IMPLEMENTASI TEXT MINING KLASIFIKASI OBJEK WISATA DENGAN METODE NAÏVE BAYES CLASSIFIER DI DINAS PARIWISATA DAN KEBUDAYAAN JAWA BARAT



DIMAS RICKY FIRMANSYAH

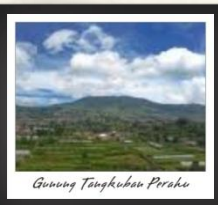
10108401

UNIKOM

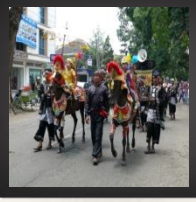
LATAR BELAKANG



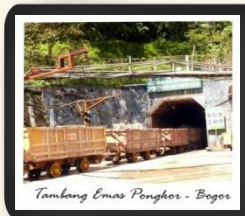
8 Karakter


Gunung Tangkuban Perahu

Karakter Hutan



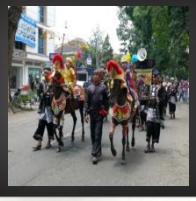
Karakter Gelar Seni &Ragam Festival


Tambang Emas Pongkor - Bogor

Karakter Religius


Gua Pawon, Gunung Masigit Cipatat - Bandung

Karakter Kepurbakalaan



Karakter Tradisi & Komunikasi Adat


Pantai Cimaja Pelabuhan Ratu Sukabumi

Karakter Pantai


Keraton Kasepuhan Cirebon

Karakter Heritage


Wisata Belanja Cihampelas Bandung

Karakter Wisata Belanja

1



Proses pengklasifikasianberdasarkan asumsi pribadi orang yang mengklasifikasikan



Membutuhkan waktu yang lama dalam mengklasifikasikannya

RUMUSAN MASALAH



Bagaimana membangun Sistem Pengklasifikasian Objek Wisata berdasarkan kategori ke dalam 9 karakter dengan metode naïve bayes classifier Di Dinas Pariwisata Jawa Barat

MAKSUD & TUJUAN



Menghasilkan aturan klasifikasi untuk
Setiap kategori objek wisata ke dalam 9 karakter



Membantu dalam proses klasifikasi objek wisata
Berdasarkan kategori ke dalam 9 karakter

BATASAN MASALAH

• Sistem yang akan dibangun adalah sistem pengklasifikasian mengenai objek wisata di Dinas Pariwisata Jawa Barat.

• Data objek wisata didapat dari Dinas Pariwisata dan Kebudayaan Jawa Barat.

• Informasi yang diberikan sistem yaitu berupa identitas objek wisata (nama objek wisata, karakter, dan kota/kab) dan deskripsi objek wisata.

• Data objek wisata yang digunakan yaitu data objek wisata yang ditulis dalam bahasa Indonesia, dan tidak dilakukan tahap *tagging* pada proses *text mining* karena tidak menangani teks yang berbahasa inggris.

•Pengklasifikasian objek wisata berdasarkan 9 karakter, yaitu proses pengelompokkan objek wisata yang ada di Jawa Barat berdasarkan kategori.

•Data Latih yang digunakan adalah 15 objek wisata yang terdapat pada buku informasi pariwisata Jawa Barat.

•Algoritma yang akan diimplementasikan pada bahasan mengenai proses pengklasifikasian objek wisata adalah *Naïve Bayes Clasifier*.

•Perangkat lunak atau aplikasi ini akan berjalan pada sistem berbasis website.

•Pendekatan analisis yang digunakan adalah *Objek Oriented* (OO).

TEXT MINING



“Proses penggalian informasi dimana seorang user berinteraksi dengan sekumpulan dokumen menggunakan tools analisis yang merupakan komponen-komponen dalam data *mining* yang salah satunya adalah kategorisasi”

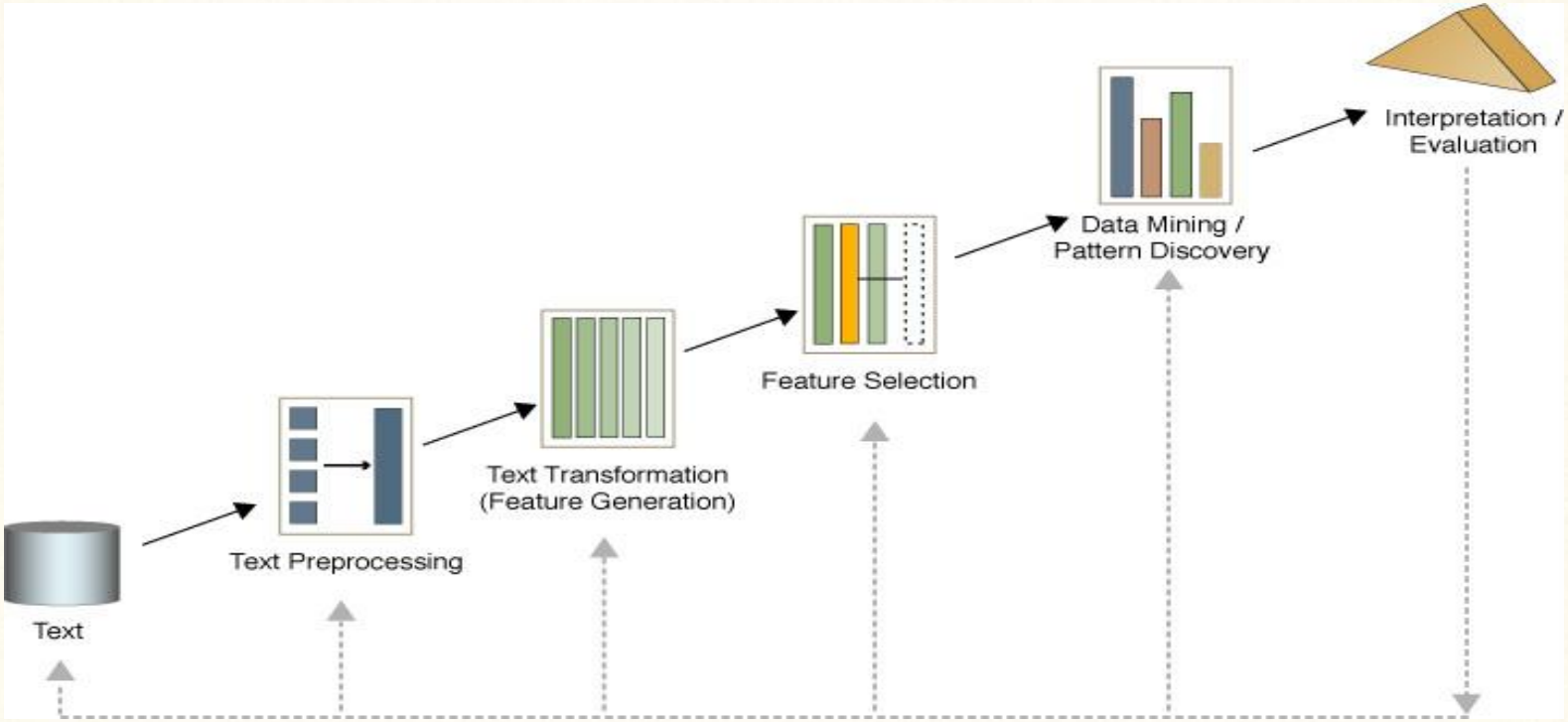
Ronen Feldman dan James Sanger[4]

2

TUJUAN

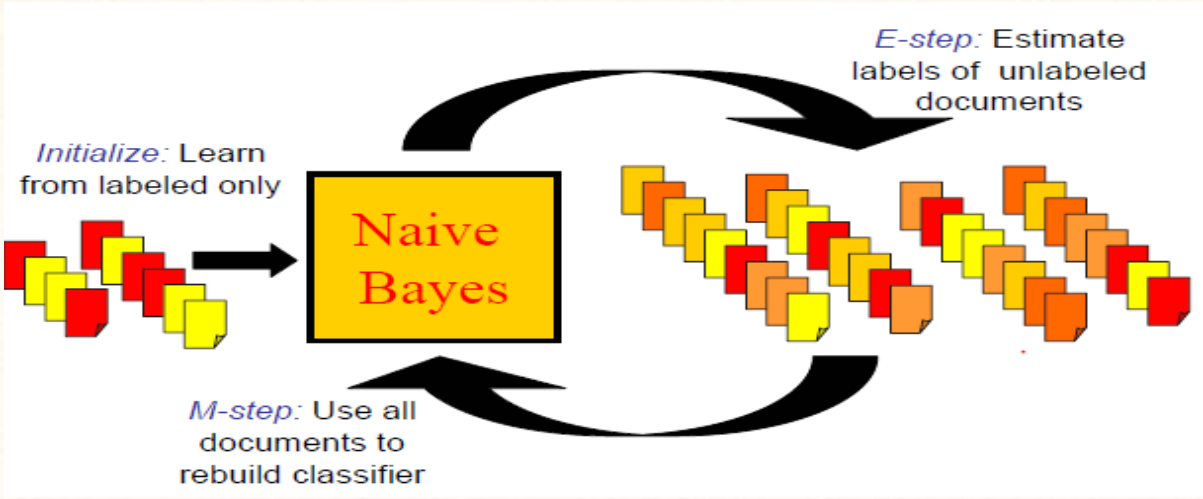
"Tujuan dari *Text Mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen tetapi Tujuan utama *text mining* adalah mendukung proses *knowledge discovery* pada koleksi dokumen yang besar"

PROSES TEXT MINING



Gambar Proses *Text Mining* menurut even,zohar [6]

NAIVE BAYES CLASSIFIER



Gambar klasifikasi *naïve bayes classifier*.[9]

"pengklasifikasian statistik yang dapat digunakan untuk Memprediksi probabilitas keanggotaan suatu class"

Mitchell, Tom A. *Machine Learning*. McGraw-Hill.1997

KONSEP NVB

1

1. Pembelajaran(*Learning*)

1. Bentuk kosakata pada setiap dokumen data pelatihan.

2. Untuk setiap kategori *vj* :

   a. Tentukan *Docs j* (himpunan dok. Dalam kategori *vj*).

   b.Hitung probabilitas pada setiap kategori *P(vj).*

   c. Hitung *P(wk|vj)* untuk setiap kata *wk* dalam kosakata pada kategori *vj*.

2

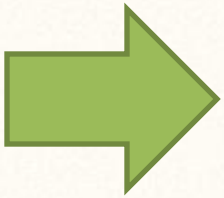
Klasifikasi(*Classify*)

1. Hitung $P(vj) \prod_{i} P(ai \mid vj)$
   untuk setiap kategori *vj*.

2. Tentukan kategori dengan nilai
   $P(vj) \prod_{k} P(ai \mid vj)$
   maksimalnya sebagai hasil kategorisasi.

ANALISIS KEBUTUHAN DATA

DATA LATIH



No	Nama Kategori dari 9 karakter	Nama Data Latih Objek Wisata
1.	Karakter Hutan	1. Taman Nasional Gunung Pangrango. 2. Kawah Putih.
2.	Karakter Perkebunan.	1. Perkebunan Teh Gunung Mas. 2. Perkebunan Kopi pangalengan.
3.	Karakter Pantai.	1. Pantai Cimaja Sukabumi.
4.	Karakter Unik.	1. Tambang Emas Pongkor.
5.	Karakter Tradisi & Komunikasi Adat.	1. Kampung Naga.
6.	Karakter Kepurbakalaan.	1. Gua Pawon Cipatat 2. Situs Gunung Padang
7.	Karakter Gelar Seni dan Ragam Festival.	1. Kampung adat cigugur. 2. Atraksi seni domba garut.
8.	Karakter Heritage.	1. Keraton Kesepuhan Cirebon. 2. Gedung Sate Jawa Barat.
9.	Karakter Wisata Belanja.	1. Pasar Baru Bandung 2. Wisata Belanja Tajur

