

**PENERAPAN NAMED ENTITY RECOGNITION UNTUK MENGENALI
FITUR PRODUK PADA E-COMMERCE MENGGUNAKAN RULE
TEMPLATE DAN HIDDEN MARKOV MODEL**

SKRIPSI

Untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun oleh:

M. Yusron Syauqi Dirgantara

NIM: 135150207111105



PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK IFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BRAWIJAYA
MALANG
2018

PENGESAHAN

PENERAPAN NAMED ENTITY RECOGNITION UNTUK MENGENALI FITUR PRODUK PADA E-COMMERCE MENGGUNAKAN RULE TEMPLATE DAN HIDDEN MARKOV MODEL

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan
memperoleh gelar Sarjana Komputer

Disusun Oleh:

M. Yusron Syauqi Dirgantara

NIM: 135150207111105

Skripsi ini telah diuji dan dinyatakan lulus pada
16 Januari 2018

Telah diperiksa dan disetujui oleh:

Dosen Pembimbing I

Dosen Pembimbing II

Mochammad Ali Fauzi, S.Kom, M.Kom

NIK. 201502 890101 1 001

Rizal Setya Perdana, S.Kom, M.Kom

NIK. 201603 910118 1 001

Mengetahui

Ketua Jurusan Teknik Informatika

Tri Astoto Kurniawan, S.T, M.T, Ph.D

NIP. 19710518 200312 1 001

PERNYATAAN ORISINALITAS

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu perguruan tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis disitasi dalam naskah ini dan disebutkan dalam daftar pustaka.

Apabila ternyata didalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur plagiasi, saya bersedia skripsi ini digugurkan dan gelar akademik yang telah saya peroleh (sarjana) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No. 20 Tahun 2003, Pasal 25 ayat 2 dan Pasal 70).

Malang, 16 Januari 2018



M. Yusron Syauqi Dirgantara

NIM: 135150207111105

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Allah SWT, karena dengan rahmat, hidayah dan karunia-Nya, penulis dapat menyelesaikan skripsi dengan judul “Penerapan Named Entity Recognition Untuk Mengenali Fitur Produk Pada E-commerce Menggunakan Rule Template Dan Hidden Markov Model” dengan baik.

Skripsi ini disusun sebagai salah satu persyaratan untuk menyelesaikan studi di Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Brawijaya. Melalui kesempatan ini, penulis ingin menyampaikan rasa hormat dan banyak terima kasih kepada berbagai pihak yang telah memberikan bantuan dan dukungan selama proses pengerjaan skripsi, diantaranya:

1. Kedua orang tua terutama ayah saya Sigit Dewantoro yang tidak henti-hentinya mendo'akan, memberi kasih sayang serta memberikan dorongan dari awal hingga akhir, semoga Rahimahullah dan mendapatkan maghfirah Allah.
2. Bapak Mochammad Ali Fauzi, S.Kom, M.Kom selaku dosen pembimbing I yang telah banyak sekali meluangkan waktu, memberikan ilmu, membantu, membimbing dan memberikan saran dalam pengerjaan skripsi.
3. Bapak Rizal Setya Perdana, S.Kom, M.Kom selaku dosen pembimbing II yang telah banyak memberikan bimbingan, arahan, ilmu dan saran selama penyusunan laporan skripsi ini.
4. Seluruh dosen Fakultas Ilmu Komputer Universitas Brawijaya yang telah membimbing, mengarahkan dan mengajarkan ilmu yang menunjang dalam pengerjaan skripsi.
5. Terima Kasih kepada Sabrina Nurfadilla yang selama 4 tahun ini menemani, membantu, memarahi dan memotivasi selama perkuliahan.
6. Teman seperjuangan, Rian Bayu Pambudi, Alif Robby Alcafi, Anandita Azharunisa Sasmito dan Anandhi Tristiaratri terima kasih telah menemani, membantu dan memotivasi selama penyusunan skripsi.
7. Kelas TIF-I semester satu terutama team MURSSKRIPSKRUP terima kasih telah menemani, memberikan ilmu, bantuan dan motivasinya.
8. Ibu Agustien Rahmawati selaku pemilik kost yang ramah dan baik memberikan tempat tinggal yang nyaman.
9. Takmir Masjid Al-Muhajirin GSE yang sering mengadakan kajian, sarapan pagi dan membantu mengingatkan para mahasiswa-mahasiswa untuk selalu ingat kepada Allah SWT.

