

PROSIDING

PENINGKATAN DAYA SAING BANGSA MELALUI
PENGEMBANGAN DAN PEMANFAATAN
TEKNOLOGI INFORMASI



PROSIDING

SeTISI 2013 Seminar Teknik Informatika dan Sistem Informasi Fakultas Teknologi Informasi, Universitas Kristen Maranatha

Editor: **Robby Tan**

Desain Sampul: **Laurentius Risal**

Penerbit:

Maranatha University Press (MUP)

Jl. Prof. Drg. Suria Sumantri, MPH No. 65

Bandung 40164

Cetakan pertama, 2013

Hak cipta dilindungi undang-undang

Perpustakaan Nasional: Katalog Dalam Terbitan (KDT)

Prosiding SeTISI 2013: Prosiding Seminar Teknik Informatika dan Sistem Informasi 2013

Peningkatan Daya Saing Bangsa Melalui Pengembangan dan Pemanfaatan Teknologi Informasi / editor: Robby Tan, Bandung, MUP, 2013

299 hlm, 21 × 29,7 cm

ISBN 978-602-98685-3-1

KOMITE

Pelindung

Rektor Universitas Kristen Maranatha

Penanggung Jawab

Dekan Fakultas Teknologi Informasi
Universitas Kristen Maranatha

Ketua Pelaksana

Dr. Andi Wahyu Rahardjo Emanuel, BSEE., MSSE.

Komite Program

Dr. Andi Wahyu Rahardjo Emanuel, BSEE., MSSE. (UKM)

Ir. Dana Indra Sensuse, MLIS, Ph.D. (UI)

Dr. Ir. Husni Setiawan Sastramihardja, M.T. (ITB)

Ito Wasito, Ph.D. (UI)

Ir. Kridanto Surendro, M.Sc., Ph.D. (ITB)

Dr. Ir. Mewati Ayub, M.T. (UKM)

Dr. dr. Oerip S. Santoso, M.Sc. (ITB)

Drs. Retantyo Wardoyo, M.Sc., Ph.D. (UGM)

Dr. Ir. Rila Mandala, M.Eng. (ITB)

Dra. Sri Hartati, M.Sc., Ph.D. (UGM)

Yenni M. Djajalaksana, Ph.D. (UKM)

Komite Pelaksana

Radiant Victor Imbar, S.Kom., M.T.

Doro Edi, S.T., M.Kom.

Tanti Kristanti, S.T., M.T.

Hendra Bunyamin, S.Si., M.T.

Hapnes Toba, M.Sc.

Yenni M. Djajalaksana, Ph.D.

Robby Tan, S.T., M.Kom.

Maresha Caroline Wijanto, S.Kom.

Laurentius Risal, S.T.

Meliana Christianti J., S.Kom., M.T.

Daniel Jahja Surjawan, S.Kom., M.T.

Diana Trivena Yulianti, S.Kom., M.T.

Tjatur Kandaga, S.Si., M.T.

Sendy Ferdian, S.Kom.

