



Sistem Informasi Klasifikasi Hadits Berdasarkan Algoritma *Naive Bayes Classifier*

Indri Mada Aprilia*¹, Isnandar Agus², Kurnia Muludi³

*^{1,2,3} Program Studi Teknik Informatika, Institut Informatika dan Bisnis Darmajaya, Bandar Lampung, Indonesia

*Email Penulis Korespondensi: indrimada12@gmail.com

Abstrak

Kemajuan teknologi informasi telah berkembang pesat dan banyak dimanfaatkan sebagai media pembelajaran serta pencarian informasi, termasuk dalam pencarian hadits berdasarkan label dan perawinya. Hadits, sebagai panduan hidup umat Islam setelah Al-Qur'an, telah dibukukan, namun penyajian hadits dalam bentuk buku membuat banyak orang kesulitan menemukan hadits yang dicari karena pencarian masih dilakukan secara manual dengan jumlah data yang besar, sehingga membutuhkan banyak tenaga dan waktu. Penelitian ini bertujuan untuk membantu masyarakat dalam mencari hadits sesuai dengan label dan perawinya. Metode pengembangan perangkat lunak yang digunakan adalah metode waterfall, yang mencakup tahapan seperti analisis kebutuhan, desain sistem, penulisan kode program, pengujian sistem, dan pemeliharaan, dengan menerapkan algoritma Naive Bayes Classifier. Proses klasifikasi ini melibatkan beberapa tahapan Preprocessing Text, seperti case folding, tokenisasi, filtering/Stopword Removal, stemming, training, dan prediction, yang pada akhirnya menghasilkan keluaran hadits terkait beserta label dan perawinya. Hasil dari penelitian ini adalah sistem informasi dan pencarian hadits dengan label dan perawinya yang menerapkan algoritma Naive Bayes Classifier berbasis web.

Kata kunci—Naive Bayes Classifier, Klasifikasi Hadits, Teknologi Informasi.

Abstract

The rapid growth of information technology has significantly enhanced its use as a medium for learning and information retrieval, including searching for hadiths based on specific labels and narrators. Hadith, being the second primary guide for Muslims after the Qur'an, is often difficult to access due to the manual search process that is time-consuming and inefficient, especially with the vast amount of data available. This research aims to assist the public in finding hadiths that correspond to specific labels and narrators. The software development method used in this study is the waterfall model, which includes stages such as needs analysis, system design, coding, system testing, and maintenance. The Naive Bayes Classifier algorithm is applied in the classification process, involving steps like text preprocessing (case folding, tokenization, stopword removal, stemming), training, and prediction, ultimately generating outputs related to

hadiths, their labels, and narrators. The outcome of this research is a web-based information system for searching hadiths by labels and narrators using the Naïve Bayes Classifier algorithm.

Keywords—*Naïve Bayes Classifier, Hadits Classification, Information Technology.*

1. PENDAHULUAN

Perkembangan internet di Indonesia telah meningkat pesat yang mengakibatkan banyak muncul portal *online* yang bertujuan membantu orang menemukan informasi di berbagai situs, terutama bagi para aktivis yang sangat membutuhkan berita dan informasi terkini [1]. Seiring dengan kemajuan teknologi, internet pun mulai berkembang dan kini digunakan secara luas di berbagai aspek kehidupan di Indonesia [2].

Kemajuan teknologi informasi yang pesat telah membawa dampak positif di berbagai bidang [3]. Namun, dengan semakin banyaknya dokumen dan informasi yang tersebar di dunia maya masalah seperti ketidak efektifan dalam pencarian dan penyajian informasi menjadi lebih menonjol. Banyak orang kini memanfaatkan teknologi media belajar dan sumber informasi, termasuk dalam pencarian hadits berdasarkan label dan perawinya [4]. Hadits yang disajikan dalam bentuk buku seringkali sulit ditemukan karena pencariannya masih dilakukan secara manual, sementara jumlah data yang besar membuat proses ini memakan waktu dan tenaga. Oleh karena itu diperlukan metode yang lebih cepat dan tepat untuk menemukan hadits yang diinginkan.

Dalam bidang ilmu komputer, terdapat suatu area yang dikenal sebagai kecerdasan buatan (*Artificial Intelligence*) atau AI [5]. Teknologi kecerdasan buatan ini telah banyak digunakan dalam kehidupan sehari-hari, termasuk dalam tugas pengklasifikasian teks. Salah satu teknik dalam pengklasifikasian teks adalah metode *Naïve Bayes* [6]. Konsep dari klasifikasi teks menggunakan algoritma ini adalah dengan menentukan label pada teks baru yang belum diketahui, melalui pelatihan pada kumpulan teks yang labelnya sudah diketahui [7]. Oleh karena itu, dilakukan kajian literatur mengenai topik penelitian teks yang mengumpulkan berbagai metode atau algoritma dengan menganalisis teks hadits. Algoritma *Naïve Bayes* memiliki keunggulan dibandingkan algoritma lain karena lebih mudah dipahami, diimplementasikan dan sangat cocok untuk klasifikasi teks.

