

Mobile Blog untuk Citizen Journalism dengan Pengkategorian Berita Menggunakan Metode Inner Product

Afrida Helen dan Yuliana Setiowati

Abstrak—Dalam makalah ini dibahas tentang sebuah aplikasi yang mendukung kegiatan *citizen journalism* berupa telepon genggam yang bisa mengirimkan berita berupa gambar dan teks yang akan ditampilkan pada sebuah blog menurut kronologi waktu dan diklasifikasikan berdasarkan kategori berita politik, ekonomi, olahraga, kriminal, hiburan. Informasi berita bisa disampaikan oleh siapa saja, berupa peristiwa yang terjadi bisa kapan saja dan di mana saja, inilah yang disebut *citizen journalism*. Untuk memudahkan pembacaan maka berita akan dikategorikan secara otomatis. Input berita akan melalui proses *text mining*, yaitu *tokenizing*, *filtering* dan *stemming*, kemudian hasilnya akan diproses dengan kumpulan kata kunci tiap kategori berita menggunakan metode *inner product*. Metode *inner product* ini memerlukan kata kunci untuk setiap kategori berita. Kata kunci adalah kata yang sering muncul dalam suatu kategori berita. Hasil uji coba menunjukkan bahwa rata-rata tingkat keakuratan hasil klasifikasi berita dengan metode *inner product* sekitar 92%. Metode ini tergantung pada banyaknya kata kunci tiap kategori berita, semakin banyak kata kunci tiap kategori berita, semakin baik hasil pengkategorian.

Kata Kunci—*citizen journalism*, *text mining*, *tokenizing*, *filtering*, *stemming*, *inner product*

1 PENDAHULUAN

INFORMASI bisa diketahui cepat oleh masyarakat melalui beberapa media pendukung, seperti koran, web atau perangkat mobile.

Informasi pada umumnya ditulis oleh seorang wartawan. Tapi dengan berkembangnya teknologi, maka setiap orang bisa mengabadikan suatu peristiwa sekaligus memberikan informasi berita. Informasi yang disampaikan oleh siapa saja biasanya disebut dengan *citizen journalism*. Namun tidak banyak redaksi berita yang menerima berita-berita dari *citizen journalism* dengan pertimbangan aturan jurnalis, biasanya redaksi berita membuat sebuah blog atau halaman khusus yang menampung berita-berita dari

citizen journalism dengan mengusung aturan kebebasan dalam menulis tanpa diatur oleh aturan-aturan pemberitaan.

Disisi lain perkembangan teknologi telepon genggam begitu cepat, banyak fasilitas yang ditawarkan misal kamera, gps, browser dan lain-lain, semuanya mendukung manusia dalam mendapatkan dan mengirim informasi. Pada penelitian sebelumnya telah dilakukan pengiriman berita *emergency* disertai gambar ke server, tapi di server tidak dilakukan pengkategorian berita [8].

Citizen journalism mungkin tidak mengetahui aturan jurnalis, maka dalam penerimaan berita, editor blog untuk *citizen journalism* akan mengecek berita yang disampaikan beritidak baik atau tidak dan mengelompokkan secara manual berita-berita yang masuk sesuai dengan kategori-kategori berita.

Penelitian ini mengembangkan dari penelitian sebelumnya yaitu dibuat sebuah aplikasi telepon genggam yang mengirimkan berita berupa gambar dan teks [2], selanjutnya berita-berita yang dikirimkan akan dikategorikan secara otomatis menggunakan metode *inner prod-*

- Afrida Helen, Jurusan Teknik Informatika, Politeknik Elektronika Negeri Surabaya, Jl. Raya ITS Keputih Sukolilo, Telp: 031-5947280. E-mail: helen@eepis-its.edu.
- Yuliana Setiowati, Jurusan Teknik Informatika, Politeknik Elektronika Negeri Surabaya, Jl. Raya ITS Keputih Sukolilo, Telp: 031-5947280. E-mail: yuliana@eepis-its.edu.

uct kemudian ditampilkan berdasarkan kronologis waktu pada sebuah blog [4]. Sehingga proses pengkategorian berita akan dilakukan secara cepat.

Dalam proses pengkategorian berita dengan metode *inner product*, diperlukan informasi berupa kata-kata yang sering muncul dalam sebuah kategori berita. Kata-kata yang sering muncul ini biasanya disebut dengan kata kunci. Untuk mendapatkan kata kunci ini dilakukan secara otomatis dengan melakukan *text mining* pada berita-berita dengan kategori yang sama, selanjutnya diproses dan diperoleh kata kunci.

