

# PENERAPAN TF-IDF DAN COSINE SIMILARITY PADA QUESTION AND ANSWERING SYSTEM TERKAIT INFORMASI KONSOL DAN GAME PLAYSTATION

Adryanus Rinaldi <sup>1)</sup> Lely Hiryanto <sup>2)</sup> Viny Christanti Mawardi <sup>3)</sup>

<sup>1)2)3)</sup> Teknik Informatika Universitas Tarumanagara

Jl. S. Parman No. 1, Grogol, Jakarta 11440 Indonesia

email : [adryanus.535210110@stu.untar.ac.id](mailto:adryanus.535210110@stu.untar.ac.id), [lelyh@fti.untar.ac.id](mailto:lelyh@fti.untar.ac.id), dan [vinym@fti.untar.ac.id](mailto:vinym@fti.untar.ac.id)

## ABSTRAK

Dengan pertumbuhan pesat informasi digital, menjadi sulit bagi pengguna untuk menemukan jawaban yang tepat dan relevan. Untuk mengatasi masalah ini, sistem Question Answering (QA) berbasis web disarankan untuk memberikan respons langsung terhadap pertanyaan pengguna yang khusus terkait dengan konsol dan permainan PlayStation. Sistem ini menemukan jawaban yang paling relevan terhadap pertanyaan yang dimasukkan melalui analisis term frekuensi-invers dokumen TF-IDF yang dikombinasikan dengan Cosine Similarity. Pengujian menggunakan dataset 50 pertanyaan dan jawabannya. Pengujian yang dilakukan merupakan pengujian pertanyaan sama persis dengan dataset dan pengujian dengan 50% kemiripan dengan pertanyaan didataset yang menggunakan nilai batasan diatas 0 dan 0,6 . pengujian pertanyaan sama persis dengan dataset menghasilkan nilai akurasi, presisi, recall, dan nilai F1 dimana semua nilai memiliki nilai 1. Percobaan dengan nilai batasan diatas 0 menghasilkan nilai akurasi 0,86 , nilai presisi 0,85 , nilai recall 0,93 , dan nilai F1 0,89. Sedangkan, pengujian dengan nilai batasan 0,6 menghasilkan nilai akurasi 0,94 , nilai presisi 0,98 , nilai recall 0,95 , dan nilai F1 0,96. Pengujian dengan nilai batasan 0,6 dibuktikan lebih baik daripada pengujian dengan nilai batasan diatas 0 dan pertanyaan yang sama persis dengan dataset akan pasti mendapat jawaban yang tepat.

## Key words

Cosine Similarity, PlayStation, Question Answering, TF-IDF

## 1. Pendahuluan

### 1.1. Latar Belakang

Akses ke informasi melalui Internet saat ini sangat mudah. Seiring dengan jumlah data yang meningkat,

semakin sulit untuk menemukan informasi yang relevan dan andal. Bagi pengguna, pencarian jawaban untuk pertanyaan mereka seringkali menjadi suatu tantangan, terutama ketika informasi tersebar di berbagai sumber. Sistem QA, juga dikenal sebagai sistem penjawab pertanyaan pada umumnya [1][2][3], dirancang sebagai suatu solusi untuk menyelesaikan masalah ini dengan memberikan pengguna jawaban langsung dalam bentuk bahasa alami untuk pertanyaan mereka, mengurangi waktu dan usaha yang diperlukan untuk mendapatkan informasi yang mereka butuhkan.

Konsol game PlayStation [4], yang merupakan salah satu produk teknologi paling populer, telah menarik banyak perhatian pada berbagai tempat. Oleh sebab itu, banyak data tentang konsol game PlayStation tersebar secara luas. Sementara video game adalah permainan elektronik yang melibatkan interaksi antarmuka pengguna atau perangkat input, seperti joystick, pengontrol, keyboard, atau sensor gerak, untuk menghasilkan umpan balik visual, konsol game merupakan mesin elektronik yang dirancang khusus untuk memainkan video game. Informasi tentang PlayStation seringkali terlalu banyak hingga membingungkan bagi pengguna dengan berbagai model dan genre game yang tersedia. Oleh sebab itu, pembangunan sistem QA yang dapat memberikan informasi yang akurat dan relevan tentang konsol permainan ini sangat penting. Tujuan dari makalah paper ini adalah untuk membangun sistem QA yang dapat menemukan jawaban atas pertanyaan yang terkait dengan PlayStation karena sistem pemeriksaan bertujuan untuk menemukan jawaban atas pertanyaan yang berasal dari kumpulan teks yang tidak terstruktur dalam bentuk bahasa manusia [5].

Agar dapat mencapai hasil yang baik, pengembangan sistem QA diperlukan untuk menggunakan metode yang dapat menganalisis dan memproses data dengan baik. Metode *Term Frequency Inverse Document Frequency* (TF-IDF) adalah salah satu metode yang paling efektif dalam pemrosesan teks dan pengambilan informasi [6]. Pada dasarnya, metode ini bertujuan untuk menemukan informasi yang paling

relevan dengan pertanyaan dengan membandingkan nilai setiap kata dalam dokumen dengan seluruh kumpulan dokumen. Memasukkan metode TF-IDF ke dalam sistem QA diharapkan dapat meningkatkan kecepatan dan akurasi informasi yang diinginkan pengguna [6].