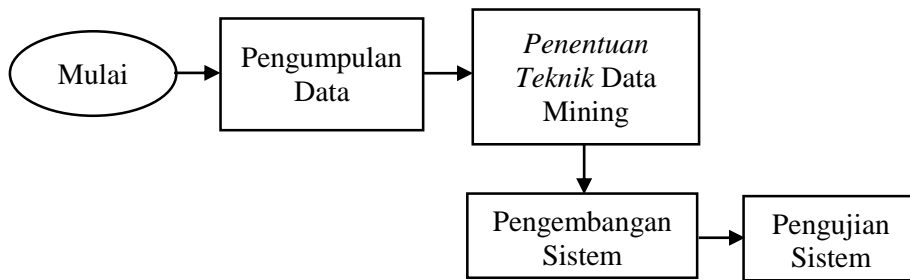
Penelitian terdahulu telah dilakukan oleh Emerald dan Kemas [8] dimana pada penelitian ini telah berhasil dilakukan pengelompokkan pengkategorian hadits. Pada penelitian lain algoritma *Naïve Bayes Classifier* berhasil mengklasifikasikan dokumen berita dengan sesuai ke dalam masing-masing kategori yang telah ditentukan [9]. Untuk penelitian pengklasifikasian yang lain juga sebelumnya pernah dilakukan untuk mengoptimisasikan pencarian pada aplikasi perpustakaan dan berhasil terklasifikasikan berdasarkan judul, kategori, serta pengarangnya [10]. Dyah dan Kurnia Iswardani [11] juga menggunakan algoritma *Naïve Bayes Classifier* untuk penelitiannya pada klasifikasi keluhan masyarakat dan berhasil mengklasifikasikannya pada tiap-tiap instansi di pemerintah Kota Probolinggo. Penelitian sebelumnya juga dilakukan [4] untuk pengklasifikasian dokumen terjemahan Al-Qur'an dengan menghasilkan 5 kategori klasifikasi dengan tepat. Dengan demikian, peneliti berharap dari penelitian ini diharapkan dapat membantu masyarakat mengenai informasi dalam pencarian hadits berdasarkan label, perawinya serta hadits terkait.

2. METODE PENELITIAN

Naïve Bayes Classifier, sebagaimana dijelaskan oleh [12] adalah metode klasifikasi sederhana yang mampu menghitung berbagai kemungkinan dengan menggabungkan berbagai kombinasi dan frekuensi nilai dari data yang diperoleh. Algoritma ini memanfaatkan teorema *Bayes* untuk memperkirakan seluruh atribut yang dianggap independen dan berhubungan dengan nilai tertentu pada variabel kelas. *Naïve bayes* adalah metode klasifikasi berbasis probabilitas yang dikembangkan oleh seorang ilmuwan asal Inggris, Thomas Bayes [13] yaitu memprediksi probabilitas masa depan berdasarkan data dan pengalaman sebelumnya.

2.1 Tahapan Penelitian

Tahapan penelitian ini dapat dilihat pada Gambar 1.



Gambar 1 Tahapan Penelitian

Dasar dari penerapan metode *Naive Bayes* dalam pemrograman adalah rumus *Bayes*, yang dirumuskan dalam formula (1).

$$P(A|B) = (P(B|A) \times P(A))/P(B) \quad (1)$$

Peluang terjadinya A sebagai B dihitung berdasarkan peluang terjadinya B ketika A terjadi, peluang terjadinya A, dan peluang terjadinya B. Dalam penerapannya nanti, rumus ini akan menjadi seperti formula (2).

$$P(C_i|D) = (P(D|C_i))/P(D) \quad (2)$$

Naive Bayes Classifier, yang juga dikenal sebagai *Multinomial Naive Bayes*, adalah model yang menyederhanakan metode *Bayes* dan sangat cocok untuk klasifikasi teks atau dokumen dengan persamaan seperti formula (3).

$$V_{MAP} = \arg \max P(V_j | a_1, a_2, \dots, a_n) \quad (3)$$

Menurut persamaan (3), maka persamaan (1) dapat ditulis menjadi formula (4).

$$V_{MAP} = \underset{V_j \in V}{\arg \max} \frac{P(a_1, a_2, \dots, a_n | v_j) P(v_j)}{P(a_1, a_2, \dots, a_n)} \quad (4)$$

$P(a_1, a_2, \dots, a_n)$ bersifat konstan, sehingga dapat dihapuskan menjadi formula (5).

$$V_{MAP} = \underset{V_j \in V}{\arg \max} P(a_1, a_2, \dots, a_n | v_j) P(v_j) \quad (5)$$

Karena menghitung $P(a_1, a_2, \dots, a_n | v_j)$ sulit dilakukan, maka diasumsikan bahwa setiap kata dalam dokumen dianggap tidak saling berhubungan, seperti pada formula (6) dengan nilai P dapat dilihat pada formula (7) dan formula (8).

$$V_{MAP} = \underset{V_j \in V}{\arg \max} P(V_j) \prod_i P(a_i | v_j) \quad (6)$$

Keterangan:

$$P(v_j) = \frac{|docs_j|}{|contoh|} \quad (7)$$

$$P(w_k | v_j) = \frac{n_k}{n + |kosakata|} \quad (8)$$

Digunakan untuk:

$P(v_j)$: Probabilitas setiap dokumen dalam kaitannya dengan sekelompok dokumen.

