

**APLIKASI WEB “ANDROID APPS VERSUS” DENGAN
MENERAPKAN METODE COSINE SIMILIARITY
BERDASARKAN ULASAN GOOGLE PLAY STORE**

SKRIPSI

Diajukan Untuk Memenuhi Persyaratan Dalam Memperoleh Gelar Sarjana
Komputer Program Studi Teknik Informatika



Oleh :

BAGUS ANDREANTO

NPM. 1634010034

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
JAWA TIMUR
2022**

LEMBAR PENGESAHAN

SKRIPSI

**JUDUL : APLIKASI WEB “ANDROID APPS VERSUS” DENGAN
MENERAPKAN METODE COSINE SIMILIARITY
BERDASARKAN ULASAN GOOGLE PLAY STORE**

OLEH : BAGUS ANDREANTO

NPM : 1634010034

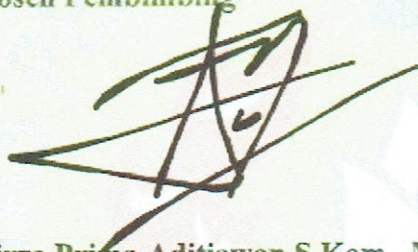
Telah Diseminarkan Dalam Ujian Skripsi Pada:

Rabu, 18 Mei 2022

Mengetahui

Dosen Pembimbing

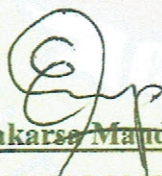
1.



Firza Prima Aditiawan, S.Kom., MTI

NIPPPK. 19860523 2021211 003

2.

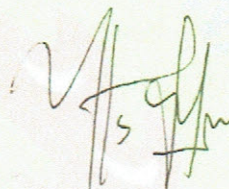


Eka Prakarsa Mandyartha, ST, M.Kom

NIP: 19880525 2018031 001

Dosen Penguji

1.



Yisti Vita Via, S.ST. M.Kom.

NIPPPK. 19860425 2021212 001

2.



Agung Mustika Rizki, S.Kom, M.Kom

NPT. 201199 30 725197

Menyetujui

Dekan

Fakultas Ilmu Komputer



Dr. Ir. Ni Ketut Sari, MTI

NIP : 19650731 199203 2001

**Koordinator Program Studi
Informatika,**

Budi Nugroho, S.Kom, M.Kom

NIPPPK: 19800907 2021211 005

09.06.2022
Pengesahan Buku Skripsi
1634010034 Bagus Andreanto

APLIKASI WEB “ANDROID APPS VERSUS” DENGAN MENERAPKAN METODE COSINE SIMILIARITY BERDASARKAN ULASAN GOOGLE PLAY STORE

Nama Mahasiswa : Bagus Andreanto

NPM : 1634010034

Program Studi : Informatika

Dosen Pembimbing : Firza Prima Aditiawan,S.Kom., MTI

Eka Prakarsa Mandyartha, ST, M.Kom

Abstrak

Smartphone adalah alat telekomunikasi pintar perkembangan dari telepon seluler yang menggunakan sinyal untuk melakukan transfer data dan informasi. Pada sebuah smartphone terdapat aplikasi–aplikasi yang dapat dipasang sesuai fungsi dan kebutuhan pengguna yang bertujuan untuk mendukung mobilitas dan pemenuhan kebutuhan sehari-harinya. Untuk menambah aplikasi pada sistem operasi Android pada umumnya menggunakan aplikasi bawaan bernama Google Play Store. Melihat akan hal itu, para pengembang aplikasi berlomba-lomba membuat aplikasi terbaik untuk memenuhi kebutuhan manusia dan memajangkannya di Google Play Store. Namun, sebagai pengguna Android tentu kita akan kesulitan untuk memilih dan membandingkan aplikasi mana yang sebaiknya kita pasang di Smartphone Android kita karena begitu banyak pilihan aplikasi yang tersedia.

Penggunaan aplikasi Android yang lebih baik dari aplikasi Android yang sejenis atau serupa berdasarkan ulasan yang diberikan oleh pengguna lain pada Google Play Store dapat dilakukan dengan mengimplementasikan algoritma Cosine Similarity untuk mengelompokkan ulasan-ulasan yang memiliki kata yang bersifat positif dan negatif.

Perbandingan antar aplikasi menggunakan sistem ini pada kategori *application* memiliki nilai akurasi sebesar 82,5%, kategori *Communication* sebesar 95%, *Business* 100%, *Education* 96%, *Entertainment* 80%, *Family* 100%, *Finance* 88% *Game* 92%, *Medical* 80% dan *Music and Audio* 100%. Dimana hasil tersebut telah diuji menggunakan pola pengujian berdasarkan data latih dan data uji yang telah disiapkan.

