

# BAB 1 PENDAHULUAN

## 1.1. Latar Belakang

Hadis adalah salah satu sumber hukum Islam setelah Al-Qur'an, memiliki peran yang sangat penting dalam kehidupan umat Islam sehari-hari. Hadis merupakan segala sesuatu yang melekat pada Rasulullah, baik berupa ucapan, perbuatan, maupun persetujuan beliau terhadap suatu tindakan atau kejadian tertentu [1]. Hadis Nabi Muhammad SAW berperan penting dalam menjelaskan dan melengkapi hukum-hukum yang terdapat dalam Al-Qur'an. Jumlah hadis yang sangat banyak, ditambah dengan beragamnya tema, menimbulkan tantangan dalam pengelolaan dan pemanfaatannya. Klasifikasi tematik manual memerlukan tenaga ahli, waktu yang panjang, dan sering kali tidak konsisten.

Dengan berkembangnya Natural Language Processing (NLP) dan deep learning, metode klasifikasi otomatis dapat menjadi solusi. CNN dikenal efektif untuk mengenali pola spasial dalam teks (Kim, 2014), sedangkan RNN, khususnya LSTM, mampu mempelajari urutan kata dan mempertahankan konteks jangka panjang (Hochreiter & Schmidhuber, 1997). Penelitian oleh Zhou et al. (2015) menunjukkan bahwa kombinasi CNN dan RNN menghasilkan performa lebih baik dibanding model tunggal [2]. Oleh karena itu, CRNN dipilih sebagai metode utama karena mampu menggabungkan keunggulan kedua arsitektur tersebut. CNN dan RNN murni juga digunakan sebagai pembandingan untuk mengevaluasi keunggulan dan kelemahan masing-masing model. Pemilihan metode ini didasarkan pada karakteristik teks hadis terjemahan yang memiliki struktur kalimat ringkas namun sarat makna, sehingga memerlukan kombinasi ekstraksi fitur lokal dan pemahaman konteks. Setiap hadis tersusun atas dua elemen pokok yang menjadi tolok ukur keaslian dan validitasnya. Elemen pertama adalah *matn*, yaitu isi atau substansi dari hadis yang berisi pesan, perintah, larangan, maupun penjelasan terkait hukum-hukum Islam. Elemen kedua adalah *sanad*, yakni rangkaian nama para perawi yang menyampaikan hadis tersebut secara berantai dari Rasulullah SAW hingga sampai kepada penyusunnya dalam kitab hadis. Sanad berfungsi sebagai bukti historis

keterhubungan hadis dengan sumber aslinya, serta menjadi penentu tingkat kepercayaan terhadap suatu hadis. Struktur ini secara keseluruhan ditampilkan dalam Tabel 1.1..

Tabel 1.1 Arsitektur hadis

|                                                                                                                                                                                                                                                                            |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| عَنْ أَبِي هُرَيْرَةَ رَضِيَ اللَّهُ عَنْهُ قَالَ قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ لَا تَحَاسَدُوا وَلَا تَنَاجَشُوا وَلَا تَبَاغَضُوا وَلَا تَحَاسَدُوا وَلَا تَنَاجَشُوا وَلَا تَبَاغَضُوا وَلَا تَذَابَرُوا وَكُونُوا عِبَادَ اللَّهِ إِخْوَانًا |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Sanad hadits ini adalah: Abu Hurairah radhiyallahu 'anhu → Rasulullah shallallahu 'alaihi wasallam

Matan hadits ini adalah: لَا تَحَاسَدُوا وَلَا تَنَاجَشُوا وَلَا تَبَاغَضُوا وَلَا تَذَابَرُوا وَكُونُوا عِبَادَ اللَّهِ إِخْوَانًا

Seiring dengan berkembangnya teknologi dan meningkatnya jumlah koleksi hadis yang tersedia, baik dalam bentuk teks cetak maupun digital, muncul tantangan besar dalam mengorganisir, mengelompokkan, dan menganalisis hadis-hadis tersebut secara efisien. Salah satu tantangan utama adalah bagaimana mengklasifikasikan hadis berdasarkan tema atau topiknya secara otomatis. Hal ini menjadi semakin kompleks mengingat satu hadis dapat memiliki lebih dari satu tema atau topik yang saling terkait. Misalnya, hadis yang berbicara tentang ibadah shalat tidak hanya berkaitan dengan fikih, tetapi juga dengan aspek akidah serta nilai-nilai adab dalam beribadah maupun informasi serta kisah.

