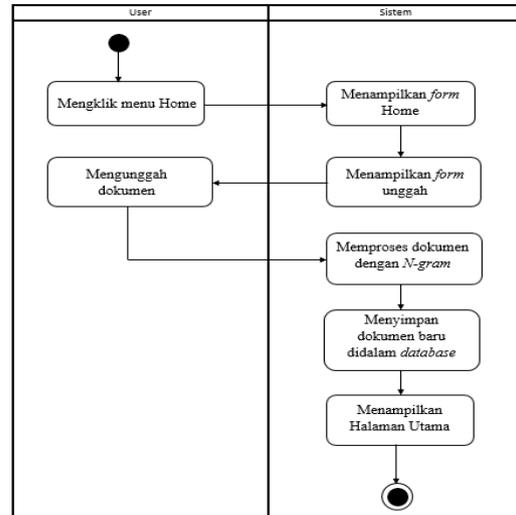


Gambar 2. Use Case Diagram

Aplikasi pendeteksian kata ini memiliki dua tingkatan hak akses yaitu admin dan user. Hak akses masing-masing dibedakan menjadi:

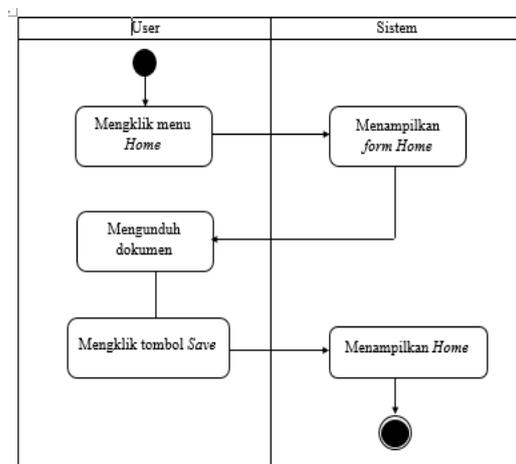
- a. Admin : Dapat melakukan pengolahan data
- b. User : Dapat melakukan penginputan pada sistem

Selain *use case diagram*, perancangan sistem juga menggunakan *activity diagram*, *Activity Diagram* menggambarkan workflow (aliran kerja) atau aktivitas dari sebuah sistem atau proses bisnis (Hendini, 2016). Berikut *activity diagram* mengunggah dokumen dan menunduh dokumen.



Gambar 3. Activity Diagram Mengunggah dokumen

Activity diagram mengunggah dokumen merupakan diagram yang menjelaskan tentang aktivitas user saat mengunggah dokumen yang ingin di deteksi oleh sistem. Sistem akan mengolah dokumen tersebut menggunakan metode *ngram* untuk menentukan deteksi kata pada dokumen yang sudah diunggah.



Gambar 4. Activity Diagram mengunduh Dokumen

Activity diagram mengunduh dokumen merupakan diagram yang menjelaskan tentang aktivitas setelah user menunggah

dokumen terlebih dahulu. Dokumen yang sebelumnya sudah diunggah akan diproses oleh sistem sehingga menghasilkan dokumen baru yang sudah terdeteksi dan dapat di unduh oleh user.

4. Implementasi Sistem

Tahapan implementasi menjelaskan tentang implementasi sistem dari awal sampai selesai. Lingkungan implementasi perangkat lunak yang dibuat meliputi spesifikasi *software* dan *hardware* yang digunakan. Berikut beberapa tampilan dari aplikasi



Gambar 5. Tampilan *input* data admin

Tampilan *input* data admin merupakan *form* untuk mengisi data diri admin baru dengan memasukkan *username* dan *password* untuk dapat digunakan sebagai akses pada *form* login. Berikut tampilan menu login



Gambar 6. Tampilan Form Login

Tampilan *form* login merupakan halaman dimana admin memasukkan *username* dan *password* untuk masuk kedalam sistem .



Gambar 7. Tampilan *form* user

Tampilan *form* user merupakan tampilan dari data admin sebagai user yang memiliki akses lebih luas pada sistem. Admin dapat membuat akses untuk admin lainnya dengan mengisi *username* dan *password* baru yang nanti akan menjadi akses untuk masuk kedalam sistem.



Gambar 8. Tampilan Halaman Kamus

Tampilan halaman Kamus merupakan kumpulan dari data-data kata yang sudah dimasukkan oleh admin sebelumnya, admin diberi akses untuk merubah atau menghapus data tersebut.



