

**ANALISIS KEPOPULARITASAN BISNIS PADA PORTAL BERITA
MENGUNAKAN ALGORITMA K-MEANS**

SKRIPSI



disusun oleh

Wahyu Hidayat

15.11.8947

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**



**ANALISIS KEPOPULARITASAN BISNIS PADA PORTAL BERITA
MENGUNAKAN ALGORITMA K-MEANS**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Wahyu Hidayat

15.11.8947

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2019**



PERSETUJUAN

SKRIPSI

ANALISIS KEPOPULARITASAN BISNIS PADA PORTAL BERITA MENGUNAKAN ALGORITMA K-MEANS

yang dipersiapkan dan disusun oleh

Wahyu Hidayat

15.11.8947

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 3 Oktober 2018

Dosen Pembimbing,



Ainul Yaqin, M. Kom

NIK. 190302255

PENGESAHAN

SKRIPSI

ANALISIS KEPOPULARITASAN BISNIS PADA PORTAL BERITA MENGUNAKAN ALGORITMA K-MEANS

yang disusun oleh

Wahyu Hidayat

15.11.8947

telah dipertahankan di depan Dewan Penguji
pada tanggal 20 Februari 2019

Susunan Dewan Penguji

Nama Penguji

Andi Sunyoto, M.Kom

NIK. 190302052

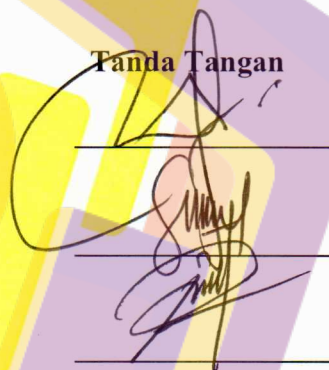
Sumarni Adi, S.Kom, M.Cs

NIK. 190302256

Ainul Yaqin, M. Kom

NIK. 190302255

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 25 Februari 2019

DEKAN FAKULTAS ILMU KOMPUTER



Krisnawati, S.Si, M.T

NIK. 190302038



PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 20 Februari 2019

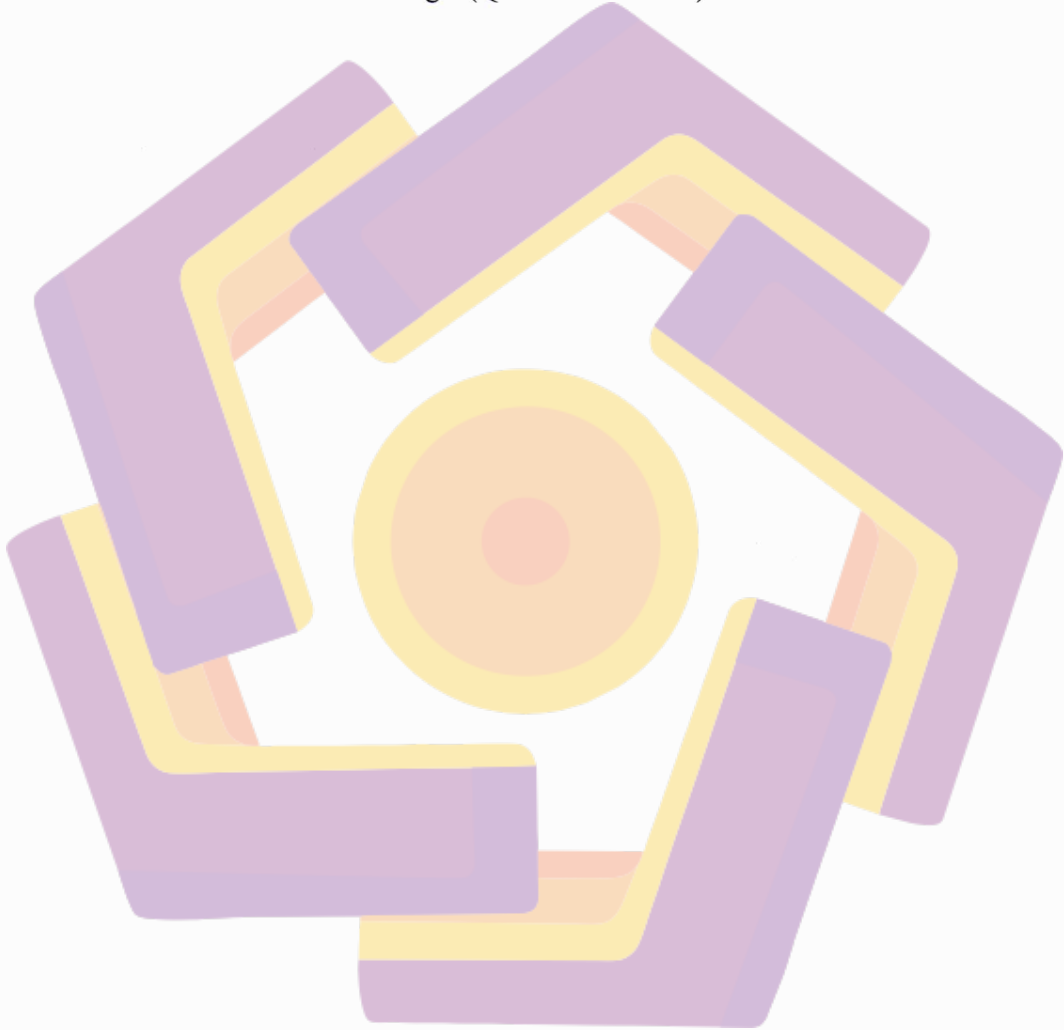


Wahyu Hidayat

NIM. 15.11.8947

MOTTO

" Hai orang-orang yang beriman, bersabarlah kamu dan kuatkanlah kesabaranmu dan tetaplah bersiap siaga dan bertakwalah kepada Allah, supaya kamu beruntung." (Qs. Al Imran: 200)



PERSEMBAHAN

Saya mempersembahkan skripsi ini kepada semua pihak yang terlibat secara langsung maupun tidak langsung dalam proses pembuatan skripsi.

1. Tuhan Yang Maha Esa yang memberikan segala nikmat dan kasih sayangNya sampai sejauh ini.
2. Orang tua dan kakak saya, yang selalu mendoakan, memberikan semangat dan fasilitas untuk menunjang kuliah.
