

# **EKSTRASI NAMA – NAMA HOTEL BERBASISKAN ATURAN**

Skripsi



oleh

**EVAN SETIAWAN**

**71130060**

# **EKSTRASI NAMA – NAMA HOTEL BERBASISKAN ATURAN**

Skripsi



Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi Informasi  
Universitas Kristen Duta Wacana  
Sebagai Salah Satu Syarat dalam Memperoleh Gelar  
Sarjana Komputer

Disusun oleh

**EVAN SETIAWAN**

**71130060**

## PERNYATAAN KEASLIAN SKRIPSI

### PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

#### **EKSTRASI NAMA - NAMA HOTEL BERBASISKAN ATURAN**

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi kesarjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar kesarjanaan saya.

Yogyakarta, 9 Januari 2018



EVAN SETIAWAN

71130060

HALAMAN PERSETUJUAN

Judul Skripsi : EKSTRASI NAMA – NAMA HOTEL  
BERBASISKAN ATURAN ~~DAN LKSIKON~~

Nama Mahasiswa : EVAN SETIAWAN

NIM : 71130060

Matakuliah : Skripsi (Tugas Akhir)

Kode : TIW276

Semester : Gasal

Tahun Akademik : 2017/2018

Telah diperiksa dan disetujui di  
Yogyakarta,  
Pada tanggal 30 November 2017

Dosen Pembimbing I

  
Lucia Dwi Krisnawati, Dr.

Dosen Pembimbing II

  
Gloria Virginia, S.Kom., MAI, Ph.D.

# HALAMAN PENGESAHAN

## HALAMAN PENGESAHAN

### EKSTRASI NAMA - NAMA HOTEL BERBASISIKAN ATURAN

Oleh: EVAN SETIAWAN / 71130060

Dipertahankan di depan Dewan Penguji Skripsi  
Program Studi Informatika Fakultas Teknologi Informasi  
Universitas Kristen Duta Wacana - Yogyakarta  
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana Komputer  
pada tanggal 14 Desember 2017

Yogyakarta, 9 Januari 2018  
Mengesahkan,

Dewan Penguji:

1. Lucia Dwi Krisnawati, Dr.
2. Gloria Virginia, S.Kom., MAI, Ph.D.
3. Rosa Delima, S.Kom., M.Kom.
4. Danny Sebastian, S.Kom., M.M., M.T.

Dekan

  
(Hudi Susanto, S.Kom., M.T.)

Ketua Program Studi

  
(Gloria Virginia, Ph.D.)

## UCAPAN TERIMA KASIH

Dalam penelitian tugas akhir ini, penulis mendapatkan bantuan, saran, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Bapak Budi Susanto, S.Kom., M.T. selaku Dekan Fakultas Teknologi Informasi Universitas Kristen Duta Wacana.
2. Ibu Gloria Virginia, S.Kom., MAI., Ph.D. selaku Ketua Program Studi Teknik Informatika Universitas Kristen Duta Wacana.
3. Ibu Dr. Lucia Dwi Krisnawati, selaku dosen pembimbing I yang telah memberikan waktu secara rutin untuk melakukan konsultasi dan memberikan saran dan masukan mengenai pemrograman sistem, dan penyelesaian masalah dengan cara yang lebih sederhana.
4. Ibu Gloria Virginia, S.Kom., MAI., Ph.D. selaku dosen pembimbing II yang telah meluangkan waktu untuk memberikan bimbingan dan memberikan saran serta masukan mengenai penulisan laporan juga analisis sistem.
5. Yedija Kadmiel Elnatan, S.Kom. yang telah membantu dalam pengerjaan dokumen uji.
6. Orangtua, saudara-saudara dan teman-teman terdekat yang selalu memberikan dukungan motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
7. Pihak – pihak lain yang tidak dapat penulis sebutkan satu per satu yang berperan secara langsung maupun tidak langsung selama pengerjaan tugas akhir.

## KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas berkat dan kasih karunia-Nya sehingga penulis dapat menyelesaikan pembuatan sistem dan laporan tugas akhir dengan judul “*EKSTRASI NAMA – NAMA HOTEL BERBASISKAN ATURAN*” dengan baik.

Penulisan laporan tugas akhir ini diajukan sebagai salah satu syarat guna mencapai gelar Sarjana Strata Satu (S1) di Fakultas Teknologi Informasi Program Studi Teknik Informatika Universitas Kristen Duta Wacana Yogyakarta.

Dalam pembuatan laporan ini, penulis menyadari bahwa masih ada kekurangan, baik dari materi maupun teknik penyajiannya. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran dari pembaca. Akhir kata penulis memohon maaf apabila dalam penulisan laporan ini, ada kalimat yang kurang berkenan. Semoga hasil dari pengerjaan tugas akhir ini dapat berguna dan bermanfaat bagi banyak pihak.

Yogyakarta, 27 November 2017

Penulis

## INTISARI

### EKSTRAKSI NAMA - NAMA HOTEL BERBASISKAN ATURAN

Ekstraksi informasi merupakan teknologi yang berkaitan dengan cara menjadikan dokumen teks yang tidak terstruktur dengan domain tertentu ke dalam sebuah struktur informasi yang relevan. Ekstraksi informasi dapat dilakukan dengan berbagai metode, yaitu metode berbasis aturan, metode statistik, dan metode *transformation based learning*.

Pada penelitian ini, penulis menggunakan metode berbasis aturan untuk menghasilkan nama – nama hotel. *Preprocessing* dalam penelitian ini menggunakan *split* kalimat, normalisasi dan tokenisasi. Proses pengecekan dilakukan dengan menggunakan aturan-aturan, aturan tersebut merupakan aturan *middle*, aturan *suffix*, aturan *prefix* dan aturan preposisi. Hasil *output* dari sistem ini berupa nama – nama hotel.

Pengujian sistem dilakukan dengan menggunakan presisi dan *recall* dengan sedikit memodifikasi perhitungan dengan menggunakan tiga level granularitas yaitu level uji, level partial dan level kata. Hasil yang didapat dari 40 dokumen uji menunjukkan bahwa nilai presisi untuk level unit sebesar 82,28%, level partial sebesar 85,57%, level kata sebesar 90% dan *recall* untuk level unit sebesar 81,75%, level partial sebesar 88,81% dan level kata sebesar 93,75%.

**Kata Kunci:** [ekstraksi informasi, berbasis aturan, hotel, presisi, *recall*]



## DAFTAR ISI

HALAMAN JUDUL.....	
PERNYATAAN KEASLIAN SKRIPSI.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN.....	v
UCAPAN TERIMA KASIH.....	vi
KATA PENGANTAR .....	vii
INTISARI.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiii
DAFTAR LAMPIRAN.....	xiv
BAB 1    PENDAHULUAN .....	1
1.1    Latar Belakang Masalah.....	1
1.2    Perumusan Masalah.....	2
1.3    Batasan Masalah.....	2
1.4    Tujuan Penelitian.....	3
1.5    Metode Penelitian.....	3
1.6    Sistematika Penulisan.....	4
BAB 2    TINJAUAN PUSTAKA .....	5
2.1    Tinjauan Pustaka .....	5
2.2    Landasan Teori .....	6
2.2.1. <i>Information Extraction</i> .....	6
2.2.2.    Tahapan <i>Information Extraction</i> .....	9
2.2.3.    Presisi dan <i>Recall</i> .....	10
2.2.4.    Hotel.....	12
BAB 3    PERANCANGAN SISTEM .....	13
3.1.    Tahapan Pembangunan Sistem.....	13
3.1.1.    Pra-pemrosesan .....	13
3.1.2.    Pemrosesan.....	13
3.1.3.    Pasca Pemrosesan.....	14

