

Pengkelompokan Komentar Pengguna Aplikasi Mobile Legends Di Google Playstore Menggunakan Gaussian Mixture Models

Silvia Malona Nainggolan¹, Harlen Gilbert Simanullang², Arina Prima Silalahi³
^{1,2,3}Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Histori Artikel:

Received, Augst 10, 2025

Revised, Sept 25 , 2025

Accepted, Nov 15, 2025

Keywords:

Text Mining, Gaussian Mixture Models, Mobile Legends, Google Playstore, TF-IDF, Clustering.

ABSTRAK

Perkembangan pesat industri game online, khususnya genre *Multiplayer Online Battle Arena* (MOBA), telah mendorong tingginya interaksi pengguna melalui komentar di platform distribusi aplikasi seperti Google Playstore. Salah satu aplikasi game yang populer adalah Mobile Legends: Bang Bang. Komentar pengguna sering kali tidak sejalan dengan penilaian bintang, sehingga menimbulkan kesulitan dalam memahami persepsi pengguna secara objektif. Penelitian ini bertujuan untuk mengelompokkan komentar pengguna Mobile Legends menggunakan metode Gaussian Mixture Models (GMM), suatu pendekatan statistik berbasis probabilitas yang efektif dalam mengidentifikasi pola tersembunyi dalam data tidak terstruktur. Proses dimulai dari pengumpulan data melalui *web scraping*, dilanjutkan dengan *preprocessing* teks (*cleansing*, *case folding*, *stopword removal*, *tokenizing*, dan *stemming*), kemudian ekstraksi fitur menggunakan TF-IDF, dan akhirnya proses klasterisasi menggunakan GMM. Hasil dari penelitian ini menunjukkan bahwa komentar dapat dikelompokkan secara otomatis ke dalam tiga kategori utama: positif, negatif, dan netral. Pendekatan ini mampu memberikan wawasan yang lebih akurat dan objektif terhadap persepsi pengguna, serta dapat digunakan sebagai dasar pengambilan keputusan untuk pengembangan dan perbaikan aplikasi. Implementasi GMM dalam penelitian ini membuktikan efektivitasnya dalam mengelola data teks besar secara efisien dan relevan

This is an open access article under the [CC BY-SA](#) license.



Penulis Koresponden:

Silvia Malona Nainggolan,
Fakultas Ilmu Komputer,
Universitas Methodist Indonesia, Medan,
Jl. Hang Tuah No.8, Medan - Sumatera Utara.
Email: Silvianainggolan72@gmail.com

1. PENDAHULUAN

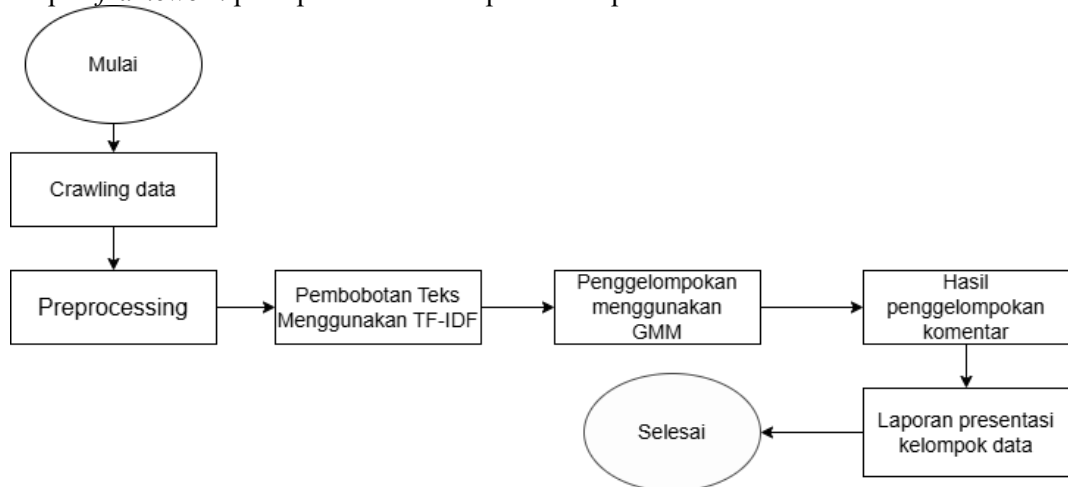
Perkembangan teknologi informasi telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk dalam industri hiburan digital[1]. Salah satu bentuk hiburan yang berkembang pesat adalah game online. Tingginya minat terhadap game online menjadikan genre *Multiplayer Online Battle Arena* (MOBA) salah satu yang paling populer di kalangan pemain. Awalnya, game MOBA hanya tersedia untuk perangkat Personal Computer (PC). Namun, seiring kemajuan teknologi, game ini mulai diadaptasi untuk perangkat ponsel pintar[2]. Di Indonesia, salah satu game MOBA yang paling populer adalah Mobile Legends: Bang Bang (MLBB). Berdasarkan data dari *Activeplayer.io*, jumlah pemain Mobile Legends di