DAFTAR ISI

KOMITE	i
KATA PENGANTAR	ii
SAMBUTAN DEKAN.....	iii
DAFTAR ISI	iv
Penggunaan Metode <i>Paper Prototype</i> untuk Melakukan Inspeksi <i>Usability</i> pada Aplikasi Berbasis Web (Studi kasus: Sistem Informasi Akademik Universitas)	
R. Sandhika Galih A.....	1
Perancangan Game “<i>Onion Boy</i>” Berbasis Android untuk Melatih Kecepatan dan Fleksibilitas	
Irene A. Lazarusli, Aditya R. Mitra, Kenny Saputra	6
Algoritma Penggantian <i>Cache</i> sebagai Optimalisasi Kinerja pada <i>Proxy Server</i>	
Suandra Eka Saputra, Timotius Witono.....	12
<i>Website</i> Perhitungan Angka Kecukupan Gizi Anak	
Pratiwi Chandraningsih, Diana Trivena Yulianti.....	18
Pengamanan Jalur Komunikasi Internet Menggunakan PPTP (<i>Point-to-Point Tunnelling Protocol</i>)	
I Made Mustika Kerta Astawa, Claudia Dwi Amanda	24
<i>Sentiment Classification</i> Menggunakan <i>Machine Learning</i>: Metode Naïve-Bayes dan <i>Support Vector Machines</i> (Studi kasus: <i>movie reviews imdb.com</i>)	
Hendra Bunyamin, Tjatur Kandaga	29
Analisis IT <i>Governance</i> pada Layanan Teknologi Informasi Perguruan Tinggi Berbasis IT <i>Service Management</i>	
Aradea.....	37
<i>Monogame Framework</i> sebagai Salah Satu <i>Framework</i> Alternatif pada Mata Kuliah Pemrograman <i>Game</i>	
Erico Darmawan Handoyo, Sulaeman Santoso	43
Penerapan SMS <i>Gateway</i> untuk Pengingat dan Rekomendasi di Rental Komik Daruma	
Teddy Marcus Zakaria, Inwan Aditya Halim	47
Penerapan Algoritma Bayesian <i>Classification</i> untuk Pemberian Harokat pada Kalimat Bahasa Arab	
Maliki Ahmad Nur, Irfan Maliki	53
<i>Website</i> Penyedia Informasi Pariwisata di Kota Bandung Menggunakan <i>Ruby on Rails</i>	
Resky Bagja Sunjaya, Robby Tan	58
<i>E-Services Customer Management System</i> Unit Pelayanan PT. XYZ	
Eka Widhi Yunarso	65
Analisis Perbandingan Unjuk Kerja Protokol TCP, UDP, dan SCTP Menggunakan Simulasi Lalu Lintas Data Multimedia	
Rinda Tri Yuniar Anggraeni, Jusak, Anjik Sukmaaji	72
<i>Best Practices for Choosing Non-Intrusive but Effective CAPTCHAs</i>	
Setia Budi	78
Deteksi Otomatis Perubahan Pustaka API dengan Solusi Sistem Repositori Kode Sumber dan Revisi API Pustaka Perangkat	

Penerapan Algoritma Bayesian *Classification* untuk Pemberian Harokat pada Kalimat Bahasa Arab

Maliki Ahmad Nur^{#1}, Irfan Maliki^{*2}

[#]Teknik Informatika, Universitas Komputer Indonesia
Jl. Dipati Ukur 112-116 Bandung 40132

¹malikiahmad11@gmail.com

^{*}Teknik Informatika, Universitas Komputer Indonesia
Jl. Dipati Ukur 112-116 Bandung 40132

²irfanmaliki007@gmail.com

Abstract — Arabic is a language used in the Qur'an and the Hadith as well as other Islamic knowledge sources. Because of grammar complexity, there are several things must be known such as grammar in writing and reading. As an initial capital to simplification of Islamic sciences study, the authors carried on a research to develop a web-based application by applying the classification method using Bayesian Classification algorithm. This method will help represents the Arabic grammar by predicting the probability of the membership of a grammar course based on training data contained in the database. Bayesian Classification was proven to have a high accuracy rate when applied into the large size database. By the testing on the simple arabic sentences, this application was applicable with up to 93,3% to give harakat of word grammatically.

Keywords — Harakat, Arabic Sentence, Bayesian Classification Algorithm.

I. PENDAHULUAN

Semakin tingginya antusias masyarakat terhadap lembaga-lembaga pendidikan agama Islam, terutama bagi para pelajar Islam, memicu mereka untuk mengkaji lebih dalam lagi mengenai makna dari Al-Qur'an dan Hadits yang lebih banyak tertuang di dalam kitab-kitab ulama terdahulu atau sering disebut kitab kuning dan hampir keseluruhannya menggunakan bahasa Arab tanpa harokat. Membaca dan memahami kitab kuning memang tidak mudah, sebab penulisannya tanpa harokat dan gramatika bahasa Arab yang terbilang rumit, serta untuk mencari arti suatu kata di dalam kamus bahasa Arab perlu diketahui kata asalnya terlebih dahulu.

Berdasarkan hal tersebut, sebagai modal awal simplikasi untuk mempelajari ilmu dasar keislaman, akan dilakukan penelitian mengenai pemberian harokat pada kalimat bahasa Arab dengan menerapkan metode klasifikasi menggunakan algoritma Bayesian Classification yang akan membantu merepresentasikan gramatika bahasa Arab dengan memprediksi probabilitas keanggotaan suatu kelas gramatika berdasarkan *data training* yang terdapat pada basis data. Menurut Kusrini dan Emha Taufiq Luthfi [1],

Bayesian Classification terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam *database* dengan data yang besar.