$P(w_k | v_j)$: Probabilitas munculnya kata w_k dalam sebuah dokumen dengan kategori class v_j .

$|docs|$: Jumlah dokumen dalam setiap kategori.

$|contoh|$: Jumlah dokumen yang ada.

N_k : Frekuensi kemunculan kata ke- k dalam setiap kategori.

Kosakata: Jumlah kata dalam dokumen uji.

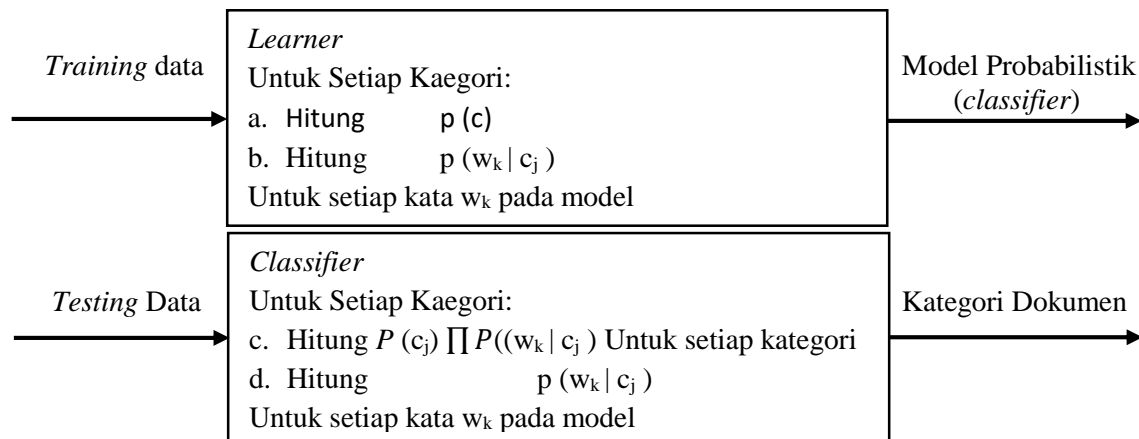
Ada penambahan angka 1 pada pembilang dalam persamaan (8) untuk mengatasi kemungkinan kata yang muncul dalam dokumen uji tetapi tidak ada di dokumen data pelatihan.

2.2 Langkah kerja Naïve Bayes Classifier

Langkah kerja *Naïve Bayes Classifier* terdiri dari dua tahap, yaitu:

- 1) Pembelajaran, yang merupakan metode pembelajaran terawasi (*supervised learning*).
- 2) Pengklasifikasian, yaitu tahap dimana data diklasifikasikan.

Langkah kerja dalam *Naïve Bayes Classifier* dapat dilihat pada Gambar 2.



Gambar 2 Langkah Kerja Naïve Bayes Classifier

Selain itu, pada penelitian ini dilakukan beberapa proses lainnya, yaitu:

1. Pengumpulan data, dalam penelitian ini dilakukan melalui dua metode, yaitu studi pustaka dan kajian literatur.
2. Penentuan teknik *data mining*, setelah data dikumpulkan untuk penelitian ini, langkah berikutnya adalah memproses data tersebut menggunakan metode *Naïve Bayes Classifier*.
3. Pengembangan perangkat lunak, metode yang akan diterapkan untuk pengembangan perangkat lunak dalam penelitian ini adalah metode *waterfall*.
4. Pengujian Sistem, setelah merancang dan mengembangkan sistem selanjutnya sistem diuji.

2.3 Pre-processing

Tahap *preprocessing* data dalam *Naive Bayes Text Classifier* melibatkan serangkaian langkah untuk menyesuaikan data dengan format yang diperlukan. Pada tahap ini, dilakukan seleksi data dari setiap dokumen yang akan diproses, dan *pre-processing* bertujuan untuk mempermudah proses klasifikasi data. Penjelasan mengenai langkah-langkah *pre-processing* yang diterapkan dalam penelitian ini dapat dilihat pada *flowchart* yang disajikan dalam Gambar 3. Tahapan tersebut yaitu sebagai berikut:

1. Case Folding

Tahap di mana semua karakter dalam dokumen diubah menjadi huruf kecil (*lowercase*) bertujuan untuk memastikan bahwa teks hadits dalam dokumen memiliki format yang konsisten dan seragam [6].