2 METODOLOGI

2.1 Blog

Blog merupakan singkatan dari "web blog" adalah bentuk aplikasi web yang menyerupai tulisan-tulisan pada sebuah halaman web umum [1]. Tulisan-tulisan ini seringkali dimuat dalam urutan terbalik (isi terbaru dahulu baru kemudian diikuti isi yang lebih lama), meskipun tidak selamanya demikian.

Blog mempunyai fungsi yang sangat beragam, dari sebuah catatan harian, media publikasi dalam sebuah kampanye politik, sampai dengan program-program media dan perusahaan perusahaan. Sebagian blog dipelihara oleh seorang penulis tunggal, sementara sebagian lainnya oleh beberapa penulis. Banyak juga weblog yang memiliki fasilitas interaksi dengan para pengunjungnya, seperti menggunakan buku tamu dan kolom komentar yang dapat memperkenankan para pengunjungnya untuk meninggalkan komentar [1].

2.2 Mobile Blog

Mobile blog [2] adalah salah satu bentuk cara *posting* pada blog dengan menggunakan media *mobile* secara langsung, contohnya dengan menggunakan telepon genggam. *Mobile blog* membantu pengguna blog yang memiliki mobilitas tinggi, sehingga memungkinkan untuk mengirim *posting* yang berupa gambar atau video, dengan media email atau MMS. Tidak hanya pada sisi pengirim, yaitu perangkat *mobile*, pada sisi server atau web yang akan

menampilkan *posting* blog ini juga harus memiliki dukungan khusus untuk menerima kiriman *posting* dari telepon genggam, sehingga bisa saling berhubungan.

2.3 Text Mining

2.3.1 Pengertian Text Mining

Text mining adalah proses menemukan informasi dari beberapa data yang berupa dokumen teks dengan tujuan mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen [3].

Text mining yang digunakan dalam proyek akhir ini terdiri dari beberapa proses, diantaranya adalah *tokenizing*, *filtering*, *stemming*.

2.3.2 Tokenizing

Tokenizing yaitu proses menghilangkan angka-angka tanda baca (, . ? !)(- * \$ # ; : % @ & + = ') dan pemotongan setiap input berdasarkan tiap kata yang menyusunnya. Contoh:

Input : konser hiburan di taman kota
menampilkan penyanyi dan
pelawak
Output: konser
hiburan
di
taman
kota
menampilkan
penyanyi
dan
pelawak

2.3.3 Filtering

Filtering yaitu mengambil kata-kata penting dari hasil *tokenizing*, menggunakan algoritma *stop list*, yaitu membuang kata-kata yang kurang penting kemudian dijadikan sebagai kata kunci untuk proses selanjutnya, dari contoh di atas dapat diambil kata-kata:

konser
hiburan
taman
kota
menampilkan
penyanyi
pelawak

kata di, dan dihilangkan

2.3.4 Stemming

Stemming yaitu tahap mencari kata dasar dari tiap kata hasil *filtering* dengan menghilangkan imbuhan-imbuhan pada kata tersebut. Untuk kata-kata dalam bahasa Indonesia digunakan pola kata sebagai berikut:

Algoritma ini didahului dengan pembacaan tiap kata dari file sampel. Sehingga input dari algoritma ini adalah sebuah kata yang kemudian dilakukan :

- 1) Pemeriksaan semua kemungkinan bentuk kata. Setiap kata diasumsikan memiliki 2 Awalan (prefiks) dan 3 Akhiran (sufiks). Sehingga bentuknya menjadi:

Prefiks 1 + Prefiks 2 + Kata dasar + Sufiks 3
+ Sufiks 2 + Sufiks 1

Seandainya kata tersebut tidak memiliki imbuhan sebanyak imbuhan di atas, maka imbuhan yang kosong diberi tanda x untuk prefiks dan diberi tanda xx untuk sufiks.

- 2) Pemotongan dilakukan secara berurutan seperti:

AW= awalan, AK=akhiran, KD= kata dasar

AW : AW

AK : AK

KD : KD

- a. AW I, hasilnya disimpan pada p1
- b. AW II, hasilnya disimpan pada p2
- c. AK I, hasilnya disimpan pada s1
- d. AK II, hasilnya disimpan pada s2
- e. AK III, hasilnya disimpan pada s3

Pada setiap tahap pemotongan di atas diikuti dengan pemeriksaan di kamus apakah hasil pemotongan itu sudah berada dalam bentuk dasar. Kalau pemeriksaan ini berhasil maka proses dinyatakan selesai dan tidak perlu melanjutkan proses pemotongan imbuhan lain-

nya.

Contoh pemenggalan kata "mempermainkannya".