3. Bapak Ainul Yaqin, M. Kom yang telah membimbing saya dari awal sampai akhir proses pembuatan skripsi.
4. Dosen Universitas Amikom Yogyakarta yang telah memberikan banyak ilmu selama kuliah. Terutama Bapak Haryoko, S.Kom, M.Cs. karena telah mengajarkan saya logika pemrograman pertama kali, Bapak Joko Dwi Santoso, M.Kom. yang sudah mengajarkan saya berpikir berbeda dari yang lainnya, Bapak Ryan Putranda Kristianto, S.Kom., M.Kom. yang mengajarkan berfikir logika kritis namun sederhana, serta Ibu Hartatik, S.T., M.Cs. yang selalu memberikan motivasi dan sebagai wali saya.
5. Semua teman kelas S1 IF 07 angkatan 2015 yang telah menemani dari awal kuliah sampai selesai. Semoga ilmu yang diperoleh bermanfaat baik bagi diri sendiri maupun masyarakat, bangsa dan negara.
6. Agustin Dwi Cahyani, yang selalu menyemangati dan memberikan motivasi kepada saya untuk mengerjakan skripsi.
7. Semua pihak yang tidak bisa saya sebutkan satu persatu yang sudah memberi semua ilmu, informasi, dan segalanya sehingga saya bisa menyelesaikan Skripsi ini.

KATA PENGANTAR

Puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat, hidayah dan kekuatan sehingga saya dapat menyelesaikan skripsi yang berjudul Analisis Kepopularitasan Bisnis pada Portal Berita Menggunakan Algoritma K-Means.

Skripsi ini saya buat guna menyelesaikan studi jenjang Strata Satu (S1) pada program studi Teknik Informatika fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang program strata 1 dan untuk memperoleh gelar Sarjana Komputer.


Dengan selesainya skripsi ini, Maka pada kesempatan ini saya mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Bapak Ainul Yaqin, M. Kom. selaku dosen pembimbing yang selalu bijaksana memberikan bimbingan, nasehat serta waktunya selama penulisan skripsi ini.
4. Dosen Penguji (Bapak Andi Sunyoto, M.Kom. dan Ibu Sumarni Adi, S.Kom, M.Cs.) dan segenap Dosen dan Karyawan Universitas AMIKOM Yogyakarta yang telah berbagi ilmu dan pengalamannya.

