



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



IMPLEMENTASI K-MEANS CLUSTERING DALAM PENGELOMPOKAN ULASAN PRODUK DI E- MARKETPLACE

SKRIPSI

Dibuat untuk Melengkapi Syarat-Syarat yang Diperlukan
untuk Memperoleh Diploma Empat Politeknik

POLITEKNIK
NEGERI
JAKARTA

Ilham Hanif

4617010030

PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK NEGERI JAKARTA

2021



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : Ilham Hanif
NPM : 4617010030
Tanggal : 29 Juli 2021

Tanda Tangan :

POLITEKNIK
NEGERI
JAKARTA



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

LEMBAR PENGESAHAN

Skripsi diajukan oleh

Nama Mahasiswa : Ilham Hanif
NIM : 4617010030
Program Studi : Teknik Informatika
Judul Skripsi : IMPLEMENTASI K-MEANS CLUSTERING DALAM
PENGELOMPOKAN ULASAN PRODUK DI E-
MARKETPLACE

Telah diuji oleh tim penguji dalam Sidang Skripsi pada hari Kamis, tanggal 5, bulan Agustus, Tahun 2021 dan dinyatakan **LULUS**.

Disahkan oleh:

Pembimbing : Eriya, S.Kom., M.T. ()
Penguji I : Euis Oktavianti, S.Si., M.T.I. ()
Penguji II : Ayu Rosyida Zain, S.ST., M.Kom. ()
Penguji III : Fachroni Arbi Murad, S.Kom., M.Kom. ()

Mengetahui:

Jurusan Teknik Informatika dan Komputer

Ketua



Mauldy Laya, S.Kom., M.Kom.

NIP. 197802112009121003



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan Skripsi ini. Penulisan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Terapan Politeknik.

Penulis menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan Skripsi ini, sangatlah sulit bagi penulis untuk menyelesaikan Skripsi ini. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Mauldy Laya, S.Kom., M.Kom. selaku ketua jurusan Teknik Informatika dan Komputer Politeknik Negeri Jakarta
2. Ibu Risna Sari, selaku ketua program studi Teknik Informatika Jurusan Teknik Informatika dan Komputer Politeknik Negeri Jakarta;
3. Ibu Eriya, S.Kom., M.T., selaku dosen pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan penulis dalam penyusunan Skripsi ini;
4. Orang tua dan keluarga penulis yang telah memberikan bantuan dukungan material dan moral; dan
5. Sahabat yang telah banyak membantu penulis dalam menyelesaikan Skripsi ini. Akhir kata, penulis berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Skripsi ini membawa manfaat bagi pengembangan ilmu.

Depok, Juli 2021

Ilham Hanif



HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Politeknik Negeri Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Ilham Hanif
NIM : 4617010030
Program Studi : Teknik Informatika
Jurusan : Teknik Informatika dan Komputer
Jenis karya : Skripsi

demikian pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Politeknik Negeri Jakarta **Hak Bebas Royalti Noneksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

IMPLEMENTASI K-MEANS CLUSTERING DALAM PENGELOMPOKAN ULASAN PRODUK DI E-MARKETPLACE

berserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Politeknik Negeri Jakarta berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di: Depok Pada tanggal: 29 Juli 2021
Yang menyatakan

(Ilham Hanif)
NIM. 4617010030

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



IMPLEMENTASI K-MEANS CLUSTERING DALAM PENGELOMPOKAN ULASAN PRODUK DI E-MARKETPLACE

Abstrak

Ulasan produk merupakan fitur pada e-marketplace untuk mengetahui kualitas dari produk atau pengalaman orang lain saat membeli produk. Semakin bagus dan banyak ulasan pada suatu produk yang dijual, semakin tinggi minat calon pembeli terhadap produk. Karenanya, perlu mengelompokkan setiap ulasan apakah ulasan tersebut mengandung makna positif atau negatif. Penelitian ini mengusulkan K-Means Clustering untuk mengelompokkan ulasan produk. Dari analisis kluster kluster 0 merupakan kluster ulasan bersentimen negatif dengan centroid di titik 0.61129469 untuk skor positif dan 0.81219006 untuk skor negatif. Sedangkan, kluster 1 merupakan kluster ulasan bersentimen positif dengan centroid di titik 0.85114994 untuk skor positif dan 0.53761561 untuk skor negatif.