## 1.2. Rumusan Masalah

Tujuan dari desain ini, seperti yang disebutkan sebelumnya, adalah untuk membuat aplikasi sistem QA berbasis web yang dapat menjawab pertanyaan pengguna tentang hal-hal yang terkait dengan PlayStation. Metode TF-IDF dan kesamaan kosinus digunakan. Pendekatan TF-IDF adalah teknik yang berguna untuk pemrosesan teks dari dokumen yang tersedia untuk meningkatkan akurasi dan efisiensi sistem QA. Metode ini akan menilai pentingnya setiap kata dalam dokumen dibandingkan dengan kata-kata lain dalam koleksi. Namun, kesamaan kosinus digunakan untuk menemukan tanggapan yang relevan yang menggabungkan sebagian besar istilah yang diminta. Untuk memberikan jawaban yang optimal, data yang dikumpulkan termasuk nama game, tanggal rilis, platform, penerbit, genre, bahasa yang tersedia, dan informasi PlayStation.

## 1.3. Komponen Desain

Ada beberapa komponen yang perlu dipertimbangkan selama proses desain untuk memastikan bahwa sistem ini berfungsi dengan baik dan mencapai tujuan yang ditetapkan. Komponen untuk desain tersebut mencakup hal-hal berikut.

1. Sumber Data  
Data yang digunakan dalam sistem QA akan diperoleh dari informasi yang tersedia di situs web dan YouTube, yang mencakup informasi seperti nama game, informasi game, tanggal rilis, platform, penerbit, genre, dan bahasa yang tersedia. Data ini dikumpulkan dan diproses untuk meningkatkan akurasi respons sistem pemeriksaan kualitas.
2. Fitur yang diterapkan  
Fungsi *Question* dan *Answering* (QA) atau tanya jawab adalah fitur utama sistem ini. Pengguna dapat mengajukan pertanyaan melalui kotak pertanyaan pada halaman web, dan fitur tanya jawab akan menghasilkan dan menyampaikan jawaban. Pengguna juga dapat mengakses fitur ini melalui halaman beranda, juga disebut sebagai *Home*, yang menampilkan setiap fitur sistem. Selain itu, ada juga fungsi yang menampilkan informasi rinci tentang sistem pengujian dan evaluasi PlayStation ini.
3. Metode TF-IDF  
Teknik penimbangan TF-IDF adalah teknik penimbangan yang banyak digunakan dalam pemrosesan teks dan Natural Language Processing (NLP). Metode ini memberikan

penjelasan tentang seberapa penting kata-kata sesuai dengan maknanya dalam dokumen [3]. Aplikasi yang sering menggunakan metode ini termasuk klasifikasi dokumen, analisis teks, dan sistem pencarian.

## 1.4. Spesifikasi Desain

Berikut ini adalah spesifikasi desain untuk setiap modul yang diperlukan untuk desain sistem QA.

1. Modul Beranda  
Modul beranda atau *Home* adalah bagian sistem QA yang pertama kali digunakan oleh pengguna dan bertujuan untuk memberikan akses langsung ke fitur online. Modul ini bermanfaat untuk memberikan pengalaman awal yang menyenangkan bagi pengguna saat menggunakan semua fitur dalam sistem.
2. Modul Tanya Jawab  
Modul Tanya Jawab adalah komponen modul yang digunakan untuk menerima pertanyaan dan memberikan jawaban yang relevan berdasarkan data yang telah diproses menggunakan metode TF-IDF. Pengguna memasukkan pertanyaan dalam kotak pertanyaan, dan sistem mengolah data untuk mengetahui apakah ada kemiripan antara pertanyaan yang diajukan dan kemudian memberikan jawaban kepada pengguna.
3. Modul Tentang Kita (About Us)  
Modul Tentang Kami merupakan komponen modul yang dirancang dengan tujuan untuk memberikan informasi umum mengenai sistem yang digunakan pengguna dan informasi dasar dari PlayStation, termasuk sejarah, fitur utama, dan spesifikasi dari berbagai model konsol. Modul ini bertujuan untuk memberikan konteks dan latar belakang kepada pengguna yang ingin memahami lebih dalam tentang PlayStation.

## 1.5. Kegunaan Sistem QA

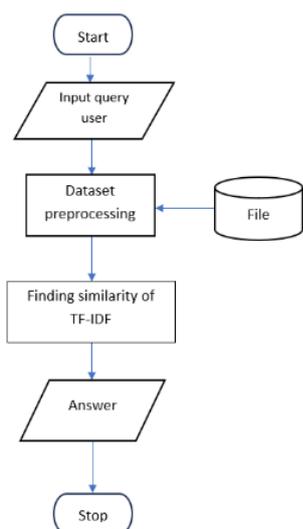
Tujuan utama dari perancangan sistem QA ini adalah memberikan jawaban yang akurat terhadap pertanyaan pengguna mengenai topik yang berkaitan dengan PlayStation. Agar tujuan tersebut dapat tercapai, sistem ini memproses data yang ada dan menganalisis pertanyaan yang diajukan. Sistem tersebut akan mencari kemiripan dari data tersebut agar mendapat dan memberi jawaban yang akurat kepada pengguna. Keakuratan respon sistem tersebut akan ditentukan dengan menggunakan metode NLP dan pemrosesan teks, yaitu TF-IDF, untuk memastikan bahwa respon yang diberikan oleh sistem seakurat mungkin untuk pertanyaan pengguna.

## 2. Landasan Teorik

### 2.1. Perancangan Sistem QA

Sistem yang dirancang bertujuan untuk memberikan informasi yang akurat relevan mengenai PlayStation. Data yang digunakan untuk system QA ini berupa informasi yang diperoleh dari ahli-ahli game yang mencakup data playstation, data game, release date, genre, dan platform game. Dengan data-data ini, system tersebut dirancang untuk menjawab pertanyaan yang ingin dijawab oleh pengguna dengan memilih salah satu kategori dalam suatu table list yang disediakan oleh system. Metode yang diterapkan merupakan metode Term Frequency-Inverse Document Frequency (TF-IDF) adalah metode yang mengevaluasi seberapa penting suatu kata (term) dalam sebuah dokumen dalam konteks koleksi dokumen yang lebih besar dengan menggunakan metode ini dalam system QA.