5. Orang tua dan kakak saya, yang selalu mendoakan, memberikan semangat dan fasilitas untuk penunjang kuliah.
6. Saudara-saudara yang senantiasa mendoakan, memberi semangat dan bantuan.
7. Teman kelas S1 IF 07 angkatan 2015 yang telah memberikan dukungan dan motivasi sehingga saya dapat menyelesaikan skripsi ini.

Semoga Allah Subhanahu wata'ala memberikan balasan yang lebih kepada semua yang telah ikut membantu saya dan menyelesaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik yang membangun akan diterima dengan senang hati dan rasa terima kasih. Semoga skripsi ini dapat bermanfaat bagi saya dan kita semua.

Yogyakarta, 20 Februari 2019



Wahyu Hidayat

DAFTAR ISI

LEMBAR JUDUL.....	i
PERSETUJUAN.....	ii
PENGESAHAN.....	iii
PERNYATAAN.....	iv
MOTTO.....	v
PERSEMBAHAN.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xiv
INTISARI.....	xviii
<i>ABSTRACT</i>	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	6
1.3 Batasan Masalah.....	6
1.4 Maksud Dan Tujuan Penelitian.....	7
1.4.1 Maksud Penelitian.....	7
1.4.2 Tujuan Penelitian.....	7
1.4.3 Manfaat Penelitian.....	8
1.5 Metode Penelitian.....	8
1.6 Sistematika Penulisan.....	9
BAB II LANDASAN TEORI	11
2.1 Tinjauan Pustaka.....	11
2.2 Dasar Teori.....	24
2.2.1 <i>Text Mining</i>	24

2.2.2	<i>Text Preprocessing</i>	24
2.2.2.1	<i>Case Folding</i>	25
2.2.2.2	<i>Tokenization</i>	25
2.2.2.3	<i>Stopword Removal</i>	25
2.2.2.4	<i>Stemming</i>	26
2.2.2.5	TF.IDF.....	26
2.2.3	Normalisasi.....	28
2.2.4	<i>Unsupervised Learning</i>	28
2.2.4.1	<i>Association</i>	29
2.2.4.2	<i>Clustering</i>	29
2.2.4.3	<i>Dimensionality Reduction</i>	29
2.2.5	<i>Clustering</i>	30
2.2.5.1	<i>Exclusive Clustering</i>	30
2.2.5.2	<i>Overlapping Clustering</i>	31
2.2.5.3	<i>Hierarchical Clustering</i>	31
2.2.5.4	<i>Density-based Clustering</i>	32
2.2.5.5	<i>Model-based Clustering</i>	33
2.2.6	Algoritma K-Means.....	34
2.2.7	Visualisasi Data (<i>Word Cloud</i>).....	38
2.2.8	Pengumpulan Data.....	38
2.2.9	<i>Web Service</i>	39
2.2.10	PHP (<i>Hypertext Preprocessor</i>).....	39
2.2.11	CSS (<i>Cascading Style Sheet</i>).....	39
2.2.12	HTML5 (<i>Hypertext Markup Language</i>).....	40
2.2.13	MySQL.....	40
2.2.14	XAMPP.....	40
2.3	Metode Analisa.....	41
2.4	Langkah-langkah Implementasi.....	42
2.4.1	<i>Flowchart</i>	42
2.4.2	<i>Entity Relationship Diagram</i>	44
2.4.3	DFD (<i>Data Flow Diagram</i>).....	45