Kata kunci: Ulasan Produk, K-Means, Clustering, Text Mining



POLITEKNIK
NEGERI
JAKARTA

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS.....	i
LEMBAR PENGESAHAN	ii
KATA PENGANTAR	iii
HALAMAN PERNYATAAN	iv
<i>Abstrak</i>	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusah Masalah.....	2
1.3 Batasan Masalah.....	3
1.4 Tujuan dan Manfaat.....	3
1.5 Metode Pelaksanaan Skripsi.....	3
1.5.1 Metode Proses <i>Data Mining</i>	3
1.5.2 Metode Pengembangan Sistem	6
BAB II TINJAUAN PUSTAKA	9
2.1 Penelitian Terdahulu.....	9
2.2 <i>E-Marketplace</i>	11
2.3 Ulasan Produk	11
2.4 <i>Text Mining</i>	11
2.5 <i>Text Preprocessing</i>	12
2.6 <i>Unsupervised Learning</i>	13
2.7 <i>Text Clustering</i>	14



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

2.8	Word2vec	14
2.9	K-Means <i>Clustering</i>	15
2.10	<i>Elbow Method</i>	15
2.11	<i>Davies-Bouldin Index</i>	16
2.12	<i>Silhouette coefficient</i>	17
2.13	Django.....	17
2.14	Vue.js	18
BAB III	PERANCANGAN DAN REALISASI	19
3.1	Proses <i>Data Mining</i>	19
3.1.1	<i>Business Understanding</i>	19
3.1.2	<i>Data Understanding</i>	19
3.1.3	<i>Data Preparation</i>	20
3.1.4	<i>Modeling</i>	31
3.1.5	<i>Evaluation</i>	34
3.2	Perancangan Program Aplikasi	34
3.2.1	Deskripsi Program Aplikasi	34
3.2.2	Analisis Kebutuhan	35
3.2.3	Cara Kerja Program Aplikasi	35
3.2.4	Rancangan Program Aplikasi.....	36
3.3	Realisasi Sistem.....	39
3.3.1	Mockup Aplikasi.....	39
3.3.2	Implementasi Model <i>Data Mining</i>	42
3.3.3	Implementasi Tampilan Aplikasi	44
BAB IV	PEMBAHASAN.....	54
4.1	Pengujian	54



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

4.2	Deskripsi Pengujian.....	54
4.2.1	Deskripsi Pengujian Algoritma K-Means	54
4.2.2	Deskripsi Pengujian Sistem.....	54
4.3	Prosedur Pengujian.....	54
4.3.1	Prosedur Pengujian Algoritma K-Means	55
4.3.2	Prosedur Pengujian Sistem.....	56
4.4	Data Hasil Pengujian.....	57
4.4.1	Hasil Pengujian Algoritma K-Means.....	57
4.4.2	Hasil Pengujian Sistem	58
4.5	Analisis Data / Evaluasi	59
4.5.1	Analisis Kluster	59
4.5.2	Evaluasi Sistem	62
BAB V	PENUTUP	63
5.1	Simpulan.....	63
5.2	Saran.....	63
DAFTAR PUSTAKA	64



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

DAFTAR TABEL

Tabel 3.1 Contoh Hasil <i>Case Folding</i>	21
Tabel 3.2 Contoh Hasil Penghapusan Karakter dan Angka.....	22
Tabel 4.1 Jumlah ulasan berdasarkan <i>rating star</i>	55
Tabel 4.2 Skenario <i>Blackbox Testing</i>	56
Tabel 4.4 Hasil Perhitungan DBI dan <i>Silhouette coefficient</i>	57
Tabel 4.5 Hasil Pengujian Halaman Utama.....	58
Tabel 4.6 Hasil Pengujian Halaman <i>Summary</i>	58
Tabel 4.7 Pengujian Halaman Evaluasi.....	59