Berdasarkan latar belakang yang telah dijelaskan, maka maksud dari penelitian ini adalah untuk membangun aplikasi pemberian harokat pada kalimat bahasa Arab, sedangkan tujuan yang akan dicapai dari penelitian ini adalah sebagai berikut:

1. Mengimplementasikan metode klasifikasi menggunakan algoritma Bayesian Classification dengan masukan berupa kalimat bahasa Arab tanpa harokat, dan akan menghasilkan keluaran berupa kalimat bahasa Arab yang diharokati sesuai gramatikanya.
2. Modal awal simplikasi para pelajar Islam dalam belajar membaca kalimat bahasa Arab tanpa harokat.

II. LANDASAN TEORI

A. Gramatika Bahasa Arab

Gramatika bahasa Arab pada dasarnya terbagi menjadi dua bidang keilmuan, yaitu ilmu *Nahwu* dan ilmu *Shorof*.

1) Ilmu Nahwu: Ilmu Nahwu secara khusus berbicara tentang kedudukan tiap elemen kalimat dan secara umum berbicara tentang aturan mengenai hubungan antar elemen tersebut.[5] kalimat dalam bahasa Arab terbagi menjadi tiga jenis, yaitu isim, fi'il, huruf ma'ani. Isim adalah kalimat yang menunjukkan makna mandiri dan tidak disertai dengan pengertian waktu.[2] Fi'il adalah kalimat yang menunjukkan makna mandiri dan disertai dengan pengertian waktu.[2] Huruf ma'ani adalah huruf yang memiliki makna tertentu, namun tidak akan sempurna maknanya kecuali jika dihubungkan dengan kata lain.[2]

2) Ilmu Shorof: Ilmu shorof berbicara tentang aturan pembentukan kata. Ia mempelajari timbangan-timbangan kata (*wazan*) berikut indikasinya, serta bentuk-bentuk perubahan yang sangat beragam.[5]

B. Algoritma Bayesian Classification

Bayesian Classification adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Bayesian Classification didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan *decision tree* dan *neural network*. Bayesian Classification terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam *database* dengan data yang besar.[1] Bayesian Classification mengasumsikan bahwa keberadaan sebuah atribut (variabel) tidak ada kaitannya dengan *n* keberadaan atribut yang lain.

Jika diketahui *X* adalah data sampel dengan kelas yang belum diketahui, *H* merupakan hipotesa bahwa *X* adalah data dengan kelas *C*, *P(H)* adalah probabilitas dari hipotesa *H* (*prior probability*), *P(X)* adalah probabilitas data sampel yang diamati, maka *P(X/H)* adalah probabilitas data sampel *X* berdasarkan kondisi pada hipotesa *H*, bila diasumsikan bahwa hipotesa *H* benar (*valid*). Karena asumsi atribut tidak saling terkait (*conditionally independent*), maka *P(X/Ci)* dapat didekati dengan cara:

$$P(X | C_i) = \prod_{k=1}^n P(X_k | C_i)$$

Gambar 1 Rumus mencari nilai probabilitas X terhadap Ci.

Jika *P(X/Ci)* diketahui maka kelas dari data sampel *X* dapat didekati dengan menghitung *P(X/Ci)*P(Ci)*. kelas *Ci* dimana *P(X/Ci)*P(Ci)* maksimum adalah kelas dari sampel *X*.

III. ANALISIS ALGORITMA

Sebelum melakukan proses klasifikasi untuk menentukan gramatika kata menggunakan algoritma Bayesian Classification, kalimat bahasa Arab sederhana tanpa harakat akan melalui beberapa tahapan *preprocessing*. *Pre processing* yang pertama dilakukan untuk memisahkan kalimat menjadi per kata yang disimpan dalam variabel *array*, seperti contoh kalimat "يكون زيد في المكتبة" menjadi \$pecah[0] = "يكون", \$pecah[1] = "زيد", \$pecah[2] = "في", \$pecah[3] = "المكتبة". Kedua, setelah kalimat dipisah menjadi per kata, setiap kata akan melewati proses pemberian harokat sesuai bentuk pola kata dan akan diidentifikasi dengan memberikan kode unik di awal kalimat. Pada Tabel I menjelaskan beberapa contoh kata berdasarkan kategorinya masing-masing. Contoh kata ini mengacu pada contoh kata yang telah dibuat [3], [4].