2.4.3.1	Kegunaan Masing-masing Simbol pada <i>Data Flow Diagram (DFD)</i>	45
2.4.3.2	Cara Menggambar DFD.....	47
2.4.4	<i>Testing</i>	48
BAB III ANALISIS DAN PERANCANGAN PENELITIAN.....		50
3.1	Identifikasi Masalah.....	50
3.2	Analisis Masalah.....	51
3.3	Hasil Analisis.....	51
3.4	Deskripsi Sistem.....	51
3.5	Perhitungan Manual.....	56
3.6	Analisis Kebutuhan.....	107
3.6.1	Analisis Kebutuhan Fungsional.....	107
3.6.1.1	Definisi Fitur.....	107
3.6.2	Analisis Kebutuhan Non-Fungsional.....	110
3.6.2.1	Kebutuhan Perangkat Keras.....	110
3.6.2.2	Kebutuhan Perangkat Lunak.....	110
3.7	Perancangan Perangkat Lunak.....	111
3.7.1	<i>Entity Relation Diagram (ERD)</i>	111
3.7.2	Diagram Konteks.....	114
3.7.3	<i>Data Flow Diagram (DFD)</i>	115
3.7.3.1	DFD Level 1 Analisis Bisnis.....	115
3.7.3.2	DFD Level 2 Analisis Bisnis.....	116
3.7.4	Rancangan Antarmuka Pengguna (<i>User Interface</i>).....	116
BAB IV HASIL DAN PEMBAHASAN.....		126
4.1	Database dan Tabel.....	126
4.2	Implementasi Antarmuka Pengguna.....	128
4.3	Pembahasan Source Code.....	136
4.4	Pengujian Sistem.....	151
4.5	Hasil dan Analisis Pengujian.....	151
4.5.1	Pengujian Analisis pada Bulan Januari 2018.....	152
4.5.2	Pengujian Analisis pada Bulan Februari 2018.....	155

4.5.3	Pengujian Analisis pada Bulan Maret 2018.....	158
4.5.4	Pengujian Analisis pada Bulan April 2018.....	161
4.5.5	Pengujian Analisis pada Bulan Mei 2018.....	165
4.5.6	Pengujian Analisis pada Bulan Juni 2018.....	168
4.5.7	Pengujian Analisis pada Bulan Juli 2018.....	171
4.5.8	Pengujian Analisis pada Bulan Agustus 2018.....	175
4.5.9	Pengujian Analisis pada Bulan September 2018.....	178
4.5.10	Pengujian Analisis pada Bulan Oktober 2018.....	181
4.5.11	Pengujian Analisis pada Bulan November 2018.....	185
4.5.12	Pengujian Analisis pada Bulan Desember 2018.....	188
4.5.13	Hasil Analisis Kepopularitasan Bisnis.....	192
4.5.14	Perhitungan Dalam Sistem.....	204
BAB V PENUTUP.....		212
5.1	Kesimpulan.....	212
5.2	Saran.....	213
DAFTAR PUSTAKA.....		214

DAFTAR TABEL

Tabel 2.1 Literatur Review dan Posisi Peneliatian	16
Tabel 2.2 Simbol – simbol <i>Flowchart</i>	43
Tabel 2.3 Interpretasi <i>Silhouette Coefficient</i>	49
Table 3.1 <i>Text Mining</i>	56
Tabel 3.2 <i>Text Prosessing</i> Data Latih	58
Tabel 3.3 TF.IDF Data Latih	75
Tabel 3.4 <i>Text Prosessing</i> Tag Data Latih	77
Tabel 3.5 TF.IDF Tag Data Latih	88
Tabel 3.6 Bobot Data Latih dan Tag	89
Tabel 3.7 Normalisasi Data Latih dan Tag	91
Tabel 3.8 <i>Centeroid Cluster</i> Awal	92
Tabel 3.9 Jarak Data Latih terhadap Pusat <i>Cluster</i>	96
Tabel 3.10 Anggota <i>Cluster</i>	96
Tabel 3.11 <i>Centeroid Cluster</i> Baru	97
Tabel 3.12 Jarak Data terhadap Pusat <i>Cluster</i>	99
Tabel 3.13 Anggota <i>Cluster</i>	100
Tabel 3.14 Jarak Data terhadap Pusat <i>Cluster</i>	100
Tabel 3.15 Anggota <i>Cluster</i>	101
Tabel 3.16 <i>Silhouette Coefficient</i> dan Interpretasi	107
Tabel 3.17 Daftar Fitur Aplikasi	107
Tabel 3.18 Daftar Fitur Sistem lanjutan	108
Tabel 3.19 Atribut Tabel Admin	112
Tabel 3.20 Atribut Tabel Member	112
Tabel 3.21 Atribut Tabel Data	113
Tabel 3.22 Atribut Tabel Proses	113
Tabel 3.23 Atribut Tabel Terkait	114
Tabel 4.1 Hasil Analisis Kepopularitasan Bisnis pada Tahun 2018	193