Gambar 9. Tampilan *Form* Dokumen

Tampilan *form* dokumen berisi dokumen-dokumen yang telah di masukkan oleh user, dalam hal ini admin juga dapat mengunduh atau menghapus dokumen dari sistem yang sebelumnya di *input* oleh user.

Berikut tampilan unggah dokumen ke sistem:



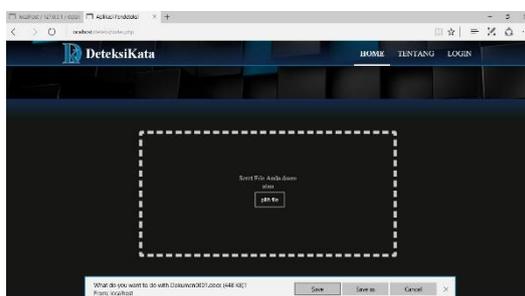
Gambar 10. Tampilan Unggah Dokumen

Gambar 10. diatas adalah tampilan unggah dokumen. Tampilan ini merupakan halaman dimana user dapat meng-inputkan dokumen atau file yang akan dideteksi kedalam aplikasi



Gambar 11. Tampilan Dokumen Siap untuk Diunduh

Tampilan unduh dokumen merupakan halaman yang menunjukkan bahwa dokumen yang sebelumnya telah di unggah sudah berhasil di proses dan siap untuk di unduh.



Gambar 12. Tampilan Unduh Dokumen

Dokumen yang sudah selesai diproses akan bisa diunduh. Gambar 5 diatas adalah tampilan untuk mengunduh dokumen yang selesai diproses.

Kesimpulan

Metode *N-gram* dapat menghitung persentasi bobot untuk menentukan apakah kata Berbahasa Indonesia akan dideteksi atau tidak. Aplikasi yang dirancang mampu

untuk mendeteksi sekitar 70% dari 10 kata yang dideteksi dan 80% keberhasilan dari pengujian karya ilmiah oleh sistem. Aplikasi dapat memberikan rekomendasi kata yang sesuai dengan ilmu komputer.

Referensi

- Afandi, R. S. (2013). Aplikasi Mobile Informasi Kafe 24 Jam Di Yogyakarta Berbasis Android Pendahuluan Landasan Teori Analisis dan Perancangan Sistem. *Dasi*, 14(04), 49–53.
- A.S, R., & Shalahuddin, M. (2013). *Rekayasa Perangkat Lunak Terstruktur Dan Berorientasi Objek*. Bandung: Informatika.
- Hamzah, A. (2010). Deteksi Bahasa Untuk Dokumen Teks. *Seminar Nasional Informatika*, 22(semnasIF), 5–13.
- Haviluddin. (2011). Memahami Penggunaan UML (Unified Modelling Language), 6(1), 1–15.
- Hendini, Ade. 2016. Pemodelan Uml Sistem Informasi Monitoring Penjualan Dan Stok Barang (Studi Kasus: Distro Zhezha Pontianak). *Jurnal Khatulistiwa Informatika*, Vol. IV, No. 2
- Lisangan, E. A., Atma, U., & Makassar, J. (2015). Implementasi n-gram Technique dalam Deteksi Plagiarisme pada Tugas Implementasi N-Gram Technique Dalam Deteksi, (May).
- Sarmadan. (2017). Pengembangan Bahan Ajar Menulis Karya Ilmiah Dalam Pengajaran Bahasa Indonesia Di Stikom Kota Jambi Sarmadan 1, 17(1), 159–171.
- Setiawati, Sulis. 2016. Penggunaan Kamus Besar Bahasa Indonesia (Kbbi) Dalam Pembelajaran Kosakata Baku dan Tidak Baku pada Siswa Kelas IV SD. *Jurnal Gramatika Jurnal Penelitian Bahasa dan Sastra Indonesia V2.i1* .Vol 2 No 1 (44-51).
- Supriadin. (2016). Identifikasi Penggunaan Kosakata Baku Dalam Wacana Bahasa Indonesia Pada Siswa Kelas VII Di SMP Negeri 1 Wera Kabupaten Bima Tahun Pelajaran 2013/2014. *JIME*, Vol. 2 No. 2 ISSN 2442 - 9511
- Zaman, B., & Hariyanti, E. (2015). Sistem Deteksi Bahasa pada Dokumen menggunakan N-Gram, 1(2), 21–26