Semoga jasa dan amal baik mendapatkan balasan dari Allah SWT. Dan semoga penulisan laporan skripsi ini bermanfaat bagi pembaca untuk pengembangan penelitian selanjutnya. Penulis menyadari bahwa skripsi ini masih jauh dari sempurna dan tidak terlepas dari kekurangan dan kesalahan karena keterbatasan materi, kemampuan, pengalaman dan pengetahuan yang dimiliki penulis. Oleh karena itu, segala kritik dan saran yang bersifat membangun, penulis harapkan untuk penyempurnaan skripsi ini.

Malang, 16 Januari 2018

M.Yusron Syauqi Dirgantara
Syauqidirgantara@gmail.com

ABSTRAK

Teknologi informasi dengan adanya internet memberikan dampak perkembangan perdagangan elektronik atau *e-commerce* yang memperoleh banyak popularitas. Data APJII tahun 2016 menyatakan sebanyak 130,8 juta penduduk Indonesia menggunakan internet untuk menawarkan barang dan jasa. Di dalam manajemen *e-commerce* terdapat *customer service* yang bertugas untuk menangani segala macam pertanyaan yang disampaikan oleh pelanggan. Penyampaian informasi oleh *customer service* biasanya melalui *call center* atau aplikasi *chatting*. Didorong kemampuan *intelligent digital assistants chatbot* banyak digunakan untuk membantu pekerjaan *customer services*. Dibutuhkan suatu analisis terhadap bahasa pelanggan pada *chatbot* agar dapat dikenali informasi apa saja yang terdapat pada pertanyaan tersebut, sehingga dibutuhkan klasifikasi dan ekstraksi informasi agar mendapatkan informasi penting yang dibutuhkan oleh *chatbot* dalam menjawab pertanyaan dari pelanggan. *Named Entity Recognition* (NER) merupakan bagian dari ekstraksi informasi yang bertugas untuk pengklasifikasi teks dari sebuah dokumen atau korpus yang dikategorikan kedalam beberapa kelas seperti nama orang, lokasi, bulan, tanggal, waktu dan sebagainya. Ekstraksi nama secara otomatis dapat berguna untuk mengatasi beberapa permasalahan seperti pada mesin terjemahan, pencarian informasi, tanya jawab dan peringkasan teks. Pada penelitian ini NER yang dilakukan menggunakan metode *Hidden Markov Model* dan *Rule Template* dengan 6 entitas yaitu *MEREK*, *TIPE*, *HARGA*, *SPEK*, *N_SPEK* dan *N_TAG*. Secara keseluruhan pengenalan entitas yang dilakukan pada penelitian ini menghasilkan nilai akurasi pada *Rule Template* sebesar 97.20% dan nilai akurasi pada *Hidden Markov Model* sebesar 92.23%.

Kata kunci: *e-commerce*, *Named Entity Recognition*, *Hidden Markov Model*, *Rule Template*, *information extraction*

ABSTRACT

Information technology with the Internet gives the impact of the development of electronic commerce or e-commerce that gained a lot of popularity. APJII data in 2016 states as many as 130.8 million Indonesians use the internet to offer goods and services. In e-commerce management there is customer service that is tasked to handle all of questions submitted by customers. Submission of information by customer service is usually through a call center or chat application. In thrust the ability of intelligent digital assistants chatbot is widely used to help the work of customer services. It takes an analysis of the customer's language on chatbot in order to be able to recognize what information is contained in the question, so it takes the classification and extracting of information in order to get important information needed by chatbot in answering questions from customers. Named Entity Recognition (NER) is part of the extraction of information assigned to the classification of text from a document or corpus categorized into classes such as person's name, location, month, date, time and so on. Automatic name extraction can be useful for addressing some issues such as translation engines, information retrieval, frequently asked questions and text summary. In this study NER is done using the method of Hidden Markov Model and Rule Template with 6 entities i.e. BRAND, TYPE, PRICE, SPEK, N_SPEK and N_TAG. Overall introduction of entities conducted in this study resulted in the accuracy value in the Rule Template of 97.20% and the accuracy value in the Hidden Markov Model of 92.23%.