3.2.	Analisis Kebutuhan Sistem .....	14
3.2.1.	Kebutuhan Fungsional .....	14
3.2.2.	Kebutuhan Non – Fungsional.....	14
3.2.3.	Kebutuhan Hardware dan Software .....	15
3.3.	Rancangan Sistem .....	15
3.3.1.	Alur Kerja Sistem.....	15
3.3.2.	Alur Kerja Sistem Preprocessing .....	16
3.3.3.	Alur Kerja Sistem <i>Rule Based</i> .....	16
3.3.4.	<i>Use Case Diagram</i> .....	19
3.3.5.	User Interface .....	20
3.4.	Rancangan Pengujian .....	23
BAB 4	IMPLEMENTASI DAN ANALISIS SISTEM.....	24
4.1.	Implementasi Sistem .....	24
4.1.1.	Implementasi Manual <i>POS Tagging</i> Dokumen .....	24
4.1.2.	Implementasi Alur Kerja Sistem Berbasis Aturan .....	25
4.1.3.	Implementasi Pembuatan Dokumen Evaluasi.....	30
4.1.4.	Implementasi <i>User Interface</i> .....	31
4.2.	Implementasi Program .....	36
4.2.1.	Proses <i>Preprocessing</i> .....	36
4.2.2.	Proses Pengecekan Aturan .....	37
4.2.3.	Proses Evaluasi .....	40
4.3.	Hasil Pengujian Sistem.....	43
4.3.1.	Presisi .....	44
4.3.2.	<i>Recall</i> .....	47
4.4.	Analisis Pengujian Sistem .....	51
4.4.1.	Pembahasan Analisis.....	51
4.4.2.	Penggunaan Leksikon .....	52
4.4.3.	Kekurangan Sistem .....	53
BAB 5	KESIMPULAN DAN SARAN .....	54
5.1.	Kesimpulan.....	54
5.2.	Saran .....	55

DAFTAR PUSTAKA .....	56
LAMPIRAN A .....	1
LAMPIRAN B .....	1
LAMPIRAN C .....	1
LAMPIRAN D .....	1

©UKDW

## DAFTAR TABEL

<i>Tabel 2.1</i> Contoh daftar dari <i>prefix</i> , <i>suffix</i> dan preposisi .....	8
<i>Tabel 2.2</i> Rumus Perhitungan Untuk Presisi, <i>Recall</i> dan <i>Accuracy</i> .....	11
<i>Tabel 3.1</i> Tabel <i>Use Case</i> Memasukkan Data .....	19
<i>Tabel 3.2</i> Tabel <i>Use Case</i> Evaluasi .....	20
<i>Tabel 3.3</i> Tabel <i>Confusion Matrix</i> .....	23
<i>Tabel 4.1</i> Tabel Penjelasan <i>Tag</i> .....	24
<i>Tabel 4.2</i> Tabel <i>Prefix</i> dan <i>Suffix</i> .....	26
<i>Tabel 4.3</i> Daftar Kata Kerja dan Kata Preposisi .....	40
<i>Tabel 4.4</i> Pengujian Presisi Level Unit Dokumen Uji 1 dengan Dokumen Evaluasi 1 .....	44
<i>Tabel 4.5</i> Pengujian Presisi Level Partial Dokumen Uji 1 dengan Dokumen Evaluasi 1 .....	45
<i>Tabel 4.6</i> Pengujian Presisi Level Kata Dokumen Uji 1 dengan Dokumen Evaluasi 1 .....	45
<i>Tabel 4.7</i> Hasil Presisi Pengujian Sistem .....	46
<i>Tabel 4.8</i> Pengujian <i>Recall</i> Level Unit Dokumen Uji 1 dengan Dokumen Evaluasi 1 .....	48
<i>Tabel 4.9</i> Pengujian <i>Recall</i> Level Partial Dokumen Uji 1 dengan Dokumen Evaluasi 1 .....	48
<i>Tabel 4.10</i> Pengujian <i>Recall</i> Level Kata Dokumen Uji 1 dengan Dokumen Evaluasi 1 .....	49
<i>Tabel 4.11</i> Hasil <i>Recall</i> Pengujian Sistem .....	50
<i>Tabel 4.12</i> <i>Macro Average</i> Presisi .....	52
<i>Tabel 4.13</i> <i>Macro Average Recall</i> .....	52