2. Tokenization

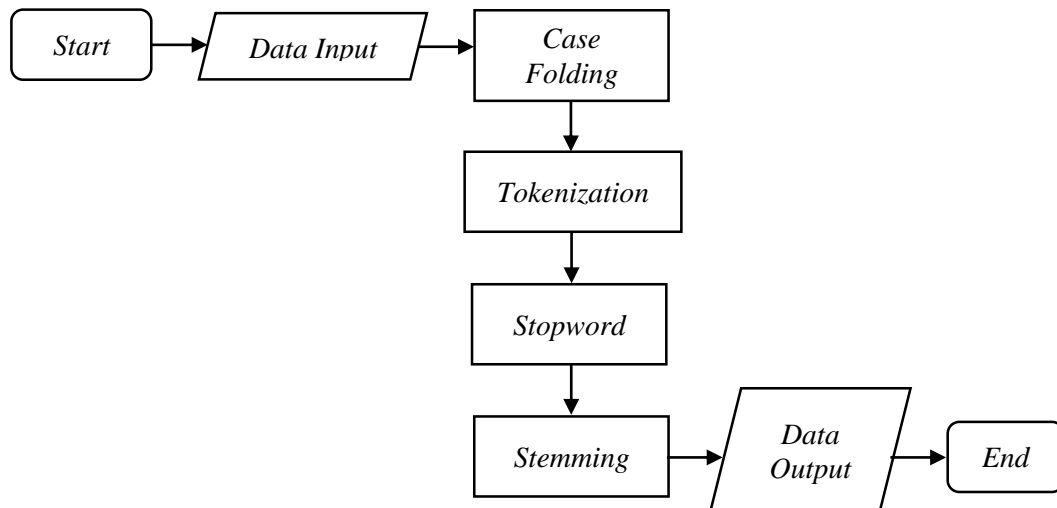
Tahap ini biasa disebut dengan tokenisasi, yaitu proses dimana kalimat akan dipecah menjadi kata atau yang disebut dengan token [14] tujuannya adalah untuk membedakan apakah sesuatu berfungsi sebagai pemisah kata atau tidak.

3. Stopword

Tahap ini digunakan untuk mengekstrak kata-kata kunci dari hasil token sebelumnya. [15]. Kata yang sering muncul dan tidak memiliki makna itulah yang disebut dengan *stopword*. Seperti penggunaan kata penghubung seperti dan, yang, serta, setelah, dan lainnya.

4. Stemming

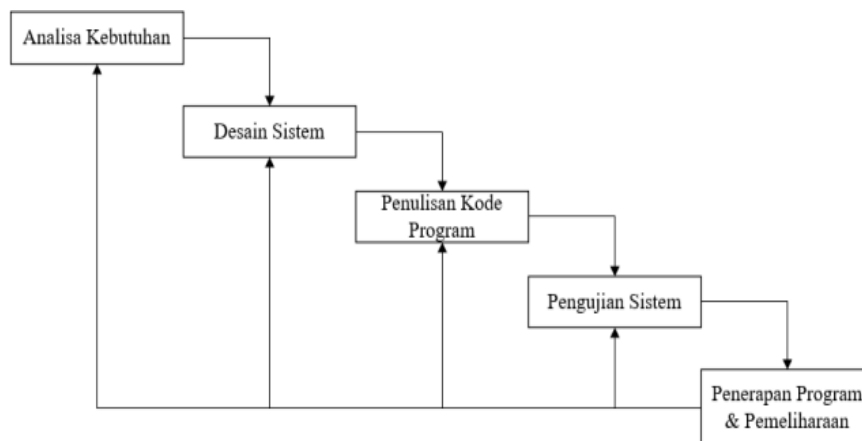
Tahap ini digunakan untuk menentukan kata dasar dari suatu kata. Teknik ini diterapkan untuk mengekstrak kata dasar dari teks yang memiliki imbuhan. Tujuannya adalah agar saat proses pembobotan, kata-kata yang berbagi kata dasar yang sama mendapatkan bobot yang serupa [7].



Gambar 3 Tahapan *Preprocessing*

2.4 Metode Pengembangan Perangkat Lunak

Metode yang diterapkan dalam pengembangan perangkat lunak pada penelitian ini adalah model *Waterfall*, yang termasuk dalam kategori metode pengembangan perangkat lunak klasik dengan mengikuti pendekatan sistematis dan berurutan [16]. Tahapan dari metode *waterfall* dapat dilihat pada Gambar 4.



Gambar 4 Tahapan Metode *Waterfall*

Berikut adalah tahapan metode *Waterfall* yang dimaksud:

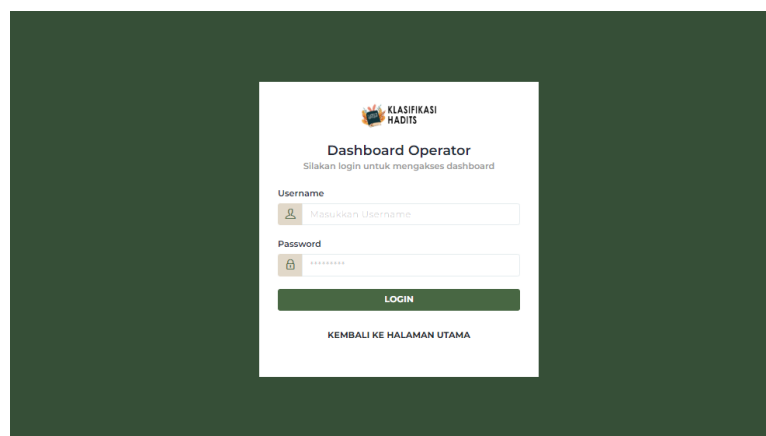
1. Analisis Kebutuhan: pada tahap ini, dilakukan analisis untuk menentukan kebutuhan sistem. Pengumpulan data dapat dilakukan melalui penelitian, wawancara, atau studi literatur.
2. Desain Sistem: fase ini bertujuan untuk merancang sistem perangkat lunak yang akan dikembangkan.
3. Pengkodean: tahap ini melibatkan penulisan kode program untuk membangun sistem perangkat lunak berdasarkan desain yang telah disiapkan.

