

**IMPLEMENTASI ALGORITMA ASOSIASI *FP-GROWTH* DAN  
ALGORITMA KLASIFIKASI *K-MEANS* TERHADAP POLA  
PEMBELIAN KONSUMEN DI *MARKETPLACE SHOPEE***

**SKRIPSI**



**Disusun Oleh :**

**MUHAMAD ARIF SAIFUDIN**

**19081010011**

**PROGRAM STUDI INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"**

**JAWA TIMUR**

**2024**

## LEMBAR PENGESAHAN SKRIPSI

Judul : IMPLEMENTASI ALGORITMA ASOSIASI *FP-GROWTH* DAN  
KLASIFIKASI *K-MEANS* TERHADAP ANALISIS POLA  
PEMBELIAN KONSUMEN DI *MARKETPLACE SHOPEE*

Oleh : MuhamadArif Saifudin

NPM : 19081010011

Telah Diseminarkan Dalam Ujian Skripsi

Pada : Hari Jumat, 05 Januari 2024

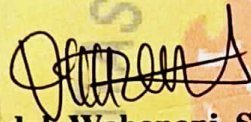
Mengetahui,

Dosen Pembimbing

Dosen Penguji

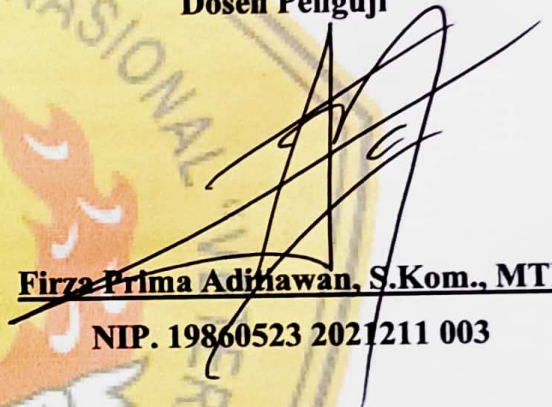
1.

1.



Henni Endah Wahanani, ST, M.Kom

NIP. 19780922 2021212 005



Firza Prima Aditawan, S.Kom., MTI

NIP. 19860523 2021211 003

2.

2.



Achmad Junaidi, S.Kom, M.Kom

NPT. 3 7811 04 0199 1



Andreas Nugroho Sihananto, S.Kom., M.Kom

NPT. 211199 00 412271

Menyetujui,

Dekan

Fakultas Ilmu Komputer

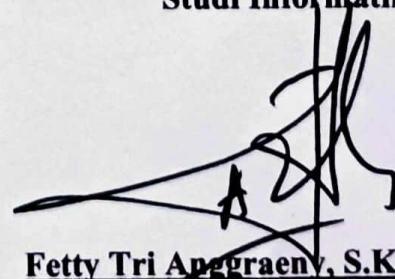


Prof. Dr. Ir. Novirina Hendrasarie, M.T

NIP. 19681126 199403 2 001

Koordinator Program

Studi Informatika



Fetty Tri Anggraeny, S.Kom, M.Kom

NIP. 19820211 2021212 005

## SURAT PERNYATAAN ORISINILITAS

Saya, mahasiswa Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur, yang bertandatangan di bawah ini :

Nama : Muhamad Arif Saifudin

NPM : 19081010011

Menyatakan bahwa judul skripsi/tugas akhir yang saya ajukan dan dikerjakan yang berjudul :

**“IMPLEMENTASI ALGORITMA ASOSIASI *FP-GROWTH* DAN  
ALGORITMA KLASIFIKASI *K-MEANS* TERHADAP POLA  
PEMBELIAN KONSUMEN DI *MARKETPLACE SHOPEE*”**

Bukan merupakan plagiat dari skripsi/tugas akhir/penelitian orang lain dan juga merupakan produk dan atau software yang saya beli dari pihak lain. Saya juga menyatakan bahwa skripsi ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam daftar pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di UPN “Veteran” Jawa Timur maupun di institusi pendidikan lain. Jika ternyata di kemudian hari pernyataan ini terbukti tidak benar, maka saya siap menerima segala konsekuensinya.

