

## **SURAT KETERANGAN**

Nomor: 350/UNUSA/Adm-LPPM/IV/2019

Lembaga Penelitian dan Pengabdian Kepada Masyarakat (LPPM) Universitas Nahdlatul Ulama Surabaya menerangkan telah selesai melakukan pemeriksaan duplikasi dengan membandingkan artikel-artikel lain menggunakan perangkat lunak **Turnitin** pada tanggal 22 April 2019.

Judul : Klasifikasi Kanker Pada Citra Mammogram  
Penulis : Miftahus Sholihin  
Identitas : Seminar Nasional Teknologi Informasi dan Komunikasi  
(SENTIKA) Maret 2016  
No. Pemeriksaan : 2019.04.23.148

Dengan Hasil sebagai Berikut:

**Tingkat Kesamaan diseluruh artikel (*Similarity Index*) yaitu 16%**

Demikian surat keterangan ini dibuat untuk digunakan sebagaimana mestinya

Surabaya, 23 April 2019

Ketua LPPM,



Dr. Istas Pratomo, S.T., M.T.

NPP. 16081074

**LPPM Universitas Nahdlatul Ulama Surabaya**

Website : [lppm.unusa.ac.id](http://lppm.unusa.ac.id)

Email : [lppm@unusa.ac.id](mailto:lppm@unusa.ac.id)

Hotline : 0838.5706.3867

# Paper

*by* Akhir Rudin 7

---

**Submission date:** 22-Apr-2019 01:43PM (UTC+0700)

**Submission ID:** 1116778222

**File name:** Miftahus\_Sholihin\_5\_-\_Biro\_Administrasi\_Umum\_Kepegawaian.pdf (793.2K)

**Word count:** 8538

**Character count:** 54117

## KLASIFIKASI KANKER PADA CITRA MAMMOGRAM

Miftahus Sholihin<sup>1</sup>

*1Program Studi Teknik Informatika, Fakultas Teknik, Universitas Islam Lamongan*

*Jl. Veteran No. 53 A Lamongan*

*Telp. (0322) 324706*

*E-mail: miftah.sholihin@gmail.com*

### ABSTRAKS

Kanker payudara merupakan salah satu jenis kanker yang paling berbahaya yang menyerang perempuan di seluruh dunia. Analisa terhadap citra mammografi yang dilakukan oleh radiologis masih dilakukan secara manual, hal ini dapat menyebabkan ketidakakuratan dan memakan waktu yang lama untuk menganalisa citra mammografi tersebut. Oleh sebab itu dibutuhkan alat bantu yang dapat membantu radiologis dalam menganalisa citra mammografi. Pada penelitian ini dirancang suatu sistem yang bisa membantu radiologis dalam melakukan klasifikasi kanker pada citra mammogram. Secara garis besar sistem yang dibangun terdiri dari 4 proses utama yaitu *preprocessing*, segmentasi, ekstraksi ciri, dan klasifikasi.

*Kata Kunci: Kanker payudara, Preprocessing, segmentasi, ekstraksi ciri, klasifikasi*

### ABSTRACT

*Breast cancer is one of the most dangerous types of cancer that affects women across the world. Analysis of mammography images performed by radiologists is still done manually, this can lead to inaccuracies and take a long time to analyze the mammography image. Therefore needed a tool that can assist radiologists in analyzing mammography images. In the study designed system that can help radiologists in the classification of cancer on mammogram image. The system built consist of 4 main process, preprocessing, segmentation, feature extraction, and classification.*

*Keywords: Breast cancer, preprocessing, segmentation, feature extraction, classification.*

12

### 1. PENDAHULUAN

Pengolahan citra telah banyak diaplikasikan diberbagai bidang termasuk bidang kedokteran. Pada bidang kedokteran pengolahan citra digunakan sebagai alat bantu untuk mendiagnosis suatu penyakit salah satu contoh pengolahan citra dibidang kedokteran adalah untuk mendeteksi keberadaan kanker payudara. Upaya untuk mendeteksi keberadaan kanker payudara dilakukan dengan alat bantu mammogram yang menggunakan sinar X sehingga menghasilkan citra mammografi. Berdasarkan citra mammografi ini seorang radiologis atau dokter ahli bisa menemukan perubahan pada payudara meskipun penderita belum merasakan gejala apa-apa (Indrati, 2009). Analisa terhadap citra mammografi yang dilakukan oleh radiologis masih dilakukan secara manual (Sjamsuhidayat, 1997). Sehingga hasil yang didapatkan masih bersifat subyektif. Seringkali radiologis merasa tidak yakin dengan hasil analisisnya sehingga memerlukan pihak kedua untuk melakukannya yang berdampak pada biaya yang harus dikeluarkan (Uyun, 2011). Oleh sebab itu dibutuhkan alat bantu yang dapat mempercepat kinerja dari radiologis dengan menggunakan teknik pengolahan citra digital.

Pada penelitian sebelumnya (Indrati, 2009) melakukan penelitian mengenai representasi bentuk dari tumor payudara. Berdasarkan bentuk dari tumor payudara tersebut bisa ditentukan jenis dari

kanker yang ada pada payudara tersebut. Proses pengambilan daerah yang dicurigai sebagai kanker dilakukan secara manual. Pada tahun yang sama (Surendiran, 2009) melakukan penelitian tentang klasifikasi kanker pada citra mammogram yang didasarkan dari bentuk daerah yang terkena kanker. Penelitian ini menggunakan fitur *shape properties* yang terdiri dari *area*, *equivdiameter*, *entropi*, *circularity*, *compactness*, dan *shape index*. Daerah yang dicurigai sebagai kanker dilakukan *cropping* manual. Penelitian lain yang dilakukan (Alolfe, 2008) melakukan klasifikasi kanker pada citra mammogram dengan menggabungkan beberapa fitur yaitu histogram, GLCM, fraktal, dan bentuk. Metode klasifikasi yang digunakan pada penelitian ini adalah KNN. Pada penelitian ini proses pengambilan daerah yang dicurigai sebagai kanker masih dilakukan secara manual.

Berdasarkan penelitian yang telah dilakukan dimana proses pengambilan daerah yang dicurigai sebagai kanker masih dilakukan secara manual, maka pada penelitian ini akan mencoba merancang suatu sistem yang mampu menentukan daerah yang dicurigai sebagai kanker tanpa melakukan *cropping* secara manual. Setelah daerah yang dicurigai didapatkan proses berikutnya yang dilakukan adalah melakukan klasifikasi dengan KNN dan menggunakan *moment invariant* sebagai ciri fiturnya.

## 11 PEMBAHASAN

### 2.1 Pengolahan Citra Digital

Secara harfiah, citra (*image*) adalah gambar pada bidang dwimatra atau dua dimensi (Munir, 2004). Ditinjau dari sudut pandang matematis, citra merupakan fungsi menerus (*continue*) dari intensitas cahaya pada bidang dwimatra. Sumber cahaya menerangi objek, objek memantulkan kembali sebagian dari berkas cahaya tersebut. Pantulan cahaya ditangkap oleh alat-alat optik, misalnya mata, kamera, scanner, dan sebagainya.

### 5 2.2 Perbaikan Citra

Perbaikan citra bertujuan untuk meningkatkan kualitas tampilan citra untuk mengkonversi suatu citra agar memiliki format yang lebih baik sehingga citra tersebut menjadi lebih mudah untuk diolah dengan mesin (komputer). Perbaikan terhadap suatu citra dapat dilakukan dengan operasi titik (*point operation*), operasi spasial (*spatial operation*), operasi geometri (*geometric operation*), dan operasi aritmatik (*arithmetic operation*).

### 17 2.3 *Contrast Limited Adaptive Histogram Equalization (CLAHE)*

CLAHE dapat digunakan sebagai alternatif pengganti histogram ekualisasi. Histogram ekualisasi bekerja pada seluruh citra, sedangkan CLAHE beroperasi pada daerah kecil pada citra yang disebut dengan blok. Setiap blok akan ditingkatkan kontrasnya, sehingga histogram dari wilayah sekitar cocok untuk histogram tertentu. Setelah proses pemerataan, CLAHE akan menggabungkan blok-blok yang sudah di cari histogramnya. CLAHE juga dapat digunakan untuk menghindari *noise* yang ada pada citra dengan membatasi kontras pada daerah *homogen*. Daerah ini dapat dicirikan sebagai puncak yang tinggi pada histogram terkait dengan daerah kontekstual oleh karena banyak piksel tergabung di dalam kisaran abu-abu yang sama. Lereng yang terkait dengan skema penempatan *gray-level* menjadi terbatas dengan CLAHE. Hal ini dapat dicapai dengan hanya memungkinkan jumlah maksimum dari piksel di setiap kelompok data yang terkait dengan histogram lokal (Zuiderveld, 1994).

### 4 2.4 *Thresholding*

Pengembangan citra (*image thresholding*) merupakan metode yang paling sederhana untuk melakukan segmentasi. *Thresholding* digunakan untuk mengatur jumlah derajat keabuan yang ada pada citra. Proses *thresholding* ini pada dasarnya adalah proses pengubahan kuantisasi pada citra. Untuk mendapatkan hasil segmentasi yang bagus, beberapa operasi perbaikan kualitas citra dilakukan terlebih dahulu untuk mempertajam batas antara objek dengan latar belakangnya (Usman, 2005).

*Thresholding* adalah suatu proses yang digunakan untuk menghasilkan citra *biner* yaitu

citra yang memiliki dua nilai tingkat keabuan, yaitu : hitam dan putih bergantung apakah nilai piksel asli tersebut lebih besar atau lebih kecil dari nilai  $T$ . Piksel akan diubah menjadi putih jika nilai tingkat keabuannya lebih besar daripada  $T$ , dan akan diubah menjadi hitam jika nilai tingkat keabuannya lebih kecil atau sama dengan  $T$ .