4. Pengujian Sistem: pada tahap ini, dilakukan pengujian untuk memastikan bahwa sistem perangkat lunak memenuhi kebutuhan pengguna.
5. Pemeliharaan: tahap ini bertujuan untuk memperbaiki kesalahan dan meningkatkan sistem perangkat lunak setelah sistem tersebut dirilis ke pengguna [17].

3. HASIL DAN PEMBAHASAN

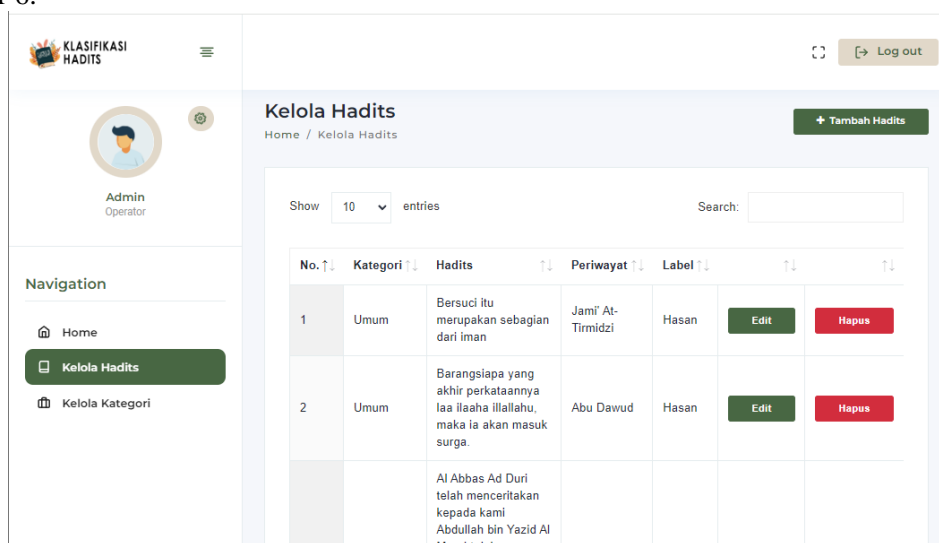
3.1 Hasil

Penelitian ini menghasilkan sebuah situs web untuk mengklasifikasikan hadits berdasarkan jenis dan perawinya menggunakan metode *Naïve Bayes*. Situs ini dirancang untuk mempermudah pengguna dalam mencari hadits, dengan menyediakan hasil pencarian yang relevan, beserta label dan nama perawinya. Akses login hanya diberikan kepada admin untuk memungkinkan pembaruan sistem secara berkala. Pengguna dapat mengakses situs *web* tanpa perlu login terlebih dahulu. Tampilan antar muka aplikasi dapat dilihat pada Gambar 5.



Gambar 5 Login Admin

Setelah berhasil login, *admin* akan diarahkan ke halaman beranda yang menampilkan jumlah kategori hadits serta total hadits yang diperoleh dari studi literatur. Selain halaman beranda, terdapat juga menu Kelola Hadits, di mana admin dapat menambah, mengedit, atau menghapus data hadits. Tampilan halaman beranda dan menu Kelola Hadits dapat dilihat pada Gambar 6.



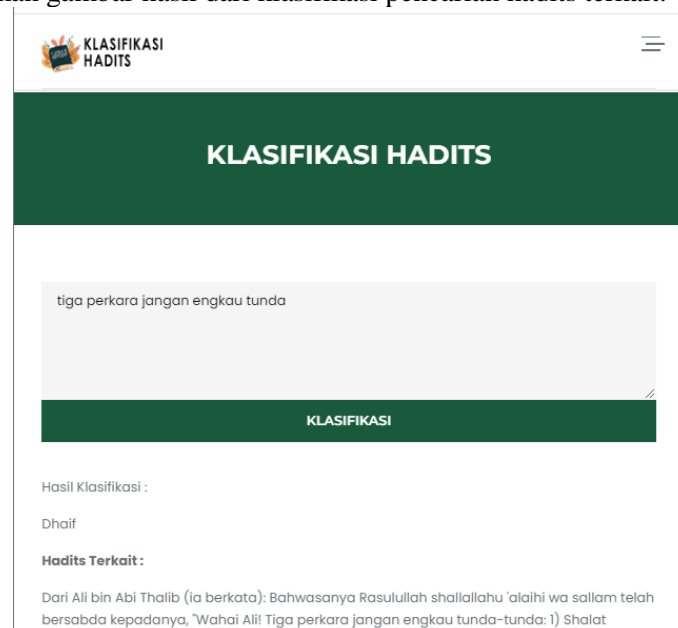
Gambar 6 Tampilan dari Halaman Utama dan Kelola Hadits pada Admin

Gambar 7 menunjukkan halaman beranda yang dirancang untuk pengguna, di mana mereka dapat mengakses kumpulan hadits yang telah dimasukkan oleh admin berdasarkan jenis hadits dan perawinya. Gambar 7 tersebut menggambarkan halaman daftar hadits.



Gambar 7 Lembar Daftar Hadits

Selesai *user* mengunjungi halaman beranda yang berisi kumpulan hadits-hadits, kemudian terdapat halaman pencarian hadits. Untuk mencari hadits yang dimaksud oleh *user*, *user* cukup hanya dengan memasukkan tiga atau lebih kata saja dan kemudian sistem tersebut akan mengeluarkan output hadits yang dimaksud oleh *user* lengkap dengan label serta perawinya. Gambar 8 merupakan gambar hasil dari klasifikasi pencarian hadits terkait.



