

**KOMPARASI SEARCH QUERY CLUSTERING
PADA E-COMMERCE MENGGUNAKAN
K-MEANS DAN ADAPTIVE DBSCAN**

TESIS

Oleh:

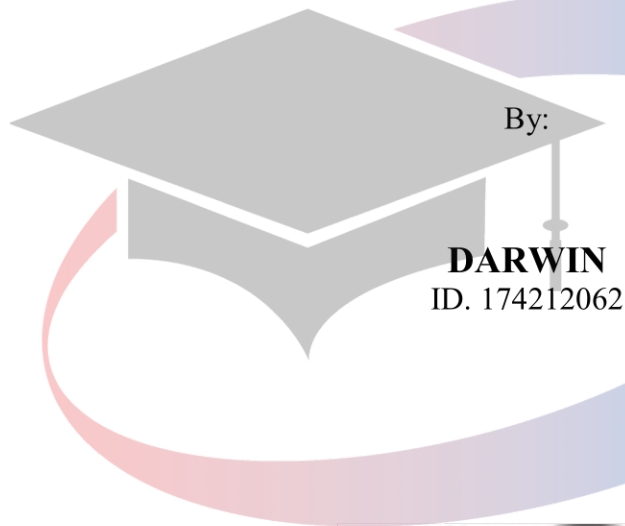
DARWIN
NIM. 174212062



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

**QUERY CLUSTERING SEARCH COMPARATION
ON E-COMMERCE USING
K-MEANS AND ADAPTIVE DBSCAN**

THESIS



**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 : Darwin
Nim : 174212062
Peminatan : IT Innovation

Saya telah melaksanakan penelitian dan penulisan Tesis dengan judul “Komparasi *Search Query Clustering* pada *E-Commerce* Menggunakan *K-Means* dan *Adaptive DBSCAN*”, 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, 20 Agustus 2019

Saya yang membuat pernyataan,


Darwin

LEMBARAN PENGESAHAN
KOMPARASI SEARCH QUERY CLUSTERING
PADA E-COMMERCE MENGGUNAKAN
K-MEANS DAN ADAPTIVE DBSCAN

TESIS

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

Oleh:

DARWIN
NIM. 174212062

Disetujui Oleh:

Dosen Pembimbing,


Dr. Ronsen Purba, M.Sc.

Medan, 20 Agustus 2019
Diketahui dan Disahkan Oleh:

Ketua Program Studi
Magister Teknologi Informasi,




Dr. Ronsen Purba, M.Sc.

ABSTRAK

Search query pada *e-commerce* sangat bervariasi tergantung kepada perilaku penggunanya, tetapi ketika preferensi pengguna dapat terealisasi melalui *query* maka dapat berimplikasi kepada minat beli. *Search query clustering* dilakukan untuk mendapatkan model pengetahuan bisnis terutama fitur *search* berdasarkan taksonomi. Dari penelitian ini didapatkan akurasi algoritma *clustering* sebesar 92.63% menggunakan *adaptive DBSCAN* dan 91.75% menggunakan *k-means*. Hasil taksonomi menjadi acuan dalam meningkatkan tampilan *e-commerce* sehingga hasil keluaran dari fitur *search* dapat menjadi daya tarik bagi pengguna untuk melakukan transaksi di *e-commerce*.

Kata Kunci: *Adaptive DBSCAN*, *k-Means*, Taksonomi, *Search Query Clustering*



UNIVERSITAS
MIKROSKIL

ABSTRACT

Search queries on e-commerce vary greatly depending on user behavior, but when user preferences can be realized through queries it can have implications for buying interest. Search query clustering is done to get business knowledge models, especially taxonomic search features. From this study, the accuracy of the clustering algorithm was 92.63% using adaptive DBSCAN and 91.75% using k-means. Taxonomic results are a reference in improving the appearance of e-commerce so that the output from the search feature can be an attraction for users to make transactions on e-commerce.

Keywords: Adaptive DBSCAN, k-Means, Taxonomy, Search Query Clustering



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 “Komparasi *Search Query Clustering* pada *E-Commerce* Menggunakan *K-Means* dan *Adaptive DBSCAN*”.