DESKRIPSI UMUM

SISTEM YANG AKAN DIBANGUN

1. **Klasifikasi *Naïve Bayes Classifier (NVB)* Pembelajaran *(Learning)***

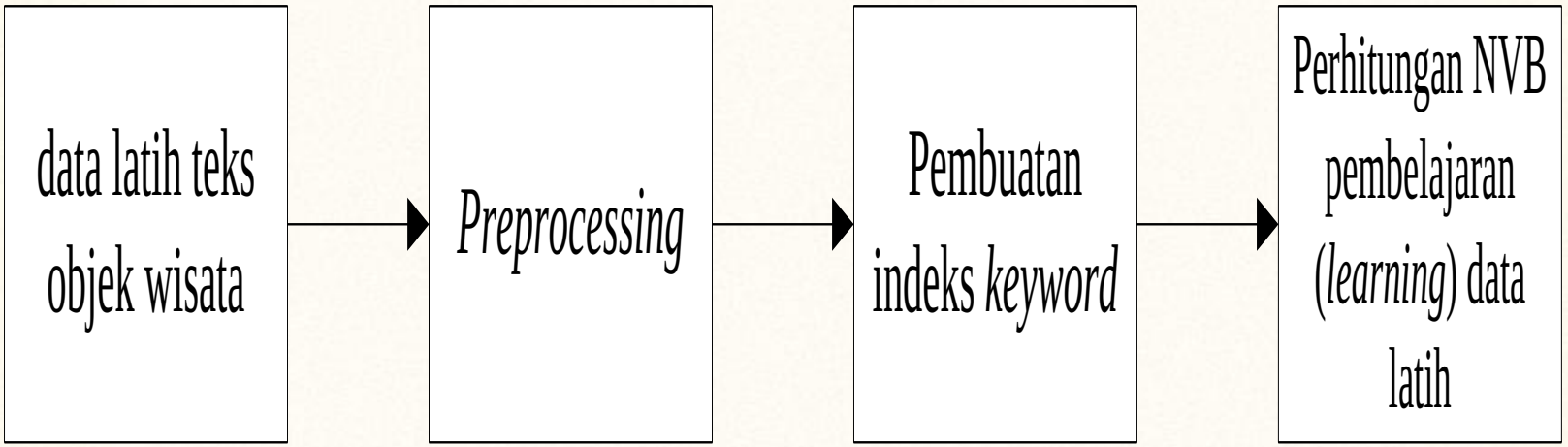
   Untuk data latih objek wisata

2. **Klasifikasi *Naïve Bayes Classifier (NVB)* Untuk Data Uji Objek Wisata**

1

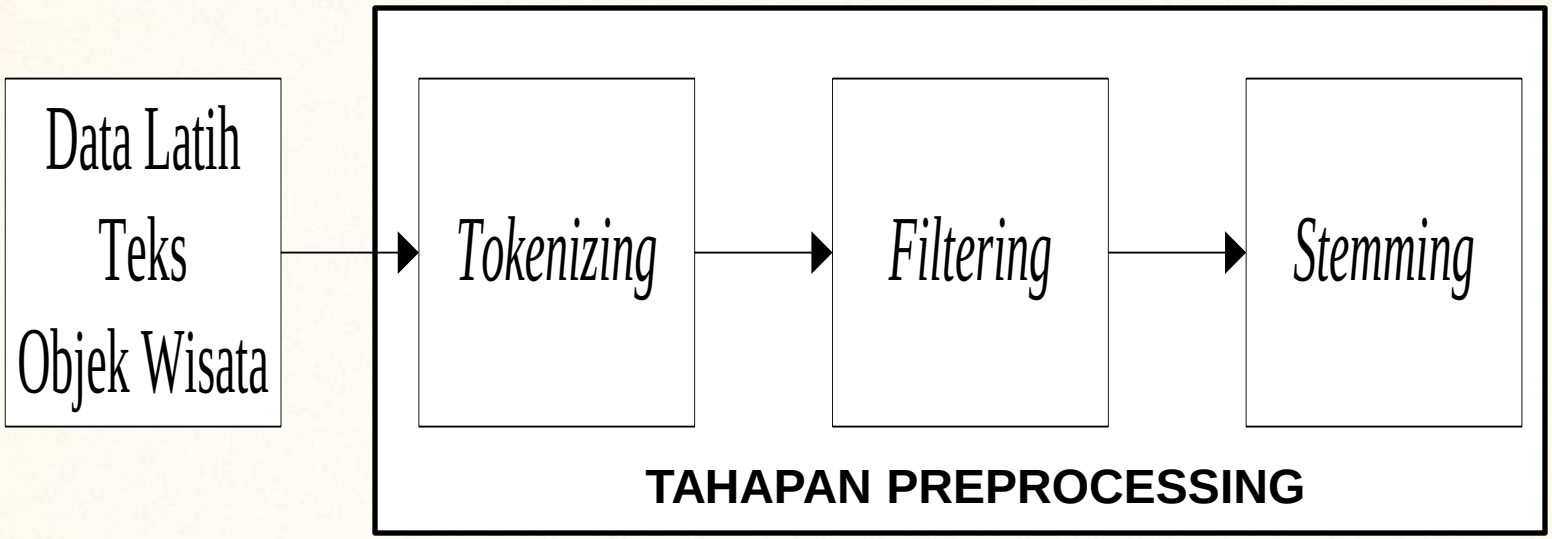
Klasifikasi *Naïve Bayes Classifier (NVB)* Pembelajaran *(Learning)*

Untuk data latih objek wisata



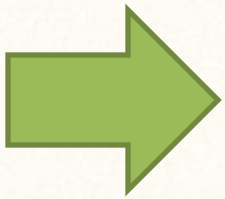
Gambar Proses Pembelajaran Data Latih Objek Wisata

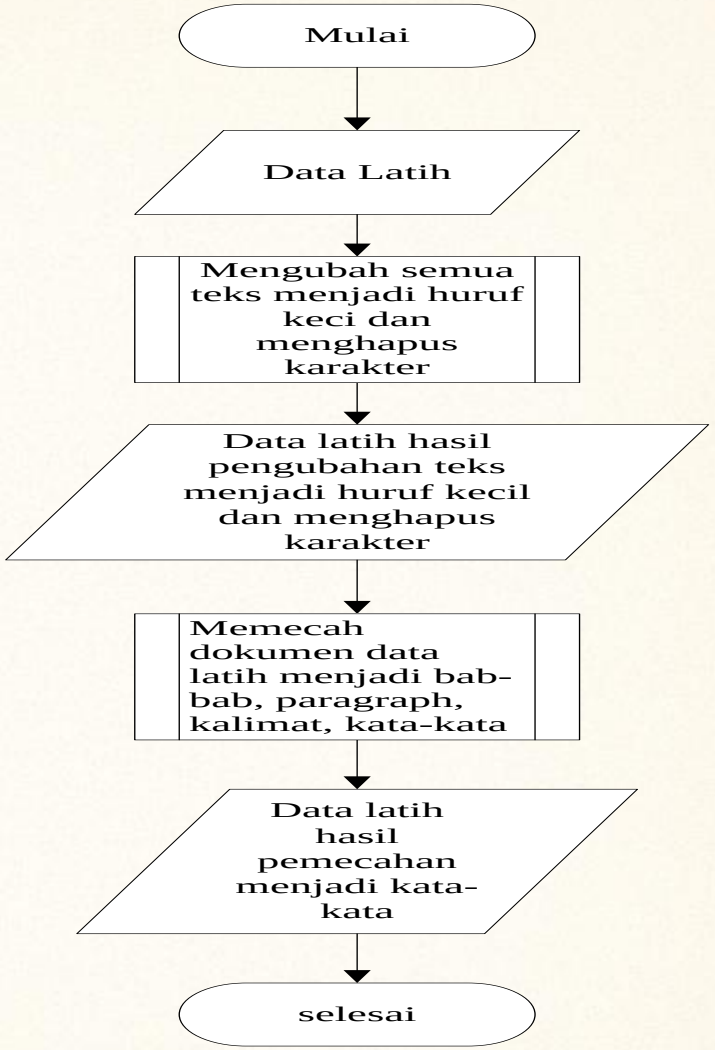
ANALISIS DATA LATIH PROSES TEXT MINING



Gambar Konsep Kerja Proses *Text mining*

A. TOKENIZING





ATURAN TOKENIZING

Tabel Aturan tahapan *Tokenizing* tahap pertama

ATURAN	HASIL
Jika inputan data latih memiliki huruf kapital [A…..Z] dan karakter simbol.	Maka akan mengubah data latih tersebut menjadi huruf kecil [a……z] semua dan membuang karakter simbol.

Tabel Daftar Simbol

Karakter Simbol		
^	,	:
@	(	;
%	)	|
$	-	]
#	<	}
‘	>	{
~	+	\
&	=	/
.	!	“

Tabel Ilustrasi *Tokenizing* tahap pertama

Deskripsi Data Latih		**Tahapan *Tokenizing* Mengubah semua teks data latih menjadi huruf kecil**	
Input	Kawasan hutan tropis di lereng Gunung Pangrango dan Gunung Gede. Gunung Gede juga memiliki keanekaragaman ekosistem.	Output	kawasan hutan tropis di lereng gunung pangrango dan gunung gede gunung gede juga memiliki keanekaragaman ekosistem

ATURAN TOKENIZING(2)

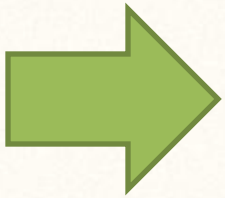
Tabel Aturan tahapan *Tokenizing* tahap kedua

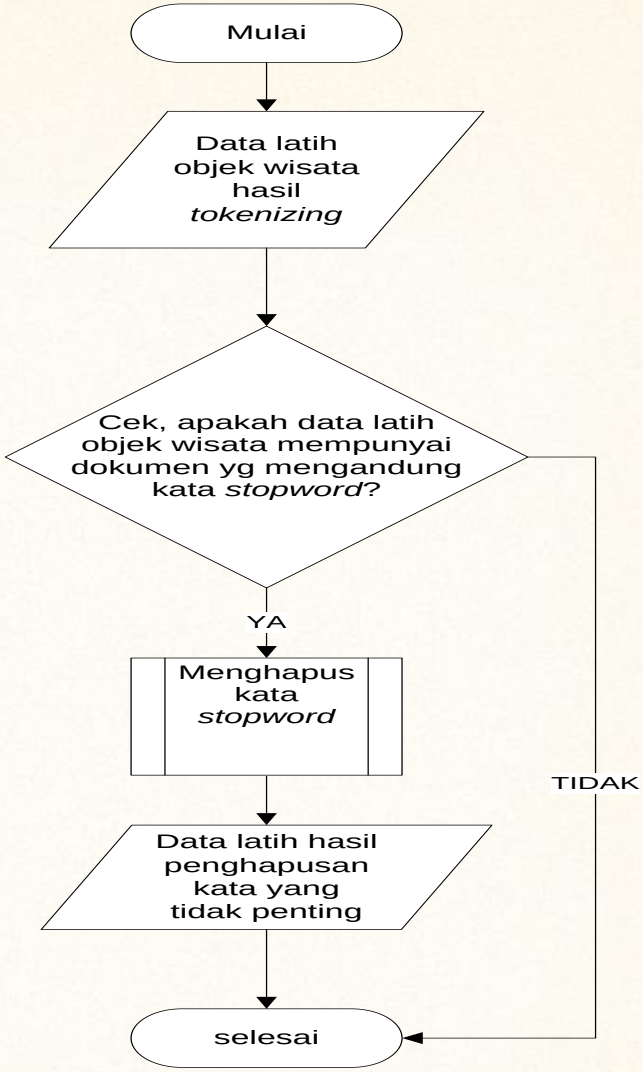
ATURAN	HASIL
Jika inputan data latih bertemu spasi.	Maka akan memecah dari deskripsi data latih menjadi bab-bab per bagian kata atau *string*.

Tabel Ilustrasi *Tokenizing* tahap pertama

Deskripsi Data Latih hasil pengubahan teks menjadi huruf kecil semua		**Tahapan *Tokenizing* Memecah dokumen data latih menjadi bab-bab, paragraf**		
Input	kawasan hutan tropis di lereng gunung pangrango dan gunung gede gunung gede juga memiliki keanekaragaman ekosistem	Output	kawasan hutan tropis di lereng gunung pangrango dan	gunung gede gunung gede juga memiliki keanekaragaman ekosistem

B. FILTERING




Mulai
Data latih objek wisata hasil *tokenizing*
Cek, apakah data latih objek wisata mempunyai dokumen yg mengandung kata *stopword*?
YA
TIDAK
Menghapus kata *stopword*
Data latih hasil penghapusan kata yang tidak penting
selesai

ATURAN FILTERING

Tabel Aturan tahapan *Filtering*

ATURAN	HASIL
Jika Inputan data latih mengandung kata pada *database stopword*	Maka akan menghapus kata atau *string* dalam data latih.

Tabel Daftar Stopword

Data *Stopword*			
telah	Punya	mendapatkan	dari
untuk	Setiap	pernah	lain
baik	Memang	lakukan	melakukannya
bisa	Tetapi	sudah	karena
jadi	Seperti	ada	antara
juga	Hampir	semua	setelah
di	tentang	mampu	yang
memiliki	dia	miliknya	bagaimana
bagaimanapun	jika	ke	dalam

