

IDENTIFIKASI MATAN HADITS MENGGUNAKAN NATURAL LANGUAGE PROCESSING DAN ALGORITMA KNUTH MORRIS PRATT BERBASIS WEB

^[1]Aris Munandar, ^[2]Victor Amrizal, ^[3]Arini

*Teknik Informatika, Fakultas Sains dan Teknologi
Universitas Islam Negeri Syarif Hidayatullah Jakarta*

Jl. Ir. H. Juanda No. 95, Tangerang Selatan – Banten 15412

^[1]aris.munandar14@mhs.uinjkt.ac.id, ^[2]victor.amrizal@uinjkt.ac.id, ^[3]arini@uinjkt.ac.id

Abstract - The process of determining the validity of the hadith is a process that requires a long period, because the process is manually. Hadith investigators need to refer from one book to another. Therefore, to facilitate the search for hadith the identification system of hadith is needed. This study uses natural language processing methods through the stages of tokenizing, filtering, and analysis as well as the Knuth Morris Pratt algorithm. Using a natural language processing system is able to understand the hadith text entered by each user. By using the knuth morris pratt algorithm, it can provide convenience in searching the text of the Matanic text which has similarities. So the system is able to produce a hadith that is in line with expectations and can find out the validity of the hadith.

Keyword : *Hadith, Natural Language Processing, Knuth Morris Pratt Algorithm.*

I. PENDAHULUAN

Hadits atau yang disebut juga dengan sunah, sebagai sumber ajaran islam yang berisi pernyataan, pengalaman, pengakuan maupun sikap persetujuan yang datang dari Nabi Muhammad SAW [2]. Sebagai dasar hukum kedua setelah Al-Qur'an bagi umat islam, maka hadits penting untuk dijaga dan dimurnikan dari segala pemalsuan. Sehingga para ulama di kalangan sahabat

berusaha menjaga hadits dengan serius melalui beberapa cara, salah satunya dengan perlawatan ke berbagai daerah islam untuk mengecek kebenaran hadits yang sampai kepada mereka baik dari segi matan atau sanad [4].

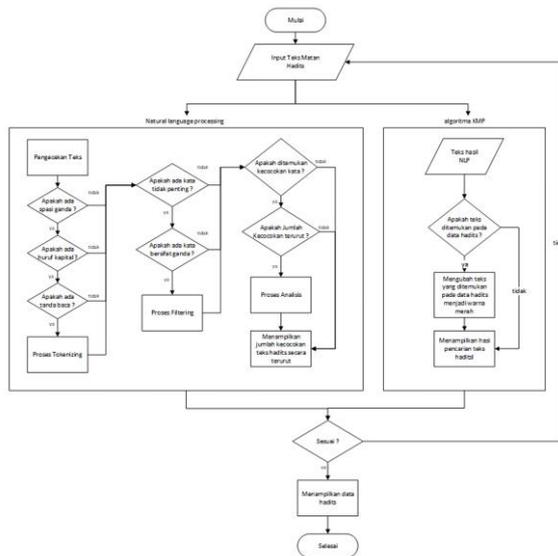
Melihat banyaknya hadits yang beredar, baik melalui tulisan maupun acara kajian dan diskusi publik, hal itu sering menjadikan masyarakat menjadi kebingungan untuk menerima atau tidak hadits yang diutarakan [1]. Menurut Penelitian (Nursyahidah, 2016) kesahihan hadits bergantung kepada sanad dan matan hadits. Proses penentuan kesahihan hadits adalah proses yang memerlukan masa yang panjang, karena proses tersebut dilakukan secara manual [6].

Peneliti melakukan riset mengenai pengembangan sistem *natural language processing* untuk identifikasi matan hadits menggunakan algoritma *knuth morris pratt* yaitu untuk mengetahui tanggapan masyarakat muslim terhadap sistem ini. Berdasarkan 43 responden yang telah mengisi kuesioner, menghasilkan data sebesar 97,7% menyatakan setuju apabila ada sistem untuk identifikasi matan hadits dan menghasilkan data sebesar 95,3% menyatakan sistem yang dikembangkan ini akan bermanfaat.

II. METODE PENELITIAN

Pada tahap ini, peneliti akan mengembangkan sistem dengan

mengaplikasikan metode *natural language processing* (NLP) menggunakan tahapan *tokenizing*, *filtering*, dan analisis serta menerapkan algoritma *knuth morris prrat* (KMP). Berikut merupakan alur kerja sistem:



Gambar 2.1 Flowchart Identifikasi Matan Hadits

Berikut ini merupakan tahapan metode *natural language processing* yang digunakan untuk membangun sistem identifikasi matan hadits:

a. *Tokenizing*

Pada tahap ini sistem akan memecah teks dari *input* pengguna dengan beberapa tahapan antara lain:

- Sistem akan menghapus spasi ganda yang terdapat dalam teks matan hadits agar setiap kata dapat dipisah dengan satu spasi.
- Sistem akan mengubah teks matan hadits menjadi huruf kecil semua agar menjadi seragam.
- Sistem akan menghapus tanda baca yang terdapat dalam teks matan hadits.