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 Muhammad Fermi Pasha, Ph.D, 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 yang telah memberikan masukan dan saran kepada penulis.
6. Bapak atau Ibu Dosen STMIK Mikroskil Medan yang telah membantu proses penulisan tesis ini.
7. Anggota keluarga, saudara-saudari, dan teman 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, 20 Agustus 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.....	2
1.2.2. Rumusan Masalah.....	2
1.3. Tujuan dan Manfaat Penelitian.....	3
1.4. Ruang Lingkup	3
1.5. Metodologi Penelitian	4
1.6. Sistematika Penulisan.....	4
BAB II KAJIAN LITERATUR.....	6
2.1. Tinjauan Pustaka	6
2.1.1. <i>E-Commerce</i>	6
2.1.2. <i>Data Mining</i>	8
2.1.2.1. <i>Classification</i>	10
2.1.2.2. <i>Association</i>	11
2.1.2.3. <i>Clustering</i>	12
2.1.3. <i>Algoritma Clustering</i>	13
2.1.3.1. <i>k-Means</i>	13
2.1.3.2. <i>DBSCAN</i>	15
2.1.3.3. <i>Adaptive DBSCAN</i>	18
2.1.4. <i>Search Query</i>	20
2.1.5. <i>Search Query Clustering</i>	21

2.1.6. <i>Confusion Matrix</i>	25
2.2. Kerangka Konsep Pemecahan Masalah.....	27
BAB III METODOLOGI PENELITIAN.....	30
3.1. Analisis Masalah	30
3.2. Data Yang Digunakan	31
3.3. Metode Penelitian	32
3.4. Alat-Alat Penelitian	34
3.5. Teknik Analisis.....	35
BAB IV HASIL DAN PEMBAHASAN.....	37
4.1. Hasil.....	37
4.1.1. Hasil <i>Clustering</i>	37
4.1.2. Hasil Pengujian.....	44
4.2. Pembahasan	49
BAB V PENUTUP	50
5.1. Kesimpulan.....	50
5.2. Saran	50
DAFTAR PUSTAKA	51
DAFTAR RIWAYAT HIDUP	54

UNIVERSITAS MIKROSKIL

DAFTAR GAMBAR

Gambar 2.1. Tahap Penemuan Pengetahuan Pada <i>Data Mining</i>	9
Gambar 2.2. Algoritma <i>k-Means</i>	13
Gambar 2.3. <i>Core</i> dan <i>Border</i>	16
Gambar 2.4. <i>Flowchart</i> Algoritma <i>Adaptive DBSCAN</i>	19
Gambar 2.5. Algoritma <i>Adaptive DBSCAN</i>	20
Gambar 2.6. Taksonomi <i>Search Query</i>	24
Gambar 2.7. <i>Confusion Matrix</i>	26
Gambar 2.8. Kerangka Konsep Pemecahan Masalah	28
Gambar 3.1. <i>Flowchart</i> Metode Penelitian.....	32
Gambar 3.2. Rancangan Implementasi	33
Gambar 3.3. <i>Confusion Matrix</i>	35
Gambar 4.1. <i>Line Chart</i> Pengujian Kemiripan Dan <i>Density</i>	45
Gambar 4.2. <i>Line Chart</i> Pengujian <i>Shallow</i>	46
Gambar 4.3. <i>Line Chart</i> Pengujian <i>Major</i>	47
Gambar 4.4. <i>Line Chart</i> Pengujian <i>Minor</i>	47
Gambar 4.5. <i>Line Chart</i> Pengujian <i>Target</i>	47
Gambar 4.6. <i>Line Chart</i> Pengujian <i>Hard</i>	48

UNIVERSITAS
MIKROSKIL

DAFTAR TABEL

Tabel 3.1. Skema Data (<i>Database PT Royal Utama Makmur</i>).....	31
Tabel 4.1. Hasil <i>Clustering</i> 1.000 <i>Records</i>	37
Tabel 4.2. Hasil <i>Clustering</i> 1.200 <i>Records</i>	38
Tabel 4.3. Hasil <i>Clustering</i> 1.400 <i>Records</i>	39
Tabel 4.4. Hasil <i>Clustering</i> 1.600 <i>Records</i>	40
Tabel 4.5. Hasil <i>Clustering</i> 1.800 <i>Records</i>	41
Tabel 4.6. Hasil <i>Clustering</i> 2.000 <i>Records</i>	42
Tabel 4.7. Hasil <i>Clustering</i> 2.074 <i>Records</i>	43
Tabel 4.8. Pengujian Kemiripan Dan <i>Density</i>	45
Tabel 4.9. Hasil Pengujian Lanjutan	46
Tabel 4.10. <i>Confusion Matrix</i>	48
Tabel 4.11. Perbandingan Keakuratan <i>k-Means</i> dan <i>Adaptive DBSCAN</i>	48

UNIVERSITAS
MIKROSKIL

DAFTAR LAMPIRAN

Lampiran 1. Daftar Riwayat Hidup.....	56
---------------------------------------	----



UNIVERSITAS
MIKROSKIL