## DAFTAR GAMBAR

<i>Gambar 2.1</i> Ilustrasi Tokenisasi .....	10
<i>Gambar 2.2</i> Ilustrasi Perbedaan Presisi dan <i>Accuracy</i> .....	10
<i>Gambar 2.3</i> Rumus Perhitungan Untuk Presisi, <i>Recall</i> dan <i>Accuracy</i> .....	11
<i>Gambar 3.1</i> Alur Kerja Sistem .....	16
<i>Gambar 3.2</i> Tahap <i>Preprocessing</i> .....	16
<i>Gambar 3.3.</i> Proses <i>Rule Based</i> .....	17
<i>Gambar 3.4</i> <i>Use Case</i> Diagram Sistem Ekstraksi Nama – Nama Hotel .....	19
<i>Gambar 3.5</i> <i>User Interface</i> Halaman <i>Home</i> .....	21
<i>Gambar 3.6</i> <i>User Interface</i> Halaman <i>Input</i> .....	21
<i>Gambar 3.7</i> <i>User Interface</i> Halaman Hasil .....	22
<i>Gambar 3.8</i> <i>User Interface</i> Halaman <i>About</i> .....	22
<i>Gambar 3.9</i> <i>User Interface</i> Halaman Evaluasi .....	23
<i>Gambar 4.1</i> Implementasi Alur Kerja Sistem Berbasis Aturan.....	26
<i>Gambar 4.2</i> Implementasi Alur Aturan <i>Middle</i> .....	27
<i>Gambar 4.3</i> Implementasi Alur Aturan <i>Suffix</i> .....	28
<i>Gambar 4.4</i> Implementasi Alur Aturan <i>Prefix</i> .....	29
<i>Gambar 4.5</i> Implementasi Alur Aturan Preposisi.....	30
<i>Gambar 4.6</i> Contoh dokumen evaluasi 3.....	31
<i>Gambar 4.7.</i> Halaman Awal Sistem .....	31
<i>Gambar 4.8</i> Tampilan menu <i>About</i> .....	32
<i>Gambar 4.9</i> Tampilan menu <i>Input</i> .....	32
<i>Gambar 4.10</i> Tampilan hasil <i>Split</i> per kalimat .....	33
<i>Gambar 4.11</i> Tampilan hasil tokenisasi .....	34
<i>Gambar 4.12</i> Tampilan halaman hasil sistem.....	34
<i>Gambar 4.13</i> Tampilan <i>Input</i> Evaluasi.....	35
<i>Gambar 4.14</i> Tampilan menu Evaluasi .....	35
<i>Gambar 4.15</i> Grafik Perhitungan <i>Macro-Average</i> Presisi.....	47
<i>Gambar 4.16</i> Grafik Perhitungan <i>Macro-Average Recall</i> .....	51

## DAFTAR LAMPIRAN

Lampiran A.....	A-1
Lampiran B.....	B-1
Lampiran C.....	C-1
Lampiran D.....	D-1

©UKDW

## INTISARI

### EKSTRAKSI NAMA - NAMA HOTEL BERBASIS ATURAN

Ekstraksi informasi merupakan teknologi yang berkaitan dengan cara menjadikan dokumen teks yang tidak terstruktur dengan domain tertentu ke dalam sebuah struktur informasi yang relevan. Ekstraksi informasi dapat dilakukan dengan berbagai metode, yaitu metode berbasis aturan, metode statistik, dan metode *transformation based learning*.