TABEL I
CONTOH KATA BERDASARKAN KATEGORI MASING-MASING

No	Bentuk Kata	Kategori Kata	Kode Unik
1	من، الى، عن، على، في	Huruf Jar/Khofadz	Hj
2	هو، هما، هم، هي، هن، انت، انتم، انا، نحن	Isim Dhomir (kata ganti)	id
3	هذا، هذه، هؤلاء، ذلك، تلك	Isim Isyarah (kata tunjuk)	ii
4	الذي، الذين، التي، اللاتي	Isim Maushul (kata penghubung)	im
5	ان، لن، كي، حتي	Amil Nawasib	an
6	لم، لما، الم، لما، إن	Amil Jawazim	aj
7	كان، اصبح، ليس، امسى	Kaana dan kawanannya	kn
8	ان، أن، لعل، ليت، لكن	Inna dan kawanannya	in
9	استعمل، فعل، افعل	Fi'il madhi	fm
10	يستعمل، يفعل، يفعل	Fi'il mudhore	fr
11	أستعمل، أفعل	Fi'il amr	fa
12	البيت، الرجل	Isim dengan alif lam	al
13	مدرسة، مسجد	Isim tanpa alif lam	no

Metode yang digunakan dalam *preprocessing* untuk mengidentifikasi kata yaitu metode Levenshtein Distance.

A. Identifikasi Kata Menggunakan Metode Levenshtein Distance

Metode ini digunakan dalam pencarian string dengan pendekatan perkiraan (*Approximate String Matching*). Dengan pendekatan perkiraan ini, pencarian string target menjadi tidak harus sama persis dengan yang ada di dalam string sumber.[7]

Contoh identifikasi kata dengan metode Levenstein Distance:

String sumber (\$pecah[0]) = "يكون"

String target = "كان"

Proses algoritma Levenshtein Distance adalah sebagai berikut:

- Jika panjang sumber adalah 0 maka jarak sumber dengan target adalah panjang target.
- Jika panjang target adalah 0 maka jarak sumber dengan target adalah panjang sumber.
- Buat matriks s[x], t[y] dengan ukuran s[x] = (sumber+1) X t[y] = (target+1).
- Inisialisasi s[x] pertama (0..n).
- Inisialisasi t[y] pertama (0..m).

		ي	ك	و	ن
	0	1	2	3	4
ك	1				
ا	2				
ن	3				

Gambar 2 Inisialisasi matriks (n+1)x(m+1)



- Melakukan proses pencocokan dengan membandingkan setiap karakter pada string target dengan setiap karakter pada string sumber.
- Jika karakter $s[x] = t[y]$ maka beri nilai 0.
- Jika karakter $s[x] \neq t[y]$ maka beri nilai 1.

Tentukan nilai $cell [x, y]$ dengan memasukkan nilai terkecil dari:

- Nilai $cell [x-1, y-1] + \text{nilai}$
- Nilai $cell [x-1, y] + 1$
- Nilai $cell [x, y-1] + 1$

		ي	ك	و	ن
0		1	2	3	4
ك	1	1	1	2	3
ا	2	2	2	2	3
ن	3	3	3	3	2

Gambar 3. Matriks Proses Levenshtein Distance

Hingga tahap ini dapat disimpulkan bahwa jarak antara string sumber yang terdapat pada variabel $\$specah[0]$ mendekati string target yang merupakan jenis kata *amil kaana* dengan jarak 2, sehingga isi pada variabel $\$specah[0]$ akan menjadi $\$specah[0] = \text{"يَكُونُ"}$ dengan diberikan harakat sesuai bentuk kata dan ditambah kode unik untuk *amil kaana* di awal kata.

B. Proses Penentuan Gramatika Kata Menggunakan Algoritma Bayesian Classification

1) Proses Pembelajaran: Pembelajaran terhadap aplikasi dengan memasukkan data latihan berupa kode unik pada setiap kata yang didapatkan melalui tahap identifikasi. Kemudian menentukan kelas masing-masing berdasarkan hasil hipotesis terhadap kalimat bahasa Arab sederhana. Berikut ini keterangan mengenai kode unik yang akan dijadikan data latihan:

- KODE = kode yang terdapat pada kata itu sendiri
- KD1 = kode yang terdapat pada kata setelahnya
- KD2 = kode yang terdapat pada kata sebelumnya
- KD3 = kode yang terdapat pada kata dua langkah sebelumnya.