seluruh dunia mencapai 77,19 juta orang pada Mei 2023[3]. Sebagai aplikasi game yang tersedia di platform distribusi seperti Google Playstore, Mobile Legends mendapatkan penilaian dalam bentuk rating bintang dan komentar tertulis dari penggunanya. Fitur rating bintang bertujuan memberikan gambaran umum tentang kualitas aplikasi. Namun, dalam praktiknya, fitur ini sering tidak berfungsi sebagaimana mestinya. Fenomena yang sering terjadi adalah pengguna memberikan komentar negatif, seperti keluhan tentang *bug* atau *lag*, tetapi tetap memberikan rating bintang tinggi. Ketidaksesuaian antara isi komentar dan rating bintang ini menyulitkan pihak manapun yang ingin memahami pandangan pengguna secara objektif berdasarkan data yang tersedia di Google Playstore. Untuk memperoleh gambaran yang lebih jelas dari komentar pengguna, diperlukan analisis yang mampu mengelompokkan komentar-komentar tersebut berdasarkan tema atau pola tertentu. Proses pengelompokan ini bertujuan untuk memberikan struktur pada data komentar yang bersifat tidak terorganisir. Contoh kategori komentar yang dapat terbentuk meliputi keluhan teknis (*bug*, *lag*, atau *crash*), saran atau masukan, apresiasi terhadap aspek tertentu, serta masalah lainnya seperti kesulitan dalam pembaruan atau kendala konektivitas. Salah satu metode yang dapat digunakan adalah Gaussian Mixture Models (GMM). GMM bekerja dengan memodelkan data menggunakan kombinasi distribusi Gaussian, sehingga mampu menangani data dengan distribusi kompleks. GMM memiliki beberapa keunggulan, di antaranya kemampuan mengidentifikasi pola tersembunyi dalam data, fleksibilitas untuk menangani data yang tidak terdistribusi secara simetris, serta memberikan hasil pengelompokan dengan probabilitas keanggotaan sehingga setiap komentar dapat dianalisis lebih mendalam[4]. Dengan menggunakan GMM, komentar pengguna dapat dikelompokkan secara otomatis tanpa intervensi manual. Penelitian ini difokuskan pada komentar pengguna aplikasi Mobile Legends di Google Playstore yang diperoleh melalui proses scraping sebanyak 1.773 data. Data tersebut dikelompokkan menjadi tiga kategori, yaitu baik, buruk, dan netral. Pengolahan dilakukan menggunakan bahasa pemrograman Python dengan menerapkan metode Gaussian Mixture Models (GMM). Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk melakukan pengelompokan komentar pengguna Mobile Legends menggunakan GMM serta memberikan gambaran umum persepsi pengguna berdasarkan hasil analisis otomatis. Penelitian ini diharapkan bermanfaat untuk mengetahui respons pengguna aplikasi Mobile Legends di Google Playstore, mengevaluasi implementasi metode GMM dalam pengelompokan data komentar, serta menghasilkan model yang dapat dijadikan acuan dalam pengembangan aplikasi atau penelitian serupa di masa mendatang.

2. METODE PENELITIAN

2.1. Framework Penelitian

Framework penelitian adalah kerangka konseptual yang berfungsi untuk mengarahkan dan memberikan struktur dalam proses penelitian. Kerangka konseptual tersebut mencakup gambaran secara menyeluruh dari penelitian yang akan dilakukan, prosedur penelitian dan ruang lingkup yang akan dilakukan. Adapun *framework* pada penelitian ini dapat dilihat pada Gambar 1



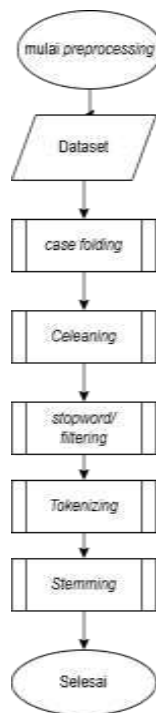
Gambar 1. *Framework* Penelitian

2.2. Teknik Pengumpulan Data

Proses pengumpulan data dilakukan untuk mendapatkan informasi berupa komentar pengguna aplikasi Mobile Legends yang tersedia di Google Playstore. Data ini akan digunakan sebagai bahan utama dalam analisis penelitian. Sumber data dalam penelitian ini adalah komentar pengguna yang tersedia secara publik di halaman ulasan aplikasi Mobile Legends di Google Playstore.

2.3. Preprocessing Text

Preprocessing bertujuan untuk mengubah teks yang berantakan atau tidak terstruktur menjadi teks yang rapi dan teratur sehingga dapat digunakan untuk analisis lebih lanjut. Tahapan preprocessing meliputi beberapa langkah berikut: *case folding*, *cleansing*, *stopword removal/filtering*, *tokenizing*, dan *stemming*[5]–[7]. Tahapan-tahapan ini ditunjukkan dalam flowchart pada Gambar 2



Gambar 2. Flowchart Preprocessing

2.4. Term Frequency

Term Frequency-Inverse Document Frequency (TF-IDF) merupakan salah satu proses untuk melakukan transformasi data dari data tekstual ke dalam data numerik untuk dilakukan pembobotan pada tiap kata atau fitur[8], [9]. Metode ini menghitung frekuensi kemunculan kata dalam dokumen (term frequency) dan memperhitungkan seberapa sering kata tersebut muncul di seluruh dokumen teks (inverse document frequency). Semakin sering kata muncul dalam dokumen tertentu, semakin tinggi skor term frequency nya, apabila semakin jarang muncul maka semakin tinggi skor inverse document frequency nya[5]. Formula TF-IDF dapat dilihat pada persamaan berikut[10]:

$$TF - IDF(t, d) = TF(t, d) \times IDF(t)$$

Keterangan:

$TF(t, d)$ = Frekuensi kemunculan *term* t dalam dokumen d

$f(t, d)$ = Banyaknya *term* t dalam dokumen

$IDF(t)$ = Frekuensi kemunculan *term* t dalam keseluruhan dokumen

N = Total keseluruhan dokumen yang ada

$df(t)$ = Banyaknya dokumen yang memuat *term* t

$TF - IDF(t, d)$ = Hasil perkalian antara nilai $TF(t, d)$ dengan $IDF(t)$

2.4. Tahapan Gaussian Mixture Models

Gaussian Mixture Model (GMM) termasuk dalam model berbasis clustering atau model distribusi, parameter GMM diestimasi menggunakan *algoritme Expectation Maximization (EM)* secara iteratif sehingga dapat terklusterkan dengan karakter yang serupa. Untuk Gaussian Mixture Model, diasumsikan bahwa sampel x yang diberikan adalah realisasi dari vektor acak yang distribusinya merupakan campuran dari beberapa distribusi [11].

