



**Penerapan Algoritma KNN dengan metode Collaborative  
Filtering dalam Mencari Produk Rekomendasi Asuransi**

**LAPORAN TUGAS AKHIR**

**MUHAMMAD FARHAN ATSAR GHIFFARI**

**UNIVERSITAS**  
**41520010199**  
**MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS MERCU BUANA**

**JAKARTA**

**2024**



**Penerapan Algoritma KNN dengan metode Collaborative  
Filtering dalam Mencari Produk Rekomendasi Asuransi**

**LAPORAN TUGAS AKHIR**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

**MUHAMMAD FARHAN ATSAR GHIFFARI**

**UNIVERSITAS**  
**41520010199**  
**MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS MERCU BUANA**

**JAKARTA**

**2024**

## HALAMAN PENYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Farhan Atsar Ghiffari  
NIM : 41520010199  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma KNN dengan metode Collaborative Filtering dalam Mencari Produk Rekomendasi Asuransi

Menyatakan bahwa Laporan Aplikatif/Tugas Akhir/Jurnal/Media Ilmiah ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

UNIVERSITAS  
MERCU BUANA

Jakarta, 22 Juli 2024



Muhammad Farhan Atsar Ghiffari.

## HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Muhammad Farhan Atsar Ghiffari  
NIM : 41520010199  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma KNN dan Pearson Correlation dengan metode Collaborative Filtering dalam Mencari Produk Rekomendasi Asuransi.

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Wawan Gunawan, S.Kom., M.T.  
NIDN : 0424108104  
Ketua Penguji : Prastika Indriyanti, S.Kom., MCS.  
NIDN : 0312089401  
Penguji 1 : Ir. Emil Robert Kaburuan, Ph.D., IPM., ASEAN Eng.  
NIDN : 0429058004  
Penguji 2 : Yuwan Jumaryadi, S.Kom, MM, M.Kom.  
NIDN : 0319078704

Jakarta, 22 Juli 2024

Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jekonowo, S.Si., M.T.I.

NIDN : 0320037002



Dr. Hadi Santoso, S.Kom., M.Kom

NIDN : 0225067701



## KATA PENGANTAR

Puji syukur saya panjatkan kepada Allah SWT, Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Laporan Skripsi ini. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan Laporan Skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., M.T.I selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Bapak Wawan Gunawan, S.Kom., M.T. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Ir. Emil Robert Kaburuan, Ph.D., IPM., ASEAN Eng. selaku Dosen Penguji 1 Tugas Akhir atas koreksi dan arahan serta masukannya.
6. Yuwan Jumaryadi, S.Kom., M.M., M.Kom. selaku Dosen Penguji 2 Tugas Akhir atas koreksi dan arahan serta masukannya.

Akhir kata, saya berharap Allah SWT, Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 22 Juli 2024

Muhammad Farhan Atsar Ghiffari



**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS  
AKHIR UNTUK KEPENTINGAN AKADEMIS**

Nama : Muhammad Farhan Atsar Ghiffari  
NIM : 41520010199  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma KNN dengan metode Collaborative Filtering dalam Mencari Produk Rekomendasi Asuransi.

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalty-Free Right) atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 22 Juli 2024

Yang menyatakan,



Muhammad Farhan atsar Ghiffari

## ABSTRAK

Nama : Muhammad Farhan Atsar Ghiffari  
NIM : 41520010199  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma KNN dengan metode Collaborative Filtering dalam Mencari Produk Rekomendasi Asuransi  
Pembimbing : Wawan Gunawan, S.kom., M.T