Ada beberapa penelitian yang melakukan klasifikasi hierarkis terhadap hadis, terutama pada teks terjemahan dalam Bahasa Indonesia. Penelitian-penelitian tersebut umumnya mengkategorikan hadis ke dalam topik Anjuran, Larangan, dan Informasi dengan menggunakan berbagai metode baik *machine learning* maupun *deep learning*. Penelitian yang akan dilakukan mencoba mengembangkan klasifikasi hierarkis dengan mengelompokkan hadis ke dalam topik-topik yang lebih spesifik dan terkait langsung dengan terminologi keislaman, seperti Akidah, Fikih, Kisah dan Adab. Pendekatan klasifikasi seperti ini diharapkan dapat memberikan gambaran yang lebih akurat terhadap kandungan hadis berdasarkan konsep-konsep dasar dalam Islam, sehingga mempermudah pemanfaatan hadis dalam berbagai konteks akademik maupun praktis.

Kemajuan dalam bidang *Natural Language Processing* (NLP) dalam beberapa tahun terakhir telah membuka peluang baru untuk menganalisis teks, termasuk teks-teks keislaman seperti hadis. Berbagai algoritma modern dalam NLP telah diterapkan untuk tugas klasifikasi teks, di antaranya *Recurrent Neural Network* (RNN), *Convolutional Neural Network* (CNN), dan *Convolutional Recurrent Neural Network* (CRNN). Selama ini, klasifikasi hadis masih banyak dilakukan secara manual melalui kitab-kitab tematik yang disusun oleh para ulama. Meskipun sistem ini sangat berharga, pendekatan manual memiliki keterbatasan dalam hal efisiensi, skalabilitas, dan konsistensi. Oleh karena itu, dibutuhkan solusi berbasis teknologi, khususnya melalui pemanfaatan kecerdasan buatan, untuk mendukung pengelompokan hadis berdasarkan tema secara lebih cepat, akurat, dan adaptif.

Deep learning sebagai bagian dari perkembangan Artificial Intelligence (AI), telah menunjukkan performa tinggi dalam tugas-tugas pemrosesan bahasa alami (*Natural Language Processing*/NLP), termasuk klasifikasi teks. Arsitektur seperti *Convolutional Neural Network* (CNN) dikenal efektif dalam mengekstraksi fitur lokal dari teks, sementara *Recurrent Neural Network* (RNN), khususnya varian *Long Short-Term Memory* (LSTM), unggul dalam menangkap konteks sekuensial. Kombinasi keduanya diyakini dapat memberikan hasil yang lebih optimal dalam tugas klasifikasi tematik yang kompleks, seperti pada hadis.

Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi hierarkis tema hadis berbasis teks terjemahan menggunakan pendekatan deep learning dengan menggabungkan arsitektur CNN dan RNN. Model ini diharapkan mampu mengklasifikasikan hadis secara bertingkat mulai dari tema utama hingga subtema secara otomatis dan akurat, sehingga dapat mendukung pengembangan sistem pencarian tematik hadis serta aplikasi pendidikan Islam berbasis AI di masa depan.

## **1.2. Rumusan Masalah**

Dalam penelitian ini, permasalahan yang akan dibahas adalah pengembangan model klasifikasi hierarkis hadis menggunakan algoritma *Recurrent Neural Network* (RNN), *Convolutional Neural Network* (CNN), dan kombinasi RNN dan CNN yaitu *Convolutional Recurrent Neural Network* (CRNN). Pengklasifikasian tema

dalam hadis merupakan tugas yang kompleks karena satu teks hadis sering kali mengandung lebih dari satu tema. Dengan demikian, dibutuhkan pendekatan yang mampu mengidentifikasi dan mengelompokkan teks hadis secara hierarki ke dalam beberapa kategori yang relevan.

### 1.3. Tujuan dan Manfaat

Pada penelitian ini akan diimplementasikan metode-metode klasifikasi RNN, CNN dan *CRNN* dalam melakukan klasifikasi teks hierarki terhadap teks terjemahan hadis dengan menggunakan dataset yang telah dikumpulkan dari ensiklopedi. Evaluasi kinerja dilakukan menggunakan metrik precision, recall, F1-score, dan confidence guna menentukan model yang paling efektif. Manfaat yang diharapkan dari penelitian ini adalah tersedianya fondasi teknis dan konseptual bagi pengembangan sistem pencarian, pengelompokan, dan analisis tematik hadis yang adaptif, efisien, dan akurat. Dengan demikian, penelitian ini dapat mempermudah peneliti, akademisi, dan masyarakat dalam mengakses, mempelajari, serta memahami kandungan hadis secara tematik, mendukung digitalisasi literatur keislaman, dan menjadi acuan untuk pengembangan aplikasi kecerdasan buatan di bidang studi Islam.