Gambar 8 Hasil Klasifikasi Pencarian Hadits terkait

3.2 Pembahasan

Data awal (data *primer*) adalah data yang akan dianalisis dalam bentuk teks, yang dikumpulkan oleh peneliti dan diperoleh dari [18] serta akan digunakan untuk menentukan label hadits dan perawinya. Beberapa *dataset* yang digunakan dalam penelitian ini dapat dilihat pada Gambar 9, Gambar 10, dan Gambar 11.

1	Hadits	Periwayat	Label
2	Segala sesuatunya terdapat kerusakan, sedangkan kerusakan di dalam agama ialah Para pemimpin (ur	Syaikhul hnam Muhammad Nashiru	Maudhu
3	Memandang kepada wanita yang cantik dan sesuatu yang hijau akan menambah (kekuatan) pada mata	Abu Nu'aim di kitabnya Al Hilyah	Maudhu
4	Memandang kepada wajah yang cantik/cakap akan membuat mata menjadi terang, sedangkan mema	Al Khatib Al Baghdadiy di kitabnya	Maudhu
5	Melihat kepada wajah yang cantik/cakap adalah ibadah	Al Imam Ali Qari'	Maudhu
6	Ketika saya di isra'kan Jibril membawaku melewati kubur Ibrahim, maka Jibril berkata: "Turunlah, mak	Al Imam Abul Fadhi Muhammad bir	Maudhu
7	Janganlah kamu mendekati Yahudi dan Nashara pada hari raya-hari raya mereka, karena sesungguhnya	Al Imam Ibnu Hibban di kitabnya Al Maudhu	
8	Maukah aku kabarkan kepada kamu orang yang paling celaka? Yaitu orang yang Allah telah mengumpu	Al Imam Ibnu Hibban di kitabnya Al Maudhu	
9	Shalat dengan memakai sorban pahalanya menyamai berperang di jalan Allah.	Al Imam Ibnu Adiy di kitabnya Al K	Maudhu

Gambar 9 Hasil Pengumpulan Dataset

209	Mutharif bin Abdillah berkata: "Aku dan Imran bin Hushain shalat di	Bukhari	Shahih
210	Anas berkata bahwa Nabi Saw., Abu Bakar, dan Umar memulai shalat	Bukhari	Shahih
211	Abu Humaid As-S ^o di e berkata: "Sahabat bertanya tentang bagaiman	Bukhari	Shahih
212	Abu Hurairah r.a. berkata: "Rasulullah n bersabda: 'Jika imam memba	Bukhari	Shahih
213	Abu Hurairah r.a. berkata: "Rasulullah Saw. bersabda: 'Jika kalian mer	Bukhari	Shahih
214	Abu Hurairah r.a. berkata: "Rasulullah Saw. bersabda: 'Jika imam sele	Bukhari	Shahih
215	(Surat) Yasin itu hati Qur'an, tidak membacanya seseorang asrena All	Imam Ahmad bin Hambal	Dhaif
216	Aku adalah kota ilmu sedangkan Ali pintunya.	Tirmidzi	Dhaif
217	Apabila engkau berjumpa dengan orang yang baru pulang dari menun	Ahmad	Dhaif
218	Apabila seorang hamba itu menikah maka telah sempurna setengah c	Imam As Sakhawiy	Dhaif

Gambar 10 Lanjutan Hasil Pengumpulan Dataset

283	Tidak seorangpun yang berumur empat puluh tahun di dalam Islam n	Imam Ibnu Hibban di kitabnya Al M	Dhaif
284	Dari Ali bin Abi Thalib (ia berkata): Bahwasanya Rasulullah shallallahu	Imam Tirmidzi dan Hakim	Dhaif
285	Zakat adalah harta kekayaan Islam.	Baihaqiyy, Ishaq bin Rahuwaih, Thal	Dhaif
286	"Dari Ibnu Umar dari Muhammad bin al-Mutsanna dari Muhammad	Abu Dawud	Hasan
287	dari Abu Hurairah bahwa Rasulullah shallallahu 'alaihi wasallam bersa	Jami' At-Tirmidzi	Hasan
288	Sesungguhnya pintu surga berada di bawah bayangan pedang	Jami' At-Tirmidzi	Hasan
289	Ahmad bin Mani telah menceritakan kepada kami Abbad bin 'Abbad	Bukhari	Hasan
290	Hadis tentang pemberian nama anak 'Ubaidullah bin Sa'd bin Ibrahim	Jami' At-Tirmidzi	Hasan

Gambar 11 Lanjutan Hasil Pengumpulan Dataset

3.3 Pemberian Bobot TF-IDF

Setelah mengumpulkan *dataset*, langkah selanjutnya adalah proses pembobotan menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Tujuan dari pembobotan ini adalah untuk mengubah nilai hitung setiap kata menjadi fitur yang berguna. Proses perhitungan TF-IDF dilakukan dengan mengalikan nilai *Term Frequency* (TF) dengan *Inverse Document Frequency* (IDF) untuk memperoleh skor TF-IDF bagi setiap kata, dengan menggunakan enam dokumen yang telah disediakan.