Pada Tabel II menjelaskan data latihan yang sudah diklasifikasi:

TABEL II
CONTOH DATA LATIHAN

ID	Kode	KD1	KD2	KD3	Kelas
1	Al	al	Fr	an	Fa'il
2	Al	hj	Fr	Aj	Fa'il
3	Al	0	Fr	0	Fa'il
4	No	0	Hj	Al	Isim majrur
5	Al	0	No	Kn	Khobar amil kaana
6	No	0	Al	0	Khobar
7	Al	No	0	0	Mubtada
8	No	No	Fm	0	Fa'il
9	No	0	No	Fm	Maf'ul
10	Al	0	Hj	No	Isim majrur

Setiap kelas akan ditentukan nilai $P(C_i)$, untuk $i=1..n$. $P(C_i)$ adalah nilai *prior probability* untuk setiap kelas berdasarkan data latihan seperti pada tabel III.

TABEL III
NILAI PRIOR PROBABILITY UNTUK SETIAP KELAS

No	P(Ci)	Nilai P(Ci)
1	P(gramatika="Mubtada")	$4/38 = 0.105$
2	P(gramatika="Khobar")	$3/38 = 0.078$
3	P(gramatika="Fa'il")	$4/38 = 0.105$
4	P(gramatika="Maf'ul")	$8/38 = 0.210$
5	P(gramatika="Isim amil kaana")	$3/38 = 0.078$
6	P(gramatika="Khobar amil kaana")	$2/38 = 0.053$
7	P(gramatika="isim amil inna")	$3/38 = 0.078$
8	P(gramatika="Khobar amil inna")	$2/38 = 0.053$
9	P(gramatika="Isim Majrur")	$5/38 = 0.131$
10	P(gramatika="Fi'il Mansub")	$2/38 = 0.053$
11	P(gramatika="Fi'il Majzum")	$2/38 = 0.053$

2) Proses Klasifikasi: Pada proses klasifikasi digunakan data testing. Data testing merupakan data masukan yang belum diklasifikasi. Data testing adalah data yang telah melewati proses pemberian kode unik dan telah diberi harokat sesuai bentuk kata. Sebagai contoh, kalimat masukan "يكون زيد في المكتبة", setelah melewati proses pemberian kode unik dan harokat secara umum sebagai berikut: $\$specah[0] = \text{"يَكُونُ"}$, $\$specah[1] = \text{"نَزِيدُ"}$, $\$specah[2] = \text{"هِيَ"}$, $\$specah[3] = \text{"المكتبة"}$. Kata dalam variabel " $\$specah[1]$ " memiliki data sebagai berikut: KODE="no" KD1="hj" KD2="kn" KD3="0". Kata tersebut belum diketahui kelasnya, sehingga harus melalui tahap klasifikasi menggunakan algoritma Bayesian Classification, proses klasifikasi kata dalam variabel " $\$specah[1]$ " adalah sebagai berikut:

- Diketahui data $X = (\text{KODE}=\text{"no"} \text{ KD1}=\text{"hj"} \text{ KD2}=\text{"kn"} \text{ KD3}=\text{"0"})$
- Memaksimalkan $P(X | C_i) P(C_i)$ untuk $i=1..11$. hasil perhitungan terdapat pada tabel IV.

TABEL IV
HASIL PERHITUNGAN PROBABILITAS X TERHADAP KELAS C_i

Kelas	Kode	KD 1	KD 2	KD3	P(X C _i)	P(X C _i) * P(C _i)
Mubtada	0,25	0,25	0	1	0	0
Khobar	1	0,33	0	1	0	0
Fa`il	0,25	0,25	0	0,5	0	0
Maf`ul	0,375	0	0	0,625	0	0
Isim Kaana	0,33	0,33	1	1	0,1089	0,0084
Khobar Kaana	0,5	0	0	0	0	0
Isim Inna	0,6	0,33	0	1	0	0
Khobar Inna	0,5	0	0	0	0	0
Isim Majrur	0,4	0	0	0,2	0	0
Fi`il Mansub	0	0	0	0,5	0	0
Fi`il Majzum	0	0	0	1	0	0

Kesimpulan dari data pada Tabel IV adalah nilai probabilitas X terhadap C_i maksimum terdapat pada kelas gramatika *isim kaana*, maka kata “ زَيْدٌ ” di dalam kalimat “ يكون زيد في المكتبة ” mempunyai kedudukan dalam gramatika berupa “*isim kaana*”.