Surabaya, 5 Januari 2024

Hormat saya,



Muhamad Arif Saifudin

NPM. 19081010011

# IMPLEMENTASI ALGORITMA ASOSIASI *FP-GROWTH* DAN ALGORITMA KLASIFIKASI *K-MEANS* TERHADAP POLA PEMBELIAN KONSUMEN DI *MARKETPLACE SHOPEE*

Nama Mahasiswa : Muhamad Arif Saifudin  
NPM : 19081010011  
Program Studi : Informatika  
Dosen Pembimbing : Henni Endah Wahanani, ST, M.Kom  
Achmad Junaidi, S.Kom, M.Kom

## ABSTRAK

*Online shop* merupakan salah satu alternatif yang lebih mudah dan praktis dalam transaksi yang merupakan salah satu dampak perkembangan teknologi dan internet. Saat ini banyak konsumen lebih memilih belanja atau transaksi secara *online* yang memiliki banyak kelebihan salah satunya di *marketplace* shopee. Shopee merupakan salah satu *e-commerce* yang paling banyak dikunjungi selama periode tahun 2020.

Saat Bulan Ramadhan menjelang Hari Raya Idul Fitri banyak konsumen yang mencari berbagai produk busana muslim atau baju koko. Kebutuhan akan busana muslim atau baju koko meningkat seiring dengan momen tersebut. Pusatbusanamuslim1 merupakan salah satu toko *online* di *marketplace* shopee yang menjual berbagai macam baju koko atau busana muslim harus memahami pola pembelian konsumen untuk mengoptimalkan penjualan mereka agar dapat bersaing dengan toko lainnya.

Oleh karena itu digunakan algoritma Klasifikasi *K-Means* dan Asosiasi *FP-Growth* untuk analisis pola pembelian konsumen di *marketplace* shopee. Tujuan utama dari implementasi dua metode tersebut untuk menghasilkan rekomendasi produk dan segmentasi pelanggan. Dalam implementasi algoritma Klasifikasi *K-Means* menggunakan model *LRFM* untuk segmentasi pelanggan kemudian dilakukan penentuan jumlah Kluster yang paling optimal dengan metode *Elbow* yang hasilnya nanti akan dilabelkan pada hasil segmentasi pelanggan dengan model

*LRFM*. Sementara untuk implementasi algoritma Asosiasi *FP-Growth* akan dibangun *FP-Tree* dan pencarian aturan asosiasi dengan nilai *minimum support* dan *minimum confidence* yang telah ditentukan.

Dari hasil pengujian dengan *K-Means* diperoleh tingkat akurasi sebesar 56% dari 3 klaster (Klaster 0, Klaster 1, dan Klaster 2) dengan Klaster 2 merupakan Klaster terbaik hasil dari segmentasi pelanggan dengan model *LRFM*. Sementara dari hasil implementasi algoritma *FP-Growth* didapatkan beberapa aturan asosiasi antara lain Nibras Sarimbit Chesa Brown Couple sering dibeli jika pelanggan membeli Nibras Sarimbit 70 Coklat dengan nilai *minimum support* sebesar 0.50 % dan nilai *lift ratio* sebesar 4.6 %. *Dataset* yang digunakan yaitu data transaksi penjualan produk sebanyak 349.

**Kata kunci** : *K-Means*, *FP-Growth*, Analisis Pola, Shopee, *LRFM*, *FP-Tree*

## KATA PENGANTAR

Puji Syukur kami panjatkan kehadiran Allah SWT atas segala limpahan Rahmat, hidayah, serta karunia-Nya, yang telah melimpahkan berkah-Nya sehingga dapat menyelesaikan penelitian ini. Dalam kesempatan ini, penulis ingin mengucapkan penghargaan yang setinggi-tingginya kepada semua pihak yang telah memberikan dukungan, bimbingan, dan sumbangan berharga dalam penulisan skripsi yang berjudul :

**“IMPLEMENTASI ALGORITMA ASOSIASI *FP-GROWTH* DAN  
ALGORITMA KLASIFIKASI *K-MEANS* TERHADAP POLA PEMBELIAN  
KONSUMEN DI *MARKETPLACE SHOPEE*”**

Semoga hasil penelitian ini dapat memberikan koontribusi bermanfaat dalam bidang yang lebih luas, serta menjadi pijakan untuk penelitian-penelitian selanjutnya. Segala keterbatasan dalam kata tak mampu mengungkapkan segenap rasa terima kasih kami. Terima kasih atas semua yang telah membantu dan mendukung kami dalam perjalanan ini.

Akhir kata, mohon maaf apabila terdapat kekurangan dalam penyusunan kata pengantar ini. Kritik, saran, dan masukan yang membangun sangat kami harapkan guna perbaikan di masa yang akan dating. Sekian kata pengantar ini kami sampaikan, atas perhatian dan pengertiannya, kami ucapkan terima kasih.

