

***CRAWLING ENGINE* PADA WEBSITE MANN,
BALDWIN, FLEETGUARD DAN PENGELOMPOKAN
PRODUK METODE ALGORITMA K-MEANS DI
PT RIVINDI ARTA MANDIRI**

Skripsi

Oleh :

Andi Rahman

201810225026



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2022**

LEMBAR PERSETUJUAN PEMBIMBING

LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : Crawling Engine pada website Mann,
Baldwin, Fleetguard dan Pengelompokan
Produk dengan Algoritma Clustering di PT
Rivindi Arta Mandiri

Nama Mahasiswa : Andi Rahman

Nomor Pokok Mahasiswa : 201810225026

Program Studi/Fakultas : INFORMATIKA/ILMU KOMPUTER

Tanggal Lulus Ujian Skripsi :

Jakarta, 20 Juli 2022

MENYETUJUI,

Pembimbing I

Pembimbing II

WOWON PRIATNA, ST., M.Ti.

Dr.Dra. TYASTUTI SRI LESTARI, M.M

NIDN : 0429118007

NIDN : 0327036701

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul Skripsi : Crawling Engine pada website Mann,
Baldwin, Fleetguard dan Pengelompokan
Produk dengan Algoritma Clustering di PT
Rivindi Arta Mandiri

Nama Mahasiswa : Andi Rahman

Nomor Pokok Mahasiswa : 201810225026

Program Studi/Fakultas : INFORMATIKA/ILMU KOMPUTER

Tanggal Lulus Ujian Skripsi : 25 Juli 2022

Bekasi, 1 Agustus 2022

MENGESAHKAN,

Ketua Tim : Dwipa Handayani, S.Kom., M.M.S.I.
penguji NIDN : 0317078008

Penguji II : Dwi Budi Srisulistiwati, S.Kom., M.M.
NIDN : 0323057701

Penguji III : Wowon Priatna, ST, M.Ti.
NIDN : 0429118007

MENGETAHUI,

Ketua Prodi
Informatika

Ahmad Fathurrozi, S.E., M.M.S.I.
NIP. 2012486

Dekan
Fakultas Ilmu Komputer

Dr.Dra. Tyastuti Sri Lestari, M.M
NIP. 1408206

LEMBAR PERNYATAAN

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa: Skripsi yang berjudul

Crawling Engine pada website Mann, Baldwin, Fleetguard dan Pengelompokan Produk dengan Algoritma Clustering di PT Rivindi Arta Mandiri.

ini adalah benar-benar merupakan hasil karya saya sendiri dan tidak mengandung materi yang ditulis oleh orang lain kecuali pengutipan sebagai referensi yang sumbernya telah dituliskan secara jelas sesuai dengan kaidah penulisan karya ilmiah. Apabila di kemudian hari ditemukan adanya kecurangan dalam karya ini, saya bersedia menerima sanksi dari Universitas Bhayangkara Jakarta Raya sesuai dengan peraturan yang berlaku. Saya mengizinkan skripsi ini dipinjam dan digandakan melalui Perpustakaan Universitas Bhayangkara Jakarta Raya. Saya memberikan izin kepada Perpustakaan Universitas Bhayangkara Jakarta Raya untuk menyimpan skripsi ini dalam bentuk digital dan mempublikasikannya melalui Internet selama publikasi tersebut melalui portal Universitas Bhayangkara Jakarta Raya.

Bekasi, 20 Juli 2022

Yang membuat pernyataan



Andi Rahman

NPM : 201810225026

ABSTRAK

Andi Rahman, 201810225026, *Crawling Engine* pada website Mann, Baldwin, Fleetguard dan Pengelompokan Produk dengan Algoritma K-Means di PT Rivindi Arta Mandiri, Pada penelitian ini akan mempunyai permasalahan adalah PT Rivindi Arta Mandiri kesulitan dalam melakukan penginputan data produk dan tidak adanya kelompok dari data produk tersebut. Dalam *crawling engine* akan sangat membantu dalam memasukan data produk secara otomatis mengambil data dari website produk tersebut, kemudian di input dalam aplikasi Odoo dengan bantuan API. Algoritma k-means klustering sendiri adalah salah satu metode data mining pengelompokan ke dalam kelompok k, di mana k merupakan parameter input. Tiap data kemudian ditetapkan pada setiap pengamatan cluster berdasarkan kedekatan pengamatan nilai rata-rata cluster. Pengelompokan ini akan sangat membantu dalam klasifikasi produk berdasarkan *cross reference*. Hasil dari penelitian ini adalah produk produk terinput secara otomatis dan data sesuai dengan website produk tersebut dan produk terkelompok sesuai dengan *cross reference*.