### 2.5 Dilasi

Dilasi adalah operasi menambah jumlah piksel-piksel dari *boundary* objek. Dilasi bisa juga diartikan sebagai operasi morfologi yang "memperbesar" atau "mengentalkan" sebuah objek dari sebuah citra. Jumlah piksel-piksel yang ditambahkan dari sebuah objek ditentukan oleh ukuran dan bentuk dari struktur elemen. Jika  $A$  dan  $B$  adalah anggota  $Z^2$  (Gonzalez, 2008), dilasi  $A$  dan  $B$  dinyatakan  $A \oplus B$  dan didefinisikan dengan Persamaan (1).

$$A \oplus B = \{z | (\bar{B})_2 \cap A \neq \emptyset\} \quad (1)$$

### 2.6 Erosi

Erosi adalah operasi menghapus piksel-piksel yang ada pada *boundary* objek. Erosi juga dapat diartikan sebagai operasi morfologi yang "menyusutkan" atau "menguruskan" sebuah objek pada citra biner. Seperti pada dilasi, "penyusutan" dan "pengurusan" juga ditentukan oleh struktur elemen yang digunakan. Jika  $A$  dan  $B$  himpunan dalam  $Z^2$  (Gonzalez, 2008), erosi  $A$  oleh  $B$  dinyatakan dengan  $A \ominus B$  dan didefinisikan dengan Persamaan (2).

$$A \ominus B = \{z | (B_2) \subseteq A\} \quad (2)$$

### 2.7 Momen

Momen (*moment*) dapat menggambarkan suatu objek dalam hal area, posisi, orientasi dan parameter terdefinisi lainnya. Untuk fungsi kontinu 2 dimensi  $f(x,y)$ , momen orde  $(p+q)$  didefinisikan dengan Persamaan (3) (Gonzalez, 2008).

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad (3)$$

dimana  $p, q = 0, 1, 2$ . Untuk implementasi dalam bentuk digital, maka Persamaan (3) akan menjadi bentuk Persamaan (4).

$$m_{pq} = \sum \sum x^p y^q f(x, y) \quad (4)$$

dengan orde dari momen adalah  $(p + q)$ .  $x$  dan  $y$  menyatakan koordinat titik, sedangkan  $f(x,y)$  menyatakan intensitas titik.

### 2.8 KNN

*K-Nearest Neighbor (KNN)* adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data *learning* yang 3 raknya paling dekat dengan objek tersebut. KNN termasuk algoritma *supervised learning* dimana *query instance* yang baru diklasifikasikan berdasarkan

mayoritas dari kategori pada KNN. Kelas yang paling banyak muncul yang akan menjadi kelas hasil klasifikasi.

Algoritma *K-Nearest Neighbor* sangatlah sederhana, bekerja berdasarkan jarak terpendek dari *query instance* ke *training sample* untuk menentukan KNN-nya. *Training sample* diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi *training sample*.

Pada umumnya, jarak antar data dihitung dengan menggunakan *Euclidean distance* pada Persamaan (5).

$$D(x, y) = \sqrt{\sum_{i=1}^d (x_i - y_i)^2} \quad (5)$$

Dimana matrik  $D(x,y)$  adalah jarak dari kedua vektor  $x$  dan  $y$  dari matrik dengan ukuran  $d$  dimensi.

## 2.9 Perancangan Sistem

Secara garis besar perancangan sistem terbagi menjadi 4 tahapan utama seperti Gambar 1 yaitu *pre-processing*, segmentasi, ekstraksi ciri, dan klasifikasi.

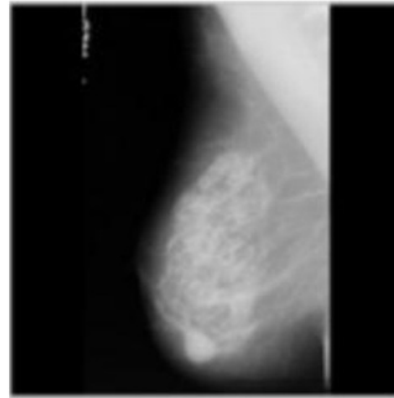


Gambar 1. Perancangan Sistem

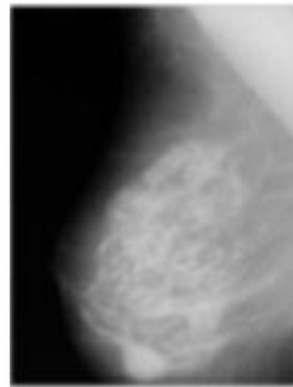
*Pre-processing* bertujuan untuk memperbaiki kualitas citra yang siap digunakan untuk proses berikutnya. *Pre-processing* dilakukan untuk membuang bagian citra yang kurang penting seperti label citra atau *background*. *Pre-processing* yang

dilakukan pada penelitian ini dibagi menjadi beberapa tahap yaitu:

- a) Penghapusan *Background* dan Label Citra. Proses ini bertujuan untuk menghilangkan *background* warna hitam dan label citra yang berada di sisi kiri maupun kanan dari citra mammogram. Gambar 2 adalah citra asli dari *database* MIAS. Hasil dari proses penghapusan *background* dan label citra ditunjukkan pada Gambar 3.



Gambar 2. Citra Mammogram MIAS



Gambar 2. Hasil Proses Penghapusan Label dan *Background*

- b) Segmentasi *Pectoral Mucle*

Segmentasi *pectoral muscle* ini bertujuan untuk mendapatkan *pectoral muscle*. Proses segmentasi ini dilakukan dengan *thresholding* sesuai dengan Persamaan (6) (Tayel, 2010).

$$T = \text{mean} + \left(\frac{2}{3}\right) * \text{standar deviasi} \quad (6)$$



Setelah *pectoral muscle* di dapatkan, maka langkah berikutnya adalah melakukan pengurangan citra hasil proses penghapusan *background* dengan *pectoral muscle*. Gambar 4 adalah hasil dari segmentasi *pectoral muscle* sedangkan Gambar 5 adalah hasil dari proses pengurang antara citra hasil penghapusan *background* dan label citra dengan citra hasil *pectoral muscle*.



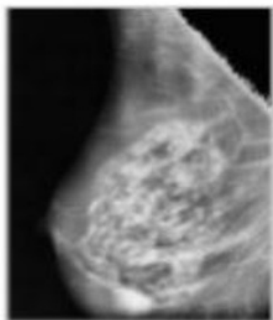
Gambar 4. Hasil Segmentasi *Pectoral Muscle*



Gambar 5. *Pectoral Muscle*

c) Peningkatan Kontras Citra

Tujuan dari proses ini adalah meningkatkan kontras dari citra, hal ini dilakukan karena kontras dari citra mammogram yang sangat rendah sehingga perlu ditingkatkan tingkat kekontrasannya.



Gambar 6. Hasil Akhir *Preprocessing*

Peningkatan kontras dilakukan dengan *contrast limited adaptive histogram equalization (CLAHE)*. CLAHE memperbaiki kontras citra pada *region-region* kecil pada citra. Ukuran *region* yang digunakan adalah [8 8] dan nilai *clip limit* adalah 0.01. Gambar 6 adalah hasil dari peningkatan kontras citra.

Proses berikutnya adalah melakukan segmentasi terhadap objek yang diduga sebagai kanker. Proses ini dilakukan dengan *thresholding* yang sesuai dengan Persamaan (6). Hasil akhir dari segmentasi terkadang masih terdapat *noise* yang harus dihilangkan agar didapatkan objek yang benar-benar sebagai kanker. Untuk mengatasi permasalahan tersebut, maka dilakukan operasi erosi terhadap citra hasil segmentasi. Erosi merupakan proses penghapusan titik-titik objek (1) menjadi bagian dari latar (0), berdasarkan dengan struktur elemen yang digunakan. Hasil operasi erosi menghasilkan citra yang kecil, sehingga citra tersebut dibalkan dengan melakukan operasi dilasi. Dilasi merupakan proses penggabungan titik-titik latar (0) menjadi bagian dari objek (1) yang didasarkan pada struktur elemen yang digunakan. Operasi dilasi dapat dilakukan dengan menggantikan setiap titik pada citra dengan struktur elemen yaitu mengganti setiap titik pada struktur elemen dengan citra. Gambar 7 adalah hasil dari proses *thresholding* sedangkan Gambar 8 adalah hasil dari proses erosi.

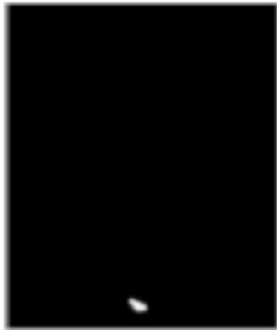
Gambar 7. Hasil Proses Segmentasi





Gambar 8. Hasil Proses Erosi

Gambar 9 adalah hasil akhir dari proses segmentasi adalah sebuah *region of interest* (ROI) yang diindikasikan sebagai objek yang terkena kanker, baik itu kanker jenis *benign* maupun kanker jenis *malignant*. ROI yang dihasilkan dari proses segmentasi inilah yang nantinya akan dilakukan proses ekstraksi ciri dan juga digunakan untuk proses klasifikasi.



Gambar 9. Hasil Akhir Proses Segmentasi

Setelah didapatkan daerah yang dicurigai sebagai kanker, maka proses berikutnya adalah melakukan ekstraksi ciri terhadap objek tersebut. Pada penelitian ini ciri yang digunakan adalah ciri *moment invariant*. Proses ekstraksi ciri diawali dengan menghitung besarnya nilai momen orde ke-1 sampai dengan momen orde ke-3. Nilai dari momen mencerminkan area atau luasan dari daerah objek kanker. Setelah nilai momen didapatkan, proses selanjutnya adalah menghitung momen pusat. Proses ini dilakukan karena momen masih berpengaruh terhadap rotasi, ukuran, dan penskalaan. Proses berikutnya adalah menghitung nilai *miu*, hal ini dilakukan karena momen pusat masih bergantung pada ukuran. Proses berikutnya adalah menghitung nilai *eta* dan proses yang terakhir dalam proses ekstraksi ciri adalah menghitung ke tujuh nilai *moment invariant* yang ada.