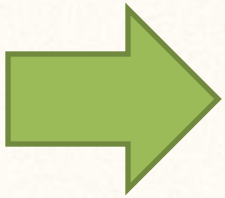
Tabel Ilustrasi Filtering

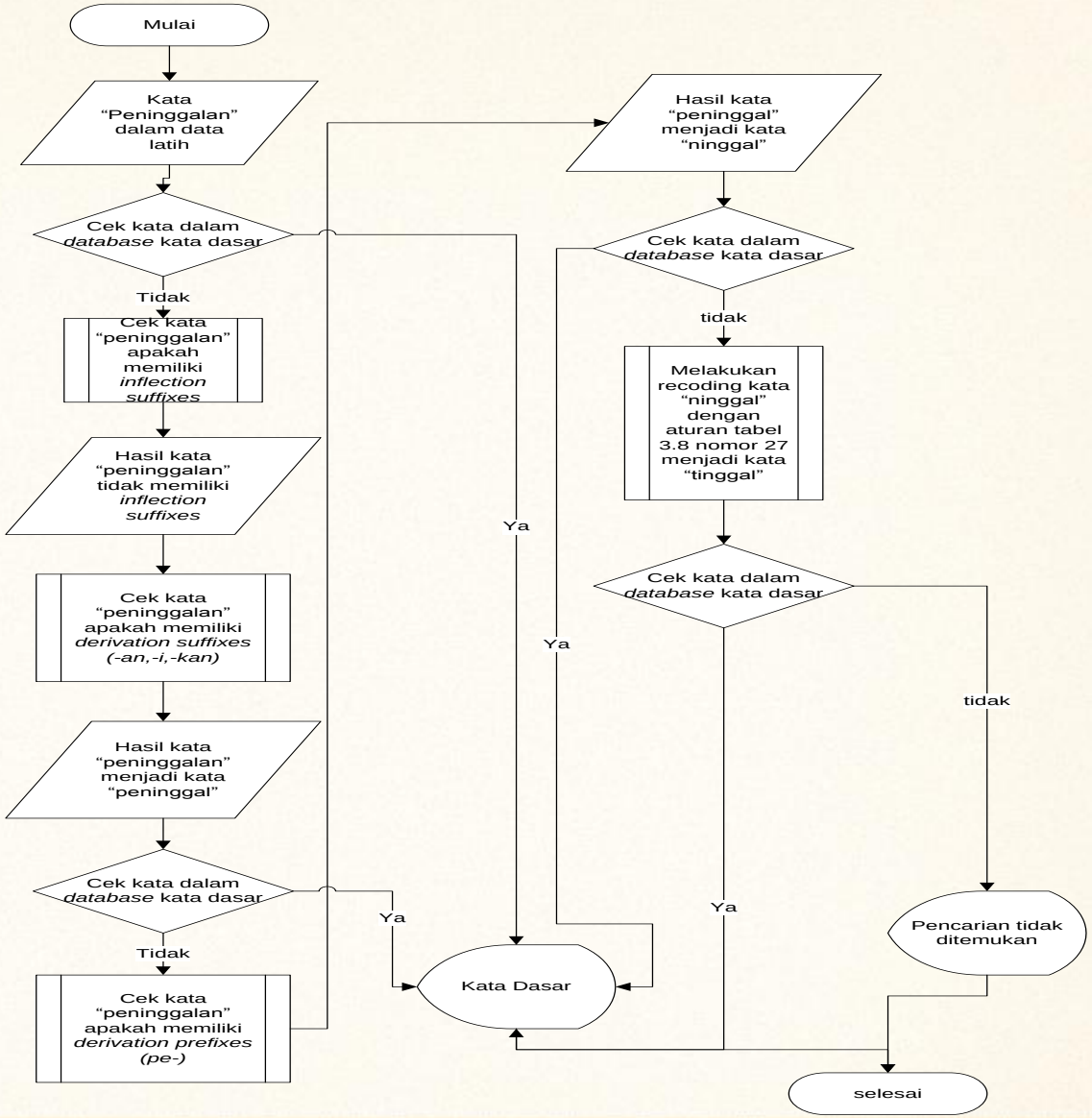
No	Deskripsi Data Latih Hasil *Tokenizing*		Apakah mengandung kata-kata *Stopword*	Deskripsi Data Latih Hasil *Filtering*	
D1	kawasan hutan tropis di lereng gunung pangrango dan	gunung gede gunung gede juga memiliki keanekaragaman ekosistem	telah punya dari untuk setiap pernah lain baik	kawasan hutan tropis lereng gunung pangrango gunung	gede gunung gede memiliki keanekaragaman ekosistem
D2	kawasan lahan yang cukup luas kawasan kebun teh ini	karpet hijau yang menutupi lereng lereng pegunungan di kawasan	memang lakukan bisa mendapatkan tetapi sudah karena jadi seperti	kawasan lahan luas kawasan kebun teh menghampar bagaikan	karpet hijau menutupi lereng lereng pegunungan kawasan puncak
	menghampar bagaikan	puncak	ada antara juga hampir semua setelah di tentang mampu yang memiliki dia miliknya bagaimana jika ke dalam itu sama paling biarkan mungkin aku sebagian besar harus saya sering pada hanya atau kita sendiri agak kata begitu mereka kemudian		
D3	peninggalan bangunan pada masa pemerintahan belanda bangunan kuno ini merupakan	bangunan pertama di bandung yang menggunakan konstruksi beton bertulang		peninggalan bangunan masa pemerintahan belanda bangunan	kuno bangunan bandung konstruksi beton bertulang
D4	pantai yang memiliki perpaduan antara pantai curam dan pantai landai pantai laut	selatan yang memiliki gelombang tinggi ini menjadi surga bagi para selancar		pantai perpaduan pantai curam pantai landai pantai	laut selatan gelombang tinggi surga selancar
D5	pertambangan emas pongkor terkenal akan kandungan emas wisatawan dapat menyaksikan suasana tambang bawah	tanah seakan merasakan sendiri para pekerja tambang saat bekerja di kedalaman perut bumi		pertambangan emas pongkor terkenal kandungan emas wisatawan menyaksikan suasana tambang	tanah merasakan pekerja tambang bekerja kedalaman perut bumi

Tabel Ilustrasi Filtering(2)

D6	perkampunga n yang dihuni sekelompok masyarakat yang masih	memegang adat istiadat peninggalan leluhurnya	ini sungguh keinginan adalah kami apa kapan mana	perkampungan dihuni sekelompok masyarakat	memegang adat istiadat peninggalan leluhurnya
D7	disini ditemukan artefak berupa alat alat budaya yang dimiliki oleh manusia praneolitik yang hidupnya mulai menetap	di gua gua dari situs ini ditemukan peninggalan historis arkeologis berbagai bentuk artefak fitur maupun ekofak	sementara siapa akan dengan belum sebuah atas menurut sesuai seberang sebenarnya sekali lagi terhadap sendirian bersama	ditemukan artefak berupa alat alat budaya dimiliki manusia praneolitik hidupnya menetap	gua gua situs ditemukan peninggalan historis arkeologis berbagai bentuk artefak fitur ekofak
D8	tradisi perayaan pokok seren tahun ditandai dengan ritual pokok ngagondang atau menumbuk padi menggunakan alu	pada lesung diiringi nyanyian ritual ngagondang ini berangkat dari kepercayaan lama masyarkat sunda yang agraris	meskipun selalu apapun saja selain muncul menghargai tepat sekitar samping terkait tersedia sangat menjadi sebab	tradisi perayaan pokok seren tahun ditandai ritual pokok ngagondang menumbuk padi alu	lesung diiringi nyanyian ritual ngagondang berangkat kepercayaan lama masyarkat sunda agraris
D9	aneka hasil kerajinan seperti tikar tas hingga topi menjadi ciri khas	tangan dari rajapolah salah satu sentra kerajinan anyaman di tasikmalaya	sebelum belakang percaya bawah terbalik lebih dating penyebab tertentu perubahan jelas	aneka hasil kerajinan seperti tikar tas hingga topi menjadi ciri khas	tangan rajapolah salah satu sentra kerajinan anyaman tasikmalaya

C. STEMMING




Mulai
Kata "Peninggalan" dalam data latih
Cek kata dalam database kata dasar
Ya
Tidak
Cek kata "peninggalan" apakah memiliki inflection suffixes
Hasil kata "peninggalan" tidak memiliki inflection suffixes
Cek kata "peninggalan" apakah memiliki derivation suffixes (-an,-i,-kan)
Hasil kata "peninggalan" menjadi kata "peninggal"
Cek kata dalam database kata dasar
Ya
Tidak
Cek kata "peninggalan" apakah memiliki derivation prefixes (pe-)
Hasil kata "peninggal" menjadi kata "ninggal"
Cek kata dalam database kata dasar
Ya
tidak
Melakukan recoding kata "ninggal" dengan aturan tabel 3.8 nomor 27 menjadi kata "tinggal"
Cek kata dalam database kata dasar
Ya
tidak
Kata Dasar
Pencarian tidak ditemukan
selesai

ATURAN STEMMING

Tabel Kombinasi Awalan Akhiran yang dilarang

Awalan (*prefix*)	Akhiran (*suffix*) yang dilarang
be-	-i
di-	-an
ke-	-i, -kan
me-	-an
se-	-i, -kan
te-	-an

Tabel Aturan Pemenggalan Kata

Aturan	Format Kata	Pemenggalan
1	berV…	ber-V... | be-rV...
2	berCAP…	ber-CAP... dimana C!=’r’ & P!=’er’
3	berCAerV...	ber-CaerV... dimana C!=’r’
4	belajar	bel-ajar
5	beC1erC2...	be-C1erC2... dimana C1!={’r’|’l’}
6	terV...	ter-V... | te-rV...
7	terCerV...	ter-CerV... dimana C!=’r’
8	terCP...	ter-CP... dimana C!=’r’ dan P!=’er’
9	teC1erC2...	te-C1erC2... dimana C1!=’r’
10	me{l|r|w|y}V...	me-{l|r|w|y}V...
11	mem{b|f|v}...	mem-{b|f|v}...
12	mempe...	mem-pe...
13	mem{rV|V}...	me-m{rV|V}... |me-p{rV|V}…
14	men{c|d|j|z}...	men-{c|d|j|z}...
15	menV...	me-nV... | me-tV
16	meng{g|h|q|k}...	meng-{g|h|q|k}...
17	mengV...	meng-V... | meng-kV...
18	menyV...	meny-sV...
19	mempV...	mem-pV... dengan V!=’e’
20	pe{w|y}V...	pe-{w|y}V...
21	perV...	per-V... | pe-rV...
22	perCAP	per-CAP... dimana C!=’r’ dan P!=’er’
23	perCAerV...	per-CAerV... dimana C!=’r’
24	pem{b|f|V}...	pem-{b|f|V}...
25	pem{rV|V}...	pe-m{rV|V}... | pe-p{rV|V}…
26	pen{c|d|j|z}...	pen-{c|d|j|z}...
27	penV...	pe-nV... | pe-tV...
28	peng{g|h|q}...	peng-{g|h|q}...
29	pengV...	peng-V... | peng-kV...
30	penyV...	peny-sV...
31	pelV...	pe-lV... kecuali “pelajar” yang menghasilkan “ajar”
32	peCerV...	per-erV... dimana C!={r|w|y|l|m|n}
33	peCP...	pe-CP... dimana C!={r|w|y|l|m|n} dan P!=’er’
34	terC1erC2...	ter-C1erC2... dimana C1!=’r’
35	peC1erC2...	pe-C1erC2... dimana C1!={r|w|y|l|m|n}

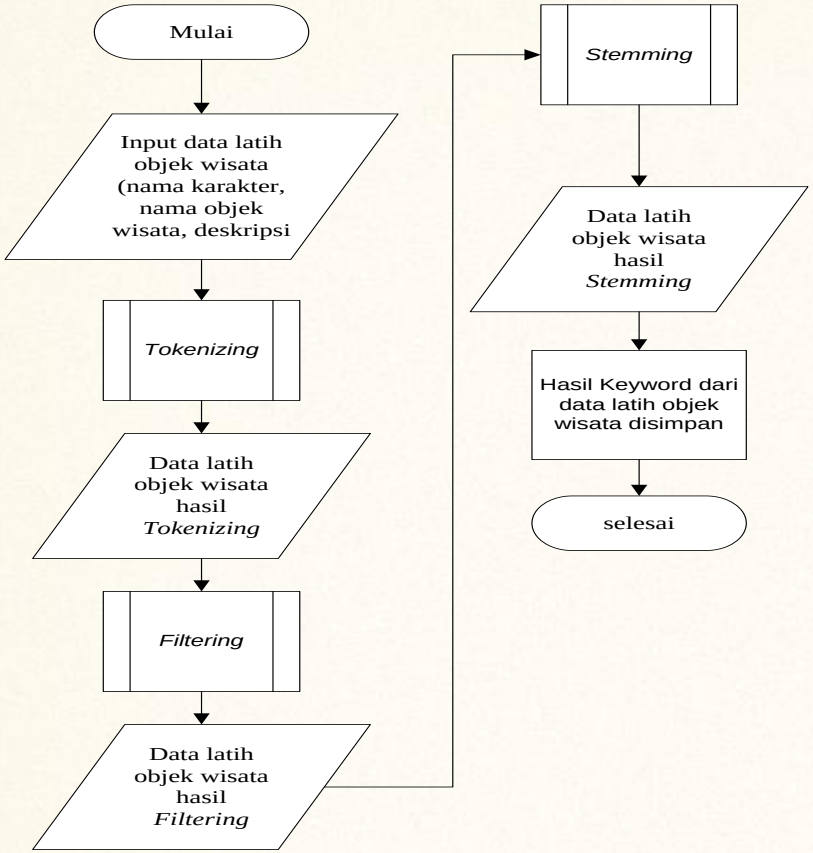
Tabel Ilustrasi Stemming

NO	Dokumen Data Latih Hasil *Filtering*		Dokumen Data Latih Hasil *Stemming*	
D1	kawasan hutan tropis lereng gunung pangrango	gunung gede keanekaragaman ekosistem	kawasan hutan tropis lereng gunung	pangrango gunung gede keanekaragaman ekosistem
D2	kawasan lahan luas kawasan kebun teh menghampar bagaikan	karpet hijau menutupi lereng lereng pegunungan kawasan puncak	kawasan lahan luas kawasan kebun teh hampar bagai	karpet hijau tutup lereng lereng gunung kawasan puncak
D3	peninggalan bangunan masa pemerintahan belanda bangunan	kuno bangunan bandung konstruksi beton bertulang	tinggal bangun masa pemerintah belanda bangun	kuno bangun bandung konstruksi beton bertulang
D4	pantai perpaduan pantai curam pantai landai pantai laut	selatan gelombang tinggi surga selancar	pantai padu pantai curam pantai landai pantai	laut selatan gelombang tinggi surga selancar
D5	pertambangan emas pongkor terkenal kandungan	tanah akan merasakan pekerja tambang	tambang emas pongkor kenal kandung	tanah akan rasa kerja tambang

Tabel Ilustrasi Stemming(2)

	emas wisatawan menyaksikan suasana tambang	bekerja kedalaman perut bumi	emas wisatawan saksi suasana tambang	kerja dalam perut bumi
D6	perkampungan dihuni masyarakat	memegang adat istiadat peninggalan leluhurnya	kampung huni masyarakat	pegang adat istiadat tinggal leluhur
D7	ditemukan artefak berupa alat alat budaya dimiliki manusia prasejarah hidupnya menetap	gua gua situs ditemukan peninggalan historis arkeologis berbagai bentuk artefak fitur ekofak	temu artefak rupa alat alat budaya milik manusia prasejarah hidup tetap	gua gua situs temu tinggal historis arkeologis bentuk artefak fitur ekofak
D8	tradisi perayaan pokok seren ditandai ritual pokok ngagondang memumbuk padi alu	lesung diiringi nyanyian ritual ngagondang berangkat kepercayaan lama masyarkat sunda agraris	tradisi raya pokok seren tanda ritual pokok ngagondang tumbuk padi alu	lesung iringi nyanyi ritual ngagondang berangkat percaya lama masyarkat sunda agraris
D9	aneka hasil kerajinan tikar tas hingga topi ciri khas buah	tangan rajapolah salah satu sentra kerajinan anyaman tasikmalaya	aneka hasil rajin tikar tas hingga topi ciri khas buah	tangan rajapolah salah satu sentra rajin anyaman tasikmalaya

Pembuatan Indeks Keyword



Gambar Flowchart pembuatan indeks keyword

Aturan

- Data latih langsung diketik di form Penambahan atau Pengubahan Data latih atau menyalin melalui media lain, seperti Notepad.