Keywords: *e-commerce, Named Entity Recognition, Hidden Markov Model, Rule Template, information extraction*

DAFTAR ISI

PENGESAHAN	2
PERNYATAAN ORISINALITAS	3
KATA PENGANTAR.....	4
ABSTRAK.....	6
ABSTRACT.....	7
DAFTAR ISI.....	8
DAFTAR TABEL	11
DAFTAR GAMBAR.....	12
KODE PROGRAM	12
BAB 1 PENDAHULUAN	Error! Bookmark not defined.
1.1 Latar belakang	Error! Bookmark not defined.
1.2 Rumusan masalah.....	Error! Bookmark not defined.
1.3 Tujuan	Error! Bookmark not defined.
1.4 Manfaat	Error! Bookmark not defined.
1.5 Batasan masalah	Error! Bookmark not defined.
1.6 Sistematika pembahasan.....	Error! Bookmark not defined.
BAB 2 LANDASAN KEPUSTAKAAN	Error! Bookmark not defined.
2.1 Kajian Pustaka.....	Error! Bookmark not defined.
2.2 E-Commerce	Error! Bookmark not defined.
2.3 Text Mining.....	Error! Bookmark not defined.
2.4 Named Entity Recognition.....	Error! Bookmark not defined.
2.5 Hidden Markov Model.....	Error! Bookmark not defined.
2.6 Algoritme Viterbi	Error! Bookmark not defined.
2.7 Additive Smoothing	Error! Bookmark not defined.
2.8 Rule Template.....	Error! Bookmark not defined.
2.9 Regular Expression.....	Error! Bookmark not defined.
2.10 Pengukuran Evaluasi.....	Error! Bookmark not defined.
BAB 3 METODOLOGI	Error! Bookmark not defined.
3.1 Tahapan Penelitian	Error! Bookmark not defined.
3.2 Studi Pustaka	Error! Bookmark not defined.
3.3 Analisis Kebutuhan	Error! Bookmark not defined.

3.3.1 Kebutuhan Data	Error! Bookmark not defined.
3.3.2 Kebutuhan Sistem	Error! Bookmark not defined.
3.4 Perancangan Sistem	Error! Bookmark not defined.
3.5 Implementasi Sistem	Error! Bookmark not defined.
3.6 Pengujian Sistem	Error! Bookmark not defined.
3.7 Penarikan Kesimpulan dan Saran	Error! Bookmark not defined.
BAB 4 PERANCANGAN	Error! Bookmark not defined.
4.1 Diskripsi Umum Sistem	Error! Bookmark not defined.
4.2 Perhitungan Manual	Error! Bookmark not defined.
4.2.1 Rule Template	Error! Bookmark not defined.
4.2.2 Hidden Markov Model	Error! Bookmark not defined.
4.3 Perancangan Sistem	Error! Bookmark not defined.
4.3.1 Rule Template	Error! Bookmark not defined.
4.3.2 Pre-processing.....	Error! Bookmark not defined.
4.3.3 Hidden Markov Model	Error! Bookmark not defined.
4.3.4 Proses Hybird NER	Error! Bookmark not defined.
4.4 Perancangan Antarmuka	Error! Bookmark not defined.
4.5 Perancangan Pengujian	Error! Bookmark not defined.
BAB 5 IMPLEMENTASI.....	Error! Bookmark not defined.
5.1 Implementasi Kode Program	Error! Bookmark not defined.
5.1.1 Proses Rule Template	Error! Bookmark not defined.
5.1.2 Proses Pre-processing Data Latih....	Error! Bookmark not defined.
5.1.3 Proses Frekuensi	Error! Bookmark not defined.
5.1.4 Proses Perhitungan Transition Probability	Error! Bookmark not defined.
5.1.5 Proses Perhitungan Emission Probability	Error! Bookmark not defined.
5.1.6 Proses Pre-processing Data uji.....	Error! Bookmark not defined.
5.1.7 Proses Algoritme Viterbi	Error! Bookmark not defined.
5.1.8 Proses Hybird NER	Error! Bookmark not defined.
5.2 Implementasi Antarmuka	Error! Bookmark not defined.
BAB 6 PENGUJIAN DAN ANALISIS	Error! Bookmark not defined.
6.1 Skenario Pengujian	Error! Bookmark not defined.
6.2 Pengujian Rule Template.....	Error! Bookmark not defined.
6.3 Pengujian Hidden Markov Model.....	Error! Bookmark not defined.