Kata Kunci : Aplikasi, *Cosine Similarity*, *Text Mining*

DAFTAR ISI

APLIKASI WEB “ANDROID APPS VERSUS” DENGAN MENERAPKAN METODE COSINE SIMILIARITY BERDASARKAN ULASAN GOOGLE PLAY STORE	i
APLIKASI WEB “ANDROID APPS VERSUS” DENGAN MENERAPKAN METODE COSINE SIMILIARITY BERDASARKAN ULASAN GOOGLE PLAY STORE	ii
LEMBAR PENGESAHAN SKRIPSI	i
SURAT PERNYATAAN ANTI PLAGIAT	ii
ABSTRAK	iii
KATA PENGANTAR	iv
UCAPAN TERIMA KASIH	v
DAFTAR ISI	vi
DAFTAR GAMBAR	ix
DAFTAR TABEL	xi
DAFTAR KODE	xii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah	2
1.4. Tujuan	3
1.5. Manfaat	3
BAB II TINJAUAN PUSTAKA	4
2.1. Penelitian Pendahulu	4
2.2. Landasan Teori	5
2.2.1. Text Mining	5
2.2.2. Vector Space Model	6
2.2.3. <i>Cosine Similiarity</i>	7
2.2.4. <i>K-Fold Cross Validation</i>	8
2.2.5. <i>Stemming Library</i>	9
BAB III METODE PENELITIAN	10
3.1. Analisis Data	10
3.2. Analisis Sistem	14

3.2.1. Flowchart Guest	14
3.2.2. Flowchart Admin	16
3.3. Pembagian Data	17
3.4. Preprocessing	17
3.4.1. <i>Tokenizing</i>	18
3.4.2. <i>Stemming</i>	18
3.4.3. Removing Stopword	19
3.5. Algoritma Cosine Similiarity	19
3.5.1. Penghitungan Term Frequency	20
3.5.2. Penghitungan Inverse Document Frequency	22
3.5.3. Penghitungan Bobot (W)	24
3.5.4. Perhitungan <i>Vector Space Model</i>	24
BAB IV HASIL DAN PEMBAHASAN	27
4.1. Implementasi Data	27
4.2. Implementasi Desain Antar Muka	27
4.2.1. Tampilan Login(Admin)	28
4.2.2. Tampilan Data Training List (Admin)	29
4.2.3. Tampilan <i>Add Data Training List</i> (Admin)	30
4.2.4. Tampilan <i>Data Test</i> (Admin)	30
4.2.5. Tampilan <i>Add Data Test</i> (Admin)	31
4.2.6. Tampilan <i>Display Accuration</i> (Admin)	32
4.2.7. Tampilan K-Fold Cross Validation (Admin)	33
4.2.8. Tampilan <i>Compare App List</i> (Guest)	34
4.2.9. Tampilan <i>Add Compare App</i> (Guest)	35
4.2.10. Tampilan <i>Compare App Result</i> (Guest)	36
4.2.11. Tampilan Compare App Review List	37
4.2.12. Tampilan Detail <i>Term Frequency</i>	38
4.2.13. Tampilan Detail <i>Inverse Document Freq</i> dan <i>Weight Term</i>	38
4.2.14. Tampilan Detail <i>Cosine Similarity</i>	39
4.2.15. Tampilan Predict Cosine Similiarity	40
4.3. Pembahasan	41
4.3.1. Pembagian Data	42
4.3.2. Perhitungan <i>Term Frequency</i>	45
4.3.3. Perhitungan <i>Inverse Document Frequency</i>	46

4.3.4. Perhitungan Bobot (<i>Weight</i>)	47
4.3.5. Perhitungan <i>Vector Space Model</i>	48
4.4. Pengujian Sistem	49
4.4.1. Pengujian Akurasi	49
4.4.2. Pengujian K-Fold Cross Validation	53
4.4.2.1. K-Fold Cross Validation kategori <i>Application</i>	53
4.4.2.2. <i>K-Fold Cross Validation</i> kategori <i>Communication</i>	55
4.4.2.3. <i>K-Fold Cross Validation</i> kategori <i>Business</i>	57
4.4.2.4. K-Fold Cross Validation kategori <i>Finance</i>	58
4.4.2.5. <i>K-Fold Cross Validation</i> kategori <i>Music and Audio</i>	59
4.4.2.6. <i>K-Fold Cross Validation</i> kategori <i>Game</i>	59
4.4.2.7. <i>K-Fold Cross Validation</i> kategori <i>Family</i>	60
4.4.2.8. <i>K-Fold Cross Validation</i> kategori <i>Medical</i>	61
4.4.2.9. <i>K-Fold Cross Validation</i> kategori <i>Entertainment</i>	61
4.4.2.10. <i>K-Fold Cross Validation</i> kategori <i>Education</i>	62
4.4.3. Pengujian Fitur <i>Compare Similiar Apps</i>	62
BAB V KESIMPULAN DAN SARAN	65
5.1. Kesimpulan	65
5.2. Saran	65
DAFTAR PUSTAKA	67
BIODATA PENULIS	69