- Jika ada poin-poin dalam deskripsi gunakan tanda " - " sebagai tanda poin dalam kalimat.

Perhitungan Klasifikasi *Naïve Bayes Classifier (NVB)* Pembelajaran *(Learning)* Untuk data latih objek wisata

a. **Bentuk kosakata pada setiap dokumen data latih objek wisata**

b. **Untuk setiap kategori *vj* :**
   1. **Tentukan *Docs j* (himpunan dok. Dalam kategori *vj*) dan hitung Probabilitas *P(vj)….***
   2. **Hitung *P(wk|vj)* untuk setiap kata *wk* dalam kosakata pada kategori *vj***

a. Bentuk kosakata pada setiap dokumen data latih objek wisata

Tabel Himpunan Data Latihan(1)

Dok	Data Latih Hasil proses *preprocessing text mining* Menjadi keyword		kemunculan *keyword*	Karakter Objek Wisata (y)
D1	kawasan hutan tropis lereng gunung	pangrango gunung gede keanekaragaman ekosistem	kawasan(1),hutan(1),tropis (1),lereng(1),gunung(3), pangrango(1),gede(2), keanekaragaman(1), ekosistem(1).	Karakter Hutan
D2	kawasan lahan luas kawasan kebun teh hampar bagai	karpet hijau tutup lereng lereng gunung kawasan puncak	kawasan(3),lahan(1),luas(1), kebun(1),teh(1),hampar(1), bagai(1),karpet(1),hijau(1), tutup(1), lereng(1),gunung(1), puncak(1).	Karakter Perkebunan
D3	tinggal bangun masa pemerintah belanda bangun	kuno bangun bandung konstruksi beton bertulang	tinggal(1),bangun(3),masa (1),Pemerintah(1),belanda(1) ,kuno(1),bandung(1), konstruksi(1),beton(1), bertulang(1).	Karakter Heritage
D4	pantai padu pantai curam pantai landai pantai	laut selatan gelombang tinggi surga selancar	pantai(5),padu(1),curam(1), landai(1),laut(1),selatan(1), gelombang(1),tinggi(1), surga(1),selancar(1)	Karakter Pantai
D5	tambang emas pongkor kenal kandung emas wisatawan saksi suasana tambang	tanah akan rasa kerja tambang kerja dalam perut bumi	tambang(3),emas(2),pongkor (1),kenal(1),kandung(1), wisatawan(1),saksi(1), suasana(1),tanah(1),akan(1), rasa(1),kerja(2),dalam (1),perut(1),bumi(1).	Karakter Unik

Tabel Himpunan Data Latihan(1)

D6	kampung huni masyarakat	pegang adat istiadat tinggal leluhur	kampung(1),huni(1), masyarakat (1),pegang(1),adat(1),istiadat (1),tinggal(1),leluhur(1).	Karakter Tradisi dan Komunikasi Adat
D7	temu artefak rupa alat alat budaya milik manusia prameolitik hidup tetap	gua gua situs temu tinggal historis arkeologis bentuk artefak fitur ekofak	temu(2),artefak(2),rupa(1), alat(2),budaya(1),milik(1), manusia(1),prameolitik(1), hidup(1),tetap(1),gua(2), situs(1),tinggal(1),historis(1) ,arkelogis(1),bentuk(1), fitur(1),ekofak(1).	Karakter kepurbakalaan
D8	tradisi raya pokok seren tanda ritual pokok ngagondang	lesung iringi nyanyi ritual ngagondang berangkat percaya lama	tradisi(1),raya(2),pokok(2), seren(1),tahun(1),tanda(1), ritual(2),ngagondang(2), tumbuk(1),padi(1),alu(1), lesung(1),iringi(1),nyanyi(1) ,berangkat(1),kepercayaan (1),lama(1),masyarakat(1), sunda(1),agraris(1).	Karakter Gelar seni dan Ragam Festival
	tumbuk padi alu	masyarkat sunda agraris		
D9	aneka hasil rajin tikar tas hingga topi ciri khas buah	tangan rajapolah salah satu sentra rajin anyaman tasikmalaya	aneka(1),hasil(2),rajin(2), tikar(1),tas(1), hingga(1),topi(1),ciri(1), khas(1),buah(1), tangan(1),rajapolah(1),salah (1),satu(1),sentra(1), anyaman(1),tasikmalaya(1).	Karakter Wisata Belanja

b. Untuk setiap kategori *vj* :

1. Tentukan *Docs j* (himpunan dok. Dalam kategori *vj*) dan hitung Probabilitas *P(vj)….*

Dengan persamaan :

$$P(vj) = \frac{|Docs\ j|}{|Contoh|} \quad (1)$$

dimana :

- $P(vj)$: menghitung probabilitas setiap dokumen terhadap sekumpulan dokumen.
- $|Docs\ j|$: banyaknya dokumen yang memiliki kategori j dalam pelatihan,
- $|Contoh|$: banyaknya dokumen dalam contoh yang digunakan saat pelatihan.

Tabel Nilai *P(vj)* untuk setiap kategori

Dokumen	kemunculan Keyword	Total |Kosakata|	Karakter Objek Wisata (vj)	$P(vj) = \frac{\|docs\ j\|}{\|Contoh\|}$
D1	kawasan(1),hutan(1), tropis(1),lereng(1), gunung(3),pangrango(1),gede(2), keanekaragaman(1), ekosistem(1).	12	Karakter Hutan	$\frac{1}{9}$
D2	kawasan(3), lahan(1), luas(1), kebun(1), teh(1),hampar(1), bagai(1),karpet(1), hijau(1),tutup(1), lereng(2),gunung(1), puncak(1).	16	Karakter Perkebunan	$\frac{1}{9}$
D3	tinggal(1),bangun(3), masa(1),pemerintah(1), belanda(1),kuno(1), bandung(1),konstruksi (1),beton(1),tulang(1).	12	Karakter Heritage	$\frac{1}{9}$
D4	pantai(4),padu(1), curam(1),landai(1), laut(1),selatan(1), gelombang(1),tinggi(1), surga(1),selancar(1)	13	Karakter Pantai	$\frac{1}{9}$
D5	tambang(3),emas(2), pongkor(1),kenal(1), kandung(1),wisatawan (1),saksi(1),akan(1), suasana(1),tanah(1), rasa(1),kerja(2), dalam(1),perut(1), bumi(1).	19	Karakter Unik	$\frac{1}{9}$

Tabel Nilai *P(vj)* untuk setiap kategori(2)

D6	kampung(1),huni(1), masyarakat(1),pegang (1),adat(1),istiadat(1), tinggal(1),leluhur(1).	8	Karakter Tradisi dan Komunikasi Adat	$\frac{1}{9}$
D7	temu(2),artefak(2), rupa(1),alat(2), budaya(1),milik(1), manusia(1), praneolitik(1),hidup(1), tetap(1),gua(2), situs(1),tinggal(1), historis(1),arkelogis(1), bentuk(1),fitur(1), ekofak(1).	18	Karakter kepurbakalaan	$\frac{1}{9}$
D8	tradisi(1),raya(1), pokok(2),seren(1), tanda(1),ritual(2), ngagondang(2), tumbuk(1),padi(1),alu (1),lesung(1),iringi(1), nyanyi(1),angkat(1), percaya(1),lama(1), masyarakat(1),sunda(1), agraris(1).	18	Karakter Gelar seni dan Ragam Festival	$\frac{1}{9}$
D9	aneka(1),hasil(2),rajin (2),tikar(1),tas(1), hingga(1),topi(1),ciri(1),kh as(1),buah(1), tangan(1),rajapolah(1), salah(1),satu(1), sentra(1),anyaman(1), tasikmalaya(1).	19	Karakter Wisata Belanja	$\frac{1}{9}$

1. **Hitung *P(wk|vj)* untuk setiap kata *wk* dalam kosakata pada kategori *vj***

$$P(wk \mid vj) = \frac{nk+1}{n+|Kosakata|} \qquad (2)$$

Dimana :

$P(wk \mid vj)$: Probabilitas kemunculan kata *wk* pada suatu dokumen jika diketahui dokumen tersebut berkategori *vj*.

nk : frekuensi munculnya kata *wk* dalam dokumen yang berkategori *vj*.

n : banyaknya seluruh kata dalam dokumen berkategori *vj*.

|**Kosakata**| : banyaknya kata dalam contoh pelatihan.

Pada persamaan 2 terdapat suatu penambahan 1 pada pembilang, hal ini dilakukan untuk mengantisipasi jika terdapat suatu kata pada dokumen uji yang tidak ada pada setiap dokumen data pelatihan.

Tabel Hasil *Training* Model Probabilitas dari data latih

Nama Doc	No	Kata	Dokumen $P(wk \mid vj) = \frac{nk+1}{n+ \mid Kosakata \mid}$								
			K1	K2	K3	K4	K5	K6	K7	K8	K9
D1	1	kawasan	2/126	4/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	2	hutan	2/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	3	tropis	2/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	4	lereng	2/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	5	gunung	4/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	6	pangrango	2/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	7	gede	3/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	8	keanekaragaman	2/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	9	ekosistem	2/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
D2	10	lahan	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	11	luas	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	12	kebun	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	13	teh	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	14	hampar	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	15	bagai	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	16	karpet	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	17	hijau	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	18	tutup	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
	19	puncak	1/126	2/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
D3	20	tinggal	1/126	1/130	2/126	1/127	1/133	1/122	1/132	1/132	1/133
	21	bangun	1/126	1/130	4/126	1/127	1/133	1/122	1/132	1/132	1/133
	22	masa	1/126	1/130	2/126	1/127	1/133	1/122	1/132	1/132	1/133
	23	perintah	1/126	1/130	2/126	1/127	1/133	1/122	1/132	1/132	1/133
	24	belanda	1/126	1/130	2/126	1/127	1/133	1/122	1/132	1/132	1/133
	25	kuno	1/126	1/130	2/126	1/127	1/133	1/122	1/132	1/132	1/133
	26	bandung	1/126	1/130	2/126	1/127	1/133	1/122	1/132	1/132	1/133
	27	konstruksi	1/126	1/130	2/126	1/127	1/133	1/122	1/132	1/132	1/133
	28	beton	1/126	1/130	2/126	1/127	1/133	1/122	1/132	1/132	1/133
	29	tulang	1/126	1/130	2/126	1/127	1/133	1/122	1/132	1/132	1/133
D4	30	pantai	1/126	1/130	1/126	5/127	1/133	1/122	1/132	1/132	1/133
	31	padu	1/126	1/130	1/126	2/127	1/133	1/122	1/132	1/132	1/133
	32	curam	1/126	1/130	1/126	2/127	1/133	1/122	1/132	1/132	1/133
	33	landai	1/126	1/130	1/126	2/127	1/133	1/122	1/132	1/132	1/133
	34	laut	1/126	1/130	1/126	2/127	1/133	1/122	1/132	1/132	1/133
	35	selatan	1/126	1/130	1/126	2/127	1/133	1/122	1/132	1/132	1/133
	36	gelombang	1/126	1/130	1/126	2/127	1/133	1/122	1/132	1/132	1/133
	37	tinggi	1/126	1/130	1/126	2/127	1/133	1/122	1/132	1/132	1/133
	38	surga	1/126	1/130	1/126	2/127	1/133	1/122	1/132	1/132	1/133
	39	selancar	1/126	1/130	1/126	2/127	1/133	1/122	1/132	1/132	1/133

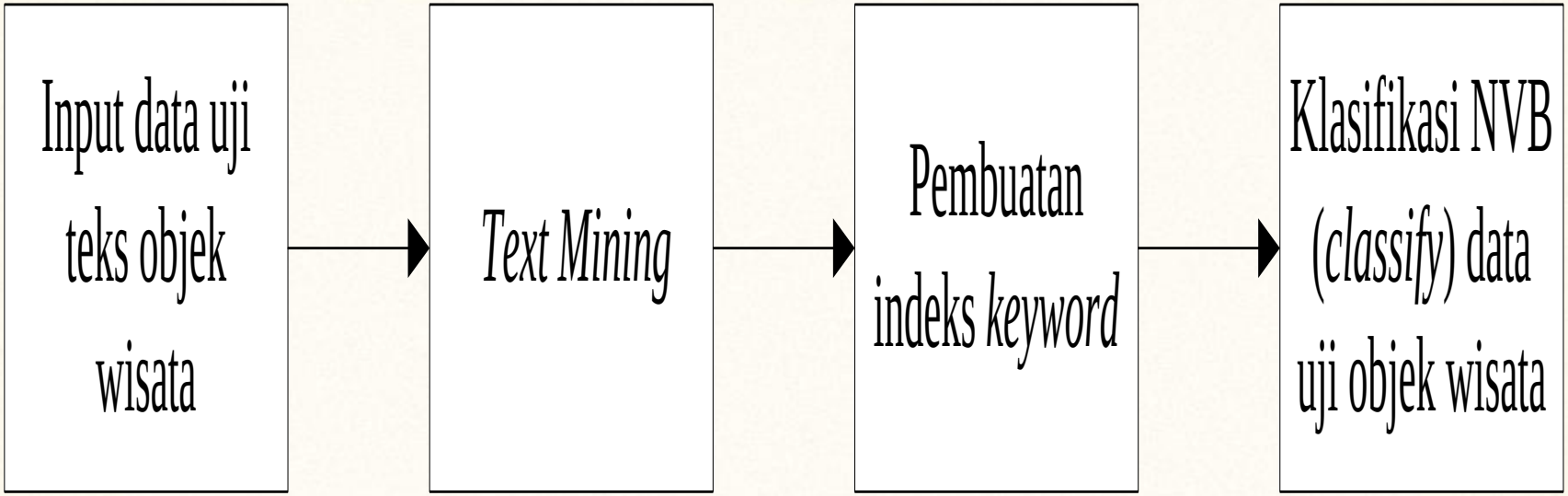
Tabel Hasil *Training* Model Probabilitas dari data latih(2)