6.4 Pengujian Hidden Markov Model dengan Additive Smoothing	Error! Bookmark not defined.
6.5 Pengujian Rule Template dengan Hidden Markov Model	Error! Bookmark not defined.
6.6 Pengujian Rule Template dan Hidden Markov Model dengan penambahan Additive Smoothing.....	Error! Bookmark not defined.
6.7 Analisis Hasil Pengujian	Error! Bookmark not defined.
BAB 7 PENUTUP	Error! Bookmark not defined.
7.1 Kesimpulan	Error! Bookmark not defined.
7.2 Saran	Error! Bookmark not defined.
DAFTAR PUSTAKA	Error! Bookmark not defined.
LAMPIRAN A Data Latih	Error! Bookmark not defined.
LAMPIRAN B Data Uji	Error! Bookmark not defined.
LAMPIRAN C Tabel Hasil Pengujian	Error! Bookmark not defined.
C.1 Data Latih.....	Error! Bookmark not defined.
C.2 Hasil Pengujian Rule Template	Error! Bookmark not defined.
C.3 Hasil Pengujian Hidden Markov Model ...	Error! Bookmark not defined.
C.4 Hasil Pengujian Hidden Markov Model dengan Additive Smoothing	Error! Bookmark not defined.
C.5 Hasil Pengujian Rule Template Dengan Hidden Markov Model	Error! Bookmark not defined.
C.6 Hasil Pengujian Rule Template dan Hidden Markov Model dengan menggunakan Additive Smoothing	Error! Bookmark not defined.

DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian Sebelumnya dengan Penelitian Penulis	Error! Bookmark not defined.
Tabel 2.2 Perbandingan Penelitian Sebelumnya dengan Penelitian Penulis (lanjutan)	Error! Bookmark not defined.
Tabel 2.3 Contoh Reguler Expression	Error! Bookmark not defined.
Tabel 2.4 Tabel Confusion Matrix	Error! Bookmark not defined.
Tabel 4.1 Contoh Entitas yang akan dikenali	Error! Bookmark not defined.
Tabel 4.2 Contoh Entitas yang akan dikenali (lanjut)	Error! Bookmark not defined.
Tabel 4.3 Data Uji Rule Template	Error! Bookmark not defined.
Tabel 4.4 Contoh pengenalan Rule Template	Error! Bookmark not defined.
Tabel 4.5 Data dokumen pertanyaan produk	Error! Bookmark not defined.
Tabel 4.6 Hasil perhitungan <i>start probability</i> data latihan	Error! Bookmark not defined.
Tabel 4.7 Hasil perhitungan <i>start probability</i> data latihan (lanjut)	Error! Bookmark not defined.
Tabel 4.8 Frekuensi entitas pada data latihan	Error! Bookmark not defined.
Tabel 4.9 Frekuensi transisi entitas pada data latihan	Error! Bookmark not defined.
Tabel 4.10 Hasil perhitungan <i>transition probability</i> data latihan	Error! Bookmark not defined.
Tabel 4.11 Hasil perhitungan <i>emission probability</i> data latihan	Error! Bookmark not defined.
Tabel 4.12 Hasil perhitungan <i>emission probability</i> data latihan (lanjutan)	Error! Bookmark not defined.
Tabel 4.13 Hasil perhitungan Viterbi	Error! Bookmark not defined.
Tabel 4.14 Hasil pemberian entitas pada data uji	Error! Bookmark not defined.
Tabel 6.1 Skenario Pengujian	Error! Bookmark not defined.
Tabel 6.2 Hasil Pengujian Rule Template	Error! Bookmark not defined.
Tabel 6.3 Hasil Pengujian Hidden Markov Model	Error! Bookmark not defined.
Tabel 6.4 Hasil Pengujian Hidden Markov Model dan Additive Smoothing	Error! Bookmark not defined.
Tabel 6.5 Hasil Pengujian Rule Template dengan Hidden Markov Model	Error! Bookmark not defined.
Tabel 6.6 Hasil Pengujian Rule template dan Hidden Markov Model dengan penambahan Additive Smoothing	Error! Bookmark not defined.
Tabel 6.7 Hasil Skenario Pengujian	Error! Bookmark not defined.