GMM digunakan dalam berbagai bidang diantaranya clustering, dimensionality reduction, pemodelan distribusi data, restorasi gambar, dan lain-lain. Salah satu kelebihan digunakannya GMM adalah bentuk fungsi densitas yang sederhana, dan membutuhkan jumlah parameter yang tidak banyak. Pada GMM terdapat beberapa model yang dapat digunakan untuk melakukan clustering. Model tersebut berdasarkan bentuk geometris yang terbentuk dari komponen Gaussian [4].

Fungsi kerapatan probabilitas (probability density function) untuk distribusi Gaussian satu dimensi adalah:

$$f(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi|\Sigma|}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Keterangan :

x = Data (bisa berupa vektor jika multivariat)

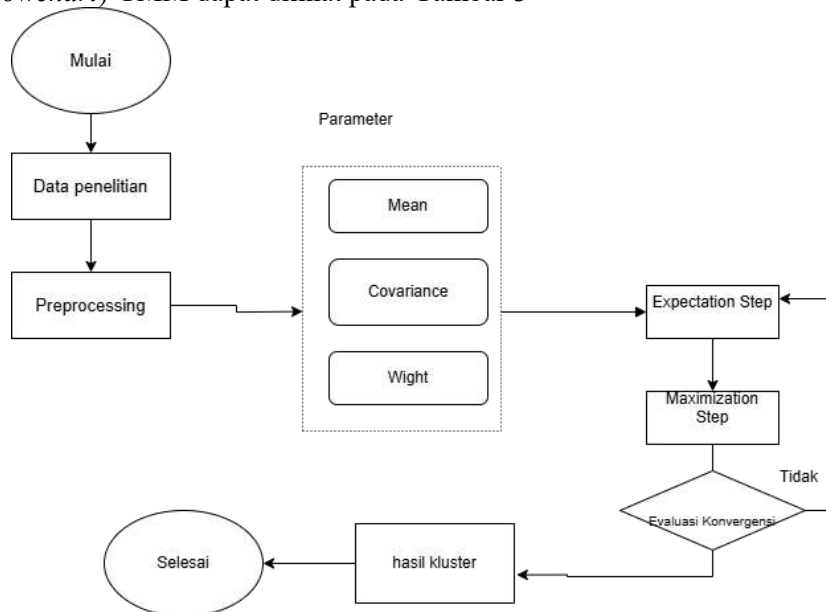
k = Jumlah kluster atau komponen Gaussian

π_k = koefisien campuran dari kluster k

μ_k = vektor mean (rata-rata) dari distribusi Gaussian ke- k

Σ_k = Matriks kovarian dari distribusi Gaussian ke- k

Diagram alur (*flowchart*) GMM dapat dilihat pada Gambar 3



Gambar 3. Alur Metode GMM

3. HASIL DAN PEMBAHASAN

3.1. Preprocessing Data

3.1.1 Cleansing

Cleaning adalah proses menghilangkan tanda baca dan kata-kata yang tidak perlu. Berikut ini merupakan penerapan cleaning dapat dilihat pada Tabel 1

Tabel 1. Hasil *Cleansing*

Komentar	<i>Cleansing</i>
Terlalu banyak dark sistem,sering terjadi bug sinyal,sering terjadi afk padahal gak ninggalin game fram rate sudah dibuat rendah tapi masih sering fram Gamenya bagus tapi min kasih musuh nya yg agak sedikit gampang masa saya mau naik tier susah banget musuhnya terus tim saya bot gak bisa baca map sama sekali tolong min please 🙏 Tolong dong montoon yang terhormat kenapa jaringan didalam game suka hilang padahal wifi saya bagus jaringanya.... Maaf saya uninstal.... Capek naikin skor credit samapai 2 minggu habis itu jaringan hilang afk lagi..... Capekkkkkkkkk oiiaiiii game nya bagus, dapet role roam mulu, nice Seruu banget main banyak event menarik tapi terkadang ada ngeleg tapi seru banyak Hero yang bisa dipilih aku suka MLBB Seru main game ini tapi pas aku main padahal sinyalku bagus tpi dari pas main tiba² jelek sinyalku. padahal sinyal ku bagus tolong di perbaiki sinyal Mobile Legends Bang bang ini supaya saya bisa melanjut kan lagi bermainnya!!! Bagus karena banyak skin gratis menyenangkan Bagus karena banyak skin gratis menyenangkan Saya sangat suka game ini jos buat ml Dan terlebih lebih saya minta kasi saya skin yang banyak dan bagus. Dan berikut saya diskon dan diamond 💎 plisss agar saya semangat bermain mobile legends Bang bang	terlalu banyak dark sistemsering terjadi bug sinyalsering terjadi afk padahal gak ninggalin game fram rate sudah dibuat rendah tapi masih sering fram gamenya bagus tapi min kasih musuh nya yg agak sedikit gampang masa saya mau naik tier susah banget musuhnya terus tim saya bot gak bisa baca map sama sekali tolong min please tolong dong montoon yang terhormat kenapa jaringan didalam game suka hilang padahal wifi saya bagus jaringanya maaf saya uninstal capek naikin skor credit samapai minggu habis itu jaringan hilang afk lagi capekkkkkkkkk oiiaiiii game nya bagus dapet role roam mulu seruu banget main banyak event menarik tapi terkadang ada ngeleg tapi seru banyak hero yang bisa dipilih aku suka mlbb seru main game ini tapi pas aku main padahal sinyalku bagus tpi dari pas main tiba jelek sinyalku padahal sinyal ku bagus tolong di perbaiki sinyal mobile legends bang bang ini supaya saya bisa melanjut kan lagi bermainnya bagus karena banyak skin gratis saya sangat suka game ini jos buat ml dan terlebih lebih saya minta kasi saya skin yang banyak dan bagus dan berikut saya diskon dan diamond plisss agar saya semangat bermain mobile legends bang

3.1.2 Case folding

Case folding merupakan tahap perubahan huruf dari huruf kapital menjadi huruf kecil. Berikut ini merupakan penerapan *case folding* dapat dilihat pada Tabel 2.