Proses terakhir dalam penelitian ini adalah melakukan klasifikasi kanker tersebut. Proses ini diawali dengan menentukan berapa jumlah ketetapan yang akan digunakan atau menentukan nilai *k*. Proses selanjutnya adalah menghitung nilai jarak antara data *testing* dengan data *training*.

## 2.10 Hasil

Data citra yang digunakan dalam penelitian ini berasal dari *database* MIAS. Uji coba dibagi menjadi 2 tahap yaitu uji coba untuk proses segmentasi dan uji coba untuk proses klasifikasi. Tabel 1 adalah hasil dari uji coba proses segmentasi citra mammogram. Hasil segmentasi untuk citra mammogram dengan kategori *benign* memberikan hasil akurasi sebesar 68% yang didapatkan dari total 62 data citra mammogram yang berhasil tersegmentasi berjumlah 42 sedangkan 20 citra lainnya hasilnya tidak sesuai dengan informasi yang ada pada *database* MIAS. Sedangkan untuk citra mammogram kategori *malignant* memberikan hasil akurasi sebesar 70% dari 51 citra yang berhasil tersegmentasi dengan benar berjumlah 36, sedangkan 15 tidak berhasil disegmentasi. Total akurasi yang didapatkan untuk proses segmentasi citra mammogram baik *benign* maupun *malignant* sebesar 69%.

Tabel 1. Hasil Segmentasi Citra Mammogram

Jenis	Jumlah	Hasil		
		Benar	Salah	Akurasi
Benign	62	42	20	68.00%
Malignant	51	36	15	70.00%
Total	113	78	35	69.00%

Tingkat akurasi dari proses segmentasi sangat dipengaruhi oleh *background* dari citra mammogram (*Fatty*, *Fatty-grandular*, atau *Dense-grandular*). Citra mammogram yang memiliki *background grandular* (*Fatty-grandular* dan *Dense-grandular*) mempunyai sifat *background* seperti kumpulan awan yang biasanya sulit dideteksi adanya jenis kanker pada citra ini. Sedangkan citra yang memiliki *background fatty* cenderung memperlihatkan adanya ketidaknormalan yang terlihat dengan jelas. Hal ini disebabkan karena citra mammogram yang memiliki jenis *background* ini memiliki kontras yang sedikit terang atau jelas, sehingga proses pengambilan ROI yang dicurigai sebagai kanker akan dapat dilakukan dengan baik.

Sedangkan Tabel 2 adalah hasil uji coba dari proses klasifikasi. Berdasarkan hasil pengujian yang ditunjukkan oleh Tabel 2 dapat disimpulkan bahwa ketika nilai *k* atau jumlah ketetanggannya 9 sistem memberikan hasil akurasi terbesar yaitu

76,92%. Dari 43 data citra mammogram sistem dapat mengenali dengan benar 31 citra mammogram, sedangkan 12 citra yang lain tidak dapat dikenali atau salah dalam pengenalannya. Citra dengan *background fatty* memberikan hasil pengenalan yang paling baik jika dibandingkan dengan citra yang memiliki *background* selain *fatty*, hal ini dikarenakan citra dengan jenis *background fatty* memiliki kontras yang lebih terang bila dibandingkan dengan citra yang lainnya.

Berdasarkan uji coba yang sudah dilakukan, sistem yang dibangun masih melakukan kesalahan dalam menentukan kelas dari citra mammogram yang dimasukkan oleh pengguna. Hal ini disebabkan oleh beberapa faktor antara lain citra mammogram yang digunakan memiliki banyak *noise*. Disamping itu hasil dari proses klasifikasi juga dipengaruhi oleh jenis dari *background* dari citra itu sendiri. Citra dengan *background fatty* memberikan hasil yang paling bagus jika dibandingkan dengan *background* yang lainnya.

Tabel 2. Hasil Klasifikasi Citra Mammogram

Nilai K	Kelas	Hasil			Akurasi
		Benar	Salah		
1	Normal	10	7	50,56%	
	Benign	7	6		
	Malignant	4	9		
2	Normal	5	12	50,56%	
	Benign	10	3		
	Malignant	4	9		
3	Normal	15	2	63,35%	
	Benign	7	6		
	Malignant	3	10		
4	Normal	12	5	56,45%	
	Benign	8	5		
	Malignant	3	10		
5	Normal	11	6	51,58%	
	Benign	8	5		
	Malignant	2	11		
6	Normal	15	2	63,35%	
	Benign	9	4		
	Malignant	1	12		
7	Normal	16	1	72,00%	
	Benign	10	3		
	Malignant	3	10		
8	Normal	13	4	57,47%	
	Benign	9	4		
	Malignant	1	12		
9	Normal	17	0	76,92%	
	Benign	11	2		
	Malignant	3	10		
10	Normal	17	0	73,08%	
	Benign	11	2		

	Malignant	1	12	
11	Normal	17	0	75,00%
	Benign	11	2	
	Malignant	2	11	
12	Normal	17	0	71,15%
	Benign	10	5	
	Malignant	1	12	
13	Normal	16	1	70,14%
	Benign	11	2	
	Malignant	1	12	
14	Normal	14	3	66,18%
	Benign	11	2	
	Malignant	2	11	
15	Normal	14	3	66,18%
	Benign	9	4	
	Malignant	4	9	

### 3. KESIMPULAN

Berdasarkan percobaan yang telah dilakukan, dapat disimpulkan bahwa: Segmentasi dengan menggunakan metode *thresholding* dan operasi morfologi dapat digunakan untuk menentukan daerah yang terindikasi terkena kanker. Disamping itu klasifikasi dengan menggunakan ciri *moment invariant* dengan nilai ketetapan atau  $k = 9$  memberikan akurasi sebesar 76,9% dalam mengelompokkan kanker payudara ke dalam dua kelas yaitu normal dan abnormal.

### PUSTAKA

- Ahmad, U. 2005. *Pengolahan Citra Digital dan Teknik Pemrogramannya*. Graha Ilmu. Yogyakarta.
- Alolfe, M.A., Youssef, A.B.M., Kadah, Y.M. & Mohamed, A.S. 2008. *Development of a Computer-Aided Classification System for Cancer Detection from Digital Mammograms*. Radio Science Conference, NRSC 2008. National, pp. 1.8, 18-20 March 2008.
- Gonzalez, R.C. & Woods, R.E. 2008. *Digital Image Processing*, Third Ed. Pearson Prentice Hall, New Jersey.
- Indrati, A., Madenda, S. & Missaoui, R. 2009. *Representasi Blok Tumor Payudara dengan Kode Rantai*. Seminar on Application and Research in Industrial Technology, SMART Yogyakarta, 22 Juli 2009.
- Munir, R. 2004. *Pengolahan citra digital pendekatan Algoritmik*. Edisi Pertama. Informatika Bandung.
- Sjamsuhidajat, R. & Jong, W. 1997. *Buku ajar Ilmu Bedah. Kedokteran EGC*. Jakarta.
- Tayel, M. & Mohsen, A. 2010. *Breast Boarder Boundaries Extraction Using Statistical Properties of Mammogram*. Signal Processing (ICSP). IEEE 10th International Conference on, pp.2468,2471, 24-28 Oct. 2010.



- Uyun, S. & Harjoko, A. 2011. *Deteksi Abnormalitas Pada Citra Mammogram dengan Gray-Level Co-Occurrence Matrix*.
- Zuiderveld, K., 1994, *Contrast Limited Adaptive Histogram Equalization*, Chapter VII.5.

## RANCANGAN PENGEMBANGAN SOSIAL INTELIJENSI BISNIS AKADEMIK MEMANFAATKAN MEDIA SOSIAL TWITTER

Meiko Pris Hadiano<sup>1</sup>, Irya Wisnubhadra<sup>2</sup>, Albertus Joko Santoso<sup>3</sup>

<sup>1,2,3</sup>Program Studi Magister Teknik Informatika, Pasca Sarjana, Universitas Atma Jaya Yogyakarta

Jl. Babarsari 43 Yogyakarta 55281

Telp. (0274) 48758

E-mail: meikohadiano@gmail.com, irya@mail.uajy.ac.id, albjoko@mail.uajy.ac.id

### ABSTRAKS

Seiring berjalannya waktu, perkembangan media sosial sangat pesat. Salah satu contoh media sosial yang terkenal adalah twitter. Kebiasaan masyarakat menuliskan perasaan ke media sosial twitter dapat dimanfaatkan dalam berbagai analisis. Intelijensi Bisnis akademik saat ini belum dapat dikatakan sebagai alat pengambil keputusan yang baik karena hanya bersumber pada data internal. Sosial Intelijensi bisnis merupakan evolusi dari intelijensi bisnis dengan memanfaatkan media sosial. Sosial Intelijensi Bisnis dapat membantu akademisi dalam mengambil keputusan. Sosial Intelijensi Bisnis ini tentu memerlukan rancangan gudang data yang memadai agar dapat dijadikan alat analisis. Permasalahan timbul ketika data twitter memiliki data yang besar dan tidak terstruktur, sehingga rancangan intelijensi bisnis perlu disesuaikan. Apache hadoop merupakan salah satu media penyimpanan data besar dan sql server 2008 r2 sebagai pengolah data intelijensi bisnis. Penelitian ini akan menghasilkan rancangan gudang data dari hasil penggalian data twitter.