IV. IMPLEMENTASI DAN PENGUJIAN

A. Batasan Implementasi

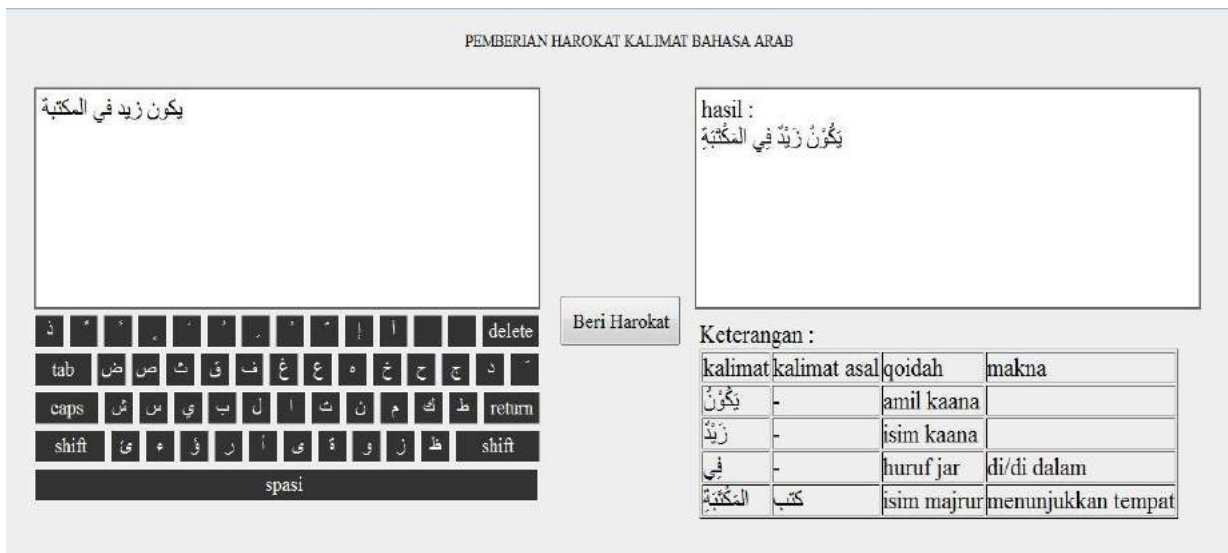
Batasan implementasi bertujuan agar dalam pengujian aplikasi pemberian harokat pada kalimat bahasa Arab dapat berjalan dengan baik, sesuai dengan data yang ada, perangkat keras dan perangkat lunak yang digunakan. Adapun batasan-batasan implementasinya adalah sebagai berikut:

1. Data yang dimasukkan admin berupa kategori kata, kosakata, pola kata, kelas gramatika, dan data latihan.
2. Parameter logika Levenstein Distance ditentukan oleh admin dengan mengambil jarak minimal antara string sumber dan string target pada basis data.
3. Data yang dimasukkan *user* berupa kalimat bahasa Arab sederhana tanpa harokat.

B. Pengujian

Antarmuka halaman pengguna pada aplikasi ini adalah seperti pada Gambar 4. Pengujian pemberian harokat ini dilakukan dengan memasukkan kalimat sederhana bahasa Arab tanpa harokat pada halaman *user*.

Untuk hasil pengujian pada aplikasi yang telah memiliki 40 data latihan yang terbagi menjadi 11 kelas, menggunakan berbagai bentuk susunan kalimat sederhana bahasa Arab tanpa harokat dengan rata-rata kalimat memiliki 4 kosakata disajikan pada Tabel V.