Berikut ini merupakan contoh menghapus spasi ganda pada teks matan hadits:

Man Soma Romadona (Teks makan hadits)



Man Soma Romadona (Hasil *tokenizing*)

b. *Filtering*

Pada tahap ini sistem akan melakukan penyaringan atau *filtering* kata dari teks matan hadits yang dimasukan. Proses *filtering* dimulai dengan sistem akan mengambil kata-kata penting dari hasil tahap sebelumnya yaitu tahap *tokenizing*. Berikut merupakan tahapan pada proses *filtering*:

- Mengambil kata-kata penting pada teks matan hadits dengan metode *stopword*.
- Menghapus kata matan hadits yang bersifat ganda pada teks matan hadits.

Dan di bawah ini merupakan contoh dari tahap *filtering* dengan mengambil kata-kata penting pada teks matan hadits:

man soma romadona imanana wa (Teks makan hadits)



soma romadona imanana (Hasil *filtering*)

c. Analisis

Pada tahap ini sistem melakukan proses analisis teks yang dimasukan oleh pengguna. Proses analisis dimulai dengan sistem akan mencari jumlah kecocokan kata matan hadits yang telah diproses sebelumnya

dengan teks matan hadits yang tersimpan di dalam *database*. Berikut merupakan tahapan pada proses analisis:

- Menghitung jumlah kecocokan kata matan hadits dengan data hadits yang tersimpan di dalam *database*.
- Mengurutkan data hadits berdasarkan jumlah kecocokan kata matan hadits yang ditemukan secara *descending*.

➤ Algoritma *Knuth Morris Pratt*

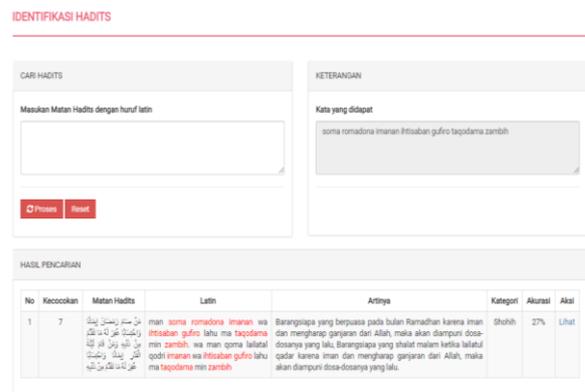
Pada tahap ini, sistem akan melakukan pencarian kata matan hadits pada setiap teks matan hadits yang tersimpan di dalam *database*. Teks matan hadits yang diproses sebelumnya dengan menggunakan *natural language processing* akan dicari kecocokannya dengan data hadits yang tersimpan di dalam *database*. Setiap kata matan hadits yang ditemukan, algoritma akan mengubah warna kata matan hadits tersebut menjadi warna merah. Sehingga pengguna tidak hanya dapat mengetahui jumlah kecocokan kata yang ditemukan, akan tetapi pengguna juga dapat mengetahui pada bagian mana saja kata matan hadits tersebut ditemukan. Dapat diketahui langkah-langkah algoritma KMP sebagai berikut:

- Algoritma *Knuth Morris Pratt* mulai mencocokkan *pattern* pada awal teks.
- Dari kiri ke kanan, algoritma ini akan mencocokkan karakter per karakter *pattern* dengan karakter di teks yang bersesuaian, sampai salah satu kondisi berikut dipenuhi:

- Karakter di *pattern* dan di teks yang dibandingkan tidak cocok (*mismatch*).
 - Semua karakter di *pattern* cocok. Kemudian algoritma akan memberitahukan penemuan di posisi ini.
- Algoritma kemudian menggeser *pattern* berdasarkan tabel next, lalu mengulangi langkah 2 sampai *pattern* berada di ujung teks.

III. HASIL DAN PEMBAHASAN

Pada tahap ini peneliti membangun sistem identifikasi matan hadits menggunakan metode *natural language processing* dengan tahapan *tokenizing*, *filtering*, dan analisis untuk memudahkan sistem dalam memahami teks yang dimasukan oleh pengguna.