Tabel 2. *Case folding*

<i>Cleansing</i>	<i>Case folding</i>
terlalu banyak dark sistemsering terjadi bug sinyalsering terjadi afk padahal gak ninggalin game fram rate sudah dibuat rendah tapi masih sering fram gamenya bagus tapi min kasih musuh nya yg agak sedikit gampang masa saya mau naik tier susah banget musuhnya terus tim saya bot gak bisa baca map sama sekali tolong min please tolong dong montoon yang terhormat kenapa jaringan didalam game suka hilang padahal wifi saya bagus jaringanya maaf saya uninstal capek naikin skor credit samapai minggu habis itu jaringan hilang afk lagi capekkkkkkkkk oiiaiiii game nya bagus dapet role roam mulu seruu banget main banyak event menarik tapi terkadang ada ngeleg tapi seru banyak hero yang bisa dipilih aku suka mlbb	terlalu banyak dark sistemsering terjadi bug sinyalsering terjadi afk padahal gak ninggalin game fram rate sudah dibuat rendah tapi masih sering fram gamenya bagus tapi min kasih musuh nya yg agak sedikit gampang masa saya mau naik tier susah banget musuhnya terus tim saya bot gak bisa baca map sama sekali tolong min please tolong dong montoon yang terhormat kenapa jaringan didalam game suka hilang padahal wifi saya bagus jaringanya maaf saya uninstal capek naikin skor credit samapai minggu habis itu jaringan hilang afk lagi capekkkkkkkkk oiiaiiii game nya bagus dapet role roam mulu seruu banget main banyak event menarik tapi terkadang ada ngeleg tapi seru banyak hero yang bisa dipilih aku suka mlbb

seru main game ini tapi pas aku main padahal sinyalku bagus tpi dari pas main tiba jelek sinyalku padahal sinyal ku bagus tolong di perbaiki sinyal mobile legends bang bang ini supaya saya bisa melanjutkan kan lagi bermainnya
 bagus karena banyak skin gratis
 saya sangat suka game ini jos buat ml dan terlebih lebih saya minta kasi saya skin yang banyak dan bagus dan berikut saya diskon dan diamond plisss agar saya semangat bermain mobile legends bang

seru main game ini tapi pas aku main padahal sinyalku bagus tpi dari pas main tiba jelek sinyalku padahal sinyal ku bagus tolong di perbaiki sinyal mobile legends bang bang ini supaya saya bisa melanjutkan kan lagi bermainnya
 bagus karena banyak skin gratis
 saya sangat suka game ini jos buat ml dan terlebih lebih saya minta kasi saya skin yang banyak dan bagus dan berikut saya diskon dan diamond plisss agar saya semangat bermain mobile legends bang bang

3.1.3 Stopword/Filtering

Stopword removal merupakan proses penghilangan kata tidak penting pada deskripsi melalui pengecekan kata-kata hasil parsing deskripsi apakah termasuk di dalam daftar kata tidak penting (stoplist) atau tidak. Berikut ini merupakan penerapan *stopword* dapat dilihat pada Tabel 3

Tabel 3. Hasil *stopword*

<i>Case folding</i>	<i>Stopword/Filtering</i>
terlalu banyak dark sistemsering terjadi bug sinyalsering terjadi afk padahal gak ninggalin game fram rate sudah dibuat rendah tapi masih sering fram gamenya bagus tapi min kasih musuh nya yg agak sedikit gampang masa saya mau naik tier susah banget musuhnya terus tim saya bot gak bisa baca map sama sekali tolong min please tolong dong montoon yang terhormat kenapa jaringan didalam game suka hilang padahal wifi saya bagus jaringanya maaf saya uninstal capek naikin skor credit samapai minggu habis itu jaringan hilang afk lagi capekkkkkkkkk oiiaiiii game nya bagus dapet role roam mulu seruu banget main banyak event menarik tapi terkadang ada ngeleg tapi seru banyak hero yang bisa dipilih aku suka mlbb seru main game ini tapi pas aku main padahal sinyalku bagus tpi dari pas main tiba jelek sinyalku padahal sinyal ku bagus tolong di perbaiki sinyal mobile legends bang bang ini supaya saya bisa melanjutkan kan lagi bermainnya bagus karena banyak skin gratis saya sangat suka game ini jos buat ml dan terlebih lebih saya minta kasi saya skin yang banyak dan bagus dan berikut saya diskon dan diamond plisss agar saya semangat bermain mobile legends bang bang	terlalu banyak dark sistemsering terjadi bug sinyalsering terjadi afk ninggalin game fram rate dibuat rendah sering fram gamenya bagus kasih musuh nya yg agak sedikit gampang masa naik tier susah banget musuhnya tim bot baca map sekali please dong montoon terhormat jaringan didalam game suka hilang wifi bagus jaringanya uninstal capek naikin skor credit samapai minggu habis jaringan hilang afk capekkkkkkkkk oiiaiiii game nya bagus dapet role roam seruu banget banyak event menarik terkadang ada ngeleg seru banyak hero dipilih suka mlbb seru game sinyalku bagus dari tiba jelek sinyalku sinyal bagus perbaiki sinyal mobile legends bang bang melanjutkan kan bermainnya bagus banyak skin gratis suka game jos ml terlebih lebih minta kasi skin banyak bagus berikut diskon diamond plisss semangat bermain mobile legends

3.1.4 Tokenizing

Tokenizing adalah proses memecah dokumen menjadi kumpulan kata. Berikut ini merupakan penerapan *tokenizing* dapat dilihat pada Tabel 4.