Langkah 1 : Cek apakah kata ada dalam kamus

Ya : Succes

Tidak : lakukan pemotongan AW I

Kata : memainkannya

Langkah 2 : Cek apakah kata ada dalam kamus

Ya : Succes

Tidak : lakukan pemotongan AW II

Kata : mainkannya

Langkah 3 : Cek apakah kata ada dalam kamus

Ya : Succes

Tidak : lakukan pemotongan AK I

Kata : mainkan

Langkah 4 : Cek apakah kata ada dalam kamus

Ya : Succes

Tidak : lakukan pemotongan AK II

Kata : main

Langkah 5 : Cek apakah kata ada dalam kamus

Ya : Succes

Tidak : lakukan pemotongan AW III

(Dalam hal ini AK III tidak ada, sehingga kata tidak diubah)

Kata : main

Langkah 6 : Cek apakah kata ada dalam kamus

Ya : Succes

Tidak : "Kata tidak ditemukan"

- 3) Jika sampai pada pemotongan AK III, belum juga ditemukan di kamus, maka dilakukan proses kombinasi. KD yang dihasilkan dikombinasikan dengan imbuhan-imbuhannya dalam 12 konfigurasi berikut:

a) KD

b) KD + AK III

c) KD + AK III + AK II

d) KD + AK III + AK II + AK I

e) AW I + AW II + KD

f) AW I + AW II + KD + AK III

g) AW I + AW II + KD + AK III + AK II

h) AW I + AW II + KD + AK III + AKII + AKI

i) AW II + KD

j) AW II + KD + AK III

k) AW II + KD + AK III + AK II

l) AW II + KD + AK III + AK II + AK I

Sebenarnya kombinasi *a, b, c, d, h,* dan *l* sudah diperiksa pada tahap sebelumnya, karena kombinasi ini adalah hasil pemotongan bertahap tersebut. Dengan demikian, kombinasi yang masih perlu dilakukan tinggal 6 yakni pada kombinasi-kombinasi yang belum dilakukan (*e, f, g, i, j, dan k*).

Tentunya bila hasil pemeriksaan suatu kombinasi adalah 'ada', maka pemeriksaan pada kombinasi lainnya sudah tidak diperlukan lagi. Pemeriksaan 12 kombinasi ini diperlukan, karena adanya fenomena *overstemming* pada algoritma pemotongan imbuhan. Kelemahan ini berakibat pada pemotongan bagian kata yang sebenarnya adalah milik kata dasar itu sendiri yang kebetulan mirip dengan salah satu jenis imbuhan yang ada. Dengan 12 kombinasi itu, pemotongan yang sudah terlanjur tersebut dapat dikembalikan sesuai posisinya [5].

2.4 Inner Product

Inner product adalah merupakan istilah lain dari *dot product* atau dapat juga disebut sebagai *scalar product* adalah perkalian antara dua vektor yang dapat berupa vektor baris atau vektor kolom, sehingga hasilnya berupa skalar dengan rumus sebagai berikut [4]:

$$a.b = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \dots + a_n b_n \quad (1)$$

dimana *a* adalah elemen vektor 1 dan *b* adalah elemen vektor 2.

Penerapan metode *inner product* ini pernah dilakukan oleh peneliti sebelumnya pada tahun 2008 [4], untuk mengklasifikasikan Tugas Akhir. Jika diilustrasikan dalam kasus ini sebagai berikut:

Kata-kata hasil *stemming* akan menjadi elemen vektor satu dengan nilai masing-masing 1, kata-kata yang termasuk dalam kata kunci tiap kategori berita sebagai elemen dari vektor dua, sehingga nilai dari elemen vektor satu akan dikalikan dengan nilai dari elemen vektor dua, di mana nilai elemen pada vektor dua akan memiliki nilai tidak sama dengan nol, apabila kata hasil *stemming* sama dengan kata kunci tersebut, contoh perkaliannya dapat dilihat pada Tabel 1.

Proses perhitungan dari Tabel 1 adalah sebagai berikut:

- Politik : (1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)=0
- Ekonomi : (1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)=0
- Kriminal : (1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)=0
- Olahraga : (1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)+(1)(0)=0
- Hiburan : (1)(1)+(1)(1)+(1)(0)+(1)(0)+(1)(0)+(1)(1)+(1)(1)=4

Tabel 1
Contoh roses *inner product*

Kategori	Kon ser	Hi bur				Nya nyi	La wak	ha sil
Politik	0	0	0	0	0	0	0	0
Ekonomi	0	0	0	0	0	0	0	0
kriminal	0	0	0	0	0	0	0	0
Olahraga	0	0	0	0	0	0	0	0
Hiburan	1	1	0	0	0	1	1	4
Input nilai	1	1	1	1	1	1	1	
Input kata	kon ser	Hi bur	ta man	ko ta	tam pil	nya nyi	la wak	

Dari proses diatas terlihat hasil tertinggi adalah hiburan, sehingga didapatkan output bahwa inputan teks berita termasuk dalam kategori hiburan.