Kata Kunci : Crawling, K-Means, Clustering, Data Mining, Odoo, API

ABSTRACT

Andi Rahman, 201810225026, *Crawling Engine on Mann, Baldwin, Fleetguard and Product Clustering websites with K-means Algorithm at PT Rivindi Arta Mandiri, In this study will have problems, PT Rivindi Arta Mandiri has difficulty in inputting product data and there is no group of product data. In crawling engines, it will be very helpful in entering product data automatically taking data from the product website, then inputting it in the Odoo application. The k-means clustering algorithm itself is an algorithm of grouping observations into k groups, where k is the input parameter. Each data is then assigned to each cluster observation based on the proximity of the observation of the cluster mean value. This grouping will be very helpful in the classification of products based on cross references. The result of this study is that the product is inputted automatically and the data is in accordance with the product's website and the product is grouped according to cross reference.*

Keywords : *Crawling, K-Means, Clustering, Data Mining, Odoo, API.*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Andi Rahman
NPM : 201810225026
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty-Free Right*), atas karya ilmiah saya yang berjudul : Crawling Engine pada website Mann, Baldwin, Fleetguard dan Pengelompokan Produk dengan Algoritma Clustering di PT Rivindi Arta Mandiri beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta. Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 26 Juli 2022
Yang Menyatakan


41515A JX983868346
Andi Rahman
201810225026



KATA PENGANTAR

Puji syukur penulis panjatkan Kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya yang telah memberikan kesempatan pada penulis untuk menyelesaikan tugas akhir dilakukan dalam rangka memenuhi salah satu prasyarat untuk memperoleh gelar sarjana komputer Program Studi Informatika. Oleh karena itu, penulis mengucapkan terima kasih kepada :

1. Bapak Irjen Pol (Purn) Drs. Bambang Karsono, S.H.,M.M selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M., selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E.,M.MSI., selaku Ketua Prodi Informatika.
4. Bapak Wowon Priatna, S.T. M.TI selaku dosen pembimbing 1 yang telah menyediakan waktunya dalam mengarahkan dan membimbing penulis dalam menyelesaikan tugas akhir.
5. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M., selaku dosen pembimbing 2 yang telah menyediakan waktunya dalam mengarahkan dan membimbing penulis dalam menyelesaikan tugas akhir.
6. Ibu Herlawati, S.Si., MM., M.Kom., selaku Dosen Wali yang telah banyak memberikan bimbingan selama perkuliahan.
7. Kepada teman teman mahasiswa Ilmu Komputer Angkatan 2018 yang telah membantu dan memberikan dukungan kepada penulis dalam menyelesaikan penulisan tugas akhir.
8. Orang Tua, Keluarga dan Nur Asrimawati selaku istri yang telah memberikan bantuan dukungan secara moral maupun material.

Penulis menyadari bahwa masih banyak kekurangan dalam penyusunan tugas akhir ini. Saran dan kritik sangat diharapkan untuk perbaikan yang akan datang.

Bekasi 25 Juli 2022

Hormat saya

Andi Rahman

201810225026



DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN	vi
ABSTRAK	vii
<i>ABSTRACT</i>	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vii
KATA PENGANTAR	viii
Daftar Isi	ix
Daftar Tabel	xii
Daftar Gambar	xiii
Daftar Lampiran	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	2
1.3 Rumusan Masalah	2
1.4 Batasan Masalah	2
1.5 Tujuan dan Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II LANDASAN TEORI	5
2.1 Penelitian Terdahulu	5
2.2 <i>Crawling</i>	7
2.2.1 Pengertian <i>Crawling</i>	7

