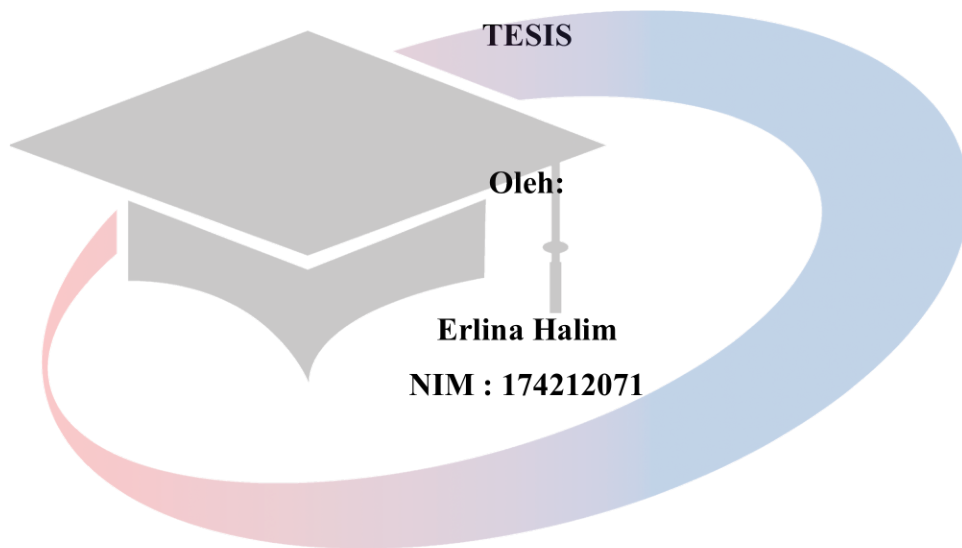
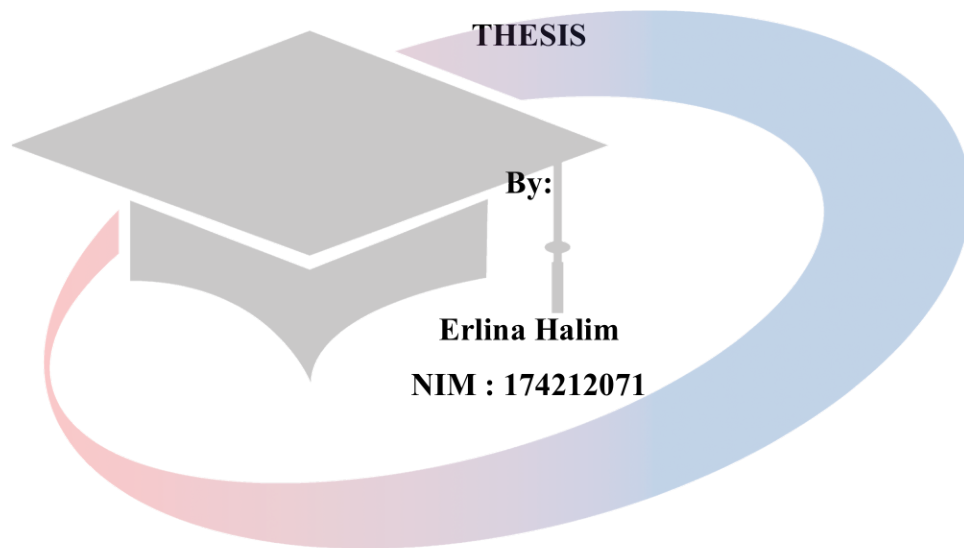


**EKSTRAKSI OPINI KONSUMEN MENGGUNAKAN TEXT
MINING UNTUK REKOMENDASI PRODUK PADA
E-COMMERCE**



**PROGRAM STUDI MAGISTER TEKNOLOGI INFORMASI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
MIKROSKIL
MEDAN
2019**

**CONSUMER OPINION EXTRACTION USING TEXT MINING
FOR PRODUCT RECOMMENDATIONS ON
E-COMMERCE**



**INFORMATION TECHNOLOGY MASTER'S PROGRAM
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
MIKROSKIL
MEDAN
2019**

LEMBAR PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Jurusan/Program Studi S-2 Magister Teknologi Informasi STMIK Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Erlina Halim
Nim : 174212071
Peminatan : Teknologi Informasi