Tabel 4. Hasil *Tokenizing*

<i>Stopword/Filtering</i>	<i>Tokenizing</i>
terlalu banyak dark sistemsering terjadi bug sinyalsering terjadi afk ninggalin game fram rate dibuat rendah sering fram gamenya bagus kasih musuh nya yg agak sedikit gampang masa naik tier susah banget musuhnya tim bot baca map sekali please	['terlalu', 'banyak', 'dark', 'sistemsering', 'terjadi', 'bug', 'sinyalsering', 'terjadi', 'afk', 'ninggalin', 'game', 'fram', 'rate', 'dibuat', 'rendah', 'sering', 'fram'] ['gamenya', 'bagus', 'kasih', 'musuh', 'nya', 'yg', 'agak', 'sedikit', 'gampang', 'masa', 'naik', 'tier', 'susah', 'banget', 'musuhnya', 'tim', 'bot', 'baca', 'map', 'sekali', 'please']

dong montoon terhormat jaringan didalam game	['dong', 'montoon', 'terhormat', 'jaringan', 'didalam', 'game',
suka hilang wifi bagus jaringanya uninstal capek	'suka', 'hilang', 'wifi', 'bagus', 'jaringanya', 'uninstal', 'capek',
naikin skor credit samapai minggu habis jaringan	'naikin', 'skor', 'credit', 'samapai', 'minggu', 'habis', 'jaringan',
hilang afk capekkkkkkkkk oiiiiiii	'hilang', 'afk', 'capekkkkkkkkk', 'oiiiiiii']
game nya bagus dapet role roam	['game', 'nya', 'bagus', 'dapet', 'role', 'roam']
seruu banget banyak event menarik terkadang ada	['seruu', 'banget', 'banyak', 'event', 'menarik', 'terkadang', 'ada',
ngeleg seru banyak hero dipilih suka mlbb	'ngeleg', 'seru', 'banyak', 'hero', 'dipilih', 'suka', 'mlbb']
seru game sinyalku bagus dari tiba jelek sinyalku	['seru', 'game', 'sinyalku', 'bagus', 'dari', 'tiba', 'jelek',
sinyal bagus perbaiki sinyal mobile legends bang	'sinyalku', 'sinyal', 'bagus', 'perbaiki', 'sinyal', 'mobile',
bang melanjut kan bermainnya	'legends', 'bang', 'bang', 'melanjut', 'kan', 'bermainnya']
bagus banyak skin geratis	['bagus', 'banyak', 'skin',
suka game jos ml terlebih lebih minta kasi skin	['suka', 'game', 'jos', 'ml', 'terlebih', 'lebih', 'minta', 'kasi', 'skin',
banyak bagus berikut diskon diamond pliss	'banyak', 'bagus', 'berikut', 'diskon', 'diamond', 'pliss',
semangat bermain mobile legends	'semangat', 'bermain', 'mobile', 'legends']

3.1.5 Stemming

Stemming bertujuan untuk menangkap kata dasar yang dimiliki oleh kata kerja yang telah menerima imbuhan kata atau keterangan lain pada kata dasar. Dalam implementasinya, hasil *stemming* ini dicek pada daftar kata dasar yang ada. Berikut ini merupakan penerapan *stemming* dapat dilihat pada Tabel 5.

Tabel 5. Hasil *Stemming*

Tokenizing	Stemming
['terlalu', 'banyak', 'dark', 'sistemsering', 'terjadi', 'bug', 'sinyalsering', 'terjadi', 'afk', 'ninggalin', 'game', 'fram', 'rate', 'dibuat', 'rendah', 'sering', 'fram']	['terlalu', 'banyak', 'dark', 'sistem', 'terjadi', 'bug', 'sinyal', 'terjadi', 'afk', 'tinggal', 'game', 'fram', 'rate', 'buat', 'rendah', 'sering', 'fram']
['gamenya', 'bagus', 'kasih', 'musuh', 'nya', 'yg', 'agak', 'sedikit', 'gampang', 'masa', 'naik', 'tier', 'susah', 'banget', 'musuhnya', 'tim', 'bot', 'baca', 'map', 'sekali', 'please']	['game', 'bagus', 'kasih', 'musuh', 'nya', 'yg', 'agak', 'sedikit', 'gampang', 'masa', 'naik', 'tier', 'susah', 'banget', 'musuh', 'tim', 'bot', 'baca', 'map', 'kali', 'please']
['dong', 'montoon', 'terhormat', 'jaringan', 'didalam', 'game', 'suka', 'hilang', 'wifi', 'bagus', 'jaringanya', 'uninstal', 'capek', 'naikin', 'skor', 'credit', 'samapai', 'minggu', 'habis', 'jaringan', 'hilang', 'afk', 'capekkkkkkkkk', 'oiiiiiii']	['dong', 'montoon', 'hormat', 'jaringan', 'dalam', 'game', 'suka', 'hilang', 'wifi', 'bagus', 'jaringan', 'uninstal', 'capek', 'naik', 'skor', 'credit', 'sampai', 'minggu', 'habis', 'jaringan', 'hilang', 'afk', 'capek', 'oiii']
['game', 'nya', 'bagus', 'dapet', 'role', 'roam']	['game', 'nya', 'bagus', 'dapet', 'role', 'roam']
['seruu', 'banget', 'banyak', 'event', 'menarik', 'terkadang', 'ada', 'ngeleg', 'seru', 'banyak', 'hero', 'dipilih', 'suka', 'mlbb']	['seru', 'banget', 'banyak', 'event', 'menarik', 'kadang', 'ada', 'ngeleg', 'seru', 'banyak', 'hero', 'pilih', 'suka', 'mlbb']
['seru', 'game', 'sinyalku', 'bagus', 'dari', 'tiba', 'jelek', 'sinyalku', 'sinyal', 'bagus', 'perbaiki', 'sinyal', 'mobile', 'legends', 'bang', 'bang', 'melanjut', 'kan', 'bermainnya']	['seru', 'game', 'sinyal', 'bagus', 'dari', 'tiba', 'jelek', 'sinyal', 'sinyal', 'bagus', 'perbaiki', 'sinyal', 'mobile', 'legend', 'bang', 'bang', 'lanjut', 'kan', 'main']
['bagus', 'banyak', 'skin', 'suka', 'game', 'jos', 'ml', 'terlebih', 'lebih', 'minta', 'kasi', 'skin', 'banyak', 'bagus', 'berikut', 'diskon', 'diamond', 'pliss', 'semangat', 'bermain', 'mobile', 'legends']	['bagus', 'banyak', 'skin', 'suka', 'game', 'jos', 'ml', 'lebih', 'lebih', 'minta', 'kasih', 'skin', 'banyak', 'bagus', 'ikut', 'diskon', 'diamond', 'pliss', 'semangat', 'main', 'mobile', 'legend']