D1 = sebaik-baik obat adalah Al-Qur'an.

D2 = sebaik-baik urusan yang sedang atau tengah-tengah.

D3 = Agama itu dibangun atas dasar kebersihan.

D4 = Imran bin Hushain berkata bahwa Nabi Saw bersabda, "Rasa malu tidak menghasilkan apa pun kecuali kebaikan."

D5 = Barang siapa yang berbicara mengenai Al-Qur'an tanpa memiliki pengetahuan, maka ia telah mempersiapkan tempat tinggalnya di neraka.

D6 = Barangsiapa yang sakit selama tiga hari akan hapus dosa-dosanya seperti pada hari dia dilahirkan ibunya.

$$w(t, d) = tf(t, d) \times idf(t, D)$$

Keterangan:

$w_{t,d}$: Bobot dari t (term) dalam satu dokumen

$tf_{t,d}$: frekuensi kemunculan t (term) dalam d (dokumen)

idf_t : *invers document frequency*

Tabel 1 Hasil Perhitungan TF-IDF *Naive Bayes Classifier*

Teks	D1	D2	D3	D4	D5	D6
Sebaik-baik	0.15051	0.10033	0	0	0	0
obat	0.22577	0	0	0	0	0

adalah	0.22577	0	0	0	0	0
al-qur'an	0.15051	0	0	0	0.04013	0
urusan	0	0.15050	0	0	0	0
yang	0	0.05016	0	0	0.00444	0.01881
sedang	0	0.15050	0	0	0	0
atau	0	0.15050	0	0	0	0
Tengah -tengah	0	0.15050	0	0	0	0
agama	0	0	0.15050	0	0	0
itu	0	0	0.07099	0.02839	0	0
dibangun	0	0	0.15050	0	0	0
atas	0	0	0.15050	0	0	0
dasar	0	0	0.15050	0	0	0
kebersihan	0	0	0.15050	0	0	0
imran	0	0	0	0.06019	0	0
bin	0	0	0	0.06019	0	0
hushain	0	0	0	0.06019	0	0
saw	0	0	0	0.06019	0	0
berkata	0	0	0	0.04013	0.04013	0
nabi	0	0	0	0.06019	0	0
bersabda	0	0	0	0.06019	0	0
malu	0	0	0	0.06019	0	0
tak	0	0	0	0.06019	0	0
mendatangkan	0	0	0	0.06019	0	0
sesuatu	0	0	0	0.06019	0	0
kecuali	0	0	0	0.06019	0	0
kebaikan	0	0	0	0.06019	0	0
barangsiapa	0	0	0	0	0.06019	0.02839
tentang	0	0	0	0	0.06019	0
tanpa	0	0	0	0	0.06019	0
ilmu	0	0	0	0	0.06019	0
maka	0	0	0	0	0.04013	0
hendaklah	0	0	0	0	0.06019	0
dia	0	0	0	0	0.02839	0.02839
mengambil	0	0	0	0	0.06019	0
tempat	0	0	0	0	0.06019	0
tinggalnya	0	0	0	0	0.06019	0
di	0	0	0	0	0.06019	0
neraka	0	0	0	0	0.06019	0
sakit	0	0	0	0	0	0.06019
selama	0	0	0	0	0	0.06019
tiga	0	0	0	0	0	0.06019
hari	0	0	0	0	0	0.120407
akan	0	0	0	0	0	0.04013
hapus	0	0	0	0	0	0.06019
Dosa-dosanya	0	0	0	0	0	0.06019
seperti	0	0	0	0	0	0.06019
pada	0	0	0	0	0	0.06019
dilahirkan	0	0	0	0	0	0.06019
ibunya	0	0	0	0	0	0.06019

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan sebagai berikut:

1. Penelitian ini menghasilkan sebuah sistem pencarian hadits berdasarkan klasifikasinya yang berbasis *website* dengan penerapan algoritma *Naïve Bayes Classifier*.
2. Sistem ini menggunakan empat tahap dalam proses klasifikasi, yaitu *case folding*, *tokenization*, *stopword removal*, dan *stemming*. Output dari sistem tersebut memberikan hasil berupa hadits yang relevan beserta label dan perawi hadits tersebut.
3. Sistem informasi yang menggunakan algoritma *Naïve Bayes Classifier* terbukti efektif dalam mengklasifikasikan hadits-hadits berdasarkan kriteria tertentu.

5. SARAN

Berdasarkan kesimpulan dan analisis yang telah disampaikan, beberapa saran untuk penelitian berikutnya adalah sebagai berikut:

1. Menerapkan algoritma lain untuk klasifikasi teks, seperti *Decision Tree*, dan *K-Nearest Neighbour*.
2. Untuk mendapatkan hasil yang lebih bervariasi, dapat digunakan bahasa pemrograman yang berbeda dan diimplementasikan pada objek lain selain yang digunakan dalam penelitian ini.
3. Menambah jumlah data hadits yang dianalisis.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Tim Redaksi Jurnal Teknik Politeknik Negeri Sriwijaya yang telah memberikan kesempatan, sehingga artikel ilmiah ini dapat dipublikasikan.