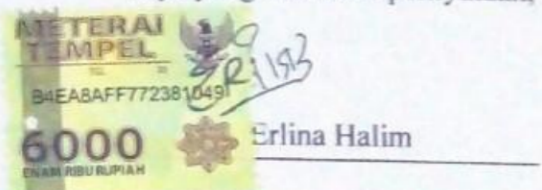
Saya telah melaksanakan penelitian dan penulisan Tesis dengan judul "EKSTRAKSI OPINI KONSUMEN MENGGUNAKAN TEXT MINING UNTUK REKOMENDASI PRODUK PADA E-COMMERCE", dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Tesis tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh STMIK Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada STMIK Mikroskil Medan Hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty Free Right*) atas Tesis saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, STMIK Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tesis saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tesis saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 12 Agustus 2019

Saya yang membuat pernyataan,



LEMBARAN PENGESAHAN
EKSTRAKSI OPINI KONSUMEN MENGGUNAKAN TEXT
MINING UNTUK REKOMENDASI PRODUK PADA E-
COMMERCE

TESIS

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Magister Strata Dua
Program Studi Magister Teknologi Informasi

Oleh:

ERLINA HALIM
NIM. 174212071

Disetujui Oleh:

Dosen Pembimbing,



Dr. Ronsen Purba, M.Sc.

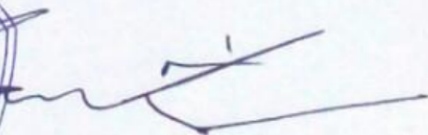
Medan, 12 Agustus 2019

Diketahui dan Disahkan Oleh:

Ketua Program Studi
Magister Teknologi Informasi,



Dr. Ronsen Purba, M.Sc.



ABSTRAK

Ekstraksi opini konsumen pada e-commerce memiliki peran dalam mempengaruhi konsumen dalam proses pembelian serta dapat menjadi rekomendasi. Namun kondisi opini yang beragam dan penggunaan kata tidak baku menjadi tantangan dalam mengolah opini. Opini perlu diolah terlebih dahulu dengan menerapkan *spelling normalization* dan perbaikan kata slang dengan menggunakan kamus slang yang berisi daftar kata slang beserta konversinya. Opini yang telah di normalisasikan akan digunakan untuk mengekstrak opini menggunakan *text mining* dengan pendekatan *lexicon*. Pendekatan ini memerlukan kamus kata yang berisi opini beserta bobot yang bernilai 1 sampai 5 untuk opini positif dan bernilai -1 sampai -5 untuk opini negatif. Bobot yang ditemukan untuk setiap opini yang ada akan digunakan untuk menentukan klasifikasi. Penentuan klasifikasi menggunakan perbandingan maksimum bobot opini positif dengan maksimum bobot opini negatif. Klasifikasi opini yang dihasilkan adalah positif, negatif atau netral. Klasifikasi opini kemudian dibandingkan dengan klasifikasi *rating* untuk mengetahui tingkat keakuratan. Perbandingan menghasilkan akurasi sebesar 80,34% dengan melengkapi kamus opini.

Kata Kunci : Text Mining, Lexicon, Klasifikasi, Opini Konsumen

UNIVERSITAS
MIKROSKIL

ABSTRACT

Consumer opinion extraction on e-commerce has a role in influencing consumers in the purchasing process and can be a recommendation. But the conditions of diverse opinions and the use of non-standard words are challenges in processing opinions. Opinion needs to be processed first by applying spelling normalization and repairs to the slang word using a slang dictionary that lists slang words and their conversions. The normalized opinion will be used to extract opinions using text mining with the lexicon approach. This approach requires a dictionary of words that contain opinions along with weight that are worth 1 to 5 for positive opinions and are worth -1 to -5 for negative opinions. The weight found for each existing opinion will be used to determine the classification. Classification uses the ratio of maximum positive opinion weight to the maximum weight of negative opinion. The resulting classification of opinions is positive, negative or neutral. The classification of opinion produced is positive, negative or neutral. Opinion classification is then compared with the rating classification to determine the level of accuracy. Comparison produces an accuracy of 80.34% by completing an opinion dictionary.

Keyword : Text Mining, Lexicon, Classification, Consumer Opinion

UNIVERSITAS
MIKROSKIL

KATA PENGANTAR

Puji dan syukur dipanjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan Tesis yang berjudul “Ekstraksi Opini Konsumen Menggunakan Text Mining untuk Rekomendasi Produk pada E-Commerce”.