3.2 TF_IDF

Berikut adalah penjelasan tahapan *TF_IDF*, Term daftar kata unik dalam dokumen. *TF* (*Term Frequency*) jumlah kemunculan setiap kata dalam setiap dokumen. *DF* (*Document Frequency*) jumlah dokumen yang mengandung kata tertentu. *IDF* (*Inverse Document Frequency*) dapat di lihat pada Table 6.

Tabel 6. Hasil *TF_IDF*

	TF							DF	IDF	W				
TOKEN	D1	D2	D3	..	D7	D8				D1	D2		D7	D8
Afk	1	0	1	..	0	0	2	0,60	0,60	0	..		0	0
Aku	0	0	0	..	0	0	2	0,60	0	0	...		0	0

Baca	0	1	0	..	0	0	1	0,90	0	0,90	...	0	0
Bagus	0	1	1	..	1	1	7	0,05	0	0,05		0,05	0,05
.....		..	...	...	..	...	...						
Banyak	1	0	0		1	1	5	0,20	0,2041	0		0,20	0,20
Tim	0	1	0		0	0	1	0,90	0	0,90	...	0	0
Tolong	0	1	1	...	0	0	3	0,42	0	0,42	...	0	0
Uninstall	0	0	1	...	0	0	1	0,90	0	0	...	0	0
Wifi	0	0	1		0	0	1	0,90	0	0	...	0	0
Total	14	21	19	...	4	19		42,64	7,21	13,86	...	1,76	7,45

3.3 Implementasi Perhitungan Manual

3.3.1 Menentukan Titik Pusat Centroid awal (Parameter)

Pusat awal *centroid* ditentukan secara random, titik pusat *centroid* yang digunakan yaitu dokumen 1 sebagai *centroid* kelompok data tidak puas (μ_1), dokumen 2 sebagai *centroid* kelompok data cukup puas (μ_2) dan dokumen 4 sebagai *centroid* kelompok data puas (μ_3). Setiap *centroid* merepresentasikan mean awal dari distribusi Gaussian untuk masing-masing kelompok. Selain itu, setiap cluster juga memiliki covarians yang menentukan bentuk dan ukuran distribusi data di sekitar *centroid* dan bobot yang menunjukkan proporsi setiap cluster dalam model.

Cluster 0 (μ_1) : (0.60, 0, 0, 0, 0.20, 0, 0.60, 0.90, 0, 0, 0, 0, 0.20, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.30, 0, 0.90, 0, 0, 0, 0, 0, 0, 1.27, 0.20, 0.90, 0, 0, 0, 0, 0, 0.20, 0, 0.90, 0, 0, 0, 0, 0)

Cluster 1 (μ_2) : (0, 0, 0.90, 0.05, 0, 0.42, 0, 0, 0, 0, 0, 0, 0.90, 0.90, 0, 0, 0, 0, 0.90, 0, 0, 0, 0.90, 0.90, 0, 0, 0, 1.20, 0.90, 0, 0, 0, 0.90, 0, 0, 0, 0, -0.19, 0.90, 0, 0, 0, 0, 0, 0, 0.90, 0.20, 0, 0, 0.90, 0.90, 0.90, 0.42, 0, 0)

Cluster 2 (μ_3) : (0, 0, 0, 0.05, 0, 0, 0, 0, 0, 0, 0, 0, 0.20, 0.90, 0.90, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)

Penentuan kovarians awal: (0.080, 0.0001, 0.181, 0.0007, 0.0092, 0.040, 0.080, 0.18, 0.0001, 0.0001, 0.0001, 0.0001, 0.0092, 0.18, 0.18, 0.0001, 0.0001, 0.0001, 0.0001, 0.0001, 0.18, 0.0001, 0.0001, 0.0001, 0.18, 0.0001, 0.0001, 0.0001, 0.18, 0.18, 0.0001, 0.0083, 0.18, 0.0001, 0.36, 0.0092, 0.18, 0.0001, 0.0001, 0.0001, 0.0001, 0.18, 0.0092, 0.0001, 0.18, 0.18, 0.18, 0.18, 0.040, 0.0001, 0.0001)

Bobot awal: $\pi_0 = \pi_1 = \pi_2 = \frac{1}{3} = 0.33$

3.3.2 Perhitungan Probabilitas Keanggotaan (Expectation Step)

Hasil perhitungan probabilitas awal setiap fitur pada dokumen didasarkan pada nilai centroid yang telah ditentukan sebelumnya. Setiap dokumen memiliki tiga probabilitas, yang masing-masing menunjukkan kemungkinan dokumen tersebut masuk ke dalam salah satu dari tiga klaster yang telah ditentukan. Penentuan klaster dilakukan dengan membandingkan nilai probabilitas terbesar pada setiap dokumen, hasil probabilitas keanggotaan dapat kita lihat pada tabel 7.