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi <i>Exclusive Clustering</i>	31
Gambar 2.2 Ilustrasi <i>Overlapping Clustering</i>	31
Gambar 2.3 Ilustrasi <i>Hierarchical Clustering</i>	32
Gambar 2.4 Contoh <i>Dendogram</i>	32
Gambar 2.5 Ilustrasi <i>Density-based Clustering</i>	33
Gambar 2.6 Ilustrasi <i>Model-based Clustering</i>	34
Gambar 3.1 <i>Flowchart</i> Proses Kepopularitasan Bisnis.....	53
Gambar 3.2 Gambar <i>Word Cloud Cluster 1</i>	101
Gambar 3.3 Gambar <i>Word Cloud Cluster 2</i>	102
Gambar 3.4 Gambar <i>Entity Relation Diagram (ERD)</i>	111
Gambar 3.5 Diagram Konteks.....	114
Gambar 3.6 <i>Data Flow Diagram (DFD) Level 1</i>	115
Gambar 3.7 <i>Data Flow Diagram (DFD) Level 2</i>	116
Gambar 3.8 Gambar <i>Form Login</i>	117
Gambar 3.9 Gambar <i>Form Pendaftaran</i>	118
Gambar 3.10 Gambar <i>Form Lupa Password</i>	118
Gambar 3.11 Gambar Halaman Utama (Admin).....	119
Gambar 3.12 Gambar Ambil Data Berita (Admin).....	120
Gambar 3.13 Gambar <i>List Data</i> (Admin).....	120
Gambar 3.14 Gambar <i>List Member</i> (Admin).....	121
Gambar 3.15 Gambar Pengaturan Akun (Admin).....	122
Gambar 3.16 Gambar Pengaturan <i>Scrapper</i> (Admin).....	122
Gambar 3.17 Gambar Hasil Analisis (Admin).....	123
Gambar 3.18 Gambar Halaman Utama (Member).....	124
Gambar 3.19 Gambar <i>List Pemrosesan</i> (Member).....	124
Gambar 3.20 Gambar Pengaturan Profil (Member).....	125
Gambar 3.21 Gambar Hasil Analisis (Member).....	125
Gambar 4.1 Gambar Tabel admin.....	126

Gambar 4.2 Gambar Tabel member.....	127
Gambar 4.3 Gambar Tabel data.....	127
Gambar 4.4 Gambar Tabel proses.....	128
Gambar 4.5 Gambar Tabel terkait.....	128
Gambar 4.6 <i>Form Login</i>	129
Gambar 4.7 <i>Form Pendaftaran</i>	130
Gambar 4.8 <i>Form Lupa Password</i>	130
Gambar 4.9 Halaman Utama (Admin).....	131
Gambar 4.10 Ambil Data Berita (Admin).....	131
Gambar 4.11 <i>List Data (Admin)</i>	132
Gambar 4.12 <i>List Member (Admin)</i>	132
Gambar 4.13 Pengaturan Akun (Admin).....	133
Gambar 4. 14 Pengaturan <i>Scrapers</i> (Admin).....	133
Gambar 4.15 Hasil Analisis (Admin).....	134
Gambar 4.16 Halaman Utama (Member).....	134
Gambar 4.17 <i>List Pemrosesan (Member)</i>	135
Gambar 4.18 Pengaturan Profil (Member).....	135
Gambar 4.19 Hasil Analisis (Member).....	136
Gambar 4.20 Source Code Fungsi Case Folding.....	136
Gambar 4.21 Source Code Fungsi Tokenizing.....	137
Gambar 4.22 Source Code Fungsi Stopword Removal.....	138
Gambar 4.23 Source Code Fungsi Stemming.....	138
Gambar 4.24 Source Code Fungsi TF.IDF.....	141
Gambar 4.25 Source Code Fungsi Normalisasi.....	142
Gambar 4.26 Source Code Fungsi Rule of Thumb.....	142
Gambar 4.27 Source Code Fungsi Mencari Pusat Cluster Awal.....	143
Gambar 4.28 Source Code Fungsi Perhitungan Jarak.....	144
Gambar 4.29 Source Code Fungsi Mencari Pusat Cluster Baru.....	146
Gambar 4.30 Source Code Fungsi Perhitungan Jarak Baru.....	147
Gambar 4.31 Source Code Fungsi Word Cloud.....	148
Gambar 4.32 Source Code Fungsi Silhouette Coefficient.....	151
Gambar 4.33 <i>Word Cloud</i> pada Bulan Januari 2018.....	152