Tesis ini dibuat untuk melengkapi persyaratan kurikulum pada Program Studi Magister Teknologi Informasi Strata Dua, STMIK Mikroskil Medan.

Penulis menyampaikan banyak terima kasih kepada:

1. Bapak Dr. Ronsen Purba, M.Sc., selaku Dosen Pembimbing dan Ketua Program Studi Magister Teknologi Informasi yang telah membimbing penulis dalam proses penyelesaian tesis ini.
2. Bapak Andri, S.Kom., M.T.I. selaku Dosen Pendamping Pembimbing yang telah membimbing penulis dalam proses penyelesaian tesis ini.
3. Bapak Dr. Pahala Sirait, S.T., M.Kom., selaku Ketua STMIK Mikroskil Medan.
4. Bapak Djoni, S.Kom., M.T.I., selaku Wakil Ketua I STMIK Mikroskil Medan.
5. Bapak atau Ibu Dosen Penguji atas masukan/saran yang diberikan.
6. Bapak atau Ibu Dosen STMIK Mikroskil Medan yang telah membantu proses penulisan tesis ini.
7. Anggota keluarga, teman, serta saudara-saudari yang terus memberikan dukungan penuh kepada penulis selama proses penulisan tesis ini.
8. Semua pihak yang telah membantu penulis dalam penulisan tesis ini.

Penulis menyadari bahwa masih terdapat banyak kekurangan yang ada. Oleh sebab itu, kritik dan saran yang bersifat membangun akan sangat diterima. Akhir kata, semoga Tesis ini dapat bermanfaat bagi masyarakat. Terima kasih.

Medan, 26 Juli 2019

Penulis

DAFTAR ISI

ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI	iv
DAFTAR GAMBAR	vi
DAFTAR TABEL	vii
DAFTAR LAMPIRAN	viii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Masalah Penelitian	2
1.2.1. Identifikasi Masalah	3
1.2.2. Rumusan Masalah	3
1.3. Tujuan dan Manfaat Penelitian	3
1.3.1. Tujuan Penelitian	3
1.3.2. Manfaat Penelitian	3
1.4. Ruang Lingkup	4
1.5. Metodologi Penelitian	4
1.6. Sistematika Penulisan	5
BAB II KAJIAN LITERATUR	6
2.1. Tinjauan Pustaka	6
2.1.1. <i>E-commerce</i>	6
2.1.2. <i>Opinion Mining</i>	8
2.1.3. <i>Text Mining</i>	9
2.1.4. <i>Klasifikasi</i>	14
2.1.5. <i>Sentistrength</i>	17
2.1.6. <i>Ukuran Evaluasi Model Klasifikasi</i>	20
2.1.7. <i>Penelitian Terkait</i>	22
2.2. Kerangka Konsep	23
BAB III METODOLOGI PENELITIAN	26
3.1. Analisis Masalah	26
3.2. Metode Penelitian	29
3.2.1. <i>Proses Persiapan Kamus</i>	29

3.2.2.	Proses Ekstraksi Opini	30
3.2.3.	Proses Perbandingan Rating vs Opini	38
3.3.	Metode Pemilihan Sampel	38
3.4.	Metode Pengumpulan Data	39
3.5.	Alat-alat Penelitian	41
3.6.	Teknik Analisis dan Pengujian Model	41
3.6.1.	Teknik Analisis Data	41
3.6.2.	Pengujian Model	42
BAB IV	HASIL DAN PEMBAHASAN	43
4.1.	Hasil	43
4.1.1.	Hasil Pengumpulan Data	43
4.1.2.	Hasil Konversi Kamus	47
4.1.3.	Hasil Pembobotan Opini	48
4.2.	Pembahasan	53
BAB V	PENUTUP	57
5.1.	Kesimpulan	57
5.2.	Saran	57
DAFTAR PUSTAKA	59
LAMPIRAN-LAMPIRAN	62
DAFTAR RIWAYAT HIDUP	66