Tabel 7. Probabilitas awal

Dokumen	Prob 1	Prob 2	Prob 3	Klaster
D1	1141,45	1089,30	1106,61	Klaster 0
D2	1089,30	1141,45	1113,13	Klaster 1
D3	667,54	678,12	691,48	Klaster 2
D4	1106,61	1113,37	1140,14	Klaster 2
D5	876,49	863,96	1141,45	Klaster2
D6	825,33	825,06	831,36	Klaster 2
D7	1028,50	1035,25	1052,57	Klaster 2
D8	709,95	710,70	731,50	Klaster 2

3.3.3 Menghitung Jarak Pusat Centroid Baru (Maximization Step)

Hasil pengelompokan data pada iterasi 1 terhadap setiap cluster selanjutnya menentukan nilai dari pusat centroid baru dengan cara menghitung Nilai tertinggi dari hasil probabilitas setiap dokumen pada klaster. Menghitung kembali data untuk mendapatkan pusat cluster yang baru berdasarkan nilai cluster 0, cluster 1 dan cluster 2 tertinggi dari keseluruhan probabilitas pada dokumen. Berikut merupakan perhitungan penentuan nilai pusat centroid baru pada cluster 0, cluster 1 dan cluster 2.

Centroid baru cluster 0 (C0): (0.601, 0, 0, 0, 0.20, 0, 0.60, 0.90, 0, 0, 0, 0, 0.20, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.30, 0, 0.90, 0, 0, 0, 0, 0, 0, 1.27, 0.20, 0.90, 0, 0, 0, 0, 0, 0.20, 0, 0.90, 0, 0, 0, 0, 0)

Centroid baru cluster 1 (C1): (0, 0, 0.90, 0.05, 0, 0.42, 0, 0, 0, 0, 0, 0, 0, 0.90, 0.90, 0, 0, 0, 0, 0.90, 0, 0, 0, 0.90, 0.90, 0, 0, 0, 1.20, 0.90, 0, 0, 0, 0.90, 0, 0, 0, 0, -0.19, 0.90, 0, 0, 0, 0, 0, 0, 0, 0.90, 0.20, 0, 0, 0.90, 0.90, 0.90, 0.42, 0, 0)

Centroid baru cluster 2 (C2): (0, 0, 0, 0.05, 0, 0, 0, 0, 0, 0, 0, 0, 0.20, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.90, 0.90, 0)

Covarians baru = (1099.94, 1122.88, 1088.75, 1118.40, 1115.02, 1106.58, 1099.94, 1088.75, 1122.88, 1122.88, 1122.88, 1122.88, 1107.16, 1088.75, 1088.75, 1122.88, 1122.88, 1122.88, 1122.88, 1122.88, 1122.88, 1077.73, 1088.75, 1122.88, 1122.88, 1111.32, 1088.75, 1088.75, 1088.75, 1088.75, 1122.88, 1130.41, 1088.75, 1122.88, 1075.06, 1115.02, 1088.75, 1122.88, 1122.88, 1122.88, 1122.88, 1088.75, 1107.16, 1122.88, 1088.75, 1088.75, 1088.75, 1088.75, 1106.58, 1122.88, 1122.88)

Penentuan nilai bobot baru dengan persamaan :

$$\pi_0 = \frac{y1_{cluster0}}{8} = 19,34$$

$$\pi_1 = \frac{y1_{cluster1}}{8} = 19,34$$

$$\pi_2 = \frac{y1_{cluster2}}{8} = 19,34$$

Hitung probabilitas dari semua data ke *centroid* baru seperti yang dilakukan pada iterasi pertama. setelah hasil perhitungan didapatkan, kemudian bandingkan hasilnya. Jika hasil posisi cluster pada iterasi kedua sama dengan posisi iterasi pertama maka proses dihentikan, tetapi jika tidak maka proses dilanjutkan ke iterasi ketiga. Tabel probabilitas baru dapat kita lihat pada Tabel 8.

Tabel 8. probabilitas baru

Dokumen	Prob 1 baru	Prob 2 baru	Prob 3 baru	Iterasi 1	Iterasi 2
D1	0,7071	0,7071	0,7071	0	0
D2	0,7070	0,7071	0,7070	1	1
D3	0,7070	0,7070	0,7071	2	2
D4	0,7071	0,7070	0,7071	2	2
D5	0,7071	0,7070	0,7071	2	2
D6	0,7071	0,7070	0,7071	2	2
D7	0,7071	0,7070	0,7071	2	2
D8	0,7071	0,7070	0,7071	2	2

Pada iterasi kedua yang terdapat pada tabel diatas didapatkan kesimpulan bahwa anggota C0 adalah (D1), anggota C1 adalah (D2) sedangkan anggota C2 adalah (D3, D4, D5, D6, D7 dan D8), Hasil posisi cluster pada iterasi kedua sama dengan posisi iterasi pertama, maka proses dihentikan karena telah dinyatakan konvergen.