Gambar 4 Antarmuka Halaman User

TABEL V
PENGUJIAN AKURASI

No	Kalimat Masukan	Hasil Keluaran	Hasil Yang Benar	Status
1	جاء علي إلى المدرسة	جَاءَ عَلِيٌّ إِلَى الْمَدْرَسَةِ	جَاءَ عَلِيٌّ إِلَى الْمَدْرَسَةِ	Benar
2	ذهب الفلاح إلى المزرعة	ذَهَبَ الْفَلَّاحُ إِلَى الْمَزْرَعَةِ	ذَهَبَ الْفَلَّاحُ إِلَى الْمَزْرَعَةِ	Benar
3	اكرم زيد محمد	أَكْرَمَ زَيْدٌ مُحَمَّدًا	أَكْرَمَ زَيْدٌ مُحَمَّدًا	Benar
4	إن محمد رسول الله	إِنَّ مُحَمَّدًا رَسُولُ اللَّهِ	إِنَّ مُحَمَّدًا رَسُولُ اللَّهِ	Benar
5	إن الله غفور رحيم	إِنَّ اللَّهَ غَفُورٌ رَحِيمٌ	إِنَّ اللَّهَ غَفُورٌ رَحِيمٌ	Benar
6	ليس رجال ضعيف	لَيْسَ رِجَالٌ ضَعِيفٌ	لَيْسَ رِجَالٌ ضَعِيفٌ	Benar
7	كتب الكاتب الرسالة	كَتَبَ الْكَاتِبُ الرَّسَالََةَ	كَتَبَ الْكَاتِبُ الرَّسَالََةَ	Benar
8	وضع محمد في الكتاب في المكتبة	وَضَعَ مُحَمَّدٌ فِي الْكِتَابِ فِي الْمَكْتَبَةِ	وَضَعَ مُحَمَّدٌ فِي الْكِتَابِ فِي الْمَكْتَبَةِ	Benar
9	لم يخرج البخيل المال	لَمْ يُخْرِجِ الْبَخِيلُ الْمَالَ	لَمْ يُخْرِجِ الْبَخِيلُ الْمَالَ	Benar
10	يتقى القتال	يَتَّقَى الْقِتَالَ	يَتَّقَى الْقِتَالَ	Benar
11	لن نقاتل الجيش	لَنْ نُقَاتِلَ الْجَيْشَ	لَنْ نُقَاتِلَ الْجَيْشَ	Benar
12	يكون زيد عالم	يَكُونُ زَيْدٌ عَالِمًا	يَكُونُ زَيْدٌ عَالِمًا	Benar
13	اشرق الكون ابتهاج	أَشْرَقَ الْكَوْنُ ابْتِهَاجًا	أَشْرَقَ الْكَوْنُ ابْتِهَاجًا	Benar
14	يكثر النخيل في الحنوط	يُكْتَبِرُ النَّخِيلُ فِي الْحَنُوطِ	يُكْتَبِرُ النَّخِيلُ فِي الْحَنُوطِ	Benar
15	هو جميل	هُوَ جَمِيلٌ	هُوَ جَمِيلٌ	Benar
16	من جد وجد	مَنْ جَدَّ وَجَدَّ	مَنْ جَدَّ وَجَدَّ	Benar
17	عطش الثعلب	عَطَشَ الثَّعْلَبُ	عَطَشَ الثَّعْلَبُ	Benar
18	احتاج الأمير إلى شيء	إِحْتَاكَ الْأَمِيرُ إِلَى شَيْءٍ	إِحْتَاكَ الْأَمِيرُ إِلَى شَيْءٍ	Benar
19	ذلك طريق واسع	ذَلِكَ طَرِيقٌ وَاسِعٌ	ذَلِكَ طَرِيقٌ وَاسِعٌ	Benar
20	أنا أحب إليك	أَنَا أُحِبُّ إِلَيْكَ	أَنَا أُحِبُّ إِلَيْكَ	Benar
21	انشرح المفتش	إِنشَرَخَ الْمُفْتَشُ	إِنشَرَخَ الْمُفْتَشُ	Benar
22	عيادة المريض	عِيَادَةُ الْمَرِيضِ	عِيَادَةُ الْمَرِيضِ	Benar
24	ركب خالد الحمار	رَكَبَ خَالِدٌ الْحِمَارَ	رَكَبَ خَالِدٌ الْحِمَارَ	Benar
25	انطلق المستوطى إلى المدينة	إِنطَلَقَ الْمُسْتَوطِيُّ إِلَى الْمَدِينَةِ	إِنطَلَقَ الْمُسْتَوطِيُّ إِلَى الْمَدِينَةِ	Benar
26	يشاهد مشاهد الجميلة	يُشَاهِدُ مَشَاهِدَ الْجَمِيلَةِ	يُشَاهِدُ مَشَاهِدَ الْجَمِيلَةِ	Benar
27	دعى عبد الغفار صاحب في المدرسة	دَعَى عَبْدُ الْغَفَارِ صَاحِبًا فِي الْمَدْرَسَةِ	دَعَى عَبْدُ الْغَفَارِ صَاحِبًا فِي الْمَدْرَسَةِ	Salah
28	من أرد أن يحافظ على النظر	مَنْ أَرَدَ أَنْ يُحَافِظَ عَلَى النَّظَرِ	مَنْ أَرَدَ أَنْ يُحَافِظَ عَلَى النَّظَرِ	Salah
29	صلى مسلم في	صَلَّى مُسْلِمٌ فِي	صَلَّى مُسْلِمٌ فِي	Benar