UNIVERSITAS MIKROSKIL

DAFTAR GAMBAR

Gambar 2.1. Sistem Rekomendasi	9
Gambar 2.2. Alur proses analisis sentimen dengan lexicon based (Nurfalah & Suryani, 2016).....	14
Gambar 2.3. Formula dalam Penentuan Sentimen (Nurfalah & Suryani, 2016) ..	18
Gambar 2.4. Rumus dalam Penentuan Sentimen (Pamungkas & Putri, 2017)	19
Gambar 2.5. Rumus dalam Penentuan Sentimen (Thelwal, 2012)	19
Gambar 2.6. Confusion Matrix	23
Gambar 2.7. Kerangka konsep penelitian	27
Gambar 3.1. Penilaian Produk Shopee.....	29
Gambar 3.2. Halaman penilaian produk	29
Gambar 3.3. Opini konsumen	30
Gambar 3.4. SentiStrength dengan input opini bahasa Inggris.....	31
Gambar 3.5. SentiStrength dengan input opini bahasa Indonesia.....	31
Gambar 3.6. Flowchart Penerjemahan Kamus.....	33
Gambar 3.7. Flowchart Proses Ekstraksi Opini	34
Gambar 3.8. Tahapan <i>Preprocessing</i>	35
Gambar 3.9. <i>Flowchart Filtering</i>	36
Gambar 3.10. <i>Flowchart Case Folding</i>	36
Gambar 3.11. <i>Flowchart Spelling Normalization</i>	37
Gambar 3.12. <i>Flowchart Stopword Removal</i>	38
Gambar 3.13. <i>Flowchart Stemming</i>	38
Gambar 3.14. Ilustrasi cara kerja SentiStrength.....	39
Gambar 3.15. Aturan Pengklasifikasian SentiStrength.....	40
Gambar 3.16. Penilaian Produk Shopee.....	42
Gambar 3.17. Proses Pengumpulan Data	43
Gambar 4.1. Hasil Proses <i>Filtering</i>	52
Gambar 4.2. Hasil Proses <i>Case Folding</i>	52
Gambar 4.3. Hasil Proses <i>Spelling Normalization</i>	52
Gambar 4.4. Hasil Proses <i>Stopword Removal</i>	52

Gambar 4.5. Hasil Proses <i>Stemming</i>	52
Gambar 4.6. Hasil Pembobotan Kata	54
Gambar 4.7. Diagram Batang Distribusi Kelas Opini dengan Kamus Konversi ..	56
Gambar 4.8. Distribusi Kelas Per Item	56
Gambar 4.9. Diagram Batang Distribusi Kelas Opini dengan Kamus Konversi + Penambahan Manual	57
Gambar 4.10. Grafik Perbandingan Kelas Rating dengan Kelas Opini Pengujian Pertama.....	58
Gambar 4.11. Grafik Perbandingan Kelas Rating dengan Kelas Opini Pengujian Kedua	58



UNIVERSITAS MIKROSKIL

DAFTAR TABEL

Tabel 2.1. Aturan Klasifikasi <i>Rating</i> (Xia dan Jiang, 2018).....	17
Tabel 3.1. Contoh Kamus Kata.....	36
Tabel 3.2. Perbedaan Model Penelitian Sebelumnya dengan Model Usulan	37
Tabel 3.3. Atribut Hasil Parsing.....	41
Tabel 4.1. Atribut File JSON Per Penilaian	43
Tabel 4.2. Data Opini Hasil Konversi JSON ke MySQL	46
Tabel 4.3. Kata <i>Entry</i> Manual Kamus <i>EmotionLookUp</i>	47
Tabel 4.4. Hasil Pembobotan Opini	50
Tabel 4.5. Distribusi Kelas Opini dengan Kamus Konversi	51
Tabel 4.6. Distribusi Kelas Opini dengan Kamus Konversi + Penambahan Manual	53
Tabel 4.7. <i>Confusion Matrix</i> Pengujian Pertama	55
Tabel 4.8. Presisi dan <i>Recall</i> Pengujian Pertama.....	55
Tabel 4.9. <i>Confusion Matrix</i> Pengujian Kedua.....	55
Tabel 4.10. Presisi dan <i>Recall</i> Pengujian Kedua.....	56

UNIVERSITAS
MIKROSKIL

DAFTAR LAMPIRAN

Lampiran 1. Flowchart Stemming (Agusta, 2009).....	62
Lampiran 2. Listing Program <i>Crawling Data</i>	63
Lampiran 3. Listing Program Konversi File JSON ke Database MySQL	64
Lampiran 4. Listing Program Konversi Kamus SentiStrength	65



UNIVERSITAS MIKROSKIL