*Kata Kunci: Intelijensi Bisnis, Media Sosial, Sosial Intelijensi Bisnis*

## 1. PENDAHULUAN

### 1.1 Latar Belakang

Seiring berjalannya waktu media sosial terus berkembang. Pada tahun 2014, pengguna media sosial di dunia mencapai 1,79 milyar jiwa. Berbagai fitur baru muncul sebagai pendukung atau penarik minat pengguna. Media Sosial memiliki beberapa fungsionalitas yaitu *identity*, *presence*, *sharing*, *conversation*, *groups*, *reputation*, dan *relationship*. Dengan fungsi tersebut media sosial dapat dikonfigurasi ke tingkatan lebih tinggi seperti alat analisis (Kietzmann, Hermkens, McCarthy and Silvestre, 2011) dan dapat menggunakan media sosial untuk memprediksi kejadian di masa depan (Zeng, Chen, Lusch and Li, 2010).

Intelijensi Bisnis atau biasa disebut dengan intelijensi bisnis saat ini sudah banyak digunakan oleh perusahaan. Selain berguna dalam mengambil keputusan, intelijensi bisnis juga mampu melakukan prediksi. intelijensi bisnis memanfaatkan data-data transaksional terdahulu yang diolah sehingga mampu digunakan kembali. Saat ini banyak vendor yang sudah membuat sebuah intelijensi bisnis dengan berbagai model. intelijensi bisnis termasuk dalam DSS sistem karena intelijensi bisnis memiliki model matematika dan memiliki metodologi analisis untuk mengenerate informasi dan pengetahuan yang berguna untuk pengambilan keputusan yang kompleks (Vercellis, 2009). Beberapa industri di Indonesia seperti PT Bursa efek Surabaya dan PT Kliring Penjaminan efek Indonesia sudah menggunakan intelijensi bisnis dengan cara masing-masing (Badan Pengawas Pasar Modal Dan Lembaga Keuangan, 2007). Hal ini bertujuan untuk meningkatkan keuntungan persaingan antar

perusahaan (Williams and Williams, 2007). Dengan perlakuan yang berbeda tiap perusahaan dapat membuat kesempatan berkembangnya intelijensi bisnis lebih jauh.

Perkembangan intelijensi bisnis yang pesat memicu beberapa tantangan dalam menggali data. Beberapa poin mendasar dalam tantangan intelijensi bisnis antara lain berbagai jenis sumber data mempengaruhi perbedaan representasi data dan kebutuhan akan data besar untuk menggali lebih dalam keuntungan (Ljubljana, Turk and Jaklič, 2010). Sumber data yang digunakan merupakan data transaksi sehingga kurang mampu berinteraksi dengan pengguna, sehingga menyebabkan intelijensi bisnis kurang bisa disebut sebagai alat pengambil keputusan (Meredith and O'Donnell, 2011). Universitas merupakan sebuah intstusi dalam bidang pendidikan. Berbagai universitas sudah menerapkan intelijensi bisnis dalam melakukan analisis. Kebutuhan analisis yang lebih lengkap merupakan hal yang dibutuhkan universitas saat ini. Dengan memanfaatkan media sosial, kebutuhan data universitas akan menjadi lebih lengkap. Universitas dapat menggunakan data twitter seperti opini mahasiswa sebagai alat analisis tambahan.

Sosial intelijensi bisnis merupakan gabungan dari media sosial dengan intelijensi bisnis. Beberapa ciri dari sosial intelijensi bisnis adalah memiliki sumber yang banyak dan tidak terbatas, memiliki data tidak terstruktur dan jenis yang berbeda (Heijnen, 2012).

Salah satu alat mengatur data yang banyak adalah dengan menggunakan hadoop. Hadoop merupakan salah satu alat pengolah data yang biasa disebut dengan *big data*. Permasalahan yang akan

muncul adalah bagaimana mempersiapkan rancangan inteligensi bisnis dengan media sosial. Persiapan ini tentu memerlukan beberapa perubahan dalam inteligensi bisnis. penelitian ini akan merancang sosial inteligensi bisnis akademik agar dapat merumuskan opini dalam twitter menjadi bahan analisis.

## 1.2 Tinjauan Pustaka

Inteligensi Bisnis merupakan kegiatan yang mengumpulkan, menganalisis data sehingga dapat digunakan dalam melakukan pengambilan keputusan yang lebih baik. inteligensi bisnis sering digunakan perusahaan maupun organisasi dalam pengambilan keputusan atau perancangan strategis. Menurut Turban inteligensi bisnis merupakan kerangka kerja konseptual untuk mendukung keputusan bisnis, inteligensi bisnis terdiri dari arsitektur, data warehouse, tool analisis dan aplikasi (Efraim Turban, 2011)

Beberapa peneliti melihat inteligensi bisnis dari sudut pandang keputusan yang diambil membuat pengertian bahwa inteligensi bisnis sistem bertujuan untuk menyediakan pengetahuan dengan tools dan metodologi yang memberikan fasilitas untuk membuat keputusan yang efektif dan *Timely*. Efektif berarti membuat seorang pengambil keputusan dapat bergantung pada informasi dan pengetahuan yang dihasilkan. Sebagai hasilnya mereka dapat membuat keputusan yang lebih baik dan dapat merancang rencana agar objective dapat dicapai. *Timely* berarti keputusan harus tepat waktu dan tidak terlambat terbaru saat ini. intelligen dalam bisnis adalah suatu proses yang sistematis yang dapat dipastikan update dan berkaitan dengan perusahaan saingan. Sistem intelligence mengacu pada set program yang digunakan oleh manager untuk akses informasi harian dunia marketing (Bahrami, Arabzad and Ghorbani, 2012)

Beberapa fitur yang ada dalam inteligensi bisnis *Organizational Learning*. Sebuah proses yang mencakup penemuan pengetahuan baru dan penyebarannya ke organisasi yang membutuhkan. Pengetahuan tersebut digunakan untuk improve internal dan external proses perusahaan

*Processing of smart*. Sebuah proses yang kompleks mencakup analisis dan penilaian dari informasi dan *decision support* dan kerjasama penuh dari berbagai keputusan yang secara langsung mempengaruhi kinerja kedepan yang merupakan kinerja keputusan terbaik dalam organisasi

inteligensi bisnis terdiri dari beberapa komponen diantaranya adalah data source (secara internal dan external, database transaksional beberapa faktor lain yang mempengaruhi data), data mart (termasuk deskripsi dari tiap data yang akan digunakan), dan alat generate reports (digunakan untuk menggabungkan skill analisis dan data yang dapat memprediksikan keputusan yang dibutuhkan bisnis proses). Proses yang ada dalam inteligensi bisnis

dikategorikan sebagai berikut (Azma and Mostafapour, 2012)

- a. *Planning and conducting* : planning dari awal hingga akhir proses intelegen. Bagian ini merupakan bagian dari manager yang memutuskan awal dari proses dengan merumuskan beberapa pertanyaan.
- b. *Obtaining information* : data dikumpulkan sesuai dengan kebutuhan. Data biasanya merupakan data transaksional atau pengetahuan tambahan. Data tersebut harus diproses kembali agar terlihat jelas.
- c. *Processing Information* : Data yang telahdikumpulkanakandihubungan, identifikasi, dan analisis hubungan tiap data dan kemudian data dapat disimpan dalam bentuk statis. Proses ini meliputi proses transform dan load data.
- d. *Analysis and produce information* :dengan menggunakan beberapa teknik, intelligence dapat dihasilkan. Bagianiniakan merespon berupa report data, diagram yang digunakan sebagai penjawab pertanyaan dari fase pertama dari berbagai sudut pandang.

Media sosial didefinisikan oleh banyak ahli yang kemudian disimpulkan sebagai "situs internet dimana seseorang dapat berinteraksi secara bebas, berinteraksi, dan mendiskusikan informasi (kebanyakan tentang kehidupan masing-masing) dengan menggunakan campuran multimedia dari kata personal, gambar, video, dan audio. Pernyataan tersebut dikemukakan oleh Curtis pada tahun 2013 untuk memperbaharui definisi lama yang dikemukakan oleh Evans pada tahun 2008 tentang media sosial sebagai kumpulan aplikasi berbasis internet yang dibangun berdasarkan ideology dan teknologi web 2.0, juga mengizinkan untuk pembuatan dan pertukaran konten sendiri (Mcintyre, 2014). Pada umumnya orang jaman sekarang lebih mengandalkan media sosial untuk belajar dari peristiwa yang mempengaruhi orang-orang dan mempelajarinya dari pengalaman. Akibatnya saat ini media sosial menjadi peran utama dalam membentuk opini publik (Patil, 2010). Saat ini sudah banyak media komunikasi secara online seperti chatting, email, blog, website, dan sosial networking seperti facebook, twitter, linkedln, google+ dan yang lainnya. Munculnya media sosial tidak lepas dari pengembangan web 2.0 dan jaringan internet yang mudah untuk didapatkan sehingga dapat digunakan untuk berkomunikasi (Ahuja and Medury, 2010). Beberapa peranan dari media sosial adalah media sosial dapat digunakan sebagai modal informasi, yang dapat mendeteksi potensi bisnis dengan cara penyebaran informasi dengan mudah dan tepat, dapat digunakan sebagai media transfer pengetahuan secara virtual, dan sebagai alat yang memberi kontribusi pembentukan jaringan positif (Georgescu and Popescu, 2015).