2.2.2	Web Crawler	8
2.2.3	Tahapan Crawling	8
2.3	Data Mining	9
2.3.1	Pengertian Data Mining	9
2.3.2	Preprosesing Data	10
2.3.3	Data	11
2.3.4	Clustering	11
2.4	Algoritma K-Means	12
2.4.1	Pengertian K-Means dan langkah langkah Clustering	12
2.4.2	Tahapan K-Means	12
2.5	Odoo ERP	13
2.6	API	13
2.7	Pengertian Produk	14
2.8	Pengertian Merek	14
2.8	Pengertian Unified Modelling Language	14
2.8.1	Activity Diagram	15
2.8.2	Use Case Diagram	15
BAB III	METODELOGI PENELITIAN	17
3.1	Tempat dan Waktu Penelitian	17
3.2	Kerangka Pikir Peneliti	17
3.3	Metode Pengumpulan Data	19
3.3.1	Metode Pengumpulan Data Primer	19
3.3.2	Metode Pengumpulan Data Sekunder	19
3.3.3	Studi Pustaka	20
3.4	Metode Analisis	20
3.5	Pengumpulan Data	20
3.6	Preprocessing Data	21
3.7	Pengelolaan Data	22

3.7 Tahap Perhitungan K-Means	22
3.8 Analisis Sistem Berjalan	24
3.9 Analisis Usulan	25
BAB IV PERANCANGAN SISTEM DAN IMPLEMENTASI	27
4.1 Hasil Pengujian Sistem	27
4.1.1 Pengujian <i>Crawling</i>	27
4.1.2 Hasil Pengujian <i>Crawling</i>	31
4.1.3 Pengolahan Data	35
4.1.4 Proses Data	37
4.1.4 Penentuan Jumlah Cluster	38
4.1.5 Perhitungan Jarak Terdekat	38
4.1.6 Iterasi	40
4.1.7 Representasi Aplikasi	44
4.2 Pembahasan	45
BAB V PENUTUP	46
5.1 Kesimpulan	46
5.2 Saran	46
Daftar Pustaka	47
PLAGIARISME	70
BIODATA MAHASISWA	71
KARTU BIMBINGAN	72
Kartu Binbingan	73

DAFTAR TABEL

Tabel 2.1 Tabel Perbandingan Studi Literatur	5
Tabel 2.2 Simbol-simbol <i>Activity Diagram</i>	15
Tabel 2.3 <i>Use Case Diagram</i>	16
Tabel 3.1 Inisialisasi <i>cross reference</i>	20
Tabel 3.2 Data produk dengan <i>cross reference</i>	21
Tabel 3.3 Contoh data produk dengan <i>cross reference</i>	22
Tabel 3.4 Hasil perhitungan dengan C1, C2, C3	23
Tabel 4.1 Hasil Pengujian <i>Crawling</i>	31
Tabel 4.2 Hasil pengelompokan produk	36
Tabel 4.3 Tabel Kelompok produk yang memiliki <i>cross reference</i>	36
Tabel 4.4 Data penentuan jumlah cluster	38
Tabel 4.5 Hasil perhitungan dengan titik centroid	38
Tabel 4.6 Hasil perhitungan jarak pusat	38
Tabel 4.7 Euclidian	40
Tabel 4.8 Hasil Iterasi ke-2	41
Tabel 4.9 Hasil Euclidian ke 2	42
Tabel 4.10 Hasil Iterasi ke-3	42
Tabel 4.10 Hasil Iterasi ke-3	42

DAFTAR GAMBAR

Gambar 2.1 Logo Software ERP Odoo.....	13
Gambar 3.1 Kerangka Pikir Peneliti.....	18
Gambar 4.1 Website Mann.....	27
Gambar 4.2 Website Baldwin.....	28
Gambar 4.3 Website Fleetguard.....	29
Gambar 4.4 Pemanggilan Library Selenium.....	29
Gambar 4.5 Odoo API.....	29
Gambar 4.6 Pencarian Data Produk.....	29
Gambar 4.7 Pembuatan data produk pada Odoo.....	29
Gambar 4.8 Contoh hasil <i>crawling</i> data produk.....	30
Gambar 4.9 Tampilan produk BF 1018/1 pada website Mann.....	30
Gambar 4.10 Tabel charter hasil pengelompokan produk.....	44
Gambar 4.11 Tampilan login pada sistem aplikasi Odoo.....	44
Gambar 4.12 Tampilan hasil pengelompokan produk pada aplikasi Odoo.....	45

DAFTAR LAMPIRAN

Lampiran 1: Surat Rekomendasi Tempat Penelitian.....	50
Lampiran 2 : Lembar Wawancara.....	51
Lampiran 3 : Lampiran Data produk.....	52
Lampiran 4 : <i>Source Code</i>	55
Code K-Means Clustering.....	55
Code <i>Crawling</i>	59