Pada penelitian ini, penulis menggunakan metode berbasis aturan untuk menghasilkan nama – nama hotel. *Preprocessing* dalam penelitian ini menggunakan *split* kalimat, normalisasi dan tokenisasi. Proses pengecekan dilakukan dengan menggunakan aturan-aturan, aturan tersebut merupakan aturan *middle*, aturan *suffix*, aturan *prefix* dan aturan preposisi. Hasil *output* dari sistem ini berupa nama – nama hotel.

Pengujian sistem dilakukan dengan menggunakan presisi dan *recall* dengan sedikit memodifikasi perhitungan dengan menggunakan tiga level granularitas yaitu level uji, level partial dan level kata. Hasil yang didapat dari 40 dokumen uji menunjukkan bahwa nilai presisi untuk level unit sebesar 82,28%, level partial sebesar 85,57%, level kata sebesar 90% dan *recall* untuk level unit sebesar 81,75%, level partial sebesar 88,81% dan level kata sebesar 93,75%.

**Kata Kunci:** [ekstraksi informasi, berbasis aturan, hotel, presisi, *recall*]

# **BAB 1**

## **PENDAHULUAN**

### **1.1 Latar Belakang Masalah**

Perkembangan teknologi yang pesat mempermudah akses terhadap informasi tekstual yang sangat besar jumlahnya, baik yang terdapat pada Internet maupun pada koleksi dokumen yang spesifik. Akan tetapi, kemampuan manusia untuk membaca informasi tersebut dan memahami isinya tidak bertambah cepat dari sebelumnya. Oleh karena itu, dibutuhkan suatu mekanisme untuk menjadikan informasi tersebut ke dalam format yang terstruktur, sebagai contohnya adalah ke dalam basis data relasional.

Ekstraksi informasi (*Information Extraction* atau IE) merupakan teknologi yang berkaitan dengan cara menjadikan dokumen teks tidak terstruktur dengan domain tertentu ke dalam sebuah struktur informasi yang relevan. Secara garis besar, proses ekstraksi informasi terdiri dari dua tahap, yaitu mengidentifikasi informasi yang relevan di dalam sebuah dokumen teks, dan menyimpannya ke dalam bentuk yang terstruktur untuk digunakan kemudian. Secara umum hasil ekstraksi informasi dapat berguna untuk aplikasi – aplikasi yang lain seperti, Question-Answering System, Info Integration dan Text Mining (Turmo, Ageno & Catala, 2006).

Dengan ini maka peneliti ingin menggunakan ekstraksi informasi untuk mengenali nama – nama hotel. Ekstraksi nama – nama hotel ini dapat dimanfaatkan di text mining dan hasilnya difungsikan sesuai kebutuhan pengguna. Sebagai contoh hasil ekstraksi nama – nama hotel dapat dianalisis untuk keperluan lebih lanjut untuk bisnis seperti promosi, saran pemilihan hotel, dll. Maka dengan adanya ekstraksi informasi nama – nama hotel ini diharapkan dapat membantu pengguna untuk memberi saran hotel di tempat tujuan wisata.



## 1.2 Perumusan Masalah

Bedasarkan latar belakang masalah diatas, rumusan masalah dalam penelitian ini adalah :

- Bagaimana membangun sebuah sistem berbasis aturan untuk mengenali nama – nama hotel?
- Berapa besar akurasi yang dihasilkan sistem ekstraksi informasi berbasis aturan dalam melakukan ekstraksi nama – nama hotel?

## 1.3 Batasan Masalah

Adapun batasan masalah yang akan dibuat dalam penelitian ini adalah sebagai berikut :

1. Dokumen teks yang dijadikan dokumen uji diambil dari internet seperti artikel berita atau pariwisata.
2. Dokumen uji dan dokumen evaluasi yang dipakai total berjumlah 80 dokumen. Masing-masing 40 dokumen untuk dokumen uji dan 40 dokumen untuk dokumen evaluasi.
3. Dokumen evaluasi yang akan digunakan telah ditag secara manual oleh peneliti.
4. Dokumen uji yang digunakan berekstensi \*.txt sedangkan dokumen evaluasi berekstensi \*.xml.
5. Nama – nama hotel yang diambil hanya berasal dari Indonesia.
6. Sistem yang akan dibangun menggunakan bahasa pemrograman PHP yang diterapkan dalam bentuk web.