Semakin meningkatnya inflasi medis di Indonesia, asuransi telah menjadi sangat penting untuk memberikan perlindungan dan keamanan finansial bagi individu dan keluarga. Meskipun kesadaran akan pentingnya kesehatan semakin tinggi di masyarakat Indonesia, banyak orang masih memiliki pandangan yang beragam terhadap asuransi. Hal ini seringkali menyebabkan kebingungan dalam memilih jenis asuransi yang sesuai dengan kebutuhan mereka. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi asuransi yang memanfaatkan model IndoBERT, Word2Vec, dan metode Collaborative Filtering. Data yang digunakan dalam penelitian ini diambil dari opini yang diungkapkan di media sosial Twitter. Variabel yang digunakan meliputi teks opini, produk\_id, user\_id, dan rating. Dengan menggunakan teknologi ini, diharapkan dapat memberikan rekomendasi yang lebih personal dan akurat bagi pengguna. Hasil penelitian menunjukkan bahwa sistem rekomendasi ini mampu mencapai Mean Absolute Error (MAE) sebesar 0.6370 dan Root Mean Square Error (RMSE) sebesar 0.8688 pada  $k=17$ , serta akurasi klasifikasi teks sebesar 43%. Sistem ini dirancang berbasis website, sehingga pengguna dapat langsung melihat hasil rekomendasi asuransi yang sesuai dengan kebutuhan mereka. Dengan adanya sistem rekomendasi ini, dapat membantu masyarakat Indonesia dalam memilih asuransi yang tepat. Hal ini sangat penting mengingat biaya medis yang terus meningkat, sehingga asuransi yang sesuai dapat memberikan perlindungan optimal.

**Kata Kunci : Collaborative Filtering, KNN, Sistem Rekomendasi, SVM, Word2vec.**

## ABSTRACT

Name : Muhammad Farhan Atsar Ghiffari  
NIM : 41520010199  
Study Program : *Informatics Engineering*  
Title Thesis : Penerapan Algoritma KNN dengan metode Collaborative Filtering dalam Mencari Produk Rekomendasi Asuransi  
Counsellor : Wawan Gunawan, S.kom., M.T

*With medical inflation on the rise in Indonesia, insurance has become very important to provide protection and financial security for individuals and families. Despite the growing awareness of the importance of health in Indonesian society, many people still have mixed views on insurance. This often leads to confusion in choosing the type of insurance that suits their needs. This research aims to develop an insurance recommendation system that utilizes IndoBERT model, Word2Vec, and Collaborative Filtering method. The data used in this research is taken from opinions expressed on Twitter social media. The variables used include opinion text, product\_id, user\_id, and rating. By using this technology, it is expected to provide more personalized and accurate recommendations for users. The results showed that this recommendation system was able to achieve a Mean Absolute Error (MAE) of 0.6370 and Root Mean Square Error (RMSE) of 0.8688 at k=17, as well as text classification accuracy of 43%. This system is designed to be web-based, so that users can directly see the results of insurance recommendations that suit their needs. With this recommendation system, it is hoped that it can help the Indonesian people in choosing the right insurance. This is very important considering the increasing medical costs, so that the appropriate insurance can provide optimal protection. The use of advanced technologies such as IndoBERT and KNN shows great potential in improving the accuracy of opinion-based recommendations, providing more personalized and appropriate solutions for users. This system not only simplifies the insurance selection process, but also increases public satisfaction and trust in insurance products in the market.*

**Keyword:** *Collaborative Filtering, KNN, System Recommendation, SVM, Word2vec.*



## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>HALAMAN PERNYATAAAN.....</b>	<b>ii</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>iii</b>
<b>KATA PENGANTAR.....</b>	<b>iv</b>
<b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....</b>	<b>v</b>
<b>ABSTRAK .....</b>	<b>vi</b>
<b>ABSTRACT .....</b>	<b>vii</b>
<b>DAFTAR ISI.....</b>	<b>viii</b>
<b>DAFTAR GAMBAR .....</b>	<b>xi</b>
<b>DAFTAR TABEL .....</b>	<b>xii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xv</b>
<b>BAB I.....</b>	<b>1</b>
<b>PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	6
1.3 Tujuan Penelitian.....	6
1.4 Batasan Penelitian .....	7
<b>BAB II .....</b>	<b>8</b>
<b>TINJAUAN PUSTAKA .....</b>	<b>8</b>
2.1 Penelitian Terdahulu.....	8
2.2 Criticak Review.....	13
2.2.1 Summerize.....	15
2.2.2 Synthesize .....	16
2.2.3 Comparison .....	17
2.2.4 Kontribusi Penelitian.....	18
2.3 Teori Pendukung .....	19
2.3.1 Collaborative Filtering .....	19
2.3.2 Word2vec .....	20