3 PERANCANGAN SISTEM

3.1 Perancangan Database

Untuk menyimpan kata kunci tiap kategori berita dibutuhkan 5 tabel, yaitu tabel untuk kata kunci politik, ekonomi, olahraga, kriminal, dan hiburan. *Field* pada tabel yaitu kata dan jumlah. *Field* kata untuk menyimpan kata kunci, sedangkan frekuensi kata kunci dalam dokumen-dokumenkategori berita disimpan dalam *field* jumlah. Struktur tabel kata kunci ditunjukkan pada Tabel 2.

Tabel 2
Stuktur tabel kata kunci

Nama Field	Type Data
Kata	varchar (2)
Jumlah	Int(10)

Berita yang dikirimkan oleh user, berupa Id berita, judul, isi berita dan gambar. Struktu tabel ditunjukkan pada Tabel 3.

Data-data *user* terdaftar akan disimpan pada tabel akun seperti pada Tabel 4.

Tabel 3
Stuktur tabel kata kunci

Nama Field	Tipe Data
Id (primary key)	Int (10)
Judul	varchar (200)
Isiberita	varchar (2000)
Gambar	varchar(15)
Kategori	varchar(15)
Tanggal	varchar(30)
Iduser	varchar(10)

Tabel 4
Stuktur tabel akun

Nama Field	Tipe Data
iduser (primary key)	int (11)
Nama	varchar (20)
Username	varchar(10)
Pass	varchar (10)
Email	varchar(20)

3.2 Pengumpulan Data

Data yang dikumpulkan untuk kata kunci tiap kategori berita, yaitu berita kategori politik, ekonomi, olahraga, kriminal dan hiburan di dapat dari situs berita (cetak.kompas.com dan www.kabarindonesia.com) kemudian berita-berita ini akan dilakukan proses *text mining* dan akan didapatkan kata kunci tiap kategori berita sekaligus jumlah kata tersebut di semua berita yang dikumpulkan.

Gambaran proses pengumpulan kata kunci diilustrasikan pada Gambar 1.

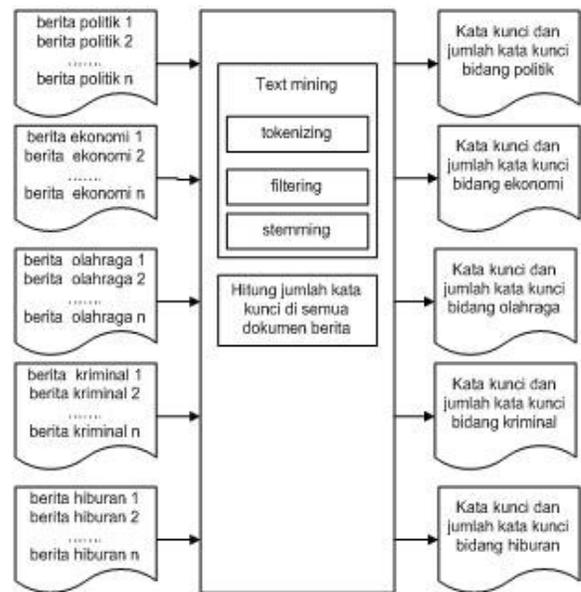
Pada Tabel 5 merupakan contoh kata kunci yang dihasilkan dari proses *text mining* berita politik.

3.3 Diagram Sistem yang Diusulkan

Secara umum sistem yang akan dibuat dapat ditunjukkan pada Gambar 2.

Dari Gambar 2 dapat dijelaskan proses secara keseluruhan sebagai berikut:

- 1) *User*, dalam hal ini adalah *citizen journalism*, menulis teks berita dan insert gam-



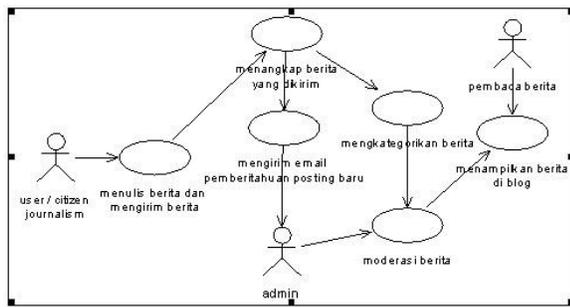
Gambar 1. Proses penentuan kata kunci setiap kategori berita