DAFTAR GAMBAR

Gambar 2.1 Alur kerja <i>Vector Space Model</i>	6
Gambar 2.2 Alur kerja <i>Cross Validation</i>	9
Gambar 3.1 Ulasan pada <i>Google Playstore</i>	10
Gambar 3.2 <i>Flowchart Guest</i>	14
Gambar 3.3 Flowchart Pencarian Aplikasi sejenis terbaik	15
Gambar 3.4 Flowchart Cosine Similiarity	16
Gambar 3.5 <i>Flowchart Admin</i>	17
Gambar 3.6 <i>Dot Product</i>	25
Gambar 3.7 <i>Magnitude</i>	25
Gambar 3.8 <i>Cosine Similiarity</i>	26
Gambar 4. 1 Tampilan Login	28
Gambar 4. 2 Tampilan Data <i>Training List</i>	29
Gambar 4. 3 Tampilan <i>Add Data Training List</i>	30
Gambar 4. 4 Tampilan Data <i>Test List</i>	31
Gambar 4. 5 Tampilan <i>Add Data Test</i>	31
Gambar 4. 6 Tampilan <i>Display Accuration</i>	32
Gambar 4.7 <i>Display All Category Accuration</i>	33
Gambar 4.8 Tampilan <i>K-fold Cross Validation</i>	34
Gambar 4. 9 Tampilan <i>Compare App List</i>	35
Gambar 4. 10 Tampilan <i>Add Compare App</i>	36
Gambar 4.11 Tampilan <i>Compare App Result</i>	36
Gambar 4. 12 Tampilan <i>Compare App Review List</i>	37
Gambar 4. 13 Tampilan Detail <i>Term Frequency</i>	38
Gambar 4. 14 Tampilan <i>Detail Inverse Document Freq dan Weight Term</i>	39
Gambar 4. 15 Tampilan Detail <i>Cosine Similarity</i>	40
Gambar 4. 16 Predict Cosine Similiarity	41
Gambar 4. 17 Hasil pengujian kategori <i>application</i>	50
Gambar 4.19 Contoh lain hasil pengujian kategori <i>application</i>	51
Gambar 4.20 Hasil pengujian kategori <i>communication</i>	51
Gambar 4.21 Hasil pengujian kategori <i>communication</i>	52

Gambar 4.22 Hasil akurasi seluruh kategori	52
Gambar 4.23 K-Fold 4 kategori Application	53
Gambar 4.24 K-Fold 5 kategori Application	54
Gambar 4.26 K-Fold 4 kategori <i>Communication</i>	56
Gambar 4.27 K-Fold 5 kategori <i>Communication</i>	56
Gambar 4.28 Penambahan aplikasi oleh pengguna	63
Gambar 4.29 Hasil akhir perbandingan aplikasi sejenis	64

DAFTAR TABEL

Tabel 3.1 Daftar Aplikasi Dataset	11
Tabel 3.2 Data Training dan Preprocessing	19
Tabel 3.3 Term Frequency	21
Table 3.4 <i>Inverse Document Frequency</i>	23
Table 4.1 K-Fold Cross Validation Application	55
Tabel 4.2 K-Fold Cross Validation Communication	57
Tabel 4.3 <i>K-Fold Cross Validation Business</i>	58
Tabel 4.4 <i>K-Fold Cross Validation Finance</i>	58
Tabel 4.5 <i>K-Fold Cross Validation Music and Audio</i>	59
Tabel 4.6 <i>K-Fold Cross Validation Game</i>	59
Tabel 4.7 <i>K-Fold Cross Validation Family</i>	60
Tabel 4.8 <i>K-Fold Cross Validation Medical</i>	61
Tabel 4.9 K-Fold Cross Validation Entertainment	61
Tabel 4.10 <i>K-Fold Cross Validation Education</i>	62
Tabel 4.11 Hasil perbandingan ulasan aplikasi sejenis	63

DAFTAR KODE

Kode 4.1 TF IDF dari data train	45
Kode 4.2 TF IDF dari query	46
Kode 4.3 Kode Pembobotan	47
Kode 4.4 Perhitungan cosine similiarity	48