Gambar 1 adalah diagram alur dari proses QA yang dirancang. Proses dimulai dengan input query. Input query dilakukan melalui penyediaan tabel pilihan kategori dari sistem. Kemudian, sistem tersebut akan melakukan dataset preprocessing yang bertujuan untuk meningkatkan kualitas analisis dengan menghapus karakter khusus yang tidak relevan seperti tanda baca, melakukan proses pemecahan teks menjadi kata-kata dalam proses yang bernama tokenizing, menghapus kata-kata umum yang tidak memberikan informasi relevan, dan mengubah kata menjadi bentuk dasarnya dengan mengubah huruf besar menjadi huruf kecil. Sistem QA menarik dataset dari dokumen file yang akan dianalisis dan dipreprocessing agar data tersebut menjadi lebih mudah untuk melakukan perhitungan TF-IDF.



Gambar 1 Diagram alur sistem yang dirancang

Pada tahap berikutnya, mencari kemiripan TF-IDF yang dimulai dengan menghitung Term Frequency (TF) dan Inverse Dokumen Frequency (IDF). Perhitungan TF memiliki tujuan untuk mengukur seberapa sering suatu

kata muncul dalam dokumen tertentu. Sedangkan, perhitungan IDF bertujuan mengukur seberapa penting suatu kata dalam keseluruhan koleksi dokumen. Setelah menghitung nilai TF-IDF, setiap dokumen direpresentasikan sebagai vektor. Setiap elemen dalam vektor mewakili bobot TF-IDF dari kata tertentu dalam dokumen tersebut. Dalam dokumen ini, Cosine Similarity digunakan untuk mengukur kesamaan antara vektor pertanyaan yang dimasukkan oleh pengguna dengan vektor dokumen dalam dataset. Cosine similarity dihitung untuk menemukan dokumen yang paling relevan berdasarkan kueri yang diajukan.

### 2.2. Game Konsol Playstation.

PlayStation (PS) adalah merek konsol permainan video rumahan yang mencakup layanan daring, pusat media, pengontrol permainan, konsol genggam, dan ponsel pintar. Diproduksi oleh Sony Interactive Entertainment, divisi dari Sony, permainan pertama kali dirilis untuk konsol permainan video rumahan PlayStation di Jepang pada Desember 1994, dan kemudian dirilis di seluruh dunia pada tahun berikutnya. Playstation asli memiliki beberapa penerus yang memiliki spec dan grafis yang berbeda. Secara total, konsol Playstation ada 5 generasi:

1. PlayStation 1  
PlayStation (PlayStation 1 atau PS1) merupakan konsol video game rumahan yang dikembangkan dan dipasarkan oleh Sony Computer Entertainment. Sebagai konsol generasi kelima, PlayStation terutama bersaing dengan Nintendo 64 dan Sega Saturn. [7]
2. PlayStation 2  
PlayStation 2 (PS2) adalah konsol penerus dari PlayStation dan angsuran kedua dari merek konsol PlayStation. Sebagai konsol generasi keenam, PlayStation 2 bersaing dengan GameCube milik Nintendo, Dreamcast milik Sega, dan Xbox milik Microsoft. [8]
3. PlayStation 3  
PlayStation 3 (PS3) adalah konsol permainan video rumah yang dikembangkan dan dipasarkan oleh Sony Interactive Entertainment. Penerus PlayStation 2, ini adalah bagian dari merek konsol PlayStation. Ini pertama kali dirilis pada 11 November 2006 di Jepang, 17 November 2006 di Amerika Utara, dan 23 Maret 2007 di Eropa dan Australia. PlayStation 3 bersaing terutama melawan Microsoft Xbox 360 dan Nintendo Wii sebagai bagian dari konsol permainan video generasi ketujuh [9].
4. PlayStation 4  
PlayStation 4 (PS4) adalah konsol permainan video buatan Sony Computer Entertainment yang merupakan konsol penerus PlayStation 3. Di pasaran, PlayStation 4 berkompetisi dengan Nintendo Wii U dan Microsoft Xbox One sebagai salah satu konsol video game generasi kedelapan [10].
5. PlayStation 5  
PlayStation 5 (PS5) adalah konsol video game rumah yang dikembangkan oleh Sony Interactive

Entertainment yang diumumkan sebagai penerus PlayStation 4. PS5 adalah bagian dari generasi kesembilan konsol video game, bersama dengan konsol Xbox Series X/S dari Microsoft, yang dirilis pada bulan yang sama [11].

### 2.3. Genre Game Playstation.

Dalam dunia game, ada banyak jenis genre game yang ada di pasar. Genre game diklasifikasikan berdasarkan karakteristik, tema, dan elemen tertentu yang ada dalam gameplay game tersebut. Berikut merupakan beberapa genre game yang ada:

1. Action  
Genre ini menekankan kecepatan, refleks, dan ketangkasan. Pemain sering terlibat dalam pertempuran atau situasi berbahaya [12]. Contohnya termasuk:
  - Tembak-menembak (Shooter): Fokus pada penggunaan senjata, baik dari sudut pandang pertama (FPS) atau ketiga (TPS).
  - Platformer: Pemain melompat dan berlari melalui level, seperti dalam Super Mario.
2. Fighting  
Game fighting melibatkan pertarungan satu lawan satu, dengan karakter yang memiliki kemampuan unik. Contoh terkenal termasuk Mortal Kombat dan Street Fighter.
3. Sport  
Genre ini mencakup berbagai cabang olahraga, dari sepak bola hingga balapan. Game seperti FIFA dan NBA 2K menawarkan pengalaman simulasi olahraga yang realistis. [12]
4. Racing  
Racing merupakan subgenre dari sports game yang fokus utamanya adalah balapan, seperti Super Mario Kart, yang menawarkan pengalaman balapan yang seru dengan karakter ikonik dari franchise Mario. Namun, tidak jarang game racing juga termasuk dalam kategori sport atau olahraga, seperti game F1 Racing atau MotoGP, dan kadang-kadang dikombinasikan dengan subgenre simulasi untuk membuat pengalaman bermain yang lebih realistis. [12]
5. Strategy  
Game strategi memiliki tujuan untuk mendorong pemain untuk berpikir lebih kritis dan mengatur sumber daya untuk mencapai target. [12] Ada dua jenis utama:
  - Real-time Strategy (RTS): Seperti StarCraft, di mana pemain harus membuat keputusan cepat.
  - Turn-based Strategy: Seperti Civilization, di mana pemain bergiliran.
6. Role Playing Game (RPG)  
Dalam RPG, pemain mengendalikan karakter dalam dunia fiksi dan sering kali terlibat dalam cerita yang mendalam. [13] Ada subgenre seperti:
  - Action RPG: Menggabungkan elemen action dan RPG, seperti Dark Souls.

- MMORPG: Game online massal seperti World of Warcraft.
7. Adventure  
Genre ini fokus pada eksplorasi dan pemecahan teka-teki. Game seperti The Legend of Zelda dan Monkey Island sangat menonjol dalam kategori ini. [13]
  8. Simulation  
Game simulasi menciptakan pengalaman mendekati kenyataan, seperti mengelola kota (SimCity) atau kehidupan sehari-hari (The Sims). [13]
  9. Horror  
Genre ini dirancang untuk menciptakan ketakutan dan ketegangan. Elemen seperti jump scares dan atmosfer gelap sering digunakan. Contoh termasuk Resident Evil dan Silent Hill.
  10. Puzzle  
Game puzzle menantang pemain untuk memecahkan teka-teki atau masalah logika. Contoh termasuk Tetris dan Candy Crush.

### 2.4. Sistem QA

Sistem QA merupakan pengembangan dari Natural Language Processing (NLP) yang merupakan sebuah teknologi yang memfasilitasi interaksi antara manusia dan komputer melalui bahasa manusia. QA memberikan jawaban langsung terhadap pertanyaan tanpa perlu membuka halaman web yang terkait. Pengguna dapat berkomunikasi dengan sistem seperti dalam percakapan. Hal ini membuatnya lebih intuitif dan mudah digunakan. [14]

### 2.5. Metode TF-IDF dan Similarity

Metode Term Frequency-Inverse Document Frequency (TF-IDF) adalah salah satu teknik yang digunakan dalam pengolahan teks dan pemodelan bahasa alami. Tujuan utama dari metode TF-IDF adalah untuk mengevaluasi seberapa relevan suatu kata (term) dalam sebuah dokumen dalam konteks koleksi dokumen yang lebih besar. Metode TF-IDF adalah singkatan dari term frequency-inverse document frequency, dan bobot TF-IDF adalah bobot yang sering digunakan dalam pengambilan informasi dan penambangan teks.

Term Frequency (TF) adalah ukuran seberapa sering suatu kata muncul dalam sebuah dokumen. Biasanya dihitung dengan membagi jumlah kata yang muncul tersebut dengan total kata dalam dokumen. Dalam beberapa situasi, TF dapat diubah dengan menggunakan metode penimbangan yang lebih kompleks [15]. Karena masing-masing dokumen memiliki panjang yang berbeda, ada kemungkinan bahwa istilah tertentu akan muncul lebih sering dalam dokumen yang lebih panjang daripada yang lebih pendek. Akibatnya, frekuensi istilah biasanya dibagi dengan jumlah dokumen secara normal [6]:

$$TF(t, d) = \text{kata}_t / |d|. \quad (1)$$

Keterangan:

Katad = jumlah term atau kata ke-i yang muncul di

dokumen ke-d  
 $|d|$  = jumlah semua kata di dokumen ke-d

Inverse Document Frequency (IDF) adalah ukuran seberapa penting kata dalam koleksi dokumen yang lebih besar. Kata-kata yang muncul lebih jarang cenderung memiliki IDF yang lebih tinggi [15]. Menghitung IDF adalah dengan membagi total jumlah dokumen dalam koleksi dengan jumlah dokumen yang mengandung kata tersebut. Untuk memperhalus skala, hasilnya kemudian dihitung dengan logaritma.[5] Semua istilah dianggap sama pentingnya saat menghitung TF. Istilah-istilah tertentu, seperti "is", "of", dan "that", dikenal sering muncul tetapi tidak signifikan. Dengan demikian,

$$IDF(t, D) = \log(|D| / \sum_{kata_t \in d, d \in D} 1) \quad (2)$$

Keterangan:

$|D|$  = jumlah dokumen  
 $\sum_{kata_t \in d, d \in D} 1$  = jumlah dokumen yang mengandung term atau kata ke-t

Setelah perhitungan TF dan IDF, rumus selanjutnya adalah mencari TF-IDF. Berikut merupakan rumus tersebut.

$$TFIDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (3)$$

Semakin tinggi nilai TF-IDF, semakin relevan kata tersebut dalam dokumen. [16]

Metode kesamaan kosinus digunakan untuk menemukan kesamaan vektor fitur dengan kumpulan data yang tersedia. Kueri yang memiliki nilai kesamaan kosinus yang lebih besar dengan vektor fitur akan digunakan untuk memilih solusi masalah. Sistem ini menggunakan kesamaan kosinus untuk mengevaluasi ukuran kesamaan antara vektor kueri dan vektor matriks kesamaan [5].

$$\cos \theta = \frac{A \cdot B}{\|A\| \times \|B\|} \quad (4)$$

A dan B adalah vektor TF-IDF dari pertanyaan dan dokumen.