Tabel 5
5 contoh kata kunci yang dihasilkan dari proses *text mining* berita politik

Kata	jumlah	Kata	jumlah
Partai	174	Komisi	40
Caleg	117	pemerintah	38
Pemilu	114	Undang	36
Suara	114	Rakyat	36
KPU	102	Presiden	34
politik	93	Kabupaten	34
anggota	71	Kampanye	33
DPR	67	Lembaga	33
masyarakat	59	Parpol	32
Calon	58	RUU	28
Partai	174	Komisi	40
Caleg	117	pemerintah	38
Pemilu	114	Undang	36
Suara	114	Rakyat	36

bar pada aplikasi yang ada di telepon genggam

- 2) Teks dan gambar berita dikirimkan melalui koneksi internet ke blog
- 3) Blog menerima berita kemudian



Gambar 2. Use case gambaran sistem keseluruhan

melakukan pengkategorian berita

- 4) Ketika ada posting baru sistem secara otomatis akan mengirim email pemberitahuan ada posting baru kepada admin, sehingga admin bisa segera melakukan moderasi untuk menampilkan berita
- 5) Admin akan melakukan moderasi, yaitu berhak menampilkan atau tidak menampilkan berita terkirim
- 6) Berita yang telah dikategorikan dan telah diijinkan untuk ditampilkan oleh admin akan ditampilkan dalam sebuah blog menurut kronologis waktu pengiriman

Prosesnya, setiap kata pada teks berita akan dicek, dibandingkan dengan kata kunci tiap kategori berita, apabila kata tersebut ada dalam daftar, maka akan diberikan bobot, untuk judul diberikan bobot 3 kali jumlah kata dalam semua berita, untuk isi berita akan diberikan bobot 1 kali jumlah kata dalam semua berita.

Dalam berita terdapat judul dan isi berita. Sebelum membaca berita, seseorang akan melihat judul yang mewakili isi berita. Sehingga judul mempunyai bobot yang lebih tinggi dibandingkan dengan isi berita, oleh sebab itu dalam penelitian ini judul diberikan bobot 3, sedangkan isi berita diberikan bobot 1.

Contoh salah satu berita yang dikategorikan adalah:

Judul : penjambret dihajar massa

Isi berita : seorang pemuda dikeroyok warga hingga babak belur, dia dihajar massa karena diduga telah menjambret dua wanita di jl. Dr. Mansyur Medan. Tersangka dibawa ke markas kepolisian Medan Sunggal.

Selanjutnya ditokenizing menjadi :

Judul : penjambret dihajar massa

Isi berita : seorang pemuda dikeroyok warga hingga babak belur dia dihajar massa karena diduga telah menjambret dua wanita di jl dr Mansyur Medan Tersangka dibawa ke markas kepolisian Medan Sunggal

Selanjutnya membuang kata-kata yang kurang penting, kata-kata yang tersisa sebagai kata kunci untuk proses selanjutnya.

Judul : penjambret dihajar massa

Isi berita : pemuda dikeroyok warga babak belur dihajar massa diduga menjambret wanita Mansyur Medan Tersangka dibawa markas kepolisian Medan Sunggal

Selanjutnya proses stemming

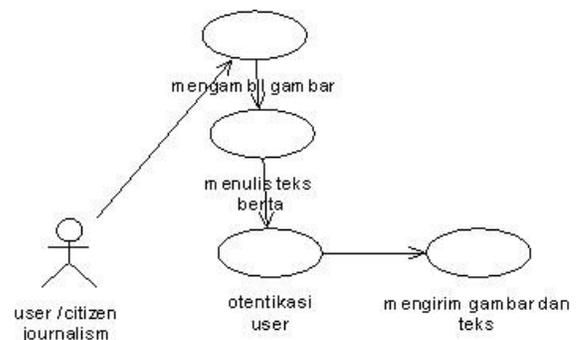
Judul : jambret hajar massa

Isi berita : muda keroyok warga babak belur hajar massa duga jambret wanita Mansyur medan sangka bawa markas polisi medan Sunggal

Selanjutnya proses pengkategorian berita ditunjukkan pada Tabel 6.