A.    CBOW (Continuous Bag of-Word).....	21
2.4    indoBERT.....	22
2.5    SVM (support vector machine).....	25
2.6    KNN.....	26
2.7    Pearson Correlation.....	27
2.8    Cosine Similarity.....	28
2.9    MAE.....	28
2.10   RMSE.....	29
<b>BAB III.....</b>	<b>30</b>
<b>METODE PENELITIAN.....</b>	<b>30</b>
3.1    Pendekatan Penelitian.....	30
3.2    Desain Penelitian.....	30
3.3    Subjek Penelitian.....	30
3.4    Teknik Pengumpulan Data.....	30
3.4.1    Perangkat yang digunakan.....	31
3.4.2    Data Sekunder.....	31
3.4.3    Data Primer.....	31
3.5    Analisis Data.....	32
3.6    Prosedur Penelitian.....	33
3.7    Evaluasi Hasil Penelitian.....	34
3.8    Timeline Penelitian.....	35
<b>BAB IV.....</b>	<b>36</b>
<b>HASIL DAN PEMBAHASAN.....</b>	<b>36</b>
4.1    Use Case Diagram.....	36
4.2    Activity Diagram.....	36
4.3    Sequence Diagram.....	39
4.4    User Interface.....	39
4.5    Dataset.....	41
4.6    Pre-processing.....	41
4.6.1    Case Folding.....	42
4.6.1    Cleaning Text.....	42
4.6.2    Lemmatize.....	43

4.6.3	Tokenisasi.....	43
4.7	Analisis Sentimen .....	44
4.7.1	indoBERT.....	44
4.8	Klassifikasi.....	52
4.9	Collaborative Filtering .....	57
4.9.1	Rekomendasi berdasarkan user id.....	57
4.9.2	Rekomendasi berdasarkan review pengguna .....	60
<b>BAB V</b>	.....	<b>68</b>
<b>PENUTUP</b>	.....	<b>68</b>
5.1	Kesimpulan .....	68
5.2	Saran.....	68
<b>DAFTAR PUSTAKA</b>	.....	<b>69</b>
<b>LAMPIRAN</b>	.....	<b>74</b>



UNIVERSITAS  
MERCU BUANA

## DAFTAR GAMBAR

Gambar 1.1 Peningkatan Inflasi Medis (Sumber : BPS).....	2
Gambar 1.2 Media Sosial untuk Mencari Informasi Tentang Asuransi Swasta.....	4
Gambar 2. 1 Critical Review.....	13
Gambar 2. 2 Referensi Publish & Perish Penelitian Terdahulu .....	14
Gambar 2. 3 Vos Viewer Penelitian.....	15
Gambar 2. 4 Metode cara kerja model Word2vec.....	21
Gambar 2. 5 Metode cara kerja metode CBOW .....	21
Gambar 3. 1 Soal Kusioner yang diberikan .....	32
Gambar 3. 2 Prodsedur Penelitian.....	33
Gambar 4.1 Use Case User .....	36
Gambar 4.3 Activity Diagram Rekomendasi user_id .....	37
Gambar 4. 4 mencari rekomendasi berdasarkan review .....	38
Gambar 4. 5 Sequence Diagram .....	39
Gambar 4. 6 UI Upload Dataset & View .....	40
Gambar 4. 7 Rekomendasi User ID .....	40
Gambar 4. 8 Rekomendasi Review teks .....	41
Gambar 4. 9 Tampilan Hasil Rekomendasi.....	63
Gambar 4. 10 nilai MAE dan RMSE pada k = 5 sampai 40 .....	65
Gambar 4. 11 nilai MAE dan RMSE pada k = 10 sampai 20 .....	66