Gambar 4.34 Hasil Pengujian pada Bulan Januari 2018.....	153
Gambar 4.35 Diagram Interpretasi pada Bulan Januari 2018.....	153
Gambar 4.36 <i>Word Cloud</i> pada Bulan Februari 2018.....	155
Gambar 4.37 Hasil Pengujian pada Bulan Februari 2018.....	156
Gambar 4.38 Diagram Interpretasi pada Bulan Februari 2018.....	156
Gambar 4.40 <i>Word Cloud</i> pada Bulan Maret 2018.....	159
Gambar 4.41 Hasil Pengujian pada Bulan Maret 2018.....	160
Gambar 4.42 Diagram Interpretasi pada Bulan Maret 2018.....	158
Gambar 4.43 <i>Word Cloud</i> pada Bulan April 2018.....	160
Gambar 4.43 Hasil Pengujian pada Bulan April 2018.....	162
Gambar 4.44 Diagram Interpretasi pada Bulan April 2018.....	163
Gambar 4.45 <i>Word Cloud</i> pada Bulan Mei 2018.....	165
Gambar 4.46 Hasil Pengujian pada Bulan Mei 2018.....	166
Gambar 4.47 Diagram Interpretasi pada Bulan Mei 2018.....	166
Gambar 4.48 <i>Word Cloud</i> pada Bulan Juni.....	169
Gambar 4.49 Hasil Pengujian pada Bulan Juni 2018.....	169
Gambar 4.50 Diagram Interpretasi pada Bulan Juni 2018.....	170
Gambar 4.51 <i>Word Cloud</i> pada Bulan Juli 2018.....	172
Gambar 4.52 Hasil Pengujian pada Bulan Juli 2018.....	172
Gambar 4.53 Diagram Interpretasi pada Bulan Juli 2018.....	173
Gambar 4.54 <i>Word Cloud</i> pada Bulan Agustus 2018.....	175
Gambar 4.55 Hasil Pengujian pada Bulan Agustus 2018.....	176
Gambar 4.56 Diagram Interpretasi pada Bulan Agustus 2018.....	176
Gambar 4.57 <i>Word Cloud</i> pada Bulan September 2018.....	178
Gambar 4.58 Hasil Pengujian pada Bulan September 2018.....	179
Gambar 4.59 Diagram Interpretasi pada Bulan September 2018.....	179
Gambar 4.60 <i>Word Cloud</i> pada Bulan Oktober 2018.....	182
Gambar 4.61 Hasil Pengujian pada Bulan Oktober 2018.....	182
Gambar 4.62 Diagram Interpretasi pada Bulan Oktober 2018.....	183
Gambar 4.63 <i>Word Cloud</i> pada Bulan November 2018.....	185
Gambar 4.64 Hasil Pengujian pada Bulan November 2018.....	186
Gambar 4.65 Diagram Interpretasi pada Bulan November 2018.....	186