## 1.4 Tujuan Penelitian

Melalui penelitian ini, tujuan yang ingin dicapai oleh penulis adalah untuk membangun sebuah sistem ekstraksi informasi yang mampu mengekstraksi nama – nama hotel berdasarkan aturan – aturan yang didefinisikan.

## 1.5 Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah :

### 1. Perancangan dan Pembangunan Sistem

Tahap perancangan sistem merupakan tahap yang berisi perancangan definisi aturan dan *user interface* yang akan dibangun. Sedangkan tahap pembangunan sistem merupakan tahap pembuatan program dan sistem yang sesuai blok diagram dengan rancangan yang telah dibuat. Tahap pertama adalah membagi sebuah dokumen teks menjadi sebuah token. Kemudian dilakukan ekstraksi informasi berdasarkan aturan dan leksikon.

### 2. Analisis dan Evaluasi Sistem

Tahap analisis dan evaluasi adalah tahap yang dilakukan untuk menganalisis dan mengevaluasi program yang dibangun dalam ekstraksi informasi teks. Tahap evaluasi yang digunakan adalah menggunakan presisi dan *recall*.

## 1.6 Sistematika Penulisan

Penulisan laporan tugas akhir ini dibagi menjadi lima bab, yaitu :

Bab 1, PENDAHULUAN, yang akan memberikan gambaran penelitian ini melalui latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

Bab 2, TINJAUAN PUSTAKA, yang akan dibagi menjadi dua bagian, yaitu tinjauan pustaka, dan landasan teori. Tinjauan pustaka akan menguraikan teori-teori tentang ekstraksi informasi yang menggunakan metode leksikon dan aturan dari berbagai sumber pustaka yang digunakan. Sementara itu, landasan teori akan memuat semua konsep yang akan dijadikan dasar untuk alur penelitian.

Bab 3, ANALISIS DAN PERANCANGAN SISTEM, akan memberikan penjelasan tentang perancangan sistem yang akan dibangun. Perancangan sistem ini adalah kebutuhan *hardware* dan *software*, perancangan *database*, *use case diagram*, algoritma yang akan digunakan, dan perancangan antarmuka yang akan dibangun.

Bab 4, IMPLEMENTASI DAN ANALISIS SISTEM, akan berisi hasil implementasi dan pembahasan tentang proses pengujian sistem yang dibuat berdasarkan perancangannya. Selanjutnya, akan ada pembahasan dan analisis dari sistem yang telah dibuat dan dijalankan.

Bab 5, KESIMPULAN DAN SARAN, berisi kesimpulan dari hasil penelitian dan pembatasan sistem yang didapat, dan saran untuk mengembangkan sistem menjadi sistem yang lebih kompleks dan lebih baik sesuai dengan tuntutan zaman.

## BAB 5

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Berdasarkan hasil analisis penelitian yang telah dibahas pada Bab 4, maka dapat ditarik kesimpulan sebagai berikut:

1. Dengan menggunakan sistem evaluasi otomatis, sistem sudah baik dalam mengenali nama – nama hotel. Pada hasil *Macro Average* Presisi level unit sistem dapat mencapai 82,28%, untuk level partial 85,57%, untuk level kata 90%. sedangkan *Macro Average Recall* level unit didapatkan 81,75%, untuk level partial 88,81%, untuk level kata 93,75%. Hal ini menunjukkan bahwa metode berbasis aturan baik dalam mengenali nama – nama hotel.
2. Semakin halus level uji evaluasi yang digunakan, maka tingkat nilai presisi dan *recall* pun juga akan mendapatkan nilai yang tinggi.
3. Kelemahan dari sistem ini adalah sistem masih belum bisa menemukan nama hotel yang tidak diawali dengan kata *prefix*, *suffix* dan preposisi yang didefinisikan. Sistem juga belum dapat bisa melakukan ekstraksi nama hotel dengan benar jika terdapat *token* yang memiliki *tag N* dan diawali huruf kapital, tetapi token tersebut bukan merupakan nama hotel. Sebagai contoh pada kasus dokumen uji nomor 32. Di dokumen tersebut terdapat nama hotel “Sapulidi Sawah Cafe”, akan tetapi sistem tidak dapat mendeteksi nama hotel tersebut dikarenakan nama hotel tersebut tidak mengandung *prefix* atau *suffix*.
4. Pada penelitian ini telah dilakukan pembangunan leksikon yang berisi nama – nama hotel. Akan tetapi setelah mendapatkan hasil dari evaluasi sistem, metode leksikon ini tidak memberikan hasil nama hotel untuk sistem ini. Oleh karena itu, leksikon nama hotel ini dihapus dari sistem sehingga metode yang dilakukan berdasarkan metode aturan saja.

## 5.2. Saran

Sistem ini sangat memungkinkan untuk dilakukan pengembangan lebih lanjut sesuai kebutuhan yang terus bertambah, sehingga dapat meningkatkan akurasi sistem. Saran yang diajukan penulis dalam pengembangan sistem kedepannya adalah sebagai berikut :

1. Sistem belum mampu untuk mengekstraksi nama hotel yang tidak diawali dengan kata prefix, suffix atau preposisi yang sudah didefinisikan. Oleh karena itu, penulis mengharapkan agar dapat dikembangkan aturan yang dapat mengatasi kelemahan tersebut. Sehingga sistem dapat melakukan ekstraksi nama – nama hotel lebih baik lagi.
2. Pada penelitian ini, metode yang dilakukan berbasiskan pada aturan. Penulis mengharapkan agar sistem dapat dikembangkan dengan untuk aplikasi ekstraksi informasi lainnya. Salah satunya dengan menggunakan pembelajaran mesin(*Machine Learning*). Sehingga nantinya sistem dapat menjadi aplikasi ekstraksi nama – nama hotel yang tepat dan akurat.

## DAFTAR PUSTAKA

- Agrawal, I. T. (1993). Mining Association rules between sets of items in large databases.
- Alfred, R. L. (2014). Malay Named Entity Recognition Based on Rule-Based Approach . *International Journal of Machine Learning and Computing*.
- Appelt, D. &. (1999). Introduction to Information Extraction Technology.
- Budi, I. B. (2005). Named Entity Recognition for the Indonesian Language: Combining Contextual, Morphological and Part-of-Speech Features into a Knowledge Engineering Approach.
- Daelmans, W. &. (2005). Memory Based Language Processing.
- Elnatan, Y. K. (2017). POS TAGGING BAHASA INDONESIA BERBASIS ATURAN DAN LEXICON (KATA KERJA + KATA KETERANGAN).
- Muslea, I. (1999). Extraction Patterns for Information Extraction Tasks: A Survey.
- Pengertian Hotel dan Definisi Hotel*. (2014, Mei 22). Retrieved September 13, 2016, from <http://jenishotel.info/perngertian-hotel>
- Raharjo, B. (2011, Maret 17). *Presisi dan Akurasi*. Retrieved Oktober 10, 2016, from <http://www.raharjo.org/presisi-dan-akurasi/>
- Sarawagi, S. (2008). Information Extraction. *Foundation and Trends*, 261-377.
- Siniakov. (2008). *Rule-Based Approach to Information Extraction*.
- Turmo, J. A. (2006). Adaptive Information Extraction. *ACM Computing Surveys*, 4.