## DAFTAR TABEL

Tabel 2.1 Review Jurnal " Collaborative filtering and kNN based recommendation to overcome cold start and sparsity issues: A comparative analysis "	8
Tabel 2.2 Review Jurnal " Recommendation system using the k-nearest neighbors and singular value decomposition algorithms"	8
Tabel 2.3 Review Jurnal " E-Learning Course Recommender System Using Collaborative Filtering Models"	9
Tabel 2.4 Review Jurnal " Music Recommender System based on Sentiment Analysis Enhanced with Natural Language Processing Technics "	9
Tabel 2.5 Review Jurnal "An Approach to Integrating Sentiment Analysis into Recommender Systems"	9
Tabel 2.6 Review Jurnal "Sentiment analysis based distributed recommendation system"	10
Tabel 2.7 Review Jurnal "Movie recommendation and sentiment analysis using machine learning"	10
Tabel 2.8 Review Jurnal " Sentiment Analysis in Indonesian Healthcare	10
Tabel 2.9 Review Jurnal "Model Klasifikasi Calon Mahasiswa Baru Untuk Sistem Rekomendasi Program Studi Sarjana Berbasis Machine Learning"	11
Tabel 2.10 Review Jurnal "Recommender system using BERT sentiment analysis"	11
Tabel 2.11 Review Jurnal "An approach to improve the accuracy of rating prediction for recommender systems"	11
Tabel 2. 12 Review Jurnal "An approach to improve the accuracy of rating prediction for recommender systems"	12
Tabel 2.13 Review Jurnal "Integrating contextual sentiment analysis in collaborative recommender systems"	12
Tabel 2.14 Review Jurnal "Movie Recommender System Using K-Nearest Neighbors Variants"	12
Tabel 2. 15 Review Jurnal "Collaborative Filtering based Hybrid Music Recommendation System"	13
Tabel 3. 1 Hardware yang digunakan	31
Tabel 3. 2 Perangkat lunak yang digunakan	31



Tabel 3. 3 Timeline Penelitian.....	35
Tabel 4. 1 Proses case folding .....	42
Tabel 4.3 Proses Lemmatizer .....	43
Tabel 4.4 Proses Tokenisasi .....	43
Tabel 4. 5 Hasil Tokenisasi .....	44
Tabel 4. 6 Hasil penambahan Token khusus .....	45
Tabel 4. 7 Dictionary Proses input ID .....	45
Tabel 4. 8 Hasil Proses input ID.....	45
Tabel 4. 9 Hasil Proses input ID menjadi tensor .....	46
Tabel 4. 10 hasil floating point dari kata allian .....	46
Tabel 4. 11 Hasil Positional Encoding dari kata allian .....	48
Tabel 4. 12 Hasil embedding kata allian .....	48
Tabel 4. 13 Hasil penjumlahan PE + embedding .....	49
Tabel 4. 14 Proyeksi Q dari kata allian .....	49
Tabel 4. 15 Proyeksi K dari kata allian .....	49
Tabel 4. 16 Proyeksi V dari kata allian .....	49
Tabel 4. 17 Proses perhitungan Attention .....	50
Tabel 4. 18 Hasil Attention.....	50
Tabel 4. 19 Multihead dari kata allian.....	51
Tabel 4. 20 Hasil FFNN setelah add & norm.....	51
Tabel 4. 21 Hasil Softmax.....	52
Tabel 4. 22 Hasil kelas rating dari argmax.....	52
Tabel 4. 23 sample Text pada dataset.....	52
Tabel 4. 24 setelah tokenisasi.....	53
Tabel 4. 25 Hasil Vector menggunakan Word2vec kata allianz.....	54
Tabel 4. 26 Hasil Vector menggunakan Word2vec kata memberikan.....	54
Tabel 4. 27Hasil Vector menggunakan Word2vec kata prima .....	54
Tabel 4. 28 Hasil v dari sample teks .....	55
Tabel 4. 29 sample dataset .....	56
Tabel 4. 31 Vector A dari teks di dataset .....	61
Tabel 4. 32 Vector B dari teks di dataset.....	61
Tabel 4. 34 Ilustrasi prediksi .....	63

Tabel 4. 33 tampilan Dataset yang digunakan .....	64
Tabel 4. 35 hasil MAE dan RMSE $k = 10$ sampai 20 .....	66



## DAFTAR LAMPIRAN

Lampiran 1 Bimbingan SIA.....	74
Lampiran 2 Kartu Asistensi TA.....	75
Lampiran 3 Cek Turnitin.....	76
Lampiran 4 CV.....	77
Lampiran 5 Sertifikat BNSP .....	78
Lampiran 6 Naskah Jurnal .....	79
Lampiran 7 Pengalihan Hak Cipta .....	80
Lampiran 8 Surat Pernyataan .....	81
Lampiran 9 Form Revisi Penguji 1 .....	82
Lampiran 10 Form Revisi Penguji 2 .....	83



UNIVERSITAS  
MERCU BUANA