Sosial Intelijen Bisnis merupakan sebuah pengembangan dari Intelijensi bisnis yang mengandalkan web 2.0. Bohringer dalam penelitiannya tahun 2010 mengatakan bahwa internet di masa depan memiliki kemungkinan dalam menggali data lebih pintar. Saat ini internet sedang berkembang berorientasi pada service seperti cloud computing atau open API. Hal ini yang memicu perkembangan intelijensi bisnis ke tahap penggalian data sosial (Böhringer, Gluchowski, Kurze and Schieder, 2010). Keberadaan intelijensi bisnis saat ini masih belum dikatakan sebagai *decision support system* karena belum memenuhi kriteria rumusan dss. intelijensi bisnis kurang dalam hal Komunikasi, Koordinasi, dan kemampuan dalam filter informasi secara aktif. Dengan memanfaatkan web 2.0 maka ketiga kriteria tersebut bisa ditutup (Meredith and O'Donnell, 2011). Beberapa langkah yang perlu diperhatikan dalam melakukan persiapan intelijensi bisnis dengan media sosial dirumuskan oleh Scott yaitu (Walters, 2013) :

- Memahami posisi dari perusahaan. Beberapa hal yang perlu diperhatikan seperti bagaimana cara perusahaan memanfaatkan media sosial, apakah perusahaan mampu menggabungkan sosial media, dan meninjau ulang infrastruktur IT perusahaan
- Identifikasi tujuan yang akan dicapai. Meyakinkan bahwa media sosial sejalan dengan bisnis proses.
- Membuat rencana sosial bisnis seperti target, sasaran, definisi dari keadaan dan detail aksi yang akan dilakukan

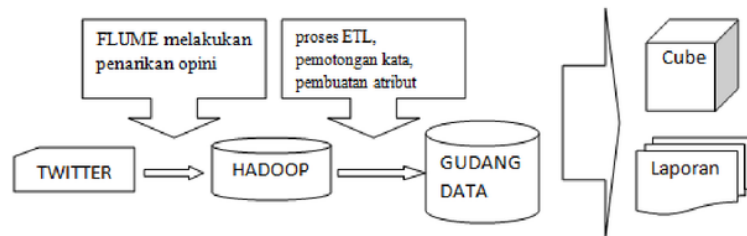
- Membuat desain sosial intelijensi bisnis dan rencana transformasi IT. Hal ini mencakup kebutuhan untuk membuat proses baru, termasuk perkiraan waktu dan biaya
- Implementasi sosial intelijensi bisnis. Kegiatan yang dilakukan mencakup, menggabungkan data terstruktur dan data tidak terstruktur yang dapat melihat perkembangan pelanggan, menganalisa potensi terbesar dari pelanggan, dan sebisa mungkin *real-time* analisis atau mendekati *real-time*.
- Membuat kebijakan untuk memantau kinerja organisasi.

### 1.3 Metode Penelitian

Penelitian ini memiliki beberapa konsep pertanyaan yang menjadi dasar pembuatan *cube* yaitu :

- Berapa banyak opini atau keluhan dari mahasiswa pada suatu mata pelajaran tertentu pada waktu, angkatan, kategori tertentu melalui twitter.
- Berapa banyak mahasiswa yang sering melakukan keluhan dilihat dari letak, waktu, dan angkatan tertentu melalui twitter.
- Berapa banyak opini yang ditujukan ke dosen atau matakuliah tertentu melalui twitter.

Penelitian ini menggunakan metologi pengamatan dan percobaan. Pengamatan dilakukan dalam hal penarikan data twitter menjadi json. Hadoop yang digunakan adalah



Gambar 1. Skema penelitian

*Hortonworks* (<http://hortonworks.com>). Hasil data twitter yang dihasilkan hadoop menggunakan *flume* salah satu service yang digunakan untuk mengumpulkan, menggabungkan, dan memindahkan log data yang besar. Percobaan yang dilakukan adalah memetakan hasil json data twitter hadoop menjadi bentuk skema gudang data. basis data yang digunakan adalah sql server 2008 r2. Kasus contoh yang diambil adalah Universitas Atma Jaya Yogyakarta bidang pengajaran.

## 2. PEMBAHASAN

Alasan penggunaan hadoop pada penelitian ini adalah data opini twitter memiliki data yang besar dan cukup rumit. Pertanyaan dasar yang dijelaskan diatas jika diuraikan, maka fakta yang diperlukan adalah jumlah opini yang ditulis melalui twitter. Pertanyaan pertama membutuhkan dimensi waktu, angkatan, dan kategori. Pertanyaan kedua membutuhkan fakta banyak opini yang dilihat dari dimensi mahasiswa, letak, waktu, dan identitas mahasiswa seperti prodi dan fakultas. Pertanyaan



ketiga membutuhkan fakta opini dengan dimensi dosen dan mata kuliah. Dari beberapa pertanyaan diatas, dapat diambil garis besar dimensi yaitu mahasiswa, dosen, matakuliah, letak, waktu, kategori, dan sentimen. Pengambilan data dari hadoop akan dilakukan melalui linked server dan bantuan *odbc connector* menuju staging tabel. Gambaran umum terdapat pada gambar 1.

## 2.1 Pengolahan Bentuk JSON

Penarikan data dari *twitter* akan memanfaatkan beberapa *hashtag* diantaranya #UAJY sebagai penanda bahwa ditujukan ke uajy, #NM(nomor nim mahasiswa) sebagai penanda identitas dari mahasiswa, #MK(mata kuliah) sebagai penanda mata kuliah yang akan digunakan. Hasil dari pengamatan yang dilakukan pada hadoop berupa struktur data *twitter* yang dihasilkan *flume*.

```
{ "extended_entities": { "media": [ { "display_url": "pic.twitter.com/eso  
rtI0dAB", "indices": [ 97, 119 ], "size  
s": { "small": { "w": 340, "h": 512, "res  
ize": "fit" }, "large": { "w": 679, "h":  
1024, "resize": "fit" }, "thumb": { "w":  
150, "h": 150, "resize": "crop" }, "me  
dium": { "w": 600, "h": 904, "resize": "  
fit" } } ] }
```

Gambar 2. Bentuk raw dari json hasil penggalian hadoop flume

Gambar 2 merupakan bentuk dari json *twitter* data. Hasil ini akan dibaca dengan *hive* dan akan diolah dalam *sql server*. Bentuk dari media video atau foto dalam *twitter* berbentuk sebuah link. Dalam penelitian ini tidak menggunakan video dan foto sebagai alat analisis sehingga akan dibuang sebelum masuk gudang data.

Dalam mengolah data json ke dalam *sql server*, diperlukan data staging untuk melakukan perhitungan sentimen dan mengubah data dari string akan digunakan dalam gudang data akan dipilih sesuai dengan kebutuhan. Data akan dihubungkan menggunakan linked server. Sebelum menambah gudang data, text *twitter* harus terlebih dahulu disaring setiap kata sehingga dapat dicari sentimen dari tiap opini *twitter*.

Permasalahan dalam mengolah data JSON adalah tipe data berupa string. SQL server memiliki banyak tipe data sebagai dasar data. Agar JSON dapat diakses ke dalam SQL server, tipe data string harus diolah dan diubah menjadi data yang dibutuhkan. Beberapa data yang harus diperhatikan adalah tipe data *Datetime*. Tipe data ini diperlukan dalam pembentukan dimensi

waktu sehingga string dalam json harus dipotong dan diubah menjadi *Datetime*. Permasalahan berikutnya adalah bentuk array dalam JSON yang dibaca sebagai string oleh *sql server*.

```
{ "screen_name": "nametosee",  
  "name": "realname", "friends_count": 55  
2, "followers_count": 841, "statuses_co  
unt": 13126, "verified": false, "utc_off  
set": "32400", "time_zone": "Alaska" }
```

Gambar 3. Struktur array user dalam JSON

Contoh array dalam JSON (gambar 3) berbentuk tipe data string. Phil Factor melakukan pengujian terhadap array string dan membentuk tipe data baru bernama hirarki dengan memanfaatkan rekursi data sehingga dapat ditampilkan dalam tabel tampilan masing-masing level array (Factor, 2010). Kode dari Factor termasuk metode sederhana karena masih dalam bentuk *sql* murni. JSON yang dihasilkan *apache flume* dapat dibaca dengan menggunakan *apache hive* dengan delimiter tertentu sehingga pemindahan data ke dalam *sql server* sudah dalam bentuk tabel. Beberapa elemen yang dapat diambil dari *twitter* menggunakan *apache flume* :

- Id *Twitter*. Id ini di generate secara otomatis dari *flume*
- Created\_at* merupakan string waktu pembuatan *twitter*.
- source* merupakan cara atau perantara opini dibuat. Beberapa aplikasi memiliki fungsi untuk share ke *twitter*
- favorited* dan *retweeted count*. Bagian ini akan mengeluarkan berapa banyak pengguna yang menjadikan opini ini sebagai favorit dan berapa banyak yang me *retweeted* ulang opini
- entities* merupakan atribut dalam *twitter* yang dapat berbentuk link, *user mention*, dan *hashtags*.
- Text* merupakan tempat opini *twitter* berada. Bagian ini berisi teks penuh opini yang dibuat beserta *entities* nya.
- User* merupakan bentuk array dari deskripsi sumber penulis opini *twitter*. Array ini terdiri dari *screen\_name*, *name*, *follower\_count*, dan format waktu pengguna

Atribut utama yang harus diambil adalah *created\_at*, *text*, dan *user*, sedangkan atribut lain dapat diambil sebagai data tambahan. Penelitian ini membatasi penggunaan array hanya pada satu level karena proses rekursi Factor terbatas pada level sehingga data yang memiliki banyak level array akan dihindari.

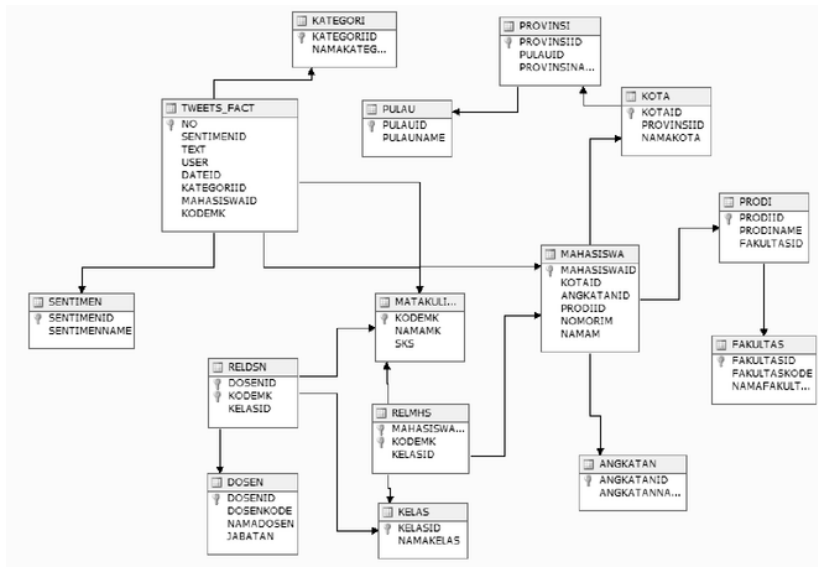
Dalam melakukan penghitungan sentimen, dibutuhkan sebuah fungsi yang dapat memisahkan kata tiap kalimat. Setelah berhasil memisahkan kata, mulai dilakukan *stemming* pada tiap kata dan pemberian bobot sederhana pada tiap kata. Perlu diperhatikan bahwa pembagian kata harus sesuai dengan kamus bahasa Indonesia dan memiliki kata sifat tiap kalimat.