D5	40	tambang	1/126	1/130	1/126	1/127	4/133	1/122	1/132	1/132	1/133
	41	emas	1/126	1/130	1/126	1/127	3/133	1/122	1/132	1/132	1/133
	42	pongkor	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	43	kenal	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	44	kandung	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	45	wisatawan	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	46	saksi	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	47	suasana	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	48	tanah	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	49	akan	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	50	rasa	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	51	kerja	1/126	1/130	1/126	1/127	3/133	1/122	1/132	1/132	1/133
	52	dalam	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	53	perut	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
	54	bumi	1/126	1/130	1/126	1/127	2/133	1/122	1/132	1/132	1/133
D6	55	kampung	1/126	1/130	1/126	1/127	1/133	2/122	1/132	1/132	1/133
	56	huni	1/126	1/130	1/126	1/127	1/133	2/122	1/132	1/132	1/133
	57	masyarakat	1/126	1/130	1/126	1/127	1/133	2/122	1/132	1/132	1/133
	58	pegang	1/126	1/130	1/126	1/127	1/133	2/122	1/132	1/132	1/133
	59	adat	1/126	1/130	1/126	1/127	1/133	2/122	1/132	1/132	1/133
	60	istiadat	1/126	1/130	1/126	1/127	1/133	2/122	1/132	1/132	1/133
	61	leluhur	1/126	1/130	1/126	1/127	1/133	2/122	1/132	1/132	1/133
D7	62	temu	1/126	1/130	1/126	1/127	1/133	1/122	3/132	1/132	1/133
	63	artefak	1/126	1/130	1/126	1/127	1/133	1/122	3/132	1/132	1/133
	64	rupa	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	65	alat	1/126	1/130	1/126	1/127	1/133	1/122	3/132	1/132	1/133
	66	budaya	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	67	milik	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	68	manusia	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	69	preneolitik	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	70	hidup	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	71	tetap	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	72	gua	1/126	1/130	1/126	1/127	1/133	1/122	3/132	1/132	1/133
	73	situs	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	74	historis	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	75	arkeologis	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133

Tabel Hasil *Training* Model Probabilitas dari data latih(3)

	76	bentuk	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	77	fitur	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
	78	ekofak	1/126	1/130	1/126	1/127	1/133	1/122	2/132	1/132	1/133
D8	79	tradisi	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	80	raya	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	81	pokok	1/126	1/130	1/126	1/127	1/133	1/122	1/132	3/132	1/133
	82	seren	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	83	tanda	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	84	ritual	1/126	1/130	1/126	1/127	1/133	1/122	1/132	3/132	1/133
	85	nggagondang	1/126	1/130	1/126	1/127	1/133	1/122	1/132	3/132	1/133
	86	tumbuk	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	87	padi	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	88	alu	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	89	lesung	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	90	iring	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	91	nyanyi	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	92	angkat	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	93	percaya	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	94	lama	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	95	masyarakat	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	96	sunda	1/126	1/130	1/126	1/127	1/133	1/122	1/132	2/132	1/133
	97	agraris	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	1/133
D9	98	aneka	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	99	hasil	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	100	rajin	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	3/133
	101	tikar	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	102	tas	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	103	hingga	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	104	topi	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	105	ciri	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	106	khas	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	107	buah	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	108	tangan	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	109	rajapolah	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	110	salah	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	111	satu	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	112	sentra	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	113	anyaman	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
	114	tasikmalaya	1/126	1/130	1/126	1/127	1/133	1/122	1/132	1/132	2/133
Total Seluruh Kata Dalam Data latih (*n*)			**114**								

Klasifikasi *(Classify) Untuk Data Uji Objek Wisata*



Gambar Proses pengklasifikasian NVB (*Classify*) Data Uji Objek Wisata

1. ANALISIS DATA UJI OBJEK WISATA

Dokumen	Nama Objek Wisata	Lokasi	Nama Deskripsi
Data Uji	Gedung Sate	Bandung, Jawa Barat	Peninggalan bangunan kuno yang mempunyai nilai sejarah berupa

2. ANALISIS PROSES TEXT MINING PADA DATA UJI

Dokumen	Nama objek wisata	Lokasi	Deskripsi	Proses *Text Mining* tahapan *preprocessing*		
				Tokenizing	***Filtering***	***Stemming***
Data Uji	Gedung Sate	Bandung, Jawa Barat	Peninggalan bangunan kuno yang mempunyai nilai sejarah berupa	peninggalan bangunan kuno yang mempunyai nilai sejarah berupa	peninggalan bangunan kuno nilai sejarah berupa	tinggal bangun kuno sejarah rupa

3. ANALISIS PEMBUATAN INDEKS KEYWORD

Dokumen	Nama Objek Wisata	Lokasi	Inputan Deskripsi teks	Hasil Proses *Text Mining* Dibuat indeks *Keyword*
Data Uji	Gedung Sate	Bandung, Jawa Barat	Peninggalan bangunan kuno yang mempunyai nilai sejarah berupa	tinggal(1), bangun(1), kuno(1), nilai(1), sejarah(1), rupa(1)

4. ANALISIS PERHITUNGAN NVB UNTUK KLASIFIKASI DATA UJI

a. Bentuk Hitung $P(vj) \prod_i P(ai \mid vj)$ untuk setiap kategori *vj*.

b. Tentukan kategori dengan nilai $P(vj) \prod_i P(ai \mid vj)$ maksimalnya sebagai hasil kategorisasi.

a. Bentuk Hitung *P(vj) II P(ai | vj)* untuk setiap kategori *vj*.

i =

Dokumen	Nama Objek Wisata	Lokasi	Inputan Deskripsi teks	Hasil Proses *Text Mining* Dibuat indeks *Keyword*	Total Kosakata	Nama Karakter
Data Uji	Gedung Sate	Bandung, Jawa Barat	Peninggalan bangunan kuno yang mempunyai nilai sejarah berupa	tinggal(1), bangun(1), kuno(1), nilai(1), sejarah(1), rupa(1)	6	Belum diketahui

V_{MAP}

$$V_{MAP} = \arg\max_{vj=v} P(vj) \prod_{i} P(ai \mid vj)$$

Hasil Perhitungan NVB untuk klasifikasi data uji

$$V_{MAP} = \underset{vj \in \{\text{"hutan"},\text{heritage"},\text{"perkebunan"},\text{"pantai"},\text{"unik"},\text{"tradisi dan komunikasi adat"},\text{"kepurbakalaan"},\text{"gelar seni dan ragam festival"},\text{"wisata belanja"}\}}{\arg\max} P(vj) \prod_{i \in \text{positions}} P(ai|vj)$$

$$P(vj) \prod_{i} P(\text{"tinggal"} | vj) P(\text{"bangun"} | vj) P(\text{"kuno"} | vj) P(\text{"nilai"} | vj) P(\text{"sejarah"} | vj) P(\text{"rupa"} | vj)$$

- **Hitung klasifikasi *Naïve Bayes* untuk setiap kategori :**

1. nilai V_{MAP} untuk **Karakter Hutan**

   P("data uji" | "Hutan") = P("Hutan") x P("tinggal" | "Hutan") x P("bangun" | "Hutan") x P("kuno" | "Hutan") x P("rupa" | "Hutan") = 1/9 x 1/126 x 1/126 x 1/126 x 1/126 x 1/126 => $3,4986843710e^{-12}$

2. nilai V_{MAP} untuk **Karakter Perkebunan**

   P("data uji" | "Perkebunan") = P("Perkebunan") x P("tinggal" | "Perkebunan") x P("bangun" | "Perkebunan") x P("kuno" | "Perkebunan") x P("rupa" | "Perkebunan") = 1/9 x 1/130 x 1/130x 1/130 x 1/130x 1/130 => $2,9925452704e^{-13}$

3. nilai V_{MAP} untuk **Karakter Heritage**

   P("data uji" | "Heritage") = P("Heritage") x P("tinggal" | "Heritage") x P("bangun" | "Heritage") x P("kuno" | "Heritage") x P("rupa" | "Heritage") = 1/9 x 2/126 x 4/126 x 2/126 x 1/126 x 1/126 => $5,5978949936305e^{-11}$

4. nilai V_{MAP} untuk **Karakter Pantai**

P("data uji" | "Pantai") = P ("Pantai") x P("tinggal" | "Pantai") x P("bangun" | "Pantai") x P("kuno" | "Pantai") x P("rupa" | "Pantai")

= 1/9 x 1/127 x 1/127 x 1/127 x 1/127 => $3,3630930674e^{-12}$

5. nilai V_{MAP} untuk **Karakter Unik**

P("data uji" | "Unik") = P ("Unik") x P("tinggal" | "Unik") x P("bangun" | "Unik") x P("kuno" | "Unik") x P("rupa" | "Unik")

= 1/9 x 1/133 x 1/133 x 1/133 x 1/133 => $2,669926375956e^{-12}$

6. nilai V_{MAP} untuk **Karakter Tradisi dan Komunikasi adat**

P("data uji" | "Tradisi dan Komunikasi adat") = P ("Tradisi dan Komunikasi adat") x P("tinggal" | "Tradisi dan Komunikasi adat") x P("bangun" | "Tradisi dan Komunikasi adat") x P("kuno" | "Tradisi dan Komunikasi adat") x P("rupa" | "Tradisi dan Komunikasi adat")

= 1/9 x 1/122 x 1/122 x 1/122 x 1/122 => $4,111102805105e^{-12}$

7. nilai V_{MAP} untuk **Karakter Kepurbakalaan**

P ("data uji" | "Kepurbakalaan") = P ("Kepurbakalaan") x P("tinggal" | "Kepurbakalaan") x P("bangun" | "Kepurbakalaan") x P("kuno" | "Kepurbakalaan") x P("rupa" | "Kepurbakalaan")

= 1/9 x 1/136 x 1/136 x 2/136x 2/136 => $9,552629634924e^{-12}$

8. nilai V_{MAP} untuk **Karakter Gelar seni dan Ragam Festival**

P ("data uji" | "Gelar seni dan Ragam Festival") = P ("Gelar seni dan Ragam Festival") x P("tinggal" | "Gelar seni dan Ragam Festival") x P("bangun" | "Gelar seni dan Ragam Festival") x P("kuno" | "Gelar seni dan Ragam Festival") x P("rupa" | "Gelar seni dan Ragam Festival")

= 1/9 x 1/136 x 1/136 x 1/136x 1/136 => $2,38815740873e^{-12}$

9. nilai V_{MAP} untuk **Karakter Wisata Belanja**

P ("data uji" | "Wisata Belanja") = P ("Wisata Belanja") x P("tinggal" | "Wisata Belanja") x P("bangun" | "Wisata Belanja") x P("kuno" | "Wisata Belanja") x P("rupa" | "Wisata Belanja")

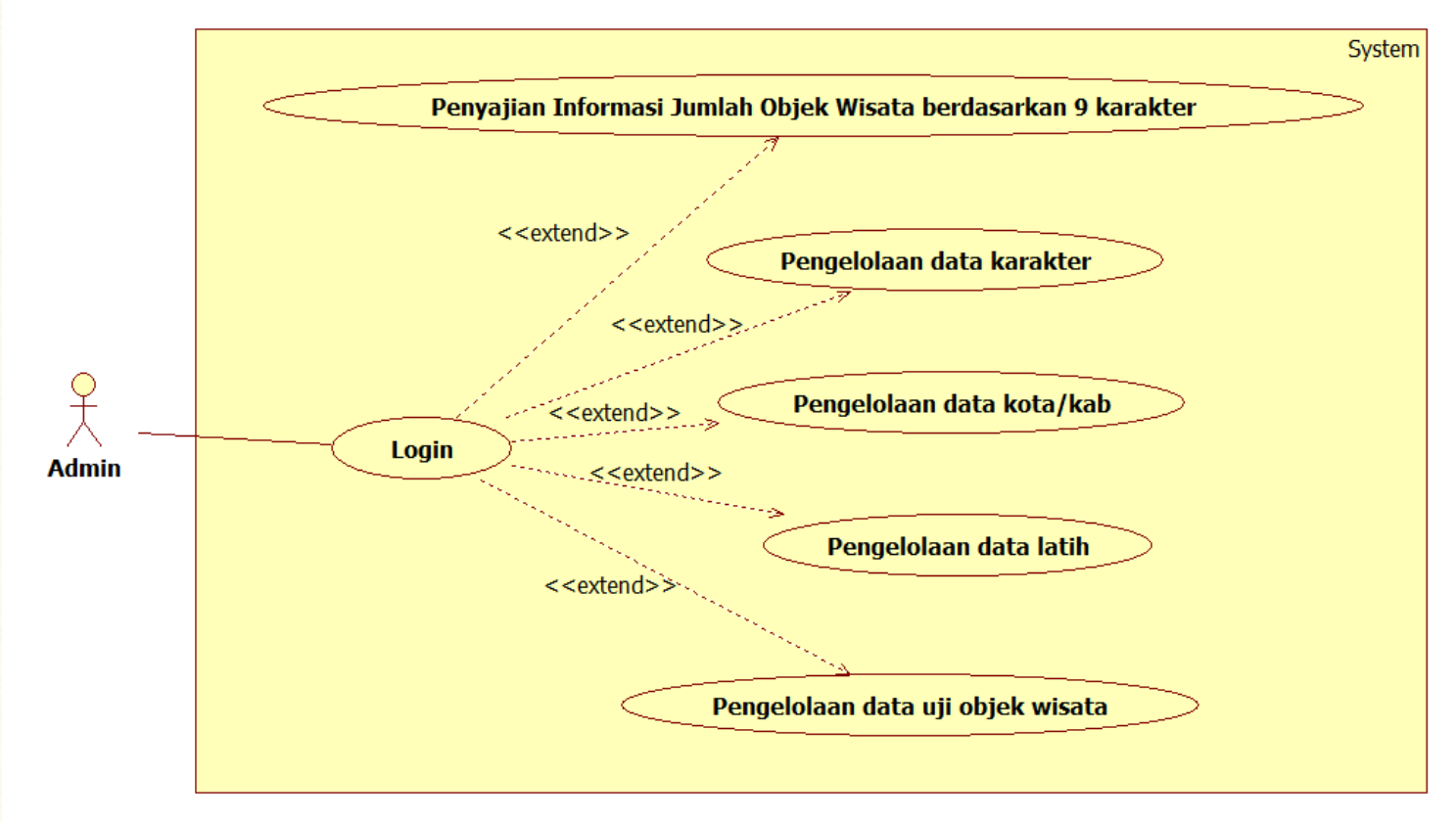
= 1/9 x 1/133 x 1/133 x 1/133 x 1/133=> $2,66992637595e^{-12}$

b. Tentukan kategori dengan nilai *P(vj) II P(ai | vj)* maksimalnya
i
sebagai hasil kategorisasi.