3.4 Aplikasi pada telepon genggam

Secara umum sistem aplikasi pada telepon genggam yang akan dibuat dapat digambarkan sebagai berikut:



Gambar 3. Use case diagram aplikasi mobile

Dari Gambar 3 dapat dijelaskan prosesnya sebagai berikut:

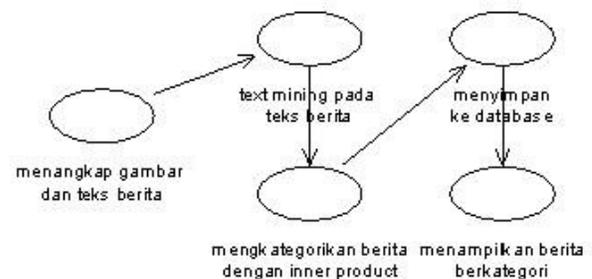
Tabel 6
Kata kunci bidang politik

Kata	Jumlah	Politik	Ekonomi	Olahraga	Kriminal	Hiburan
Judul						
jambret	1	0	0	0	110(330)	0
hajar	1	0	0	0	65(195)	0
massa	1	0	0	0	57(171)	0
Berita						
pemuda	1	0	0	0	0	0
keroyok	1	0	0	0	65(65)	0
warga	1	0	0	0	0	0
babak	1	0	0	95(95)	0	0
belur	1	0	0	0	0	0
hajar	1	0	0	0	65(65)	0
massa	1	0	0	0	57(57)	0
duga	1	0	0	0	17(17)	0
jambret	1	0	0	0	110(110)	0
wanita	1	0	0	0	0	0
Mansyur	1	0	0	0	0	0
sangka	1	0	0	0	10(10)	0
bawa	1	0	0	0	0	0
markas	1	0	0	0	9(9)	0
polisi	1	0	0	0	9(9)	0
Sunggal	1	0	0	0	0	0
Hasil						
jumlah	=	0	0	95	1038	0

- 1) Pertama kali *user* menggunakan aplikasi pada telepon genggam adalah mengambil foto dengan menggunakan kamera yang terintegrasi dengan telepon genggam.
- 2) *User* akan menulis judul dan isi dari berita
- 3) Sebelum berita dikirimkan *user* harus memasukkan *username* dan *password*, yang diperoleh ketika mendaftar sebagai penulis berita di blog. Proses otentikasi sendiri adalah proses pencocokan user dengan data di database.
- 4) Pada proses pengiriman berita akan dibuat koneksi ke internet menggunakan *HttpConnection*.

3.5 Aplikasi Blog

Aplikasi Blog pada server seperti di bawah ini:
Dari Gambar 4 dapat dijelaskan untuk aplikasi blog secara garis besar terdiri dari 5 bagian, yaitu:



Gambar 4. Activity diagram rancangan sistem aplikasi blog

- 1) Menangkap berita dari aplikasi telepon genggam
- 2) Memproses *text mining* untuk berita yang masuk
- 3) Memproses *inner product* untuk mendapatkan kategori dari berita

- 4) Berita disimpan ke database
- 5) Menampilkan berita berkategori berdasarkan kronologis waktu pengiriman

4 PENGUJIAN DAN ANALISA

Terdapat tiga uji coba yaitu penentuan kata kunci setiap kategori berita, uji coba pengiriman berita menggunakan telepon genggam dan uji coba pengkategorian berita.

4.1 Uji Coba Penentuan Kata Kunci Setiap Kategori Berita

Untuk menentukan kata kunci pada berita politik, ekonomi, olahraga, kriminal dan hiburan digunakan sebanyak 250 berita. Berita ini diambil dari situs berita (cetak.kompas.com dan www.kabarindonesia.com). Masing-masing kategori berita memerlukan sebanyak 50 berita untuk menentukan kata kunci.

Pada Tabel 7,8,9,10,11 beberapa kata kunci dari setiap kategori berita. Pada tabel 7 merupakan kata kunci dibidang politik, contoh kata partai mempunyai nilai 174, caleg mempunyai nilai 117 dan sebagainya.

Tabel 7
Kata kunci bidang politik

Kata	jumlah	Kata	jumlah
partai	174	Komisi	40
Caleg	117	pemerintah	38
Pemilu	114	Undang	36
suara	114	Rakyat	36
KPU	102	presiden	34
politik	93	kabupaten	34
anggota	71	PPP	13
DPR	67	Bawaslu	12
masyarakat	59	aktivis	11
calon	58	konstituen	10
Surat	52	politis	10
Daerah	49	Panwas	10
Golkar	48	Pansus	9
wakil	43	Wagub	8
legislatif	41	Menteri	8

Tabel 8
Kata kunci bidang ekonomi

Kata	jumlah	Kata	jumlah
Rp	228	produk	72
Usaha	188	industri	67
ekonomi	133	PT	66
Bank	131	Departemen	63
harga	122	dagang	58
pasar	107	Kredit	53
miliar	104	Nilai	52
triliun	101	infrastruktur	52
uang	95	biaya	47
ekspor	88	Transaksi	44
Produksi	86	Pertamina	43
dana	83	fiskal	42
tani	77	inflasi	41
Syariah	77	jual	41