No	Kalimat Masukan	Hasil Keluaran	Hasil Yang Benar	Status
	المسجد	الْمَسْجِدِ	الْمَسْجِدِ	
30	تعلم متعلم في المدرسة	تَعَلَّمَ مُتَعَلِّمٌ فِي الْمَدْرَسَةِ	تَعَلَّمَ مُتَعَلِّمٌ فِي الْمَدْرَسَةِ	Benar

Untuk mengevaluasi kinerja metode Bayesian Classification dalam mengklasifikasi kata pada suatu kalimat berdasarkan kedudukan gramatikanya menggunakan persentase nilai akurasi, yaitu:

$$\text{Akurasi} = \frac{\text{Jumlah klasifikasi per kata benar}}{\text{Jumlah kata}} \times 100\%$$

Gambar 5 Rumus Mencari Nilai Akurasi

Dari data pengujian akurasi pada tabel V dapat diambil kesimpulan:

1. Jumlah kalimat = 30
2. Jumlah kata = 100
3. Jumlah klasifikasi per kata benar = 97
4. Akurasi = 97%

V. SIMPULAN

Berdasarkan penelitian yang telah dilakukan, simpulan yang dapat diambil dari semua proses yang dilakukan dalam membangun aplikasi pemberian harokat pada kalimat bahasa Arab yaitu:

1. Aplikasi ini dapat melakukan pemberian harokat terhadap kalimat bahasa Arab sederhana tanpa harokat.
2. Algoritma Bayesian Classification dapat melakukan klasifikasi terhadap gramatika setiap kata pada suatu kalimat setelah kata berhasil diidentifikasi.
3. Berdasarkan hasil pengujian, aplikasi ini berhasil memberikan harokat pada setiap bentuk kata dan melakukan klasifikasi terhadap gramatika kata serta memberikan harokat akhir kata sesuai gramatikanya dengan akurasi 97 %.

DAFTAR PUSTAKA

- [1] Kusri dan Taufiq Lutfhi, Emha. Algoritma Data Mining. Andi Offset, Yogyakarta, 2009.
- [2] Muhammad Araa'ini, Syamsuddin. Ilmu Nahwu Terjemah *Mutamimmah Al-Ajurumiyah*. Sinar Baru Algensindo, Bandung, 2010.
- [3] Maksun bin Ali, Muhammad. *Al-amtsilah Attashrifiyah*. Maktabah Salim Nabhan, Surabaya, 2000.
- [4] Hakim, Taufiqul. Program Pemula Membaca Kitab Metode "Amsilati". PP Darul Falah, Jepara, 2003.
- [5] (2012) Menara Islam website. [Online]. Tersedia: <http://menaraislam.com/content/view/15/26/>.
- [6] Hamzah, Amir. Klasifikasi Teks Dengan *Naïve Bayes Classifier* (NBC), Untuk Pengelompokan Teks Berita Dan Abstrak Akademis, Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) Periode III, Nov. 03, 2012. Tersedia: http://ie.akprind.ac.id/sites/default/files/Amir%20Hamzah%20Teknik%20Informatika_.pdf.
- [7] Syaroni, Mokhammad dan Munir, Rinaldi. Pencocokan String Berdasarkan Kemiripan Ucapan (*Phonetic String Matching*) Dalam Bahasa Inggris, 2005. Tersedia: <http://journal.uui.ac.id/index.php/Snati/article/viewFile/1413/1193>