POLITEKNIK
NEGERI
JAKARTA



DAFTAR GAMBAR

Gambar 1.1 Siklus CRISP-DM.....	4
Gambar 2.1 <i>Text Preprocessing</i>	13
Gambar 2.2 Contoh Implementasi <i>Elbow Method</i>	16
Gambar 2.3 Logo Vue.js.....	18
Gambar 3.1 Contoh beberapa data ulasan.....	20
Gambar 3.2 Proses <i>Text Cleaning</i>	23
Gambar 3.3 Daftar Kata <i>Stopword</i> Sastrawi.....	23
Gambar 3.4 Proses Pengurangan dan penambahan kata <i>Stopword</i>	24
Gambar 3.5 Proses Penghapusan <i>Stopword</i>	24
Gambar 3.6 Proses Stemming.....	25
Gambar 3.7 Pembentukan Bigram dengan Kata Negasi.....	26
Gambar 3.8 Proses Pembentukan Bigram dengan <i>Phraser</i>	26
Gambar 3.9 Contoh Beberapa Hasil Pembentukan Bigram.....	26
Gambar 3.10 Flowchart Proses Pembobotan Skor Positif dan Negatif.....	27
Gambar 3.11 Pemodelan <i>Word2Vec</i>	28
Gambar 3.12 Proses <i>Training Model Word2Vec</i>	28
Gambar 3.13 100 Kata dengan Kemunculan Paling Banyak pada Data.....	28
Gambar 3.14 Proses Pencarian Kata yang Serupa untuk Kata Positif.....	29
Gambar 3.15 Proses Pencarian Kata yang Serupa untuk Kata Negatif.....	29
Gambar 3.16 Proses Pemberian Skor <i>Similarity</i>	30
Gambar 3.17 Sebaran Data Berdasarkan Skor yang Sudah Diberikan.....	30
Gambar 3.18 Pembobotan dengan TF-IDF.....	31
Gambar 3.19 Kode untuk Mencari k dengan <i>Elbow Method</i>	32
Gambar 3.20 Grafik untuk Menggunakan <i>Elbow Method</i>	32
Gambar 3.21 Pemodelan <i>Clustering</i> dengan K-Means.....	32
Gambar 3.22 Kode untuk Menampilkan Hasil Pelabelan Klaster.....	33
Gambar 3.23 Grafik Sebaran Klaster.....	33
Gambar 3.24 <i>Flowchart</i> Cara Kerja Program Aplikasi.....	36
Gambar 3.25 <i>Use Case Diagram</i> Aplikasi.....	37

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Gambar 3.26 Activity Diagram Aplikasi	37
Gambar 3.27 ERD pada Aplikasi.....	38
Gambar 3.28 <i>Mockup</i> Halaman <i>Home</i>	40
Gambar 3.29 <i>Mockup</i> Halaman <i>Summary</i>	40
Gambar 3.30 <i>Mockup</i> Halaman Evaluasi (Input).....	41
Gambar 3.31 <i>Mockup</i> Halaman Evaluasi (Output).....	42
Gambar 3.32 <i>Source code</i> untuk membersihkan teks	42
Gambar 3.33 <i>Source code</i> untuk membentuk bigram.....	43
Gambar 3.34 <i>Source code</i> stemming	43
Gambar 3.35 <i>Source code</i> untuk melakukan <i>preprocessing</i>	43
Gambar 3.36 <i>Source code</i> untuk memberikan skor <i>similarity</i>	43
Gambar 3.37 <i>Source code</i> untuk memberikan label	43
Gambar 3.38 <i>Source code</i> untuk menganalisis ulasan.....	44
Gambar 3.39 <i>Source code</i> halaman <i>Home</i> 1	45
Gambar 3.40 <i>Source code</i> halaman <i>Home</i> 2	45
Gambar 3.41 Hasil Implementasi Tampilan Halaman <i>Home</i>	46
Gambar 3.42 <i>Source code</i> halaman <i>Summary</i> 1.....	47
Gambar 3.43 <i>Source code</i> halaman <i>Summary</i> 2	47
Gambar 3.44 <i>Source code</i> halaman <i>Summary</i> 3.....	48
Gambar 3.45 Hasil Implementasi Tampilan Halaman <i>Summary</i> 1.....	48
Gambar 3.46 Hasil Implementasi Tampilan Halaman <i>Summary</i> 2.....	48
Gambar 3.47 Hasil Implementasi Halaman Evaluasi (Input)	49
Gambar 3.48 Hasil Implementasi Halaman Evaluasi (Output).....	50
Gambar 3.49 <i>Source code</i> halaman evaluasi 1.....	50
Gambar 3.50 <i>Source code</i> halaman evaluasi 2.....	51
Gambar 3.51 <i>Source code</i> halaman evaluasi 3.....	51
Gambar 3.52 <i>Source code</i> halaman evaluasi 4.....	52
Gambar 3.53 <i>Source code</i> halaman evaluasi 5.....	52
Gambar 3.54 <i>Source code</i> halaman evaluasi 6.....	53
Gambar 3.55 <i>Source code</i> halaman evaluasi 7.....	53
Gambar 4.1 <i>Source code</i> Perhitungan DBI dan <i>Silhouette coefficient</i>	56



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Gambar 4.2 *Wordcloud* Klaster 0 61
Gambar 4.3 *Wordcloud* Klaster 1 62



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB I PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah berdampak banyak terhadap kehidupan masyarakat. Aspek kehidupan yang mengalami perubahan salah satunya adalah kegiatan transaksi jual beli. Kegiatan jual beli yang sebelumnya hanya dilakukan dengan bertemu langsung antara penjual dan pembeli, kini juga dapat dilakukan dengan memanfaatkan internet melalui *e-marketplace* atau pasar maya. Pada kuartal pertama tahun 2021, tercatat dua *e-marketplace* yang dikunjungi lebih dari 100 juta pengunjung setiap bulannya yaitu, Tokopedia dan Shopee. Yang mana, jumlah kunjungan kedua *e-marketplace* ini meningkat jauh dibanding pada kuartal pertama tahun 2020 yang hanya sekitar 70 juta kunjungan per bulannya (iPrice, 2021).