Gambar 4.66 <i>Word Cloud</i> pada Bulan Desember 2018.....	189
Gambar 4.67 Hasil Pengujian pada Bulan Desember 2018.....	189
Gambar 4.68 Diagram Interpretasi pada Bulan Desember 2018.....	190
Gambar 4.69 Gambar Tabel <i>Casefolding</i>	204
Gambar 4.70 Gambar Tabel <i>Tokenization</i>	204
Gambar 4.71 Gambar Tabel <i>Stopword Removal</i>	205
Gambar 4.72 Gambar Tabel <i>Stemming</i>	205
Gambar 4.73 Gambar Tabel TF.IDF.....	205
Gambar 4.74 Gambar <i>Casefolding</i> Tag.....	206
Gambar 4.75 Gambar Tabel <i>Tokenization</i> Tag.....	206
Gambar 4.76 Gambar Tabel <i>Stopword Removal</i> Tag.....	207
Gambar 4.77 Gambar Tabel <i>Stemming</i> Tag.....	207
Gambar 4.78 Gambar Tabel TF.IDF Tag.....	207
Gambar 4.79 Gambar Tabel Normalisasi.....	208
Gambar 4.80 Gambar Tabel Pusat <i>Cluster</i> Awal.....	208
Gambar 4.81 Gambar Jarak <i>Cluter</i> Awal.....	208
Gambar 4.82 Gambar Tabel Anggota <i>Cluster</i> Awal.....	209
Gambar 4.83 Gambar Tabel Anggota <i>Cluster</i> Awal.....	209
Gambar 4.84 Gambar Tabel Jarak Pusat <i>Cluster</i> Terakhir.....	209
Gambar 4.85 Gambar Tabel Anggota <i>Cluster</i> Terakhir.....	210
Gambar 4.86 Gambar <i>Word Cloud Cluster</i> 1.....	210
Gambar 4.87 Gambar <i>Word Cloud Cluster</i> 2.....	210
Gambar 4.88 Gambar Tabel Pengujian <i>Silhoutte Coefficient</i>	210
Gambar 4.89 Gambar Tabel Interpretasi.....	211
Gambar 4.90 Gambar Diagram Interpretasi.....	211

INTISARI

Informasi mengenai bisnis dan usaha pada portal berita sangat banyak dijumpai pada portal berita daerah maupun nasional. Navigasi seperti ekonomi, dan peluang usaha merupakan navigasi untuk mempermudah pembaca dalam mencari informasi tentang perkembangan dalam bidang ekonomi dan bisnis atau usaha. Informasi pada portal berita tentang perkembangan ekonomi sering didapatkan dari masalah lapangan pekerjaan serta bisnis atau usaha pada masyarakat Indonesia, namun pembaca kesulitan menyimpulkan bisnis yang paling populer dikalangan masyarakat untuk memenuhi kebutuhan masyarakat itu sendiri.

Penggunaan *text preprocessing* dapat mengetahui bobot data berita menjadi satuan angka dari perhitungan TF.IDF selanjutnya bobot data berita tersebut diproses menggunakan Algoritma K-Means untuk membuat data berita yang mempunyai jarak terdekat dari pusat kluster akan dijadikan dalam satu kluster, sehingga berita akan terkelompokkan. Selanjutnya dari kluster-kluster yang terbentuk dapat diketahui kluster yang memiliki anggota paling banyak. Setiap kluster akan di visualisasikan menggunakan *Word Cloud* untuk mengetahui kata yang sering muncul dan setiap anggota kluster akan diuji kualitasnya menggunakan *Silhouette Coefficient* untuk mengetahui kualitas data terhadap kluster.

Hasil yang diharapkan dapat mengetahui jenis bisnis yang paling populer pada setiap bulan di tahun 2018 melalui hasil analisis setiap kluster yang terbentuk dan akan diketahui jenis bisnis dari visualisasi yang ada, serta setiap anggota berita pada kluster akan diuji kualitasnya terhadap kluster itu sendiri dan akan disimpulkan dengan interpretasi.

Kata-kunci: Data Mining, *Clustering*, Algoritma K-Means, *Silhouette Coefficient*, Analisis Bisnis

ABSTRACT

Information about business on news portals found on regional and national news portals. Navigation such as economy, and business opportunities are navigation to make it easier for readers to find information about developments in the economy and business. Information on news portals about economic development is often obtained from the problems of employment and business in Indonesian society, but readers find it difficult to conclude the most popular business among the community to meet the needs of the community itself.

Text preprocessing can determine the weight of news data into numerical units from the calculation of TF. IDF then the weight of the news data is processed using the K-Means Algorithm to make news data that has the closest distance from the cluster center to one cluster, so the news will be grouped. Furthermore, from the formed clusters, clusters that have the most members can be known. Each cluster will be visualized using Word Cloud to find out the words that often appear and each cluster member will be tested for quality using the Silhouette Coefficient to determine the quality of the data against the cluster.

The results are expected to be able to find out the most popular business types in each month in 2018 through the analysis of each cluster formed and the type of business from the existing visualization will be known, and each news data member on the cluster will be tested for the cluster itself and will conclude with interpretation.

Keywords: Data Mining, Clustering, K-Means Algorithm, Silhouette Coefficient, Business Analysis