## 2.2 Pengolahan data dalam SQL Server

Proses pemisahan kata, penyusunan *datetime* dan penghitungan sentimen kalimat, dilakukan pada area *staging*. Area ini menampung data yang berasal dari hadoop sebelum dimasukkan ke dalam gudang data. Beberapa fungsi yang dibutuhkan dalam area ini antara lain :

- Fungsi untuk memotong *hashtags* nomor induk mahasiswa, kategori opini dan matakuliah yang menjadi objek. Kedua atribut tersebut akan digunakan sebagai referensi tambahan.
- Fungsi mengubah string *created\_at* menjadi bentuk *Datetime*. Fungsi ini memiliki proses pemotongan kata dan penyusunan ulang kata.
- Fungsi sentimen digunakan untuk mencari sentimen opini. Opini akan dibagi menjadi dua yaitu positif dan negatif.

Pemanggilan fungsi tersebut diatur tiap harinya dengan metode yang cepat, karena data yang dihasilkan memiliki jumlah yang besar.



Gambar 4. Skema Sosial Intelijensi Bisnis menggunakan data twitter

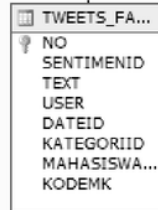
Skema pada gambar 4 memiliki dua bagian yaitu tiga tabel media sosial dan sembilan tabel gudang data.

- Tabel *tweets\_fact* merupakan tabel fakta utama. Ukuran dapat ditentukan dari banyaknya opini yang masuk dan seberapa banyak opini yang dihasilkan dari dimensi tertentu
- Tabel kategori merupakan tabel yang berisi dimensi kategori. Kategori dalam penelitian ini dibagi menjadi dua yaitu cara mengajar dosen dan materi kuliah. Dalam masa pengembangan kategori dapat diperluas lebih dalam sesuai dengan permintaan.

- Tabel sentimen merupakan tabel dimensi sentimen opini yaitu positif dan negatif.
- Tabel dosen, mahasiswa, fakultas, mata kuliah dan beberapa tabel sisanya merupakan tabel tambahan dengan memanfaatkan gudang data yang sudah ada. Universitas Atmajaya Yogyakarta memiliki tabel yang dibutuhkan seperti diatas, sehingga penelitian ini tidak membutuhkan tabel tambahan.

Pemanfaatan tabel dalam gudang data yang sudah dibuat digunakan untuk melakukan relasi antara dosen, mata kuliah, dengan mahasiswa dan fakultas, daerah asal dengan mahasiswa.

Beberapa tabel tambahan digunakan sebagai penghubung antara tabel-tabel utama sehingga dapat dilakukan pencarian.



NO	SENTIMENID	TEXT	USER	DATEID	KATEGORIID	MAHASISWA...	KODEMK
----	------------	------	------	--------	------------	--------------	--------

Gambar 5. Data dalam tabel fakta

Penggunaan data fakta dalam penelitian ini hanya berfokus pada beberapa atribut penting (gambar 5). Penggunaan *source* dapat digunakan dalam beberapa permasalahan seperti media mahasiswa dalam mengirimkan opini twitter.

Pengambilan data dari *hadoop* dilakukan menggunakan SQL Server Agent yang dijadwalkan sehari sekali penarikan.

Pengintegrasian data menggunakan *SQL Server Integration Service* (SSIS). Proses ini mencakup proses refresh yang digunakan untuk *real-time integration* dan integrasi data fakta secara bertahap.

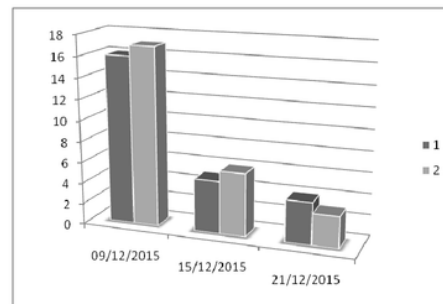
### 2.3 Pembentukan *Cube* Intelijensi Bisnis

Penelitian ini menyusun *cube* menggunakan sembilan dimensi yaitu :

1. Dimensi mahasiswa. Dimensi ini memiliki hubungan dengan beberapa tabel yaitu prodi, fakultas, dan tabel asal mahasiswa. Dimensi ini digunakan untuk mengukur keaktifan mahasiswa dalam menuliskan opini di media sosial
2. Dimensi angkatan. Dimensi ini digunakan untuk melihat fakta opini berdasarkan angkatan mahasiswa
3. Dimensi prodi. Dimensi ini digunakan untuk melihat fakta opini berdasarkan prodi mahasiswa. Dimensi prodi memiliki hirarki fakultas, prodi dan kelas
4. Dimensi dosen. Dimensi ini digunakan untuk melihat fakta opini berdasarkan dosen.
5. Dimensi kota. Dimensi ini bertujuan untuk mengelompokkan mahasiswa berdasarkan daerah asal yang ikut berpartisipasi dalam opini twitter
6. Dimensi matakuliah. Dimensi ini digunakan untuk melihat dari sudut pandang matakuliah yang terikat dengan mahasiswa, dosen dan kelas
7. Dimensi kategori. Dimensi ini merupakan dimensi utama dalam melihat kategori yang mendapat opini dalam pembelajaran.

8. Dimensi sentimen. Dimensi yang ada secara umum dalam sosial intelijensi bisnis tentang kepuasan atau emosi dari opini tersebut. Penelitian ini melakukan penghitungan sentimen dilihat dari perkata dan menggunakan metode klasifikasi sederhana.
9. Dimensi *Time*. Dimensi yang digunakan untuk melihat sudut pandang waktu pembuatan opini twitter. Dimensi ini digenerate secara langsung di server oleh *SQL Server Analysis Service* (SSAS)

Pembentukan *cube* dapat dibagi menjadi beberapa bagian sesuai dengan kebutuhan universitas. Dalam penelitian ini semua kebutuhan yang disebutkan sebelumnya dapat dipenuhi dengan menggunakan sebuah *cube*. Pengujian dilakukan terhadap 50 opini dari twitter.



Gambar 6. Grafik tabel fakta menggunakan dimensi waktu, sentimen, dan kategori tertentu

Hasil dari gambar 6 menunjukkan bahwa relasi dari *cube* yang dibentuk menggunakan gudang data tersebut berhasil digambarkan. Grafik ini menunjukkan pada waktu tertentu terdapat opini positif dan negatif dengan kategori pertama.

### 3. KESIMPULAN

Penelitian ini memiliki beberapa hasil yaitu :

1. Penarikan data twitter yang banyak dapat memanfaatkan *hadoop* sebagai data tampung dan *flume* sebagai penggalan log data twitter. Penggunaan *hadoop* akan menghasilkan data twitter berbentuk json raw.
2. Pengolahan gudang data untuk pembentukan *cube* Intelijensi Bisnis dapat dilakukan dengan berbagai platform. Dalam penelitian ini UAJY menggunakan sql server sebagai gudang data.
3. Rancangan sosial intelijensi bisnis dapat disusun dengan skema bintang dengan dimensi utama yaitu sentimen, dan kategori



- sebagai jenis pembeda opini. Pembuatan gudang data memanfaatkan data inteligensi bisnis yang sudah ada dengan mengambil tabel mahasiswa, dosen, dan matakuliah sebagai tambahan analisa.
4. Terdapat sembilan dimensi yang digunakan yaitu dim mahasiswa, dim dosen, dim mata kuliah, dim angkatan, dim prodi, dim kota, dim sentimen, dim kategori.
  5. Hasil dari rancangan gudang data berhasil diujikan sederhana dan menghasilkan grafik tertentu.

Penelitian ini masih memiliki beberapa teori dasar, sehingga masih dapat dikembangkan sesuai dengan inteligensi bisnis yang sudah berjalan. Penelitian ini juga dapat dikembangkan lebih jauh mengenai banyak kategori sehingga mendapatkan data lebih besar. Proses penghitungan sentimen juga dapat dikembangkan lagi menggunakan metode-metode yang lebih akurat.