Karena interaksi jual-beli pada *e-marketplace* berbeda dengan interaksi jual-beli konvensional, cara pembeli memilih dan menilai barang pun berbeda untuk di *e-marketplace*. Di *e-marketplace*, terdapat fitur *rating* dan *review* untuk setiap produk. Dari *review* atau ulasan produk, calon pembeli atau konsumen dapat mengetahui kualitas dari produk atau pengalaman orang lain saat membeli produk. Ulasan produk juga berpengaruh pada minat beli konsumen di *e-marketplace*. Semakin bagus dan banyak ulasan pada suatu produk yang dijual, semakin tinggi minat calon pembeli terhadap produk tersebut (Servanda *et al.*, 2019). Oleh sebab itu, ulasan produk juga dianggap penting bagi penjual karena dapat memengaruhi minat pembeli.

Untuk mengetahui pendapat setiap pembeli terhadap suatu produk melalui ulasan, perlu untuk mengelompokkan setiap ulasan apakah ulasan tersebut mengandung makna positif atau negatif. Setelah itu, penjual dapat mengetahui seperti apa pendapat pembeli terhadap produk yang dijual dan apa saja yang perlu diperbaiki atau dipertahankan dari produk tersebut. Di sisi lain, calon pembeli juga dapat



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

terbantu untuk memutuskan membeli produk tersebut atau tidak karena dapat mengetahui pendapat pembeli-pembeli yang sudah pernah membeli produk tersebut.

Sebelum mengelompokkan data ulasan produk menggunakan algoritma *clustering*, perlu untuk mengubah data ulasan yang berbentuk teks menjadi data-data numerik. Hal ini penting karena algoritma *clustering* tidak bisa langsung mengolah data teks yang bersifat tidak terstruktur. Salah satu cara untuk mengubah data teks menjadi data numerik yang dapat diolah adalah dengan menghitung *word similarity*. *Word similarity* digunakan untuk menghitung dan membandingkan kesamaan antara kata-kata. *Word similarity* sering digunakan pada penelitian *natural language processing* (NLP), *text clustering* atau pengelompokan teks dan *Text Categorization* atau kategorisasi teks (Mao, Zhang and Zhang, 2020). Pada penelitian ini, *library* yang digunakan untuk menghitung *word similarity* antar kata adalah *word2vec* dari Gensim.

Algoritma yang digunakan untuk melakukan pengelompokan pada penelitian ini adalah K-Means. Algoritma K-means adalah salah satu algoritma pengelompokan paling populer yang digunakan saat ini karena relatif cepat namun sederhana untuk dipahami dan diterapkan dalam praktik (Raykov *et al.*, 2016). Dengan menggunakan algoritma ini, ulasan produk pada *e-marketplace* dapat dikategorikan ke dalam dua kelompok atau lebih.

Berdasarkan uraian di atas, maka diperlukan aplikasi *website* untuk menampilkan informasi yang merepresentasikan ulasan produk. Aplikasi ini akan menampilkan ulasan setiap pembeli terhadap suatu produk dan mengkategorikannya ke dalam ulasan positif atau negatif. Selain itu, aplikasi juga akan menampilkan presentase dari setiap ulasan positif dan negatif sebagai ringkasan.

1.2 Perumusan Masalah

Dari paparan latar belakang di atas, maka rumusan masalah dalam skripsi ini adalah sebagai berikut.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

- a. Bagaimana cara pengelompokan ulasan produk yang ada di *e-marketplace* menggunakan algoritma K-Means
- b. Bagaimana cara membuat *website* yang menampilkan hasil pengelompokan ulasan produk yang ada *e-marketplace*.

1.3 Batasan Masalah

Batasan masalah dalam pengerjaan skripsi ini antara lain:

- a. Data yang digunakan untuk mempelajari pengelompokan ulasan produk ini adalah data ulasan produk yang didapat dari *e-marketplace*.
- b. Data yang diolah adalah ulasan berbahasa Indonesia.

1.4 Tujuan dan Manfaat

Tujuan dari skripsi ini adalah melakukan pengelompokan ulasan produk pada *e-marketplace* menggunakan algoritma K-Means dan mengimplementasikannya ke dalam bentuk *website*. Adapun manfaatnya adalah sebagai berikut.