3.4 Implementasai Berbasis Python

Penelitian ini diimplementasikan menggunakan bahasa pemrograman *Python* dan dijalankan melalui *Jupyter Notebook*, yang merupakan sebuah *executable document* atau dengan kata lain berbentuk *notebook* yang dapat digunakan untuk menyimpan, menulis, dan membagikan program yang telah ditulis. Hasil perhitungan dengan menggunakan python dapat di lihat pada gambar 4.

```

Cluster Distribution:

cluster
0      1214
2       244
1        13
Name: count, dtype: int64

```

Gambar 4. Hasil Perhitungan *python*

Dari hasil pengelompokan menggunakan Gaussian mixture Models Pada nilai $k = 3$ di dapatkan hasil pada kelompok pertama terdapat 1214 komentar dan pada kelompok kedua terdapat 244 komentar sedangkan untuk kelompok ketiga terdapat 13 komentar, Berdasarkan analisis terhadap isi komentar, klaster 0 memuat banyak keluhan, kekecewaan, serta ungkapan frustrasi dari pengguna. Kata-kata seperti “tidak disarankan”, “ga keluar-keluar”, dan “game bikin emosi” menandakan bahwa komentar yang tergolong dalam klaster ini mengandung sentimen negative, Komentar-komentar dalam klaster 1 memiliki kecenderungan untuk menyampaikan penilaian positif terhadap aplikasi atau game. Kata-kata seperti "Good" dan "good game" merupakan indikator umum dari sentimen positif, sehingga klaster ini ditetapkan sebagai klaster komentar positif, Klaster 2 mencakup komentar yang cenderung campuran atau tidak menunjukkan emosi yang sangat kuat, baik secara negatif maupun positif. Beberapa komentar bersifat netral, informatif, atau ringan. Karena tidak menunjukkan kecenderungan sentimen yang tegas, maka klaster ini dikategorikan sebagai klaster komentar netral.

4. KESIMPULAN

Hasil pengelompokan komentar pengguna aplikasi Mobile Legends di Google Play Store menggunakan metode Gaussian Mixture Models menunjukkan bahwa sekitar 82,7% komentar tergolong negatif, 16,4% merupakan komentar positif, dan 0,82% termasuk dalam kategori komentar netral. Sebagian besar pengguna menyampaikan keluhan dalam komentarnya, yang menunjukkan bahwa tingkat kepuasan terhadap aplikasi Mobile Legends masih rendah.

REFERENSI

- [1] N. S. Lubis and M. I. P. Nasution, “Perkembangan Teknologi Informasi Dan Dampaknya Pada Masyarakat,” *KOHESI J. Multidisplin Saintek*, vol. 1, no. 12, pp. 41–50, 2023.
- [2] D. Sinaga and C. Jatmoko, “ANALISIS SENTIMEN UNTUK MENGETAHUI KESAN PLAYER GAME MOBILE LEGENDS MENGGUNAKAN NAÏVE BAYES CLASSIFIER,” vol. 0, pp. 540–547, 2020.
- [3] M. A. Rizqi and M. R. Maulana, “User-Friendly as the Main Pillar of Success Mobile Legends Game in Indonesia User-Friendly sebagai Pilar Utama Kesuksesan Game Mobile Legends di Indonesia User-Friendly as the Main Pillar of Success Mobile Legends Game in Indonesia,” no. June, 2024.
- [4] D. Y. Faidah, A. M. Hudzaifa, N. Theresia, and C. E. Widiatoro, “Optimalisasi Strategi Pengelompokan Potensi Padi Sebagai Solusi Efektif Kelangkaan Beras Di Jawa Barat,” *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 5, no. 1, pp. 529–537, 2024, doi: 10.46306/lb.v5i1.592.
- [5] B. Of and C. Science, “BULLETIN OF COMPUTER SCIENCE RESEARCH Rekomendasi Berita Berkaitan dengan Menerapkan Algoritma Text Mining dan TF-IDF,” vol. 3, no. 4, pp. 276–282, 2023, doi: 10.47065/bulletincsr.v3i4.266.
- [6] K. Laembalno Berutu, H. Rumapea, and H. G. Manullang, “Analisa Sentimen Penggunaan Permainan Tradisional Lato-Lato Pada Sosial Media Twitter Menggunakan Metode K-Means,” *J. Ilm. Tek. Inform.*, vol. 4, no. 1, pp. 42–51, 2024, [Online]. Available: <https://ejurnal.methodist.ac.id/index.php/methotika>
- [7] H. G. Simanullang, A. P. Silalahi, and G. P. Pangaribuan, “PENGELOMPOKAN KOMENTAR PENGGUNA JKN MOBILE MENGGUNAKAN METODE K-MEDOIDS,” 2025.
- [8] H. G. Simanullang, S. Wijono, B. Kristianto, W. Sulisty, and A. P. Silalahi, “Detecting Cyberbullying Using Machine Learning Approaches,” in *2024 Ninth International Conference on Informatics and Computing (ICIC)*, 2024, pp. 1–8. doi: 10.1109/ICIC64337.2024.10956842.
- [9] L. M. Barus, H. G. Simanullang, and A. P. Silalahi, “Classification of English-Language Racial Hate Speech

-
- Using Random Forest ”,” *J. Intell. Comput. Inf. Syst.*, vol. 1, no. 1, pp. 1–17, 2025.
- [10] P. Studi *et al.*, “ANALISIS SENTIMEN REVIEW APLIKASI MOBILE LEGENDS DI GOOGLE PLAYSTORE MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN TF-IDF DI GOOGLE PLAYSTORE MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE DAN TF-IDF,” 2024.
- [11] R. Ummami and B. Winarno, “Gaussian Mixture Model dengan Algoritme Expectation Maximization untuk Pengelompokan Data Distribusi Air Bersih di Jawa Barat,” *Prism. Pros. Semin. Nas. Mat.*, vol. 6, pp. 745–750, 2023.