#### PUSTAKA

- Ahuja, V., and Medury, Y., 2010. Corporate blogs as e-CRM tools – Building consumer engagement through content management. *Journal of Database Marketing & Customer Strategy Management*, [online] 17(2), hlm.91–105.
- Azma, F., and Mostafapour, M.A., 2012. Business intelligence as a key strategy for development organizations. *Procedia Technology*, [online] 1, hlm.102–106.
- Badan Pengawas Pasar Modal Dan Lembaga Keuangan, D.K.R., 2007. Laporan tim studi tentang Implementasi Business Intelligence. *Business*, hlm.48.
- Bahrami, M., Arabzad, S.M., and Ghorbani, M., 2012. Innovation In Market Management By Utilizing Business Intelligence: Introducing Proposed Framework. *Procedia - Social and Behavioral Sciences*, [online] 41, hlm.160–167.
- Böhringer, M., Gluchowski, P., Kurze, C., and Schieder, C., 2010. A Business Intelligence Perspective on the Future Internet. In: *AMCIS 2010 Proceedings*. [online] p.267.
- Efrain Turban, L.V., 2011. *Information Technology in the Digital Economy*. [online] *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*.
- Factor, P. (2010, November 15). *Consuming JSON Strings in SQL Server*. Dipetik November 30, 2015, dari <https://www.simple-talk.com>: <https://www.simple-talk.com/sql/t-sql-programming/consuming-json-strings-in-sql-server/>
- Georgescu, M., and Popescu, D., 2015. Social Media – The New Paradigm of Collaboration and Communication for Business Environment. *Procedia Economics and Finance*, [online] 20(2012), hlm.277–282.
- Heijnen, J., 2012. *Social Business Intelligence*. Delft University of Technology, Thesis.
- Kietzmann, J.H., Hermkens, K., McCarthy, I.P., and Silvestre, B.S., 2011. Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, [online] 54(3), hlm.241–251.
- Ljubljana, A.P., Turk, T., and Jaklič, J., 2010. Conceptual Model of Business Value of Business Intelligence Systems. *Management: Journal of Contemporary Management*, 15(1), pp.5–29.
- Mcintyre, K., 2014. The Evolution of Social Media from 1969 to 2013: A Change in Competition and a Trend Toward Complementary, Niche Sites. *The Journal of Social Media in Society*, 3(2).
- Meredith, R., and O'Donnell, P., 2011. A Framework for Understanding the Role of Social Media in Business Intelligence Systems. *Journal of Decision Systems*, 20(3), hlm.263–282.
- Patil, M.Y., 2010. Social Media and Customer Relationship Management . *Journal of Business and Management (IOSR-JBM)*, hlm.27–32.
- Vercellis, C., 2009. *Business Intelligence: Data Mining and Optimization for Decision Making. Business Intelligence: Data Mining and Optimization for Decision Making*.
- Walters, S. (2013). Beyond Listening: Six Steps for Integrating and Acting on Social Media. *Business Intelligence Journal*, 13.
- Williams, S., and Williams, N., 2007. *The Profit Impact of Business Intelligence*. [online] *The Profit Impact of Business Intelligence*.
- Zeng, D., Chen, H., Lusch, R., and Li, S.H., 2010. Social media analytics and intelligence. *IEEE Intelligent Systems*, 25(6), hlm.13–16.



## SISTEM PENGOLAHAN DATA ADMINISTRASI POSYANDU MELATI 1 DAN MELATI 2 KELURAHAN RAMBUTAN JAKARTA TIMUR

Za'imatun Niswati<sup>1</sup>, Ni Wayan Parwati<sup>2</sup>, Mei Lestari<sup>3</sup>, Adhi Susano<sup>4</sup>

<sup>1,2,3,4</sup> Teknik Informatika Universitas Indraprasta PGRI Jakarta

Jl Nangka No.58c, Tanjung Barat, Jagakarsa, Jakarta Selatan

E-mail: zaimatunnis@gmail.com, wayan.parwati@gmail.com, mei.lestari6@gmail.com, adhi.susano@gmail.com

### 13 STRAKS

Posyandu merupakan wadah pemeliharaan kesehatan yang dilakukan dari, oleh dan untuk masyarakat yang dibimbing oleh petugas terkait. Saat ini posyandu tidak hanya dikenal sebagai pos penimbangan anak dan balita tetapi semakin diperluas dengan mencakup bidang kesehatan secara umum, pendidikan, ekonomi hingga lingkungan/pemanfaatan pekarangan. Dengan sistem pengolahan data administrasi maka proses pengolahan data pada posyandu menjadi lebih cepat dan akurat. Metode yang dilakukan dalam merancang sistem pengolahan data administrasi posyandu adalah metode analisis dan perancangan sistem berorientasi objek (Object Oriented Modelling) menggunakan tools pengembangan system Unified Modelling Language (UML). Pengujian yang dilakukan adalah pengujian validasi, pengujian kuesioner dan pengujian aplikasi. Hasil penelitian adalah sistem pengolahan data administrasi posyandu dapat meningkatkan produktivitas posyandu di lihat dari laporan bulanan yang dapat dicetak tepat waktu dan lengkap sesuai kebutuhan petugas posyandu.

Kata Kunci: posyandu, data administrasi, unified modelling language (uml), object oriented modelling.

### ABSTRACT

Posyandu is a health care organization that organized of, by and for the community and guided by the relevant officers. Currently, Posyandu not only known as a post to weigh infants and young children but increasingly cover general health care, education, economics and also including environment/utilization of the yard. By using data administration system, performance of data administration process are increase. The method is performed in designing administration data processing system Posyandu is a method of analysis and design of object-oriented system (Object Oriented Modelling) using system development tools Unified Modeling Language (UML). In this paper shows performance of data administration system by using validation test, quitionaire test and application test. It shows that data administration system increase posyandu's productivity. Monthly report can be completed and printed on schedule to the needs of posyandu officers.

Kata Kunci: posyandu, administration data, unified modelling language (uml), object oriented modelling.

## 1. PENDAHULUAN

### 1.1 Latar Belakang

Posyandu adalah wadah pemeliharaan kesehatan yang dilakukan dari, oleh dan untuk masyarakat yang dibimbing petugas terkait. (Departemen Kesehatan RI, 2006). Posyandu adalah pusat kegiatan masyarakat dalam upaya pelayanan kesehatan dan keluarga berencana (Effendy, 1998).

Pada masa lalu posyandu hanya dikenal sebagai pos penimbangan anak dan balita, kini peran posyandu diperluas mencakup bidang kesehatan secara umum, pendidikan, ekonomi hingga lingkungan/pemanfaatan pekarangan.

Masalah yang ada pada posyandu melati 1 dan melati 2 kelurahan Rambutan adalah minimnya kegiatan yang dilakukan, data pasien posyandu (ibu dan balita) dikelola secara manual, sehingga data mudah rusak, tidak tersusun secara teratur, proses temu kembali (retrieval) data lama sulit dilakukan. Hal tersebut menyebabkan sulitnya melakukan evaluasi kegiatan pada posyandu melati 1 dan melati

2, selain itu lambatnya proses pelaporan kegiatan yang dilakukan di kedua posyandu tersebut.

Menurut Departemen Kesehatan RI (2001) meningkatkan kualitas pelayanan posyandu merupakan tujuan dari revitalisasi posyandu. Tujuan dari revitalisasi posyandu yaitu untuk meningkatkan kemampuan atau pengetahuan dan keterampilan teknis serta dedikasi kader di posyandu, meningkatkan kualitas dan kuantitas, meningkatkan peran serta masyarakat dan kemitraan dalam penyelenggaraan dan pembiayaan kegiatan posyandu, pembinaan dan pendamping teknis, tenaga professional dan tokoh masyarakat, termasuk unsur Lembaga Swadaya Masyarakat (LSM).

Dengan Sistem Pengolahan Data Administrasi Posyandu ini, maka akan tersusun data-data yang akurat yang sangat penting untuk menentukan kebijakan. Sebagai contoh data tentang kesehatan ibu, bayi dan balita, akurasi data ini akan mendukung pemerintah dalam membuat program

yang dibutuhkan sesuai kondisi ibu bayi dan balita serta lansia.

Selain data balita, ibu dan lansia, **6** Sistem Pengolahan Data Administrasi Posyandu ini juga bermanfaat untuk meningkatkan koordinasi, pembinaan, fasilitasi, advokasi, dan bantuan yang berkaitan dengan fungsi dan kinerja posyandu.

Dengan latar belakang masalah di atas penulis mengambil judul Sistem Pengolahan Data Administrasi Posyandu Melati 1 dan Melati 2 Kelurahan Rambutan Jakarta Timur.

## 1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah merancang sistem pengolahan data administrasi Posyandu berbasis *java* yang dapat memudahkan proses pengolahan, pencarian, dan pembuatan laporan pada Posyandu Melati 1 dan Melati 2.

## 1.3 Referensi

Sistem adalah kumpulan dari elemen-elemen yang berinteraksi untuk mencapai suatu tujuan tertentu, suatu sistem merupakan suatu jaringan kerja dari prosedur-prosedur yang saling berhubungan, berkumpul bersama-sama untuk melakukan suatu kegiatan atau untuk menyelesaikan suatu sasaran yang tertentu (Jogiyanto, 2005)

Posyandu adalah singkatan dari Pos Pelayanan Terpadu yang merupakan suatu wadah komunikasi alih teknologi dalam pelayanan kesehatan masyarakat dari masyarakat, oleh masyarakat dan untuk masyarakat dengan dukungan pelayanan serta pembinaan teknis dari petugas kesehatan dan keluarga berencana (Sembiring, 2004 : 1)

Posyandu memiliki nilai strategis untuk proses pengembangan sumber daya manusia sejak dini, yaitu peningkatan mutu manusia di masa yang akan datang akibat dari proses pertumbuhan dan perkembangan manusia.

Terdapat posyandu antara lain :

1. Menurunkan angka kematian bayi, ibu hamil dan melahirkan
2. Meningkatkan peran serta masyarakat untuk mengembangkan kegiatan kesehatan dan KB serta kegiatan lain yang menunjang tercapainya masyarakat sehat sejahtera
3. Sebagai wahana reproduksi keluarga sejahtera, gerakan ketahanan keluarga dan gerakan ekonomi keluarga sejahtera. **3**

Manfaat Sistem Pengolahan Data Posyandu antara lain sebagai berikut:

1. Menjadi bahan acuan bagi seluruh kader **3** maupun semua instansi yang terlibat dalam Posyandu untuk memahami permasalahan yang ada sehingga dapat mengembangkan kegiatan yang tepat sasaran dan disesuaikan dengan kebutuhan sasaran.
2. Menyediakan informasi yang tepat waktu mengenai pengelolaan Posyandu, agar

berbagai pihak yang berperan dalam pengelolaan Posyandu dapat menggunakannya untuk membina Posyandu bagi kepentingan masyarakat (Sembiring, 2004: 5).