$A \cdot B$  = Perkalian dot product antara dua vektor.  
 $\|A\|$  dan  $\|B\|$  = Panjang vektor masing-masing.

## 2.6. Metode Evaluasi

Metode evaluasi yang digunakan untuk system ini adalah metode evaluasi pengguna. Metode evaluasi pengguna (user evaluation) adalah pendekatan yang digunakan untuk menilai efektivitas, kegunaan, dan kepuasan pengguna terhadap suatu sistem, produk, atau layanan. Metode ini sangat penting dalam pengembangan perangkat lunak, desain antarmuka pengguna, dan sistem informasi, karena memberikan wawasan langsung dari

pengguna akhir tentang bagaimana mereka berinteraksi dengan produk. Metode ini melibatkan pengguna yang diminta untuk menyelesaikan tugas tertentu menggunakan sistem. Pengamat mencatat kesulitan yang dihadapi pengguna, waktu yang dibutuhkan, dan kesalahan yang terjadi. Ini bertujuan untuk mengidentifikasi masalah dalam desain dan meningkatkan pengalaman pengguna.

## 3. Pembuatan Aplikasi

### 3.1. Bahasa Pemrograman

Dalam pembuatan website, beberapa bahasa pemrograman perlu digunakan. Bahasa pemrograman yang digunakan merupakan HTML, CSS, Javascript, dan PHP. Berikut merupakan penjelasan mengenai Bahasa tersebut.

1. HTML  
 HTML adalah bahasa markup standar yang digunakan untuk membuat dan menyusun konten di web. HTML mengatur struktur halaman dengan elemen-elemen yang ditandai dengan tag. Tag ini menentukan berbagai bagian dari halaman, seperti judul, paragraf, link, gambar, dan banyak lagi. [17]
2. CSS  
 CSS adalah bahasa yang digunakan untuk mendesain dan memformat tampilan halaman web yang ditulis dalam HTML. CSS mengontrol layout, warna, font, dan elemen visual lainnya, memungkinkan pemisahan antara konten dan presentasi. Dengan CSS, pengembang dapat membuat halaman web yang responsif dan menarik secara visual. [18]
3. Javascript  
 JavaScript adalah bahasa pemrograman yang digunakan untuk membuat halaman web interaktif. Dengan JavaScript, pengembang dapat menambahkan elemen dinamis, seperti animasi, interaksi pengguna, dan pemrosesan data di sisi klien. JavaScript dapat digunakan bersamaan dengan HTML dan CSS untuk meningkatkan fungsionalitas dan pengalaman pengguna di web. [19]
4. PHP  
 PHP adalah bahasa pemrograman sisi server yang digunakan untuk pengembangan web. PHP dapat menghasilkan konten HTML, menangani formulir, dan berinteraksi dengan database. PHP sering digunakan untuk membangun aplikasi web yang dinamis dan dapat disesuaikan, serta mendukung pengembangan sistem manajemen konten (CMS) seperti WordPress. [20]

### 4. Pengujian

#### a. Cara Pengujian

Pengujian fungsionalitas aplikasi dilakukan pada setiap modul yang ada. Untuk model Home dan About, pengujian modul tersebut dilakukan dengan menguji fungsionalitas fungsi yang ada pada setiap model, seperti tombol untuk membawa pengguna ke halaman web yang lain.

Untuk Q&A, Tahap pengujian ini berfokus pada kesesuaian jawaban terhadap pertanyaan yang dikirim. Pengujian ini dilakukan dengan implementasi confusion matrix. Confusion Matrix diimplementasi untuk mengevaluasi model klasifikasi. Berikut merupakan kelas dari sampel pengujian model.

1. TP (True Positive) : sampel pengujian diklasifikasikan dengan benar sebagai kelas positif
2. TN (True Negative) : sampel pengujian diklasifikasikan dengan benar sebagai kelas negative
3. FP (False Positive) : sampel pengujian diklasifikasikan secara salah sebagai kelas positif
4. FN (False Negative) : sampel pengujian diklasifikasikan secara salah sebagai kelas negative

Tabel 1 Confusion Matrix

		Hasil Prediksi	
		Positif	Negatif
Hasil Sebenarnya	Positif	$\Sigma TP$	$\Sigma TN$
	Negatif	$\Sigma FP$	$\Sigma FN$

Untuk Q&A, Tahap pengujian ini berfokus pada kesesuaian jawaban terhadap pertanyaan yang dikirim. Pengujian ini dilakukan dengan implementasi confusion matrix. Confusion Matrix diimplementasi untuk mengevaluasi model klasifikasi. Berikut merupakan kelas dari sampel pengujian model.