Gambar 3.1 Hasil Identifikasi Matan Hadits

Berdasarkan gambar di atas, dengan menggunakan metode *natural language processing* melalui tahap *tokenizing* dan *filtering*, sistem mampu memahami setiap teks matan hadits yang dimasukan oleh pengguna, sehingga sistem mendapatkan kalimat yang digunakan untuk melakukan proses pencarian hadits. Dengan menerapkan tahap analisis, sistem mampu mengetahui jumlah kecocokan kata matan

hadits, dan jumlah kecocokan kata matan hadits tersebut akan diurutkan secara *descending*. Sehingga pengguna dapat mengetahui bahwa hadits yang memiliki jumlah kecocokan kata paling besar nilainya terdapat pada urutan teratas.

Pada penelitian (Dzakiyatur Rosyidah, 2014) teks yang dapat diproses yaitu teks yang memiliki jumlah maksimal 10 kata. Pada penelitian ini peneliti akan melakukan beberapa pengujian yaitu memasukan kata matan hadits dengan jumlah lebih dari 10 kata dan kurang dari 10 kata untuk mengetahui pengaruh jumlah teks matan hadits.

No	Teks yang dimasukkan	Jumlah Kata	Matan Hadits	Jumlah Kecocokan	Akurasi
1.	naumu soim ibadatun, wa sumtuhu tasbihun wa duauhu m ustajabun wa amaluhu mudoafun	12 Kata matan hadits	naumu soim ibadatun, wa sumtuhu tasbihun wa duauhu mustajabun wa amaluhu mudoafun	9	75%
2	naumu soim ibadatun, wa sumtuhu tasbihun	6 Kata matan hadits	naumu soim ibadatun, wa sumtuhu tasbihun wa duauhu mustajabun wa amaluhu mudoafun	5	42%
3	naumu wa amaluhu	3 Kata matan hadits	naumu soim ibadatun, wa sumtuhu tasbihun wa duauhu mustajabun wa amaluhu mudoafun	2	17%
4	mempelajari hadits dengan benar	5 Kata bukan matan hadits	-	-	-

Gambar 3.2 Analisa Pengaruh Jumlah Teks Matan Hadits

Berdasarkan tabel pengaruh jumlah teks matan hadits di atas, berikut deskripsi dari hasil pengujian identifikasi matan hadits:

1. Pada pengujian pertama pengguna memasukan teks sebanyak 12 kata matan hadits, sehingga dapat menghasilkan jumlah kecocokan sebanyak 9 kata matan hadits dengan tingkat akurasi sebesar 75%.
2. Pada pengujian kedua pengguna hanya memasukan teks sebanyak 6 kata matan hadits, sehingga dapat menghasilkan jumlah kecocokan sebanyak 5 kata

matan hadits dengan tingkat akurasi sebesar 42%.

3. Pada pengujian ketiga pengguna hanya memasukan teks sebanyak 3 kata matan hadits, sehingga dapat menghasilkan jumlah kecocokan sebanyak 2 kata matan hadits dengan tingkat akurasi sebesar 17%.
4. Pada pengujian keempat pengguna memasukan teks sebanyak 5 kata namun kata yang dimuskan bukanlah berdasarkan matan hadits, sehingga sistem tidak dapat menemukan hasil identifikasi matan hadits yang sesuai.

IV. KESIMPULAN

Berdasarkan tahap-tahap analisa yang telah dideskripsikan sebelumnya, maka peneliti mengambil kesimpulan bahwa merancang sistem identifikasi matan hadits menggunakan metode *natural language processing* dengan tahap *tokenizing*, *filtering*, dan analisis akan memberikan kemudahan terhadap sistem dalam melakukan pencarian hadits berdasarkan teks yang dimasukkan oleh setiap pengguna. Dan dengan menggunakan algoritma *knuth morris pratt* dapat memberikan keefektifan waktu dalam pencarian kata matan hadits baik pada teks dengan jumlah kata yang sedikit ataupun teks dengan jumlah kata yang banyak.

DAFTAR PUSTAKA

- [1] Fathullah. 2010. *Kajian Ilmu Hadis*. Jakarta.
- [2] Ghozali, M. 2015. *Analisis Sanad dan Matan Hadits Shalat Di atas Kendaraan*. Jakarta.
- [3] "Natural Language Processing" <http://socs.binus.ac.id/2013/06/22/>

- [4] Khon, Abdul Majid. 2015. *Ulumul Hadits*. Jakarta: Sinar Grafika Offset.
- [5] Matondang, Firman. 2016. *Perancangan Aplikasi Teks Editor Dengan Menerapkan Algoritma KMP*. Medan
- [6] Nursyahidah, Alias. 2016. *Aplikasi Pencarian Sanad Hadits dalam Kitab Shohih Bukhori*. Malaysia.
- [7] Rio, Alamanda. 2016. *Aplikasi Pendeteksi Plagiat Pada Karya Tulis dengan Menggunakan Natural Language Processing*. Pontianak.