Surabaya, 5 Januari 2024

Hormat saya,

Muhamad Arif Saifudin

NPM. 19081010011

## UCAPAN TERIMA KASIH

Puji syukur penulis panjatkan atas kehadiran Allah SWT yang telah memberikan rahmat, ridha, dan karunia-Nya. Dengan tulus dan penuh rasa syukur, penulis mengucapkan terima kasih yang tak terhingga atas bantuan, dukungan, dan bimbingan yang luar biasa selama proses penulisan skripsi ini. Proses penyelesaian skripsi ini juga tak lepas dari bantuan dan dukungan berbagai pihak, untuk itu penulis ingin menyampaikan ucapan terima kasih kepada :

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Dr. Novirina Hendrasarie, S.T, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom, M.Kom. selaku Koordinator Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Bapak Budi Nugroho, S.Kom. M.Kom selaku dosen wali yang membantu dalam perwalian dari awal sampai akhir perkuliahan.
5. Henni Endah Wahanani, ST, M.Kom selaku dosen pembimbing pertama yang sangat membantu dan memberikan arahan sehingga dapat menyelesaikan tugas akhir ini dengan baik
6. Achmad Junaidi, S.Kom, M.Kom selaku pembimbing kedua yang telah membimbing dalam melakukan memberikan arahan pada tugas akhir ini dengan maksimal.
7. Seluruh Dosen dan Staf Tata usaha Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur yang telah memberikan ilmu yang bermanfaat selama perkuliahan.
8. Orang Tua, Kakak, dan Keluarga yang telah memberikan dukungan secara materi dan non-materi serta memberikan doa sehingga penulis dapat menyelesaikan perkuliahan dari awal hingga akhir dengan baik.
9. Pemilik toko di marketplace shopee pusatbusanamuslim1 yang telah mau bekerja sama dalam memberikan data laporan penjualan produk baju koko selama beberapa periode penjualan.

10. Fitria Putri Nabila, S.M yang telah meluangkan banyak waktu untuk menemani serta memberikan bantuan dan doa sehingga penulis dapat menyelesaikan penelitian ini.
11. Ilham Ainur Idhana, Eriyansyah Yusuf Suwandan, Thomas Andrew Imanzaghi, Abiyan Naufal Hilmi dan Muhammad Isra Alfajri yang telah banyak membantu dan memberikan masukan serta ide-ide cemerlangnya.
12. Teman-teman seperjuangan Informatika angkatan 2019 serta seluruh pihak yang telah membantu dan mendukung penulis dalam menyelesaikan skripsi dan perkuliahan ini.

Akhir kata, penulis mengharapkan penelitian ini dapat memberikan manfaat bagi penulis khususnya dan bagi pembaca pada umumnya.. Semoga Allah Subhanahu Wa Ta'ala memberikan balasan yang berlipat ganda atas kebaikan yang telah diberikan.

Surabaya, 5 Januari 2024

Hormat saya,

Muhamad Arif Saifudin

NPM. 19081010011



## DAFTAR ISI

<b>LEMBAR PENGESAHAN SKRIPSI .....</b>	<b>i</b>
<b>ABSTRAK .....</b>	<b>iii</b>
<b>KATA PENGANTAR.....</b>	<b>v</b>
<b>UCAPAN TERIMA KASIH .....</b>	<b>vi</b>
<b>DAFTAR ISI.....</b>	<b>viii</b>
<b>DAFTAR GAMBAR .....</b>	<b>xi</b>
<b>DAFTAR TABEL .....</b>	<b>xii</b>
<b>DAFTAR KODE PROGRAM .....</b>	<b>xiii</b>
<b>BAB 1 PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah.....	5
1.4 Tujuan.....	5
1.5 Manfaat.....	6
<b>BAB 2 TINJAUAN PUSTAKA .....</b>	<b>7</b>
2.1 Penelitian Terdahulu .....	7
2.2 Perilaku Konsumen .....	16
2.3 <i>E-Commerce</i> Shopee .....	16
2.4 <i>Data Mining</i> .....	17
2.5 <i>Clustering</i> .....	18
2.6 Algoritma K-Means .....	19
2.7 <i>Model LRFM</i> .....	20
2.8 <i>Outlier</i> .....	22
2.9 <i>Elbow Curve Method</i> .....	22
2.10 <i>Association Rules</i> .....	23