Tabel 9
Kata kunci bidang olahraga

Kata	jumlah	Kata	jumlah
main	186	Federer	51
gol	118	gelar	49
menang	101	cetak	42
tim	95	Phelps	38
laga	94	Liverpool	38
Piala	79	Stadion	37
Latih	77	Milan	36
Nadal	77	musim	34
kalah	71	set	31
juara	68	serang	30
Liga	64	olimpiade	30
tanding	62	klasemen	29
poin	59	babak	29
bola	57	Real	28
final	57	emas	28

4.2 Uji Coba Pengiriman Berita Menggunakan Telepon Genggam dan Blog yang Sudah Diupload

Uji coba sistem dilakukan dilakukan dengan menggunakan telepon genggam sony ericsson tipe K630i dan blog sudah diupload yang terkoneksi dengan internet. Langkah pertama yang dilakukan adalah install file jar dari aplikasi *mobile blog* pada telepon genggam, pada

Tabel 10
Kata kunci bidang kriminal

Kata	jumlah	Kata	jumlah
polisi	110	aparatus	10
korban	65	ilegal	9
sangka	57	kabur	9
laku	44	jahat	8
duga	38	Polres	8
tangkap	31	Shabu	8
rampok	27	bunuh	8
sita	17	Satpol	8
kasus	17	gerebek	8
Tewas	16	senjata	8
bukti	11	sekap	7
razia	11	selidik	7
Aniaya	11	gasak	7
keroyok	11	Tawur	7
Markas	10	Massa	6

Tabel 11
Kata kunci bidang hiburan

Kata	jumlah	Kata	jumlah
Musik	115	tampil	18
film	90	sinetron	17
band	61	pop	16
album	50	Ajang	15
lagu	44	finalis	15
Acara	28	rilis	14
tembang	28	produser	14
TV	26	nyanyi	13
Bintang	25	akting	13
karya	23	vokal	12
artis	21	panggung	11
Rock	20	duet	11
idola	20	genre	10
musisi	19	sutradara	10
siar	19	studio	10

percobaan kali ini akan digunakan telepon Sony Erricsson dengan tipe K630i

Untuk webnya, dilakukan upload pada web server, dengan alamat URL <http://devie.or.id:8080/mobileblog>. Pertama kali dijalankan akan ditampilkan tampilan awal dari aplikasi *mobile blog* seperti yang ditunjukkan pada Gambar 5.



Gambar 5. Tampilan awal mobile blog di telepon genggam

Kemudian menuju menu kamera, dan kamera aktif, pilih menu ambil untuk mengambil gambar, dan hasil gambar akan ditampilkan seperti yang ditunjukkan pada Gambar 6.



Gambar 6. Menampilkan hasil foto pada telepon genggam

Pilih menu proses dan akan tampil halaman untuk menuliskan judul dan isi berita seperti yang ditunjukkan pada Gambar 7 dan 8.



Gambar 7. Tampilan menulis judul berita di telepon genggam



Gambar 8. Tampilan menulis isi berita di telepon genggam

Pilih menu kirim dan menuju form untuk mengisi *username* dan *password* untuk otentikasi seperti yang ditunjukkan pada Gambar 9.



Gambar 9. Tampilan otentikasi *user* pada telepon genggam

Jika otentikasi benar, dilanjutkan pada proses pengiriman berita seperti yang ditunjukkan pada Gambar 10.



Gambar 10. Tampilan proses mengirimkan berita di telepon genggam

Pengiriman berita selesai, maka muncul status pengiriman berita seperti yang ditunjukkan pada Gambar 11.



Gambar 11. Tampilan status pengiriman berita di telepon genggam

Berita akan ditampilkan di halaman blog dan ditampilkan sesuai dengan kategori beritanya seperti yang ditunjukkan pada Gambar 12.



Gambar 12. Hasil pengiriman berita menggunakan telepon genggam

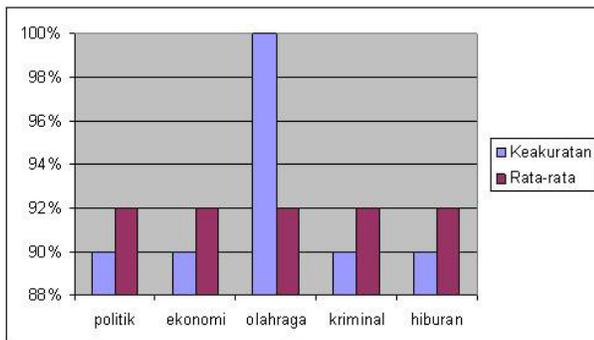
4.3 Uji Coba Pengkategorian Berita

Pada uji coba ini dilakukan analisa terhadap pengiriman 10 berita pada tiap-tiap kategori, berita diambil dari situs berita, www.kompas.com. Selanjutnya dihitung tingkat keakuratan proses pengkategorian berita. Tingkat keakuratan dihitung berdasarkan berita yang masuk terklasifikasi sesuai dengan kategori berita tersebut. Hasil uji coba ditunjukkan pada Tabel 12 dan Gambar 13.