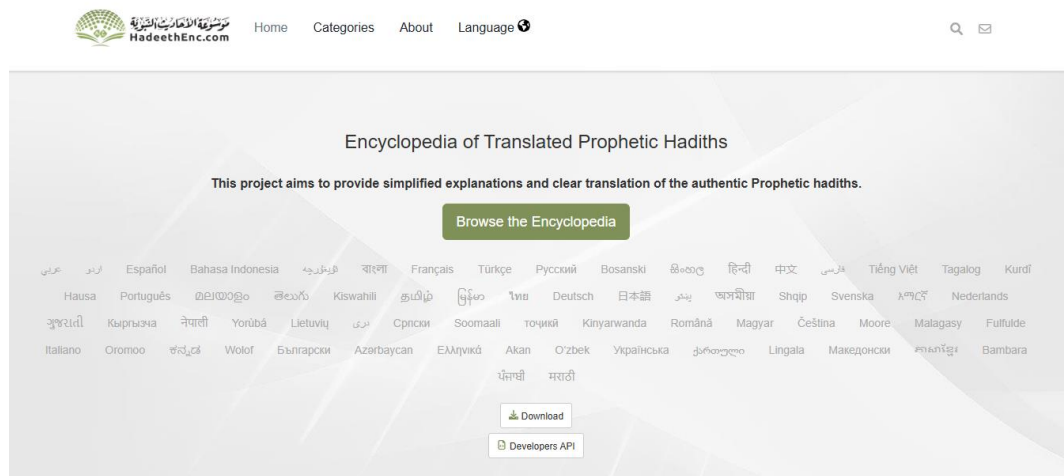
**Tabel 1.2 Rincian Dataset Hadis Terjemahan**

| Teks Hadis Terjemahan                                            |
|------------------------------------------------------------------|
| Jauhilah tujuh dosa yang membinasakan!                           |
| Timbanglah dan lebihkan!                                         |
| Ini Jibril menyampaikan salam kepadamu.                          |
| Agama itu nasihat (ketulusan)                                    |
| Janganlah engkau marah!                                          |
| Senantiasa lisanmu basah dengan zikir kepada Allah               |
| Makan sahurlah karena di dalam makanan sahur terdapat keberkahan |

### 1.4. Batasan Masalah

Batasan masalah dalam penelitian ini adalah data yang digunakan berupa teks terjemahan hadis dalam bahasa Indonesia yang diambil dari situs resmi *HadeethEnc.com* dan telah diberi label tema utama serta subtema secara manual

[3]. Proses klasifikasi dilakukan secara hierarkis pada dua level, yaitu tema utama dan subtema spesifik, tanpa mencakup klasifikasi sanad, autentisitas, atau penentuan hukum keabsahan hadis [4]. Metode yang digunakan terbatas pada pendekatan deep learning dengan arsitektur Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), dan Convolutional Recurrent Neural Network (CRNN) tanpa melibatkan metode pembelajaran mesin konvensional atau *pre-trained contextual embeddings* seperti BERT. Evaluasi performa model hanya dilakukan menggunakan metrik precision, recall, F1-score, dan confidence pada dataset yang dibagi menjadi data latih dan data uji tanpa *cross-domain testing*. Implementasi sistem dibatasi pada tahap perancangan dan pengujian model klasifikasi hierarkis tanpa pengembangan aplikasi end-user, integrasi ke sistem pencarian daring, atau pembuatan antarmuka pengguna grafis.



**Gambar 1 Homepage HadeethEnc.com [3]**

Hadis-hadis yang diperoleh dari ensiklopedi tersebut telah diklasifikasikan sebelumnya oleh para ahli hadis berdasarkan topik-topik tertentu, sebagaimana ditampilkan pada Tabel 1.3.