DAFTAR GAMBAR

- Gambar 2.1 Diagram venn text mining dan interaksinya dengan bidang lainnya **Error! Bookmark not defined.**
- Gambar 2.2 Proses *Generik* Text Mining **Error! Bookmark not defined.**
- Gambar 2.3 Teknik *text mining* dan fungsi utamanya **Error! Bookmark not defined.**
- Gambar 2.4 Diagram HMM..... **Error! Bookmark not defined.**
- Gambar 3.1 Tahapan Penelitian **Error! Bookmark not defined.**
- Gambar 4.1 Deskripsi umum sistem **Error! Bookmark not defined.**
- Gambar 4.2 Diagram alir Pengenalan NER **Error! Bookmark not defined.**
- Gambar 4.3 Diagram alir Rule Template **Error! Bookmark not defined.**
- Gambar 4.4 Diagram alir Hidden Markov Model **Error! Bookmark not defined.**
- Gambar 4.5 Diagram alir proses Rule Template **Error! Bookmark not defined.**
- Gambar 4.6 Diagram alir proses Pre-processing **Error! Bookmark not defined.**
- Gambar 4.7 Diagram alir proses Hidden Markov Model **Error! Bookmark not defined.**
- Gambar 4.8 Diagram alir proses frekuensi **Error! Bookmark not defined.**
- Gambar 4.9 Diagram alir proses *transition probability* **Error! Bookmark not defined.**
- Gambar 4.10 Diagram alir proses *transition probability* dengan *smoothing* **Error! Bookmark not defined.**
- Gambar 4.11 Diagram alir proses *emission probability* **Error! Bookmark not defined.**
- Gambar 4.12 Diagram alir proses *emission probability* dengan *smoothing* **Error! Bookmark not defined.**
- Gambar 4.13 Diagram alir proses Viterbi **Error! Bookmark not defined.**
- Gambar 4.14 Diagram alir proses Hybrid NER..... **Error! Bookmark not defined.**
- Gambar 4.15 Rancangan Antarmuka Masukan Data **Error! Bookmark not defined.**
- Gambar 4.16 Rancangan Antarmuka Hasil **Error! Bookmark not defined.**
- Gambar 5.1 Implementasi Antarmuka Masukan Data **Error! Bookmark not defined.**
- Gambar 5.2 Implementasi Antarmuka Hasil..... **Error! Bookmark not defined.**
- Gambar 6.1 Grafik Hasil Pengujian *Rule Template* ... **Error! Bookmark not defined.**
- Gambar 6.2 Grafik Hasil Pengujian *Hidden Markov Model* **Error! Bookmark not defined.**
- Gambar 6.3 Grafik Hasil Pengujian HMM dengan *Additive Smoothing* **Error! Bookmark not defined.**
- Gambar 6.4 Grafik Hasil Pengujian *Rule Template* dan *Hidden Markov Model* **Error! Bookmark not defined.**
- Gambar 6.5 Grafik Hasil Pengujian *Rule Template* dan *Hidden Markov Model* dengan penambahan *Additive Smoothing* **Error! Bookmark not defined.**

Gambar 6.6 Grafik Hasil Skenario Pengujian**Error! Bookmark not defined.**

KODE PROGRAM

Kode Program 5.1 Rule Template**Error! Bookmark not defined.**

Kode Program 5.2 Pre-processing dokumen**Error! Bookmark not defined.**

Kode Program 5.3 Proses Frekuensi Word**Error! Bookmark not defined.**

Kode Program 5.4 Proses Frekuensi Entitas**Error! Bookmark not defined.**

Kode Program 5.5 Proses Frekuensi Transisi**Error! Bookmark not defined.**

Kode Program 5.6 Proses Perhitungan *Transition Probability***Error! Bookmark not defined.**

Kode Program 5.7 Proses Perhitungan *Emission Probability***Error! Bookmark not defined.**

Kode Program 5.8 Proses Pre-processing Data Uji ...**Error! Bookmark not defined.**

Kode Program 5.9 Proses Algoritme Viterbi**Error! Bookmark not defined.**

Kode Program 5.10 Proses Hasil NER**Error! Bookmark not defined.**