DAFTAR PUSTAKA

- [1] R. F. Daud, "Dampak Perkembangan Teknologi Komunikasi Terhadap Bahasa Indonesia," *J. Interak. J. Ilmu Komun.*, vol. 5, no. 2, pp. 252–269, 2021.
- [2] F. Ramadanti, H. Sazali, and I. F. D. Peranginangin, "Efektivitas Penggunaan Website Pada Humas Di Kementerian Agama Provinsi Sumatera Utara," *J. Ris. Mhs. Dakwah dan Komun.*, vol. 3, no. 5, p. 288, 2022, doi: 10.24014/jrmdk.v3i5.15627.
- [3] A. Premana, G. Fitralisma, A. Yulianto, M. B. Zaman, and M. A. Wiryo, "Pemanfaatan teknologi informasi pada pertumbuhan ekonomi dalam era disrupsi 4.0," *J. Econ. Manag.*, vol. 2, no. 2, pp. 1–6, 2020.
- [4] G. F. Dimas, "Klasifikasi Dokumen Terjemahan Menggunakan Algoritma Naïve Bayes Pada Al- Qur ' an Versi Berbahasa Indonesia (Juz 1 Sampai Juz 4) Klasifikasi Dokumen Terjemahan Menggunakan Algoritma Naïve Bayes Pada Al- Qur ' an Versi Berbahasa Indonesia (Juz 1 Sampai," pp. 1–112, 2018.
- [5] M. R. Astari, M. S. A. K. Mardlian, S. Bahri, and M. U. Siregar, "RANCANGAN APLIKASI CHATBOT TELEGRAM 'TANYA ZAID' SEBAGAI MEDIA PEMBELAJARAN NAHWU," *Konf. Integr. Interkoneksi Islam dan Sains*, vol. 5, no. 1, pp. 313–323.
- [6] H. Syah and A. Witanti, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm)," *J. Sist. Inf. Dan Inform.*, vol. 5, no. 1, pp. 59–67, 2022.
- [7] A. Alfando and R. Hayami, "Klasifikasi Teks Berita Berbahasa Indonesia Menggunakan Machine Learning Dan Deep Learning: Studi Literatur," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 681–686, 2023, doi: 10.36040/jati.v7i1.6486.
- [8] Emerald and K. M. Lhaksana, "Klasifikasi Kategori Hadits Menggunakan Naive Bayes Classifier," *e-Proceeding Eng.*, vol. 6, no. 2, p. 11, 2019.
- [9] E. Y. Hidayat and M. A. Rizqi, "Klasifikasi Dokumen Berita Menggunakan Algoritma

-
- Enhanced Confix Stripping Stemmer dan Naïve Bayes Classifier,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 6, no. 2, pp. 90–99, 2020, doi: 10.25077/teknosi.v6i2.2020.90-99.
- [10] S. Lorena., “Teknik Data Mining Menggunakan Metode Bayes Classifier Untuk Optimalisasi Pencarian Aplikasi Perpustakaan,” *J. Tek. Komput.*, vol. 4, no. 2, pp. 17–20, 2016.
- [11] D. Ariyanti and K. Iswardani, “Teks Mining untuk Klasifikasi Keluhan Masyarakat Pada Pemkot Probolinggo Menggunakan Algoritma Naïve Bayes,” *J. IKRA-ITH Inform.*, vol. 4, no. 3, pp. 125–132, 2020.
- [12] R. Rachman and R. N. Handayani, “Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM,” *J. Inform.*, vol. 8, no. 2, pp. 111–122, 2021, doi: 10.31294/ji.v8i2.10494.
- [13] A. Pebdika, R. Herdiana, and D. Solihudin, “Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 452–458, 2023.
- [14] Rian Tineges, “Tahapan Text Preprocessing dalam Teknik Pengolahan Data,” *DQLab*. 2021, [Online]. Available: <https://www.dqlab.id/tahapan-text-preprocessing-dalam-teknik-pengolahan-data>.
- [15] B. Irawan and O. Nurdiawan, “Naive Bayes Dan Wordcloud Untuk Analisis Sentimen Wisata Halal Pulau Lombok,” *INFOTECH J.*, vol. 9, no. 1, pp. 236–242, 2023.
- [16] T. Pricillia and Zulfachmi, “Perbandingan Metode Pengembangan Perangkat Lunak (Waterfall, Prototype, RAD),” *J. Bangkit Indones.*, vol. 10, no. 1, pp. 6–12, 2021, doi: 10.52771/bangkitindonesia.v10i1.153.
- [17] K. Rahmat S, I. Kanedi, and J. Fredricka, “Application of the Naive Bayes Method in the Classification of Employee Status at PT. Prosperous New Light,” *J. Media Comput. Sci.*, vol. 2, no. 1, pp. 49–56, 2023, doi: 10.37676/jmcs.v2i1.3395.
- [18] M. F. Abdul Baqi, *Shahih Bukhari Muslim*. PT Elex Media Komputindo, 2017.
- [19] D. Atika, S. Styawati, and A. A. Aldino, “Term Frequency-Inverse Document Frequency Support Vector Machine untuk Analisis Sentimen Opini Masyarakat terhadap Tekanan Mental pada Media Sosial Twitter,” *J. Teknol. dan Sist. Inf.*, vol. 3, no. 4, pp. 86–97, 2022.