Berdasarkan diatas dapat diketahui masing-masing nilai V_{MAP} dari 9 karakter objek wisata. Dari nilai V_{MAP} tersebut kemudian diambil nilai terbesar yang menjadi kategori dari deskripsi objek wisata gedung sate. Pada deskripsi tersebut, didapat bahwa nilai V_{MAP} untuk karakter heritage, memiliki nilai tertinggi yaitu sebesar $5.5978949936305e^{-11}$ dibandingkan dengan kategori lainnya. Sehingga, deskripsi tersebut diklasifikasikan ke dalam **karakter Heritage**.

KEBUTUHAN FUNGSIONAL

USE CASE DIAGRAM



2

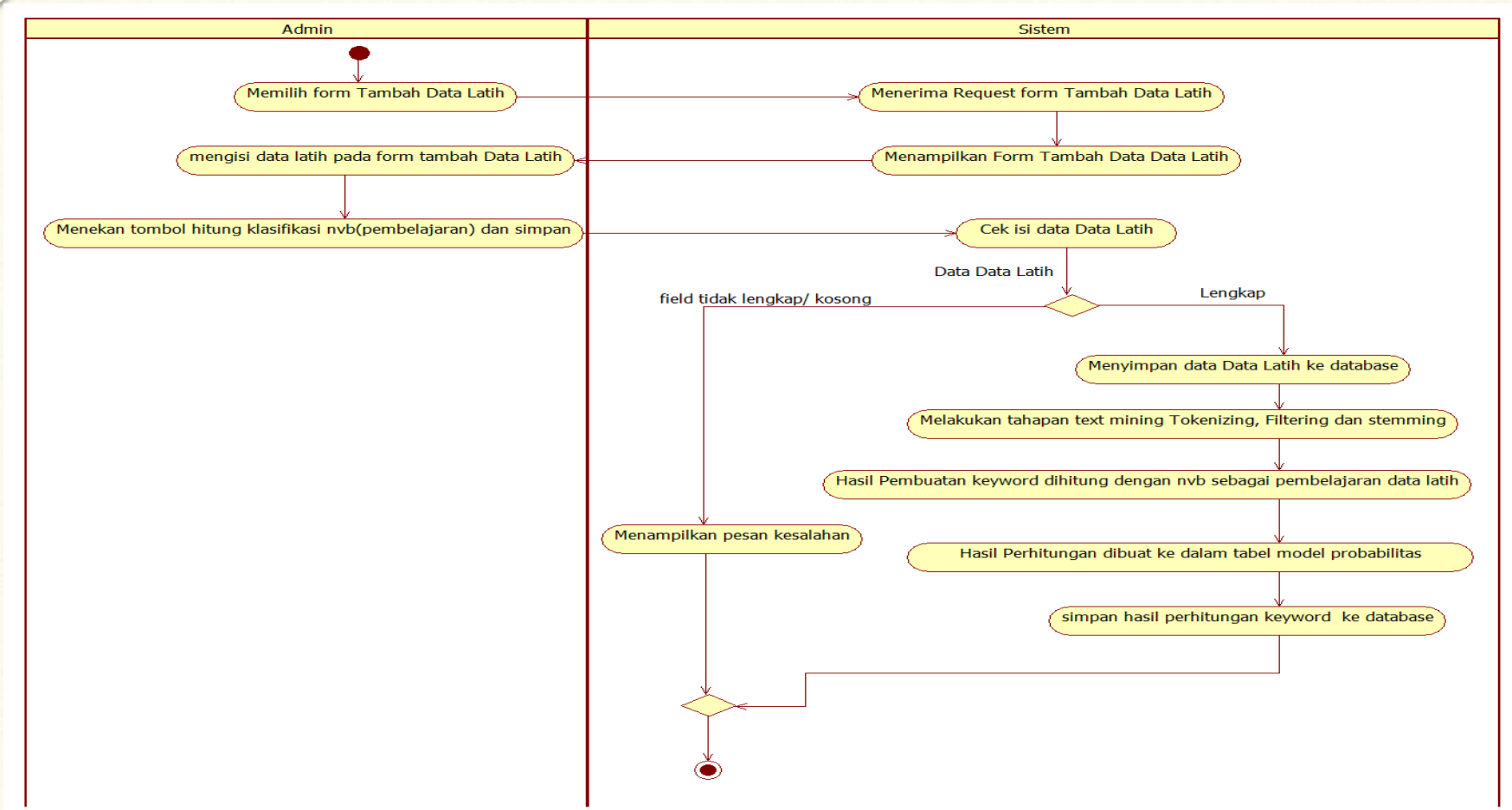
USE CASE SCENARIO TAMBAH DATA LATIH

Nama *Use Case*	Tambah data latih
Aktor	Admin
Deskripsi	Proses untuk menambah data data latih, membuat *keyword* dan menghitung kata total kata, kemunculan kata, jumlah data latih dan simpan perhitungan pada tabel probabilitas sebagai pelatihan/pembelajaran ke *database*.
Langkah- Langkah yang dilakukan	**Informasi untuk setiap langkah-langkah**
1. Admin memilih form tambah data latih	Halaman web menu data latih
2. Sistem menampilkan form tambah data latih.	Web form tambah data latih.
3. Admin mengisi data data latih	Web form tambah data latih, data latih.
4. Admin menekan tombol simpan	Web form tambah data latih, data latih.
5. Sistem mengecek dan memvalidasi data objek wisata	Web form tambah data latih, data latih.
6. Sistem melakukan tahap *text mining* preprocessing (*tokenizing*, *filtering*, dan *stemming*)	Web form tambah data latih, data latih.
7. Sistem membuat *keyword* dari data latih.	Web form tambah data data latih, data latih.
8. Sistem melakukan perhitungan kata, total kata, kemunculan kata, jumlah data latih.	Web form tambah data data latih, data latih.
9. Sistem akan menyimpan hasil perhitungan dan dibuat tabel probabilitas dari data latih sebagai tahapan klasifikasi naïve bayes sebagai tahap pembelajaran.	Web form tambah data data latih, data latih.
10. Sistem menyimpan data data latih ke *database*.	Web form tambah data data latih, data latih.
11. Halaman konfirmasi simpan dikirim ke admin.	Halaman konfirmasi tambah data latih telah disimpan.
Kondisi Awal	Data latih belum disimpan ke *database*.
Kondisi Akhir	Sistem menyimpan data latih ke *database*, membuat *keyword* dari data latih yang baru disimpan dan membuat tabel probabilitas.
Asumsi	1. Admin sudah login dengan mengisi nama pengguna dan kata kunci yang sah. 2. Sudah melakukan proses lihat data latih. 3. Sebelumnya telah melakukan proses tambah karakter, terlebih dahulu. 4. Berhasil menyimpan data latih ke *database* dan dibuat perhitungan *keyword* dari data latih ke dalam tabel probabilitas.

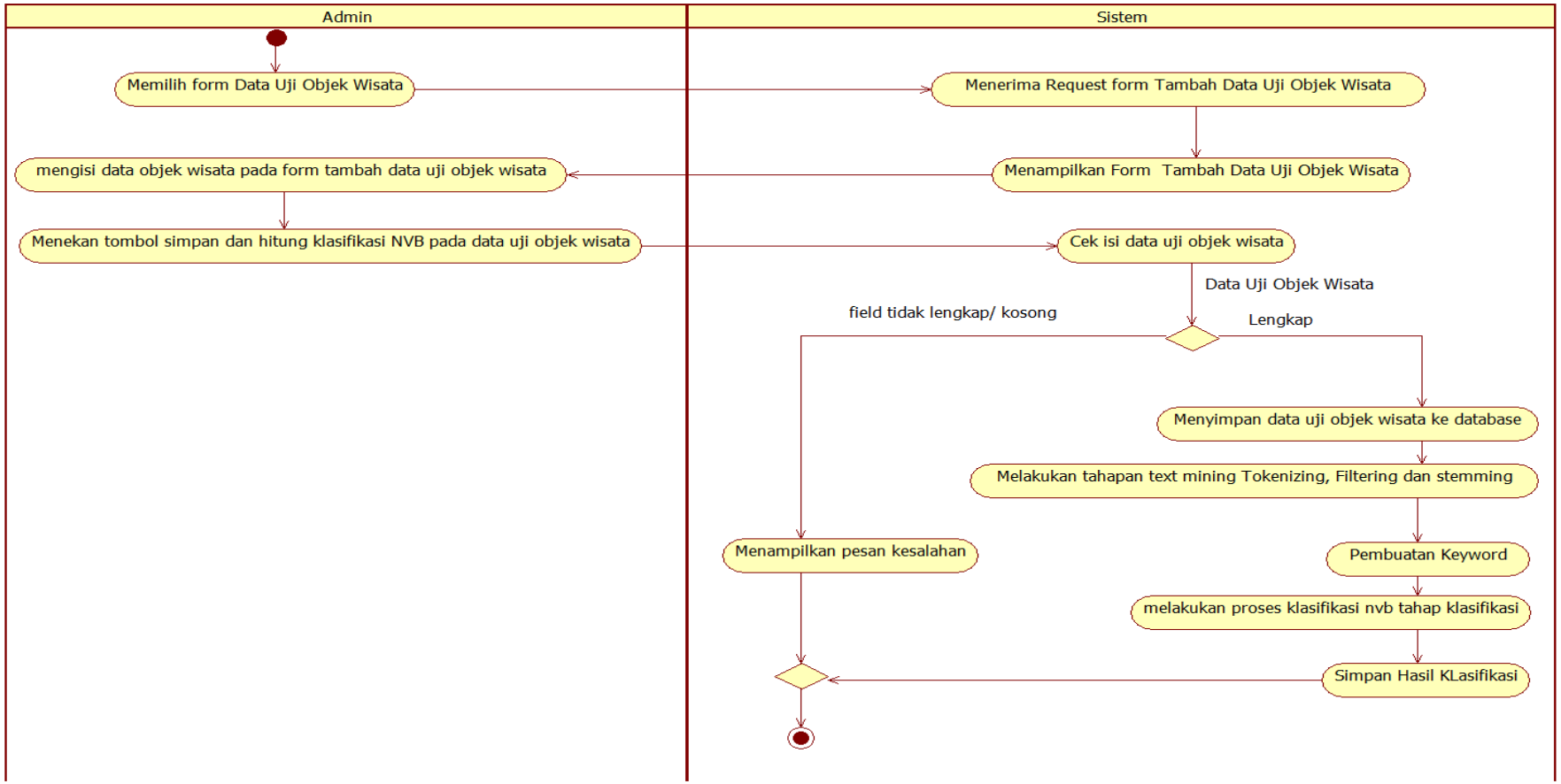
USE CASE SCENARIO TAMBAH DATA UJI

Nama *Use Case*	Tambah data uji objek wisata
Aktor	Admin
Deskripsi	Proses untuk menambah data uji objek wisata dengan membuat keyword terlebih dahulu dan melakukan perhitungan klasifikasi nvb tahap klasifikasi dengan mengklasifikasikan berdasarkan 9 karakter
Langkah- Langkah yang dilakukan	**Informasi untuk setiap langkah-langkah**
1. Admin memilih form tambah data uji objek wisata.	Web menu data uji objek wisata.
2. Sistem menampilkan form tambah data uji objek wisata.	Web form tambah data uji objek wisata
3. Admin mengisi data uji objek wisata.	Web form tambah data uji objek wisata, data uji objek wisata.
4. Admin menekan tombol simpan	Web form tambah data uji objek wisata, data uji objek wisata.
5. Sistem mengecek dan memvalidasi data data uji objek wisata.	Web form tambah data data uji objek wisata, data data uji objek wisata.
6. Sistem melakukan tahap *text mining* preprocessing (*tokenizing*, *filtering*, dan *stemming*)	Web form tambah data uji objek wisata, data uji objek wisata.
7. Sistem membuat *keyword* dari data objek wisata.	Web form tambah data uji objek wisata, data uji objek wisata.
8. Sistem melakukan klasifikasi *naïve bayes classifier* tahap klasifikasi.	Web form tambah data uji objek wisata, data uji objek wisata.
9. Sistem akan memberikan label karakter pada objek wisata setelahnya sebelumnya dilakukan proses klasifikasi *naïve bayes classifier*.	Web form tambah data uji objek wisata, data uji objek wisata.
10. Sistem menyimpan data objek wisata ke *database*.	Web form tambah data data uji objek wisata, data uji objek wisata.
11. Halaman konfirmasi simpan data uji objek wisata telah diklasifikasikan dan dikirim ke admin.	Halaman konfirmasi tambah data uji objek wisata telah disimpan.
Kondisi Awal	Data Objek Wisata objek wisata belum diklasifikasikan berdasarkan 9 karakter dan disimpan ke *database*.
Kondisi Akhir	Sistem menyimpan data objek wisata yang sudah diklasifikasikan berdasarkan karakter dengan metode *naïve bayes* ke dalam *database*.
Asumsi	1. Admin sudah login dengan mengisi nama pengguna dan kata kunci yang sah. 2. Sudah melakukan proses lihat data uji objek wisata. 3. Telah melakukan proses tambah karakter, data latih, dan kota/kab terlebih dahulu. 4. Berhasil menyimpan data uji objek wisata ke *database*.