- a. Membantu mengelompokkan ulasan produk yang dapat digunakan untuk mengevaluasi produk.
- b. Mengetahui implementasi K-Means *Clustering* dalam melakukan pengelompokan ulasan produk.

1.5 Metode Pelaksanaan Skripsi

Metode yang digunakan untuk menyelesaikan masalah pada skripsi ini terdiri dari metode proses data mining dan metode pengembangan aplikasi.

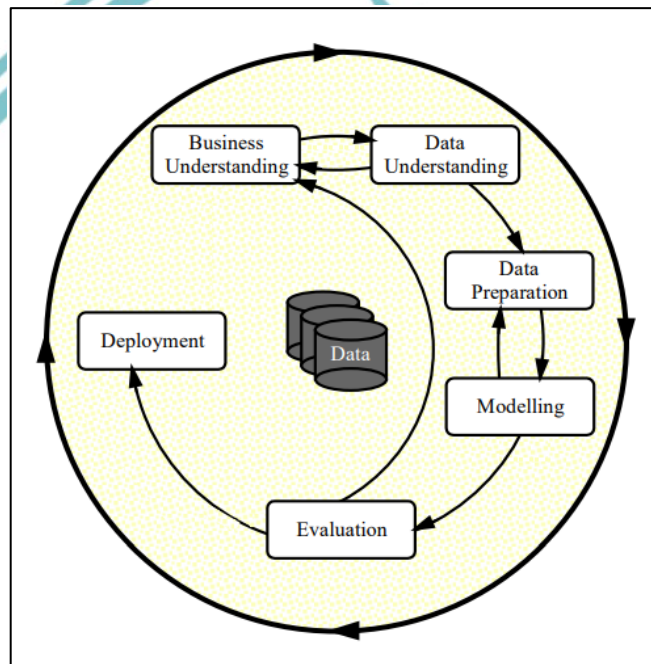
1.5.1 Metode Proses *Data Mining*



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Metode proses *data mining* yang digunakan pada skripsi ini adalah CRISP-DM (*Cross Industry Standard Process for Data Mining*). CRISP-DM merupakan standar proses siklus hidup dalam proyek *data mining* . Model proses CRISP-DM bertujuan untuk membuat proyek data mining besar, lebih murah, lebih dapat diandalkan, lebih dapat diulang, lebih mudah dikelola, dan lebih cepat. Berikut adalah tahapan-tahapan dalam siklus CRISP-DM (Wirth and Hipp, 2000).



Gambar I.1 Siklus CRISP-DM
Sumber: (Wirth and Hipp, 2000)

Meskipun pada penelitian ini *data mining* yang dilakukan adalah *text mining* yang menggunakan teks sebagai data, metode CRISP-DM tetap relevan untuk digunakan dalam prosesnya. Contohnya adalah penelitian oleh (Noviyanti and Ruldeviyani, 2020) yang melakukan klasifikasi teks terhadap data twitter, (Babaee and Sarabadani, 2018) yang melakukan penyaringan spam terhadap SMS dan (Kusuma and Suhardi, 2021) yang melakukan deteksi prostitusi online pada media social twitter. Semua penelitian tersebut menggunakan CRISP-DM untuk metode proses *data mining*.

a. *Business Understanding*



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Tahapan awal ini berfokus pada pemahaman tujuan dan kebutuhan proyek dari perspektif bisnis, dan kemudian mengubah pengetahuan ini menjadi definisi masalah data mining, dan rencana proyek awal yang dirancang untuk mencapai tujuan.

b. *Data Understanding*

Tahap *data understanding* atau pemahaman data dimulai dengan pengumpulan data awal dan dilanjutkan dengan kegiatan untuk mengenal data, mengidentifikasi masalah kualitas data, untuk menemukan wawasan pertama ke dalam data, atau untuk mendeteksi himpunan bagian yang menarik untuk membentuk hipotesis untuk informasi tersembunyi. Ada hubungan erat antara Pemahaman Bisnis dan Pemahaman Data. Perumusan masalah data mining dan rencana proyek membutuhkan setidaknya beberapa pemahaman tentang data yang tersedia.

c. *Data Preparation*

Tahap *data preparation* atau persiapan data mencakup semua kegiatan untuk membangun kumpulan data akhir (data yang akan dimasukkan ke dalam alat pemodelan) dari data mentah awal. Tugas persiapan data kemungkinan akan dilakukan beberapa kali, dan tidak dalam urutan yang ditentukan.

d. *Modeling*

Pada fase ini, berbagai teknik pemodelan dipilih dan diterapkan, dan parameternya dikalibrasi ke nilai optimal. Biasanya, ada beberapa teknik untuk tipe masalah data mining yang sama. Beberapa teknik memerlukan format data tertentu.