1. Akurasi : Jumlah sampel yang diklasifikasikan dengan benar dari semua sampel dalam data uji (lihat persamaan 5 berikut).

$$\frac{\Sigma TP + \Sigma FN}{\Sigma TP + \Sigma FP + \Sigma TN + \Sigma FN} \tag{5}$$

2. Presisi : Proporsi data yang diprediksi sebagai positif dan memang benar positif (lihat persamaan 6 berikut).

$$\frac{\Sigma TP}{\Sigma TP + \Sigma FP} \tag{6}$$

3. Recall (Sensitivitas) : Seberapa baik model dalam memprediksi kasus positif dengan benar (lihat persamaan 7 berikut).

$$\frac{\Sigma TP}{\Sigma TP + \Sigma TN} \tag{7}$$

4. Skor F1 : Skor F1 yang tinggi merupakan indikator bahwa model memiliki performa yang baik (lihat persamaan 8 berikut).

$$2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{8}$$

#### b. Hasil Pengujian

Pengujian fungsionalitas aplikasi dilakukan pada setiap modul yang ada. Untuk model Home, About dan Q&A, pengujian modul tersebut dilakukan dengan menguji fungsionalitas fungsi yang ada pada setiap model, seperti tombol untuk membawa pengguna ke halaman web yang lain.



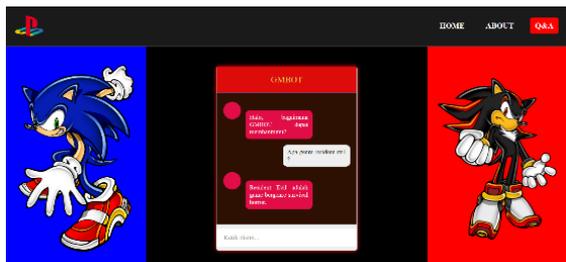
Gambar 2 Tampilan Home

Gambar 2 menunjukkan hasil pengujian modul beranda. Tujuan dari modul beranda adalah untuk menampilkan halaman utama aplikasi dengan navigasi ke modul QA dan About. Modul beranda merupakan halaman web pertama yang akan dilihat oleh pengguna saat pertama kali akses ke aplikasi web ini.



Gambar 3 Tampilan About Us

Modul ini bertujuan untuk memberikan konteks dan latar belakang kepada pengguna tentang sistem QA yang dibuat. Modul ini berisi tentang pembuatan GMBOT dan name penciptanya.



Gambar 4 Tampilan QnA

Modul QA merupakan modul yang memproses pertanyaan pengguna dan mengembalikan jawaban berdasarkan TF-IDF dan kesamaan kosinus. Halaman web QA akan menampilkan kotak chat yang menjadi media pengguna akan gunakan untuk mengirim pertanyaan ke sistem QA chatbot. Sistem tersebut akan membalas pertanyaan pengguna dengan jawaban yang memiliki kemiripan dengan pertanyaan pengguna.

Untuk pengujian modul Q&A, skenario pertama yang diuji adalah scenario dimana pertanyaan memiliki kata-kata sama dengan yang ada dalam dataset. Dalam dataset ini, apabila sesuai dengan pertanyaan yang ada dalam dataset, akan keluar jawaban tepat untuk pertanyaan tersebut.

Tabel 2 Pengujian pertanyaan yang diinputkan di GMBOT sama dengan pertanyaan di dataset

		Jawaban Hasil TF-IDF	
		Sesuai (P)	Tidak Sesuai (N)
Jawaban di dataset	Sesuai (T)	$\Sigma TP = 50$	$\Sigma TN = 0$
	Tidak Sesuai (F)	$\Sigma FP = 0$	$\Sigma FN = 0$

Berikut ini merupakan perhitungan akurasi, presisi, recall, dan nilai F1 untuk hasil percobaan ini.

$$\begin{aligned} \text{Akurasi} &= (50 + 0) / (50 + 0 + 0 + 0) = 1 \\ \text{Presisi} &= 50 / (50 + 0) = 1 \\ \text{Recall} &= 50 / (50 + 0) = 1 \\ \text{Skor F1} &= 2 * (1 * 1) / (1 + 1) = 1 \end{aligned}$$

Berdasarkan hasil perhitungan di atas, model ini sangat baik karena nilai 1 merupakan nilai tertinggi dalam rumus-rumus yang telah digunakan. Ini berarti modul sistem Q&A sangat baik dalam skenario ini, di mana semua pertanyaan akan mendapat jawaban yang benar jika pertanyaan pengguna sama persis dengan dataset.

Pada skenario berikutnya menggunakan pertanyaan yang mengandung hanya 50% kata yang sama. 50 % dari jumlah kata dalam suatu pertanyaan dapat diambil secara langsung dan menghilangkan beberapa kata dari kalimat pertanyaan tersebut dan disesuaikan untuk pengujian. Kata yang di hapus Sebagian besar merupakan kata tanya dan sebagian kata dari judul game. Kata yang disimpan merupakan kata kunci dan seringkali Sebagian dari kata kunci itu dihilangkan, ini dapat merupakan judul unik pada game yang merupakan lanjutan dari game sebelumnya, seperti God of War II, Resident Evil 2.

Apabila kalimat memiliki jumlah kata yang lebih, maka jumlah kata yang dihilangkan akan dibulatkan ke atas atau ke bawah.

Tabel 3 Pengujian pertanyaan relevan dengan 50 % kata yang sama dengan pertanyaan di dataset

		Target Jawaban Hasil TF-IDF	
		Sesuai	Tidak Sesuai
Jawaban di dataset	Sesuai	$\Sigma TP = 40$	$\Sigma TN = 3$
	Tidak Sesuai	$\Sigma FP = 7$	$\Sigma FN = 0$

Berikut ini merupakan perhitungan akurasi, presisi, recall, dan nilai F1 untuk hasil percobaan ini.

$$\begin{aligned} \text{Akurasi} &= (40 + 0) / (40 + 7 + 0 + 3) = 0,86 \\ \text{Presisi} &= 40 / (40 + 7) = 0,85 \\ \text{Recall} &= 40 / (40 + 3) = 0,93 \\ \text{Skor F1} &= 2 * (0,85 * 0,93) / (0,85 + 0,93) = 0,89 \end{aligned}$$

Dari hasil pengujian tersebut, evaluasi model klasifikasi menunjukkan bahwa model sistem ini baik karena nilai 0,89 mendekati nilai 1. Nilai tersebut lebih rendah dari hasil perhitungan percobaan yang sebelumnya.