2.11	<i>Algoritma FP-Growth</i> .....	25
2.12	<i>FP-Tree</i> .....	26
<b>BAB 3 METODOLOGI PENELITIAN .....</b>		<b>32</b>
3.1	Tahapan Penelitian.....	32
3.2	Metode Pengumpulan Data .....	33
3.3	Jenis Data .....	34
3.4	Analisis Kebutuhan.....	34
3.5	Kebutuhan Data .....	34
3.6	Kebutuhan <i>Hardware</i> dan <i>Software</i> .....	35
3.7	Pengolahan Data .....	35
3.8	Pemodelan <i>LRFM</i> .....	38
3.9	Pencarian <i>Minimum Support</i> dan <i>Lift Ratio</i> .....	41
3.10	Pembentukan <i>FP-Tree</i> .....	45
3.11	Pembuatan Model Klasifikasi.....	49
3.12	Skema Pengujian .....	51
<b>BAB 4 HASIL DAN PEMBAHASAN .....</b>		<b>53</b>
4.1	Pengumpulan Data .....	53
4.2	Implementasi Algoritma .....	54
4.3	Implementasi Algoritma <i>K-Means</i> .....	55
4.3.1	Preprocessing Data .....	55
4.3.2	Normalisasi RFM.....	60
4.3.3	Proses Penggabungan Data Hasil RFM.....	64
4.3.4	Pencarian Outlier .....	66
4.3.5	<i>Min-Max Scalling</i> .....	67
4.3.6	Proses Penentuan Jumlah Kluster dengan Elbow Curve .....	68
4.3.7	Clustering.....	70

4.3.8 Analisis Segmentasi dengan Agregasi .....	72
4.3.9 Visualisasi Hasil Clustering dengan Boxplot.....	75
4.4 Implementasi Algoritma <i>FP-Growth</i> .....	77
4.4.1 Preprocessing Data .....	77
4.4.2 Implementasi Apriori .....	78
4.4.3 Pencarian <i>Frequent Itemset</i> .....	79
4.4.4 Visualisasi dengan <i>Treemap Chart</i> .....	80
4.4.5 Implementasi <i>Association Rules</i> dengan <i>FP-Growth</i> .....	83
4.4.6 Pencarian Aturan Asosiasi Antar <i>Item</i> .....	84
4.4.7 Hasil Evaluasi Model <i>FP-Growth</i> .....	85
<b>BAB 5 PENUTUP</b> .....	<b>89</b>
5.1 Kesimpulan.....	89
5.2 Saran .....	90
<b>REFERENSI</b> .....	<b>91</b>
<b>LAMPIRAN</b> .....	<b>96</b>

## DAFTAR GAMBAR

<b>Gambar 2.1 Pembentukan FP-Tree</b> .....	28
<b>Gambar 2.2 Pembentukan FP-Tree</b> .....	29
<b>Gambar 2.3 Pembentukan FP-Tree</b> .....	29
<b>Gambar 2.4 Pembentukan FP-Tree</b> .....	30
<b>Gambar 3.1 Diagram Alur Penelitian</b> .....	32
<b>Gambar 3.2 Pembentukan FP-Tree</b> .....	48
<b>Gambar 3.3 Diagram Alir Algoritma K-Means</b> .....	49
<b>Gambar 3.4 Diagram Alir Algoritma K-Means</b> .....	50
<b>Gambar 4.1 Dataframe Output Dataset</b> .....	54
<b>Gambar 4.2 Output Hasil Proses Tokenizing dan Split Nama Barang</b> .....	58
<b>Gambar 4.3 Output Hasil Data Profiling</b> .....	59
<b>Gambar 4.4 Output Atribut Monetary</b> .....	61
<b>Gambar 4.5 Output Atribut Frequency</b> .....	62
<b>Gambar 4.6 Output Atribut Recency</b> .....	64
<b>Gambar 4.7 Dataframe Output Merge RFM</b> .....	65
<b>Gambar 4.8 Output Proses Outlier</b> .....	67
<b>Gambar 4.9 Output Min-Max Scalling</b> .....	68
<b>Gambar 4.10 Output Elbow Curve</b> .....	70
<b>Gambar 4.11 Output Hasil Pengujian K-Means Clustering</b> .....	71
<b>Gambar 4.12 Output Agregasi</b> .....	72
<b>Gambar 4.13 Output Visualisasi menggunakan Boxplot</b> .....	76
<b>Gambar 4.14 Output Split Nama Barang</b> .....	78
<b>Gambar 4.15 Output Implementasi Apriori</b> .....	78
<b>Gambar 4.16 Output Frequent Itemset</b> .....	80
<b>Gambar 4.17 Output Visualisasi Menggunakan Treemap Chart</b> .....	81
<b>Gambar 4.18 Informasi Frekuensi Item</b> .....	82
<b>Gambar 4.19 Output Pembentukan Matriks Binary</b> .....	83
<b>Gambar 4.20 Output Implementasi Algoritma FP-Growth</b> .....	84
<b>Gambar 4.21 Output Aturan Asosiasi Antar Item</b> .....	85