e. *Evaluation*

Sebelum melanjutkan ke penerapan akhir model, penting untuk mengevaluasi model secara lebih menyeluruh, dan meninjau langkah-langkah yang dijalankan untuk membangun model, untuk memastikan model tersebut mencapai tujuan bisnis dengan benar. Tujuannya adalah untuk menentukan apakah ada beberapa



masalah bisnis penting yang belum cukup dipertimbangkan. Pada akhir fase ini, keputusan tentang penggunaan hasil *data mining* harus dicapai.

f. *Deployment and Control*

Penciptaan model umumnya bukan akhir dari proyek. Biasanya, pengetahuan yang diperoleh perlu diatur dan disajikan sedemikian rupa sehingga dapat dipakai oleh klien atau pengguna. Tergantung pada kebutuhannya, fase ini dapat sesederhana membuat laporan atau serumit menerapkan proses *data mining* yang berulang. Dalam banyak kasus, pengguna, bukan analis data, yang akan melakukan langkah penerapan. Bagaimanapun, penting untuk memahami terlebih dahulu tindakan apa yang perlu dilakukan untuk benar-benar memanfaatkan model yang dibuat.

1.5.2 Metode Pengembangan Sistem

Metode yang digunakan untuk pengembangan sistem pada skripsi ini adalah SDLC (*System Development Life Cycle*). SDLC merupakan sebuah metode untuk mengembangkan sistem yang terdiri dari tahapan-tahapan yang dilakukan oleh tim pengembang sistem yang terdiri dari *planning, analysis, design, development, testing, deployment, dan maintenance* (Pressman and Maxim, 2015).

a. *Planning*

Tahap ini merupakan tahap untuk identifikasi dan penentuan ruang lingkup masalah dalam pengembangan sistem. Maka dari itu, di tahap ini dilakukan pengumpulan dan pencarian informasi yang akan dibutuhkan selama proses pengembangan. Setelah itu, direncanakan faktor-faktor penting lain yang dibutuhkan dalam pengembangan.

b. *Analysis*

Setelah dilakukan perencanaan, dilakukan analisis kebutuhan fungsional dan non-fungsional untuk sistem yang akan dikembangkan. Hal ini dilakukan agar tujuan

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

pengembangan sistem lebih jelas dan sistem yang dikembangkan sesuai dengan yang direncanakan.

c. *Design*

Setelah menentukan dan menganalisis kebutuhan sistem, dibuat desain atau rancangan sistem yang akan dikembangkan. Beberapa aspek yang dirancang adalah seperti arsitektur, tampilan antarmuka dan keamanan. Rancangan-rancangan ini kemudian akan ditinjau terlebih dahulu sebelum nanti diimplementasikan.

d. *Development*

Pada tahap ini, sistem akan dibangun berdasarkan rancangan yang sudah dibuat sebelumnya. Sistem akan dibangun dengan menulis kode menggunakan Bahasa pemrograman yang sudah ditentukan sebelumnya. Proses ini biasanya merupakan proses paling panjang dalam sebuah pengembangan sistem.

e. *Testing*

Pada tahap ini dilakukan pengujian terhadap sistem. Pengujian ini dilakukan untuk menilai apakah fungsionalitas sistem sesuai dengan yang ditentukan. Dalam pengujian ini, akan dipastikan sudah tidak ada lagi *bug* atau *error* pada sistem sehingga sistem akan berjalan stabil.

f. *Deployment*

Setelah memastikan sistem berjalan dengan baik pada tahap pengujian, maka dilakukan proses *deployment*. Pada tahap *deployment* ini, dilakukan penginstalan sistem yang telah dikembangkan. Hal ini dilakukan agar pengguna dapat menggunakan sistem yang telah dibangun.

g. *Maintenance*

Tahapan akhir dalam pengembangan sistem adalah maintenance atau pemeliharaan. Pada tahap ini, dilakukan pembaruan dalam sistem secara rutin agar sistem berjalan

dengan optimal dan stabil. Adapun pembaruan yang dilakukan adalah seperti perbaikan *bug*, peningkatan kinerja sistem dengan perangkat yang lebih baru dan penambahan fitur sesuai kebutuhan bisnis.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

BAB V PENUTUP

5.1 Simpulan

Berdasarkan penelitian, perancangan, implementasi serta evaluasi yang telah dilakukan, maka simpulan yang dapat diambil adalah sebagai berikut:

1. Telah berhasil dibuat aplikasi *website* pengelompokan ulasan produk di *e-marketplace* menggunakan algoritma K-Means. Berdasarkan hasil pengujian *clustering* menggunakan Davies-Bouldin Index dan *silhouette coefficient*, didapat pembagian kluster terbaik pada $k=2$ dengan DBI sebesar 0,705473 dan *silhouette coefficient* sebesar 0,510614.
2. Dari analisis kluster, kluster 0 merupakan kluster ulasan bersentimen negatif dengan *centroid* di titik 0.61129469 untuk skor positif dan 0.81219006 untuk skor negatif. Sedangkan, kluster 1 merupakan kluster ulasan bersentimen positif dengan *centroid* di titik 0.85114994 untuk skor positif dan 0.53761561 untuk skor negatif.
3. Dari pengujian sistem menggunakan *blackbox*, presentase keberhasilan sistem adalah 100% dari 9 item uji yang menunjukkan sistem berjalan dengan baik.

5.2 Saran

Setelah penelitian dan pengembangan sistem yang telah dilakukan, disadari bahwa masih ada kekurangan dalam penelitian ini. Karena itu, ada beberapa hal yang dapat dikembangkan pada penelitian selanjutnya, diantaranya:

1. Pada tahap *text preprocessing* dapat ditambahkan proses mengubah kata-kata non-formal menjadi formal dan perbaikan typo untuk mengurangi *noise*.
2. Menambahkan jumlah data ulasan agar model yang dibuat lebih optimal.



DAFTAR PUSTAKA

- Anandarajan, M., Hill, C. and Nolan, T. (2019) 'Text Preprocessing', in *Handbook of Natural Language Processing, Second Edition*, pp. 45–59. doi: 10.1007/978-3-319-95663-3_4.
- Babae, M. and Sarabadani, H. (2018) 'Data Analytics in Text Messages: A Mobile Network Operator Case Study', in *2018 9th International Symposium on Telecommunications (IST)*. IEEE, pp. 330–336. doi: 10.1109/ISTEL.2018.8660969.
- Capao, K. *et al.* (2018) 'Aspect Analysis of Cebu Establishments' Online Reviews using k-means Clustering and word2vec', in *2018 3rd International Conference on Computer and Communication Systems (ICCCS)*. IEEE, pp. 61–66. doi: 10.1109/CCOMS.2018.8463246.
- Fasheng Liu and Lu Xiong (2011) 'Survey on Text Clustering Algorithm', (4), pp. 901–904. doi: 10.1109/icsess.2011.5982485.
- iPrice (2021) *The Map of E-commerce in Indonesia*. Available at: <https://iprice.co.id/insights/mapofecommerce/en/> (Accessed: 24 June 2021).
- Jin, X., Zhang, S. and Liu, J. (2018) 'Word Semantic Similarity Calculation Based on Word2vec', *ICCAIS 2018 - 7th International Conference on Control, Automation and Information Sciences*, pp. 12–16. doi: 10.1109/ICCAIS.2018.8570612.
- Jo, T. (2019) *Text Mining*. 1st edn. Cham: Springer International Publishing (Studies in Big Data). doi: 10.1007/978-3-319-91815-0.
- Karo, I. M. K., MaulanaAdhinugraha, K. and Huda, A. F. (2017) 'A cluster validity for spatial clustering based on davies bouldin index and Polygon Dissimilarity function', in *2017 Second International Conference on Informatics and Computing (ICIC)*. IEEE, pp. 1–6. doi: 10.1109/IAC.2017.8280572.
- Kusuma, N. and Suhardi (2021) 'Detection of Online Prostitution in Twitter Platform Using Machine Learning Approach', *3rd 2021 East Indonesia Conference on Computer and Information Technology, EIconCIT 2021*, (October 2019), pp. 55–60. doi: 10.1109/EIconCIT50028.2021.9431864.
- Kyan, M. *et al.* (2014) 'Unsupervised Learning', in *Unsupervised Learning*. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 9–26. doi: 10.1002/9781118875568.ch2.
- Manalu, L. N. T., Arif Bijaksana, M. and Suryani, A. A. (2019) 'Analysis of the Word2Vec Model for Semantic Similarities in Indonesian Words', in *2019 7th International Conference on Information and Communication Technology (ICoICT)*. IEEE, pp. 1–5. doi: 10.1109/ICoICT.2019.8835330.
- Mao, Y., Zhang, G. and Zhang, S. (2020) 'Word Semantic Similarity Based on

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

CiLin and Word2vec', *Proceedings - 2020 International Conference on Culture-Oriented Science and Technology, ICCST 2020*, pp. 304–307. doi: 10.1109/ICCST50977.2020.00065.

Nelson, B. (2018) *Getting to Know Vue.js*. Berkeley, CA: Apress. doi: 10.1007/978-1-4842-3781-6.