Confusion matrix diatas terdiri dari 4 segmen. Nilai sesuai dan sesuai (True Positive) merupakan output dimana jawaban TF-IDF sesuai dengan jawaban yang sebenarnya didataset. Untuk yang sesuai dan tidak sesuai (False Positive), ini merupakan hasil jawaban TF-IDF tetap mengeluarkan hasil, walaupun tidak ada jawaban yang sebenarnya sesuai dengan jawaban TF-IDF. Selanjutnya, nilai tidak sesuai dan sesuai (True Negative) merupakan nilai TF-IDF tidak sesuai jawaban sebenarnya karena ada typo dipertanyaan, tetapi jawaban sebenarnya ada di dataset. Segmen terakhir (False Negative) adalah segmen nilai dimana TF-IDF tidak dapat menjawab dan tidak ada jawaban di datasetnya.

Pengujian terakhir tentang kesesuaian pertanyaan dengan respon jawaban dari TF-IDF ditingkatkan dengan memberikan nilai batasan TF-IDF untuk pertanyaan yang tidak relevan, yaitu jawaban yang sesuai tidak ada di dataset. Nilai batasan TF-IDF yang diberikan adalah 0,6.

Table 4 Pengujian pertanyaan relevan dengan 50 % kata yang sama dengan pertanyaan di dataset menggunakan nilai batasan di atas 0,6

		Target Jawaban Hasil TF-IDF	
		Sesuai	Tidak Sesuai
Jawaban di dataset	Sesuai	$\Sigma TP = 40$	$\Sigma TN = 2$
	Tidak Sesuai	$\Sigma FP = 1$	$\Sigma FN = 7$

Berikut ini merupakan perhitungan akurasi, presisi, recall, dan nilai F1 untuk hasil percobaan ini.

$$\text{Akurasi} = (40 + 7) / (40 + 2 + 1 + 7) = 0.94$$

$$\text{Presisi} = 40 / (40 + 1) = 0.98$$

$$\text{Recall} = 40 / (40 + 2) = 0.95$$

$$\text{Skor F1} = 2 * (0.98 * 0.95) / (0.98 + 0.95) = 0.96$$

Dari hasil perhitungan diatas, evaluasi model klasifikasi menunjukkan bahwa skor F1 lebih tinggi dari hasil perhitungan pada percobaan yang sebelumnya. Berdasarkan hasil tersebut, model sistem ini lebih baik karena memiliki nilai 0,96 yang mendekati nilai 1. Percobaan pada pertanyaan yang sebelumnya berada dalam kategori FP telah masuk ke dalam kategori FN. Dalam segmen FP, ada pertanyaan yang masuk ke segmen itu. Pertanyaan tersebut tidak ada jawaban, tetapi masih keluar suatu jawaban yang berhubungan dengan genre game dari pertanyaan itu.

## 5. Kesimpulan

Aplikasi berbasis web yang dibuat merupakan sistem tanya jawab atau Question-Answering (QA) yang dibuat dengan menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF) yang bernama GMBOT. Hasil dari pembuatan aplikasi ini sudah sesuai keinginan dengan memiliki tiga halaman web dengan halaman Q&A merupakan halaman utama yang memiliki sistem QA chatbotnya yang dapat dijalankan.

Pengujian utama dilakukan pada halaman yang ada sistem QA chatbotnya yang memiliki dua pengujian. Pengujian tersebut dilakukan dengan batas nilai kemiripan TF-IDF yang berbeda yang mempengaruhi apakah pertanyaan yang seharusnya tidak ada jawaban akan dibalas dengan jawaban dari dataset atau dibalas dengan jawaban yang menyatakan bahwa tidak ada jawaban untuk pertanyaan itu. Nilai batasan tersebut berupa 0,6 dan diatas nilai 0. Hasil pengujian menunjukkan bahwa nilai batasan 0,6 memiliki nilai yang lebih baik daripada nilai batasan diatas 0.

Untuk pengujian pertanyaan sama persis dengan dataset, nilai akurasi, presisi, recall dan skor F1 semua memiliki nilai 1.

Pengujian dengan 50% kemiripan dengan pertanyaan di dataset dilakukan dengan nilai batasan diatas 0 dan 0,6. Berikut adalah hasil dari pengujian dengan nilai batasan diatas 0.

1. Akurasi = 0,86
2. Presisi = 0,85
3. Recall = 0,93
4. Skor F1 = 0,89

Sedangkan, hasil dari pengujian dengan nilai batasan 0,6 sebagai berikut.

1. Akurasi = 0.94
2. Presisi = 0.98
3. Recall = 0.95
4. Skor F1 = 0.96

Hasil pengujian menunjukkan bahwa nilai batasan 0,6 memiliki nilai yang lebih baik daripada nilai batasan diatas 0, pengujian dengan pertanyaan yang sama persis dengan dataset akan selalu memunculkan jawaban yang tepat.

Sistem tanya jawab ini masih dapat dimaksimalkan dengan berbagai cara. Tampilan UI dapat dikembangkan lagi agar terlihat lebih menarik. Selain itu, sistem dapat ditambahkan dengan metode lain yang lebih efektif. Metode alternatif tersebut dapat merupakan BERT (Bidirectional Encoder Representations from Transformers) [21].