**Tabel 1.3 Daftar Klasifikasi Hadis**

| <b>No</b> | <b>Kalsifikasi Hadis</b>     |
|-----------|------------------------------|
| 1.        | Akidah                       |
| 2.        | Fikih dan <i>ushul</i> fikih |
| 3.        | Kisah dan sejarah            |
| 4.        | Al-Qur'an dan ilmunya        |
| 5.        | Hadis dan ilmu hadis         |
| 6.        | Dakwah dan <i>Hisbah</i>     |
| 7.        | Keutamaan dan Adab           |

Penelitian ini akan membahas penerapan algoritma RNN, CNN, dan CRNN dalam mengidentifikasi serta mengelompokkan hadis terjemahan bahasa Indonesia secara hierarkis pada dua level, yaitu level pertama berupa tema utama yang terdiri atas empat topik besar yaitu akidah, fikih, kisah, dan adab dan level kedua berupa subtema yang lebih spesifik di dalam masing-masing topik. Selain itu, penelitian ini juga akan membandingkan tingkat akurasi ketiga algoritma tersebut serta membahas tantangan yang dihadapi dalam klasifikasi dua level ini, seperti ketidakseimbangan distribusi data dan keragaman struktur bahasa pada teks terjemahan yang berasal dari bahasa Arab sebagai bahasa utama literatur keislaman distribusi data dan kompleksitas bahasa Arab sebagai bahasa utama literatur keislaman

### **1.5. Metode Penelitian**

Metode penelitian ini dimulai dengan tahap pengumpulan data, di mana teks terjemahan hadis dalam bahasa Indonesia diambil dari situs resmi HadeethEnc.com dan telah dilabeli secara manual berdasarkan dua level klasifikasi, yaitu tema utama dan subtema spesifik. Selanjutnya dilakukan tahap pra-pemrosesan data yang mencakup case folding untuk mengubah seluruh teks menjadi huruf kecil, tokenization untuk memecah teks menjadi kata-kata, stopword removal untuk menghapus kata-kata yang tidak memiliki makna signifikan, serta pemberian label hierarkis yang sesuai. Setelah data siap, dilakukan perancangan model deep learning dengan tiga arsitektur berbeda, yaitu Recurrent Neural Network (RNN),

Convolutional Neural Network (CNN), dan Convolutional Recurrent Neural Network (CRNN), yang masing-masing dilatih menggunakan dataset yang telah diproses. Proses pelatihan model dilakukan dengan parameter yang sama, seperti jumlah epoch, batch size, dan penerapan teknik early stopping untuk mencegah overfitting. Tahap berikutnya adalah evaluasi performa model menggunakan metrik precision, recall, F1-score, dan confidence, yang dilakukan pada dua level klasifikasi untuk menilai kemampuan model dalam memprediksi tema utama dan subtema. Hasil evaluasi kemudian dibandingkan untuk menentukan model terbaik, yang selanjutnya dianalisis untuk melihat keunggulan dan kelemahannya. Terakhir, penelitian ini ditutup dengan penyusunan kesimpulan dan saran untuk pengembangan lebih lanjut, termasuk rekomendasi penggunaan teknik embedding berbasis konteks seperti BERT, penambahan variasi data, dan penerapan attention mechanism guna meningkatkan kinerja model pada klasifikasi hierarkis hadis. Implementasi penelitian dilakukan menggunakan bahasa pemrograman *Python* pada Google Colab dengan menggunakan framework TensorFlow serta *library* tambahan seperti Pandas, NumPy, Matplotlib, dan Scikit-learn digunakan untuk analisis data dan evaluasi model. Setelah eksperimen selesai, hasil dibandingkan untuk menentukan model yang lebih unggul dalam klasifikasi hierarkis.

### 1.6. Jadwal Pelaksanaan

Rencana jadwal kegiatan yang telah disusun dapat dilihat pada Tabel 1.4.

**Tabel 1.4 Rencana Jadwal Pelaksanaan Tugas Akhir.**

| Kegiatan                                  | Bulan           |                  |               |               |             |              |
|-------------------------------------------|-----------------|------------------|---------------|---------------|-------------|--------------|
|                                           | Januari<br>2025 | Februari<br>2025 | Maret<br>2025 | April<br>2025 | Mei<br>2025 | Juni<br>2025 |
| Persiapan dan Studi Literatur             |                 |                  |               |               |             |              |
| Perancangan Model                         |                 |                  |               |               |             |              |
| Implementasi Model                        |                 |                  |               |               |             |              |
| Eksperimen dan Pengujian                  |                 |                  |               |               |             |              |
| Analisis Hasil dan<br>Penyempurnaan Model |                 |                  |               |               |             |              |
| Penulisan Laporan                         |                 |                  |               |               |             |              |