Noviyanti, C. E. and Ruldeviyani, Y. (2020) 'Risk Classification of Peer-To-Peer Lending Platform Using SVM Algorithm', *2020 International Workshop on Big Data and Information Security, IWBIS 2020*, (March), pp. 29–34. doi: 10.1109/IWBIS50925.2020.9255512.

Pressman, R. S. and Maxim, B. R. (2015) *SOFTWARE ENGINEERING: A PRACTITIONER'S APPROACH, EIGHTH EDITION*. 8th edn. New York: McGraw-Hill Education.

Radim Řehůrek (2011) *Word2Vec Model*. Available at: https://radimrehurek.com/gensim/auto_examples/tutorials/run_word2vec.html (Accessed: 28 June 2021).

Ramadandy, B. (2020) 'Pengembangan Teknologi E-Marketplace Untuk Usaha Makanan Ringan Khas Kotawaringin Timur', *Journal of Information Sistem Research (JOSH)*, 2(1), pp. 6–13.

Raykov, Y. P. *et al.* (2016) 'What to Do When K-Means Clustering Fails: A Simple yet Principled Alternative Algorithm', *PLOS ONE*. Edited by B.-J. Yoon, 11(9), p. e0162259. doi: 10.1371/journal.pone.0162259.

Rehurek, R. and Sojka, P. (2010) 'Software Framework for Topic Modelling with Large Corpora', in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Valletta, Malta: University of Malta, pp. 45–50. Available at: <http://www.fi.muni.cz/usr/sojka/presentations/lrec2010-poster-rehurek-sojka.pdf>.

Saputra, T. I. and Arianty, R. (2019) 'IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING PADA ANALISIS SENTIMEN KELUHAN PENGGUNA INDOSAT', *Jurnal Ilmiah Informatika Komputer*, 24(3), pp. 191–198. doi: 10.35760/ik.2019.v24i3.2361.

Servanda, I. R. *et al.* (2019) 'PERAN ULASAN PRODUK DAN FOT PRODUK YANG DITAMPILKAN PENJUAL PADA MARKETPLACE SHOPEE TERHADAP MINAT BELI PRIA DAN WANITA The Role of Product Reviews and Product Photos Displayed by Sellers at Shopee Marketplace Against Men ' s and Women ' s Purchase Intere', *Journals of Universitas Teknologi Sumbawa*, 2(2), pp. 69–79.

Singh, A. K. *et al.* (2020) 'Clustering Evaluation by Davies-Bouldin Index(DBI) in Cereal data using K-Means', *Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020*, (Iccmc), pp. 306–310. doi: 10.1109/ICCMC48092.2020.ICCMC-00057.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Wirth, R. and Hipp, J. (2000) 'CRISP-DM: towards a standard process model for data mining', *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, (24959). Available at: https://www.researchgate.net/publication/239585378_CRISP-DM_Towards_a_standard_process_model_for_data_mining.

Yang, J., Sarathy, R. and Lee, J. (2016) 'The effect of product review balance and volume on online Shoppers' risk perception and purchase intention', *Decision Support Systems*, 89, pp. 66–76. doi: 10.1016/j.dss.2016.06.009.

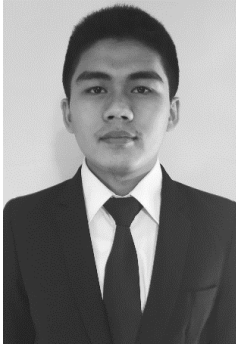
Yu, Q. and Yang, W. (2019) 'The Analysis and Design of System of Experimental Consumables Based on Django and QR code', in *2019 2nd International Conference on Safety Produce Informatization (IICSPI)*. IEEE, pp. 137–141. doi: 10.1109/IICSPI48186.2019.9095914.

Yuan, C. and Yang, H. (2019) 'Research on K-Value Selection Method of K-Means Clustering Algorithm', *J*, 2(2), pp. 226–235. doi: 10.3390/j2020016.

Zul, M. I., Yulia, F. and Nurmallasari, D. (2018) 'Social Media Sentiment Analysis Using K-Means and Naïve Bayes Algorithm', in *2018 2nd International Conference on Electrical Engineering and Informatics (ICon EEI)*. IEEE, pp. 24–29. doi: 10.1109/ICon-EEI.2018.8784326.

POLITEKNIK
NEGERI
JAKARTA

Lampiran 1 Daftar Riwayat Hidup



Ilham Hanif

Lahir di Padang Panjang, 22 Juli 2000. Lulus dari Pondok Pesantren Daar el-Qolam 2 Tangerang pada tahun 2017. Saat ini sedang menempuh pendidikan Diploma IV Program Studi Teknik Informatika di Politeknik Negeri Jakarta.



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