## REFERENSI

- [1] Zaib M, Zhang WE, Sheng QZ, Mahmood A, Zhang Y. Conversational question answering: A survey. *Knowledge and Information Systems*. 2022 Dec;64(12):3151-95. Li, Shujun., Zheng, Xuan., Mei 2002, "Cryptanalysis of a Chaotic Image Encryption Method", *Proceeding IEEE – ISACS*, Vol.2, Scottsdale -Arizona.
- [2] Zhu F, Lei W, Wang C, Zheng J, Poria S, Chua TS. Retrieving and reading: A comprehensive survey on open-domain question answering. *arXiv preprint arXiv:2101.00774*. 2021 Jan 4. Referensi memiliki kebaruan maksimal 5 tahun lamanya dari tanggal penelitian
- [3] Biancofiore GM, Deldjoo Y, Noia TD, Di Sciascio E, Narducci F. Interactive question answering systems: Literature review. *ACM Computing Surveys*. 2024 May 8;56(9):1-38.
- [4] Open Source Software Information," PlayStation, 2024. <https://www.playstation.com/en-us/oss/> (accessed Apr. 27, 2025).
- [5] Mawardi, V. C. ; J. Pragantha & Sarjono, S. (2013) "Naive bayes classification in the question and answering system," in *Proceedings of the International Conference on Information Systems for Business Competitiveness (ICISBC 2013)*, pp. 139-144.
- [6] Patil, R. and Tiwari, S. (2020). "Question answering system using a TF-IDF Method," *International Journal of Science Engineering and Technology*, pp. 1-5.
- [7] Wikipedia, "PlayStation (console)," Wikipedia, The Free Encyclopedia, [Online]. Available: [https://en.wikipedia.org/wiki/PlayStation\\_\(console\)](https://en.wikipedia.org/wiki/PlayStation_(console)). [Accessed 25 Februari 2025].
- [8] Wikipedia, "PlayStation 2," Wikipedia, The Free Encyclopedia, [Online]. Available: [https://id.wikipedia.org/wiki/PlayStation\\_2](https://id.wikipedia.org/wiki/PlayStation_2). [Accessed 25 Februari 2025].
- [9] Wikipedia, "PlayStation 3," Wikipedia bahasa Indonesia, ensiklopedia bebas, [Online]. Available: [https://id.wikipedia.org/wiki/PlayStation\\_3](https://id.wikipedia.org/wiki/PlayStation_3). [Accessed 25 Februari 2025].
- [10] Wikipedia, "PlayStation 4," Wikipedia, The Free Encyclopedia, [Online]. Available: [https://id.wikipedia.org/wiki/PlayStation\\_4](https://id.wikipedia.org/wiki/PlayStation_4). [Accessed 25 Februari 2025].
- [11] Wikipedia, "PlayStation 5," Wikipedia, The Free Encyclopedia, [Online]. Available: [https://en.wikipedia.org/wiki/PlayStation\\_5](https://en.wikipedia.org/wiki/PlayStation_5). [Accessed 25 Februari 2025].
- [12] Tempo, "Menikmati akhir pekan, mengenal berbagai macam genre game," *Tempo.co*, 11 Mei 2024. [Online].

- Available: <https://www.tempo.co/digital/menikmati-akhir-pekan-mengenal-berbagai-macam-genre-game-60172>. [Accessed 20 Februari 2025].
- [13] Axe, "10 genre game terpopuler, mana yang paling seru?," Axe, [Online]. Available: <https://www.axe.com/id/inspirasi/culture/genre-game.html>. [Accessed 25 Februari 2025].
- [14] A. T. U. B. Lubis, N. S. Harahap, S. Agustian, M. Irsyad and I. Afrianty, "Question answering system on telegram chatbot using large language models (LLM) and langchain (case study: health law)," MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. 4, no. 3, pp. 955-964, 2024.
- [15] D. Septiani and I. Isabela, "Analisis term frequency inverse document frequency (TF-IDF) dalam temu kembali informasi pada dokumen teks," SINTESIA: Jurnal Sistem dan Teknologi Informasi Indonesia, vol. 1, no. 2, pp. 81-88, 2022.
- [16] R. Madan, "TF-IDF/Term frequency technique: easiest explanation for text classification in NLP using python (Chatbot training on words)," Analytics Vidhya, 30 Mei 2019. [Online]. Available: <https://medium.com/analytics-vidhya/tf-idf-term-frequency-technique-easiest-explanation-for-text-classification-in-nlp-with-code-8ca3912e58c3>. [Accessed 20 Februari 2025].
- [17] "HTML Introduction," W3Schools, [Online]. Available: [https://www.w3schools.com/html/html\\_intro.asp](https://www.w3schools.com/html/html_intro.asp). [Accessed 1 Maret 2025].
- [18] "CSS Introduction," W3Schools, [Online]. Available: [https://www.w3schools.com/css/css\\_intro.asp](https://www.w3schools.com/css/css_intro.asp). [Accessed 1 Maret 2025].
- [19] M. D. N. (MDN), "Dynamic scripting with JavaScript - Learn web development," [Online]. Available: [https://developer.mozilla.org/en-US/docs/Learn\\_web\\_development/Core/Scripting](https://developer.mozilla.org/en-US/docs/Learn_web_development/Core/Scripting). [Accessed 1 Maret 2025].
- [20] PHP.net, "PHP Manual," [Online]. Available: <https://www.php.net/manual/en/>. [Accessed 1 Maret 2025].
- [21] M. Y. Ridho, "Apa itu BERT dan IndoBERT? Apakah sebuah Word Embedding?," Medium, May 26, 2023. [Online]. Available: <https://medium.com/@muhammadyusufidho6/apa-itu-bert-dan-indobert-apakah-sebuah-word-embedding-f870c25593e4>