ACTIVITY DIAGRAM TAMBAH DATA LATIH

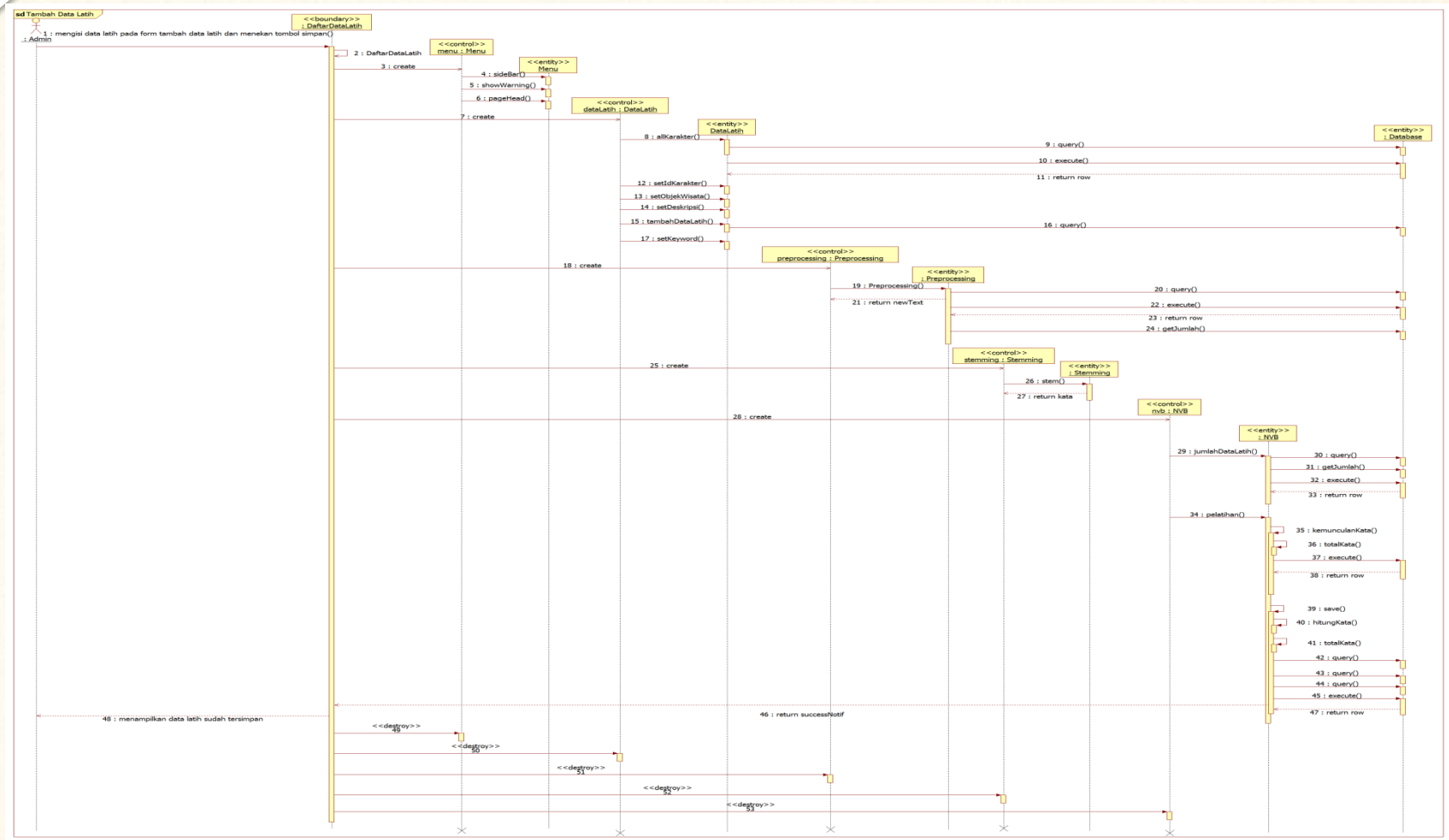


ACTIVITY DIAGRAM TAMBAH DATA UJI

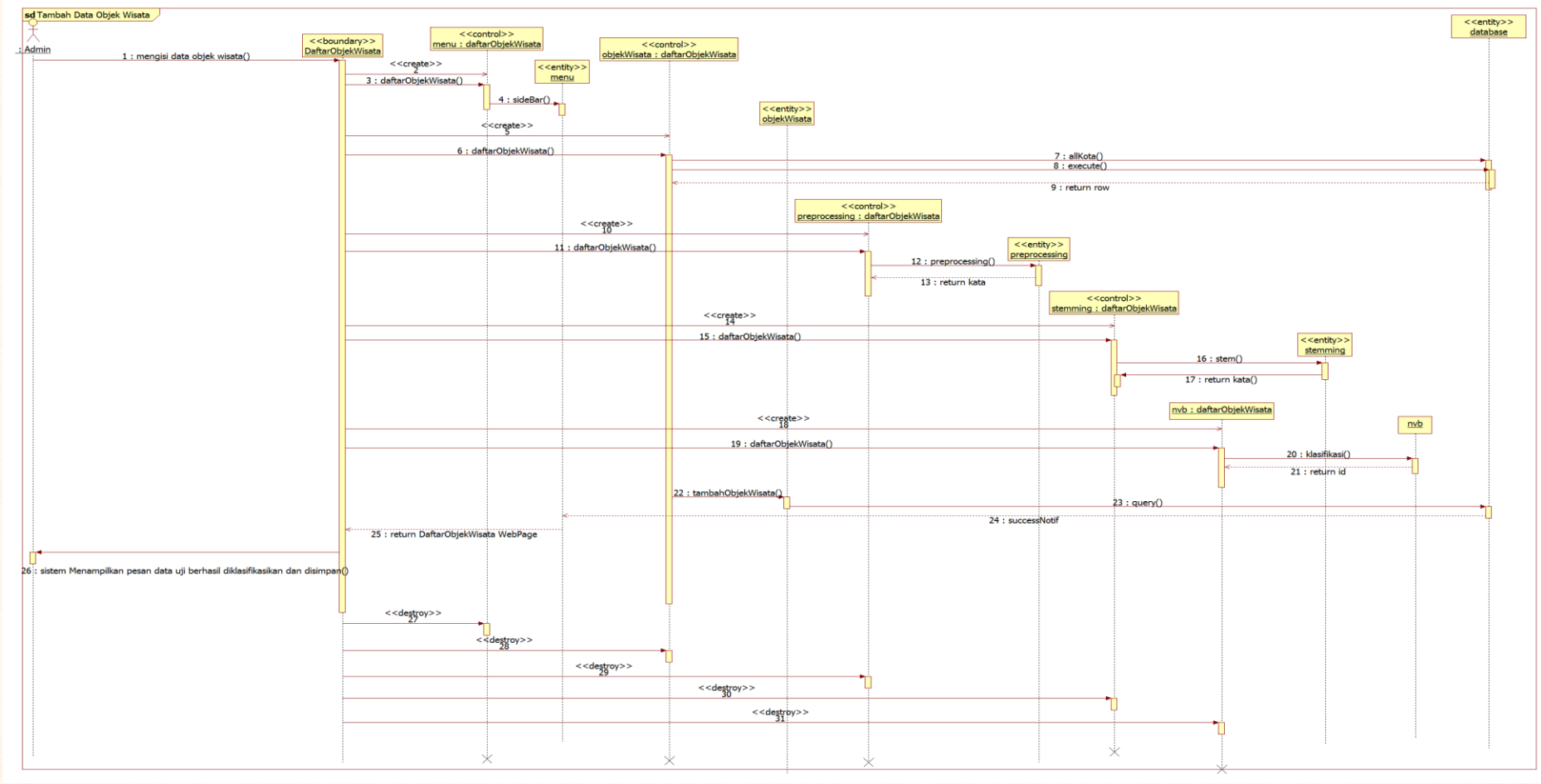


4

SEQUENCE DIAGRAM TAMBAH DATA LATIH

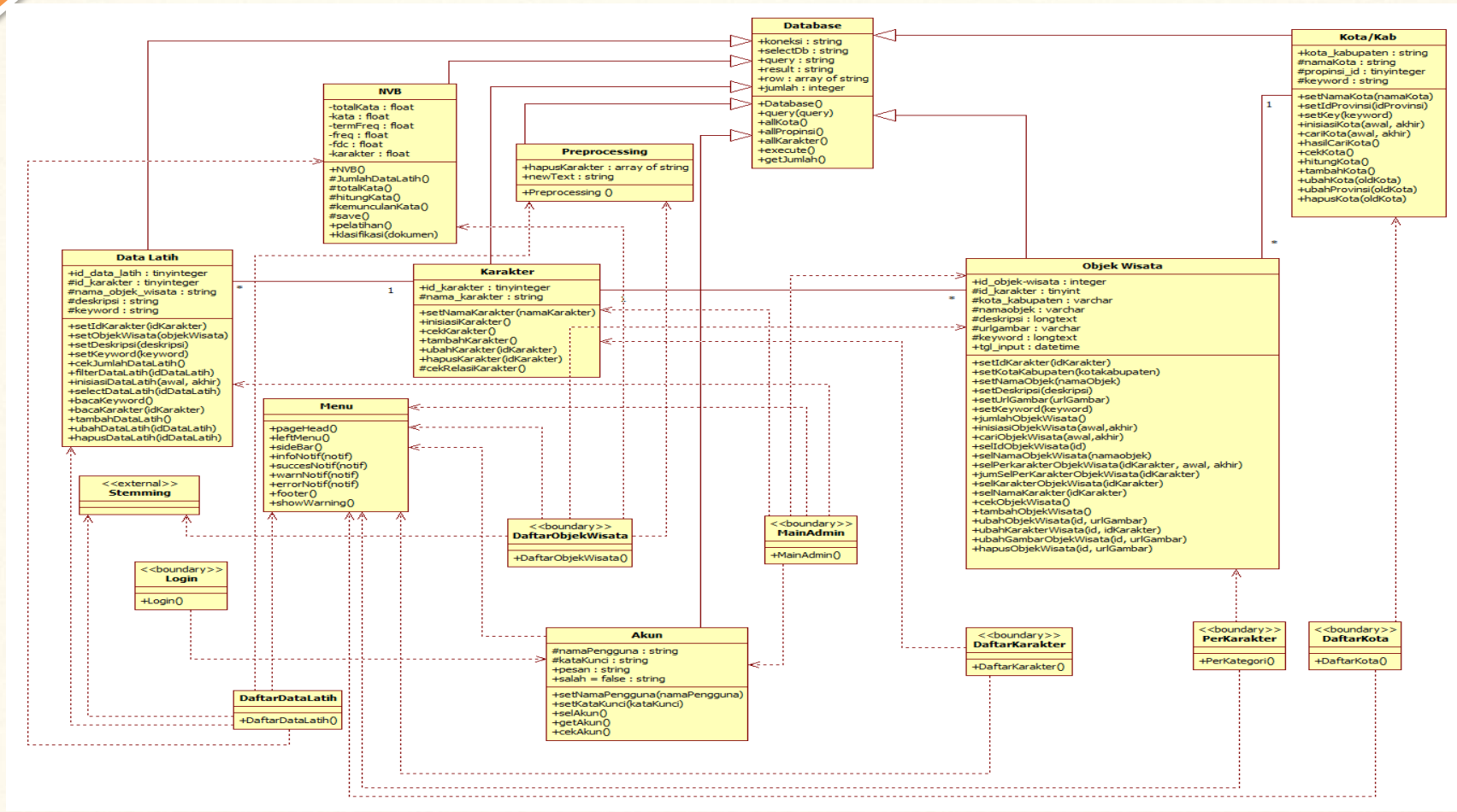


SEQUENCE DIAGRAM TAMBAH DATA UJI



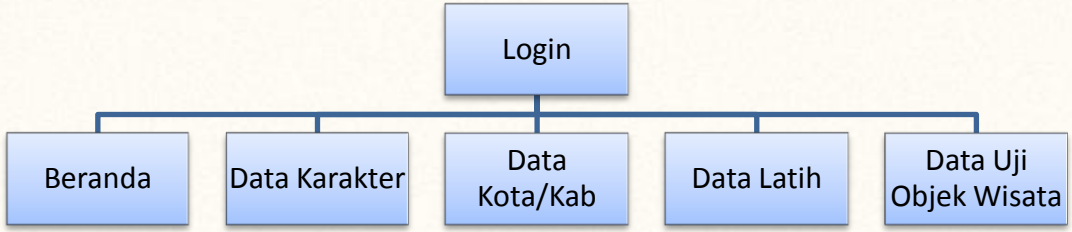
5

CLASS DIAGRAM SISTEM PENGKLASIFIKASIAN



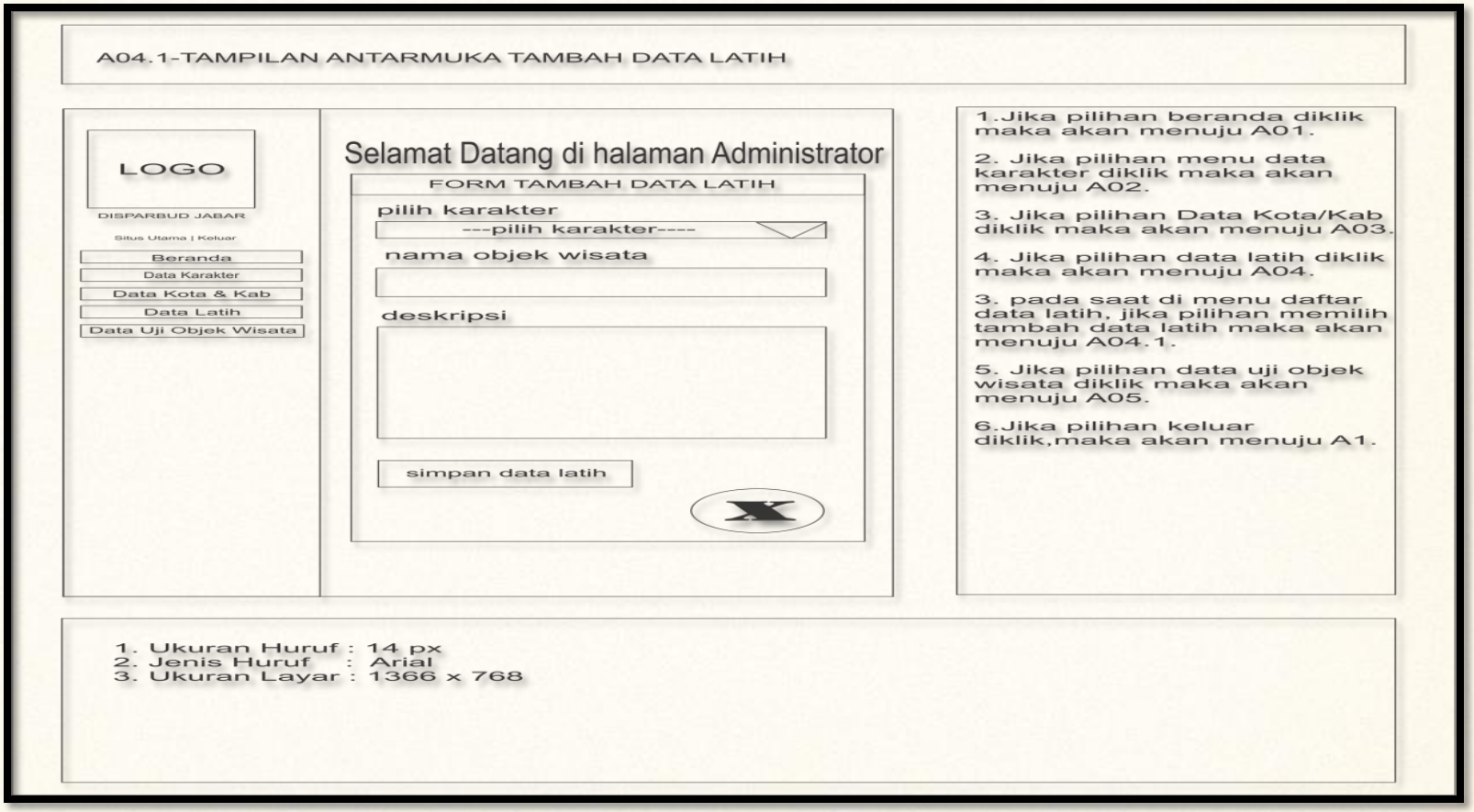
RANCANGAN STRUKTUR

MENU ADMIN

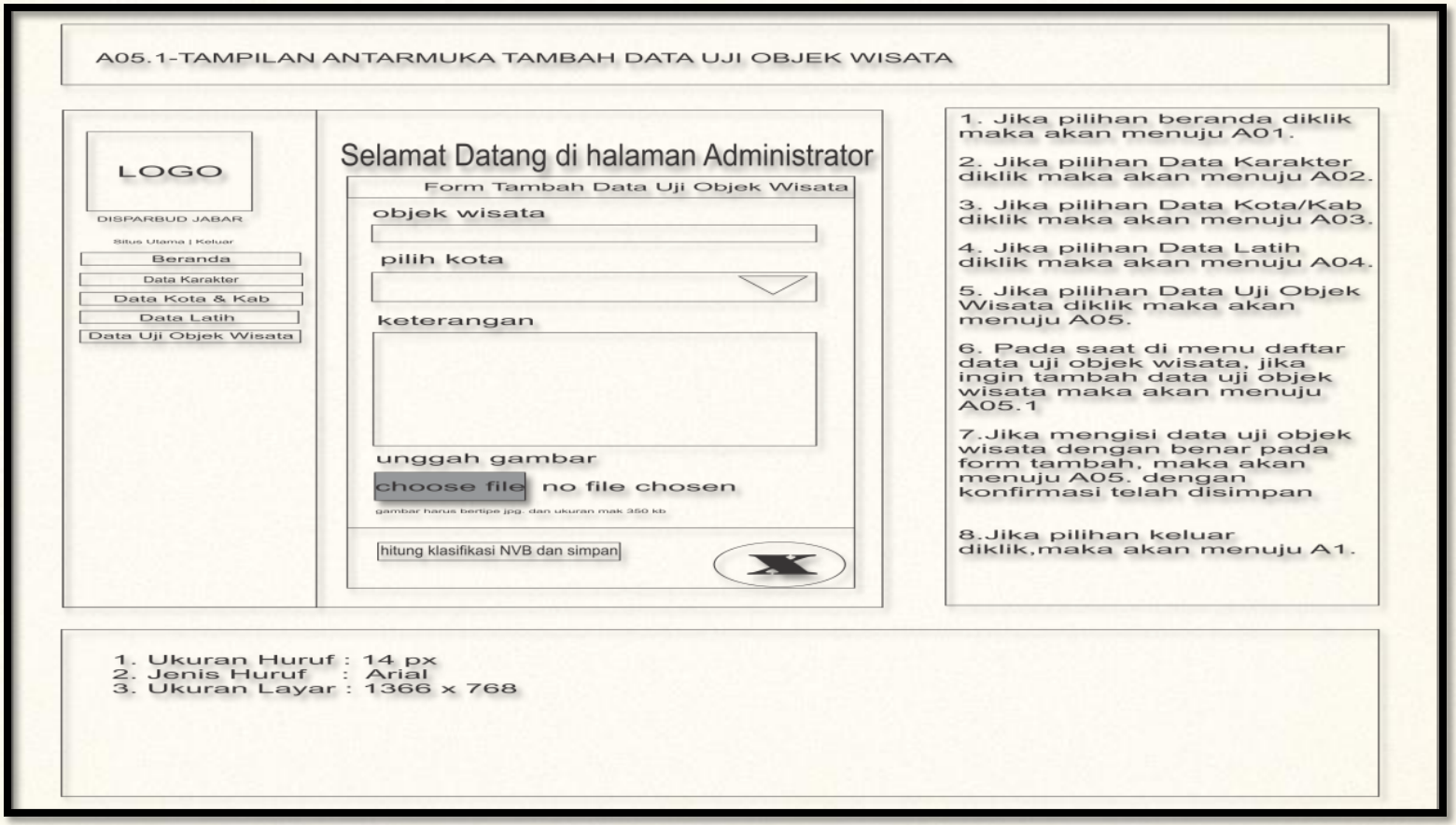

Login
Beranda
Data Karakter
Data Kota/Kab
Data Latih
Data Uji Objek Wisata

RANCANGAN ANTARMUKA

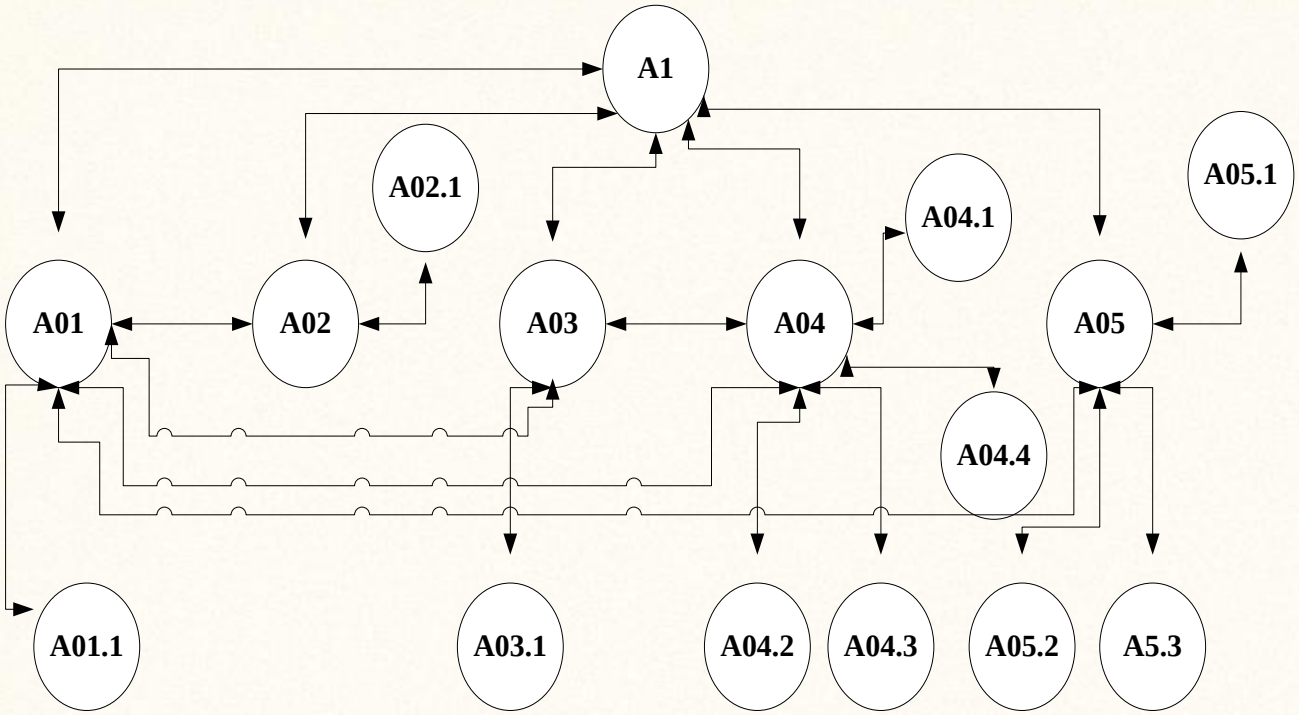
RANCANGAN ANTARMUKA TAMBAH DATA LATIH



RANCANGAN ANTARMUKA TAMBAH DATA UJI



JARINGAN SEMANTIK SISTEM PENGKLASIFIKASIAN



1. PENGUJIAN TAMBAH DATA LATIH

Kasus dan Hasil Uji (Data Normal)			
Data Masukkan	**Keluaran yang diharapkan**	**Pengamatan**	**Kesimpulan**
Mengisi data latih contoh : - Pilih Karakter : Hutan - Nama Objek Wisata : Taman Nasional Gunung Gede Pangrango - Desktipsi :" Taman Nasional Gunung Gede Pangrango ditetapkan oleh unesco sebagai cagar biosfir pada tahun 1977."	Menyimpan data latih ke *database*, dengan melakukan proses *text mining*, yaitu *tokenizing*, *filtering*, dan *stemming* lalu dibuat indeks pembuatan *keyword* dan juga melakukan klasifikasi nvb pembelajaran data latih dan dibuat ke dalam model probabilitas.	Sistem menampilkan pesan " Data latih telah berhasil disimpan dan diklasifikasikan NVB sebagai pembelajaran data latih "	[√] Diterima [ ] Ditolak

2. PENGUJIAN TAMBAH DATA UJI

Kasus dan Hasil Uji (Data Normal)			
Data Masukkan	**Keluaran yang diharapkan**	**Pengamatan**	**Kesimpulan**