**10** *Unified Modelling Language (UML)* adalah sebuah "bahasa" yg telah menjadi standar dalam industri untuk visualisasi, merancang dan mendokumentasikan sistem piranti lunak. UML menawarkan sebuah standar untuk merancang model sebuah sistem. Dengan menggunakan UML kita dapat membuat model untuk semua jenis aplikasi piranti lunak, dimana aplikasi tersebut dapat berjalan pada piranti keras, sistem operasi dan jaringan apapun, serta ditulis dalam bahasa pemrograman apapun (Darwiyanti, 2013).

*Use case diagram* menggambarkan fungsionalitas yang diharapkan dari sebuah sistem. Yang ditekankan adalah "apa" yang diperbuat sistem, dan bukan "bagaimana". Sebuah *use case* merepresentasikan sebuah interaksi antara aktor dengan sistem. *Use case* merupakan sebuah pekerjaan tertentu, misalnya login ke sistem, *create* sebuah daftar belanja, dan sebagainya. Seorang/sebuah aktor adalah sebuah entitas manusia atau mesin yang berinteraksi dengan sistem untuk melakukan pekerjaan-pekerjaan tertentu.

*Class* adalah sebuah spesifikasi yang jika diinstansiasi akan menghasilkan sebuah objek dan merupakan inti dari pengembangan dan desain berorientasi objek. *Class* menggambarkan keadaan (atribut/properti) suatu sistem, sekaligus menawarkan layanan untuk memanipulasi keadaan tersebut (metoda/fungsi). *Class diagram* menggambarkan struktur dan deskripsi *class*, *package* dan objek beserta hubungan satu sama lain seperti *containment*, pewarisan, asosiasi, dan lain-lain.

*Activity diagrams* menggambarkan berbagai alir aktivitas dalam sistem yang sedang dirancang, bagaimana masing-masing alir berawal, *decision* yang mungkin terjadi, dan bagaimana mereka berakhir. *Activity diagram* juga dapat menggambarkan proses paralel yang mungkin terjadi pada beberapa eksekusi. *Activity diagram* merupakan *state diagram* khusus, di mana sebagian besar *state* adalah *action* dan sebagian besar transisi di-trigger oleh selesainya *state* sebelumnya (*internal processing*). Oleh karena itu *activity diagram* tidak menggambarkan behaviour internal sebuah sistem (dan interaksi antar subsistem) secara eksak, tetapi lebih menggambarkan proses-proses dan jalur-jalur aktivitas dari level atas secara umum. Sebuah aktivitas dapat direalisasikan oleh satu *use case* atau lebih. Aktivitas menggambarkan proses yang berjalan, sementara *use case* menggambarkan bagaimana aktor menggunakan sistem untuk melakukan aktivitas (Pressman, 2002).



#### 1.4 Metodologi

Metode yang digunakan dalam merancang sistem pengolahan data administrasi posyandu adalah metode analisis dengan mengidentifikasi setiap tahapan dalam metodologi pengembangan sistem dan dilanjutkan dengan perancangan sistem berorientasi objek (*Object Oriented Modelling*) menggunakan tools pengembangan system *Unified Modelling Language* (UML) yang pada akhirnya terbentuk suatu rancang bangun sistem pengolahan data administrasi yang dapat digunakan untuk mendukung pelaya<sup>14</sup> Posyandu.

Implementasi digunakan untuk mengevaluasi kualitas informasi sebelum dan sesudah pengembangan sistem pengolahan data administrasi, dengan demikian bentuk desain penelitian menggunakan *pre experimental (one group pre and post test)* yaitu desain penelitian yang mengukur satu kelompok obyek sebelum dan sesudah pemberian perlakuan.

Sampel yang diambil adalah posyandu melati 1 dan melati 2. Subyek yang diamati adalah orang-orang yang berkaitan dengan pengolahan data administrasi posyandu yaitu : ketua posyandu dan petugas posyandu.

## 2. PEMBAHASAN

Saat ini pada posyandu Melati 1 dan Melati 2 kelurahan Rambutan melakukan pencatatan data balita, data penimbangan balita, data kesehatan anak masih dilakukan secara manual tanpa menggunakan program. Tujuan utama pencatatan data adalah untuk menunjang data administrasi, medis, dokumentasi sehingga memudahkan petugas posyandu dalam membuat laporan.

Kendala yang sering terjadi dalam pencatatan adalah pencatatan manual dengan media kertas menyebabkan data terarsip dengan tidak teratur, sehingga menyulitkan petugas dalam membuat rekam data, dan penyajian laporan tiap bulannya menjadi tidak tepat waktu, data yang tersaji dengan media kertas mudah rusak.

#### 2.1 Analisa Sistem Yang Sedang Berjalan

Sistem yang<sup>7</sup> berjalan pada Posyandu Melati 1 dan Melati 2 menggunakan sistem yang belum terkomputerisasi. Mulai pengelolaan data bayi dan balita, pencatatan hasil penimbangan bayi dan balita, pencatatan pemberian vitamin, pencatatan pemberian imunisasi. Berikut ini adalah prosedur pencatatan data balita pada Posyandu Melati 1 dan Melati 2<sup>7</sup>

- Ibu datang ke posyandu mendaftarkan bayi.
- Petugas mencatat data bayi kemudian<sup>7</sup> memberikan KMS pada Ibu.
- Bayi yang sudah didata kemudian diberikan pelayanan bulanan seperti penimbangan berat badan, tinggi badan, vitamin dan imunisasi.

- Petugas memberikan hasil laporan perkembangan bayi / balita menggunakan KMS kepada Ibu.
- Petugas mengarsipkan hasil penimbangan menggunakan beberapa buku.

#### 2.2 Analisis Kebutuhan Sistem

Pada posyandu Melati 1 dan Melati 2 membutuhkan sistem yang mampu mencatat dan menampilkan data balita sehingga memudahkan dalam pengolahan data administrasi dan pembuatan laporan. Adapun kebutuhan non fungsional sistem dalam merancang sistem pengolahan data administrasi yaitu, Spesifikasi komputer minimum pentium 4 dengan sistem operasi minimum Microsoft Windows XP, kebutuhan memori RAM 1 hingga 2 GB, XAMPP sebagai server basis data dan Netbeans IDE 8 sebagai editor program.

Pada sistem yang sedang berjalan data-data posyandu terdiri dari data KMS, data administrasi/registrasi, data penimbangan, data pencatatan pemberian vitamin dan data manula dicatat secara manual untuk disusun menjadi laporan posyandu.

#### 2.3 Perancangan UML

UML (*Unified Modeling Language*) adalah bahasa pemodelan untuk sistem atau perangkat lunak yang berparadigma berorientasi objek (Adi, 2010).

UML merupakan bahasa pemodelan yang memiliki perbendaharaan kata dan cara untuk merepresentasikan secara fokus pada konseptual dan fisik suatu sistem.

UML tidak hanya merupakan sebuah bahasa pemrograman visual saja, namun juga dapat secara langsung dihubungkan ke berbagai bahasa pemrograman, seperti JAVA, C++, Visual Basic, atau bahkan dihubungkan secara langsung ke dalam sebuah *object oriented database*.

UML merupakan suatu kumpulan konvensi pemodelan yang digunakan untuk menentukan atau menggambarkan sebuah system software yang terkait dengan objek (Whitten, 2004).

#### 2.4 Diagram Use Case

*Use case diagram* secara grafis menggambarkan interaksi antara sistem, sistem eksternal dan pengguna. Dengan kata lain use case diagram secara grafis mendeskripsikan siapa yang akan menggunakan sistem dan dalam cara apa pengguna mengharapkan interaksi dengan sistem itu (Henderi, 2008).

Jenis-jenis hubungan usecase (Booch, 1999) :

**Association**, garis yang menghubungkan antar aktor dengan usecase,

# Paper

---

## ORIGINALITY REPORT

---

16%

SIMILARITY INDEX

16%

INTERNET SOURCES

5%

PUBLICATIONS

9%

STUDENT PAPERS

---

## PRIMARY SOURCES

---

1	<a href="http://iqbalramon.blogspot.com">iqbalramon.blogspot.com</a> Internet Source	2%
2	Submitted to Universitas Atma Jaya Yogyakarta Student Paper	2%
3	<a href="http://repository.uinjkt.ac.id">repository.uinjkt.ac.id</a> Internet Source	1%
4	<a href="http://fr.scribd.com">fr.scribd.com</a> Internet Source	1%
5	<a href="http://etheses.uin-malang.ac.id">etheses.uin-malang.ac.id</a> Internet Source	1%
6	<a href="http://posyandu.org">posyandu.org</a> Internet Source	1%
7	<a href="http://ijns.org">ijns.org</a> Internet Source	1%
8	<a href="http://ayistiya13.blogspot.com">ayistiya13.blogspot.com</a> Internet Source	1%
9	<a href="http://rezkiblog.wordpress.com">rezkiblog.wordpress.com</a>	



Internet Source

1%

10

[www.univ-tridianti.ac.id](http://www.univ-tridianti.ac.id)

Internet Source

1%

11

[www.unisbank.ac.id](http://www.unisbank.ac.id)

Internet Source

1%

12

[repository.gunadarma.ac.id](http://repository.gunadarma.ac.id)

Internet Source

1%

13

[eprints.uny.ac.id](http://eprints.uny.ac.id)

Internet Source

1%

14

[zombiedoc.com](http://zombiedoc.com)

Internet Source

1%

15

[citisee.amikompurwokerto.ac.id](http://citisee.amikompurwokerto.ac.id)

Internet Source

1%

16

[id.scribd.com](http://id.scribd.com)

Internet Source

1%

17

[eprints.undip.ac.id](http://eprints.undip.ac.id)

Internet Source

1%

Exclude quotes  On

Exclude matches  < 1%

Exclude bibliography  Off