## DAFTAR TABEL

<b>Tabel 2.1 Contoh Data Transaksi Mentah .....</b>	<b>26</b>
<b>Tabel 2.2 Kemunculan Setiap Item.....</b>	<b>27</b>
<b>Tabel 2.3 Data Transaksi .....</b>	<b>28</b>
<b>Tabel 3.1 Contoh Dataset Sebelum Preprocessing .....</b>	<b>35</b>
<b>Tabel 3.2 Dataset Setelah Proses Tokenizing .....</b>	<b>37</b>
<b>Tabel 3.3 Dataset Sebelum Dilakukan Normalisasi RFM .....</b>	<b>38</b>
<b>Tabel 3.4 Data Transformation .....</b>	<b>39</b>
<b>Tabel 3.5 Hasil Normalisasi RFM.....</b>	<b>41</b>
<b>Tabel 3.6 Frekuensi Setiap Item Dari Semua Transaksi.....</b>	<b>41</b>
<b>Tabel 3.7 Pelabelan Data Transaksi.....</b>	<b>43</b>
<b>Tabel 3.8 Item Transaksi.....</b>	<b>45</b>
<b>Tabel 3.9 Kemunculan Setiap Item.....</b>	<b>46</b>
<b>Tabel 3.10 Frekuensi Kemunculan Setiap Item Setelah Dilakukan Penyortiran .....</b>	<b>47</b>
<b>Tabel 4.1 Perbandingan Data Sebelum dan Sesudah Tokenizing.....</b>	<b>56</b>
<b>Tabel 4.2 Perbandingan Data Sebelum dan Sesudah Splitting.....</b>	<b>57</b>
<b>Tabel 4.3 Hasil Pengujian FP-Growth.....</b>	<b>85</b>

## DAFTAR KODE PROGRAM

<b>Kode Program 4.1 Pemanggilan Dataset .....</b>	<b>53</b>
<b>Kode Program 4.2 Proses Tokenizing.....</b>	<b>55</b>
<b>Kode Program 4.3 Proses Split Nama Barang .....</b>	<b>57</b>
<b>Kode Program 4.4 Proses Data Profiling .....</b>	<b>59</b>
<b>Kode Program 4.5 Proses Pencarian Atribut Monetary .....</b>	<b>60</b>
<b>Kode Program 4.6 Proses Pencarian Atribut Frequency .....</b>	<b>61</b>
<b>Kode Program 4.7 Proses Pencarian Atribut Recency .....</b>	<b>63</b>
<b>Kode Program 4.8 Proses Merge Data Hasil RFM.....</b>	<b>64</b>
<b>Kode Program 4.9 Proses Outlier.....</b>	<b>66</b>
<b>Kode Program 4.10 Proses Min-Max Scalling.....</b>	<b>67</b>
<b>Kode Program 4.11 Pengujian dengan K-Means Clustering.....</b>	<b>71</b>
<b>Kode Program 4.12 Proses Segmentasi Menggunakan Agregasi .....</b>	<b>72</b>
<b>Kode Program 4.13 Visualisasi menggunakan Boxplot.....</b>	<b>75</b>
<b>Kode Program 4.14 Proses Split Nama Barang .....</b>	<b>77</b>
<b>Kode Program 4.15 Implementasi Apriori .....</b>	<b>78</b>
<b>Kode Program 4.16 Proses Pencarian Frequent Itemset.....</b>	<b>79</b>
<b>Kode Program 4.17 Visualisasi Menggunakan Treemap Chart .....</b>	<b>80</b>
<b>Kode Program 4.18 Proses Pembentukan Matriks Binary .....</b>	<b>82</b>
<b>Kode Program 4.19 Implementasi Algoritma FP-Growth .....</b>	<b>83</b>
<b>Kode Program 4.20 Proses Pencarian Aturan Asosiasi Antar Item .....</b>	<b>84</b>