Mengisi data Uji Objek Wisata. Contoh : -Nama Objek Wisata : Telaga Nilam -Pilih Kota/Kab : Kabupaten kuningan -Desktipsi :" Telaga Nilam masih berada dikawasan Telaga Remis." file gambar : telaga.jpg	Menyimpan data uji objek wisata ke *database*, dengan melakukan proses *text mining*, yaitu *tokenizing*, *filtering*, dan *stemming* lalu dibuat indeks pembuatan *keyword* dan tahp terakhir melakukan klasifikasi nvb data uji sebagai proses pengklasifikasian dan berhasil menghasilkan label karakter yang sesuai.	Sistem akan menampilkan pesan " Objek Wisata dengan nama objek wisata taman nastelah diklasifikasikan NVB ke dalam karakter hutan dan telah berhasil disimpan"	[√] Diterima [ ] Ditolak

KESIMPULAN
SARAN

KESIMPULAN

1. Sistem pengklasifikasian pada objek wisata dapat menghasilkan aturan klasifikasi dalam setiap kategori berdasarkan 9 karakter dengan menggunakan metode naïve bayes classifier dengan mencari nilai *Vmap* kategori yang paling maksimum.

2. Pembangunan sistem pengklasifikasian objek wisata dapat membantu khususnya admin disparbud dalam melakukan pengklasifikasian objek wisata berdasarkan kategori ke dalam 9 karakter.

3. Metode *Naïve Bayes Classifier (NVB)* dapat diimplementasikan dalam sistem pengklasifikasian objek wisata ke dalam 9 karakter.

SARAN

1. Perlu penelitian lebih lanjut untuk mengoptimalkan metode naïve bayes classifier.

2. Perlu adanya penelitian bagaimana cara mengetahui tingkat error naïve bayes classifier melakukan klasifikasi objek wisata.

3. Perlu dikembangkan lagi tidak hanya sistem mengklasifikasikan saja namun perlu ada sistem ada pencarian objek wisata memakai teknik text mining khususnya dalam bidang *information retrievalnya (IR)*.

THANK YOU

GRACIAS

Mira Kania Sabariah, S.T.,M.T.
Inne Novitasari, S.Si., M.Si.
Adam Mukharil Bachtiar,S.Kom.