Hasil percobaan pada Tabel 12 dapat dilihat rata-rata kemampuan sistem untuk mengkategorian berita cukup baik, sekitar 92%. Rata-rata

Tabel 12
Tingkat keakuratan pengkategorian berita

Kategori	Benar	Salah	Tingkat Keakuratan (%)
politik	9	1	90
ekonomi	9	1	90
olahraga	10	0	100
kriminal	9	1	90
hiburan	9	1	90
Rata-Rata			92



Gambar 13. Tingkat keakuratan pengkategorian berita

kesalahan pengkategorian disebabkan karena berita asli terkategorikan dalam kategori tertentu tapi dalam berita tersebut terkandung banyak kata kunci dari berita lain, misal berita asli dikategorikan sebagai berita kriminal, tetapi di dalam berita tersebut banyak mengandung kata-kata yang termasuk dalam kata kunci kategori ekonomi, sehingga proses perhitungan akan menentukan berita tersebut termasuk dalam berita ekonomi.

Semakin banyak berita yang dimasukkan dalam training maka kata kunci untuk tiap bidang yang didapat juga semakin banyak, hal ini akan meningkatkan tingkat keakuratan dari proses pengklasifikasian berita.

5 KESIMPULAN

Setelah dilakukan pengujian dan analisa program, maka dapat diperoleh kesimpulan sebagai berikut :

- 1) Aplikasi telepon genggam mampu digunakan untuk mengirimkan berita berupa teks dan gambar ke blog(server). Sistem

mampu mengkategorikan secara otomatis berita-berita yang dikirimkan dengan menggunakan metode inner product dengan tingkat keakuratan sekitar 92%.

- 2) Tingkat keakuratan pengkategorian berita sangat ditentukan oleh penentuan kata kunci tiap berita dan jumlahnya. Semakin banyak berita yang dimasukkan maka kata kunci dan jumlah kata kunci tersebut semakin valid.

DAFTAR PUSTAKA

- [1] id.wikipedia.org/wiki/Blog
- [2] en.wikipedia.org/wiki/Mobile_blogging
- [3] Tessa Badriyah, materi kuliah Text Mining
- [4] Entin Martiana dan Andri Permana. klasifikasi topik TA sesuai dengan kemampuan menggunakan metoda inner product. Surabaya: PENS-ITS. 2008
- [5] Arifin Zainal Agus dan Setiono Novan Ari, Klasifikasi Dokumen Berita Kejadian Berbahasa Indonesia dengan Algoritma Single Pass Clusterin. Surabaya: Teknik Informatika-ITS
- [6] Budi Raharjo, Tuntunan Pemrograman Java untuk Handphone Bandung. Informatika : 2007
- [7] M Sholahuddin, Pemrograman J2ME belajar cepat pemrograman perangkat telekomunikasi mobile. Bandung. Informatika : 2006
- [8] Leo Willyanto Santoso, Sukanto Tedjokusumo, Marcel Renaldy Soetanto. Aplikasi Pelaporan Berita Emergeni Secara Visual dan Tekstual Lewat Telepon Selular. Surabaya: Universitas Kristen Petra. 2005
- [9] Jonathan Knudsen. Introduction to mobile blogging. <http://developers.sun.com>. 2003



ditekuni adalah *Soft Computing*.

Afrida Helen lahir di Bukittinggi, memperoleh gelar Sarjana Teknik (ST) pada Jurusan Teknik Elektro pada tahun 1993 dan magister Komputer (M.Kom) pada tahun 2004, keduanya dari Institut Teknologi Sepuluh Nopember (ITS) Surabaya. Ia adalah pengajar pada jurusan Teknik Informatika, Politeknik Elektronika Negeri Surabaya. Bidang penelitian yang



Yuliana Setiowati lahir di Surabaya, memperoleh gelar Sarjana pada Jurusan Teknik Informatika ITS pada tahun 2002 dan magister pada jurusan Teknik Informatika ITS pada tahun 2007. Beliau adalah pengajar pada jurusan Teknik Informatika, Politeknik Elektronika Negeri Surabaya. Bidang penelitian yang ditekuni adalah *Soft Computing*.