

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Umat Islam yang memiliki semangat dalam mempelajari ilmu agama membutuhkan teknologi untuk mengidentifikasi ayat Al-Qur'an. Al-Qur'an terdiri dari surah dan ayat, sehingga ayat ini dapat dikenali melalui surah dan nomor ayat tersebut. Dengan demikian, seseorang dapat mengetahui letak surah dan ayat di Al-Qur'an dengan memanfaatkan teknologi identifikasi ayat. Misalnya, seperti yang disebutkan dalam kitab hadis Sahih Muslim, Ayat Kursi adalah ayat paling utama dalam Al-Qur'an dengan banyak keutamaan dan manfaat membacanya. Oleh karena itu, sudah seharusnya seorang Muslim menghafal ayat tersebut. Namun, terkadang ada Muslim yang tidak mengetahui letak surah dan ayat dari Ayat Kursi, meskipun mereka sudah hafal dan disarankan untuk membacanya. Oleh karena itu, ada kebutuhan akan teknologi yang mengidentifikasi surah dan ayat Al-Qur'an dalam bentuk suara.

Dalam konteks studi Al-Qur'an, para penuntut ilmu agama yang memiliki semangat belajar dan ingin memperdalam pemahaman tentang ayat-ayat Al-Qur'an, dapat memanfaatkan teknologi untuk mengidentifikasi ayat-ayat tertentu. Hal ini memungkinkan mereka untuk menemukan letak ayat dan surah dengan mudah, yang sangat penting untuk memahami makna ayat secara utuh. Selain itu, teknologi juga dapat membantu dalam pelafalan yang benar dari Al-Qur'an, mengingat adanya aturan dan teknik khusus dalam membacanya.

Al-Qur'an adalah kitab suci umat Islam yang ditulis dalam bahasa Arab dan sangat penting bagi kehidupan. Terdiri dari 114 surah dan 6.236 ayat, Al-Qur'an terbagi menjadi 30 juz. Allah menurunkan Al-Qur'an melalui rasul-Nya dan tetap terjaga sanad-nya baik tulisan maupun cara membacanya oleh para imam qiraah. Ilmu membaca Al-Qur'an disebut Hukum tajwid dan pembacaannya hampir sama

antar imam qiraah, meskipun mungkin ada perbedaan dalam 1 atau 2 huruf. Fitur emosional dan transisi nada juga membedakan cara membaca Al-Qur'an [2].

Namun, tidak hanya isi teksnya saja yang penting, tetapi juga teknologi yang dapat membantu mempermudah akses dan Pemahaman terhadap Al-Qur'an. Sebagai contoh, suara manusia saat membaca ayat-ayat Al-Qur'an dapat di-konversi menjadi sinyal digital dan di-ekstraksi fiturnya untuk tujuan identifikasi ayat-ayat tersebut.

Pada dasarnya, suara manusia saat membaca Al-Qur'an dapat diubah menjadi sinyal digital, yang mewakili setiap level sinyal pada sinyal waktu diskret [3]. Oleh karena itu, ekstraksi fitur harus dilakukan pada suara Untuk mengidentifikasi ayat-ayat Al-Qur'an, membuat proses ini menjadi sangat bermanfaat.

Ekstraksi fitur adalah proses mengubah suara menjadi nilai digital. Fitur yang umumnya digunakan dalam *Speech Recognition* adalah algoritma MFCC (*Mel Frequency Cepstral Coefficients*). MFCC menangkap bentuk kasar saluran vokal atau rongga mulut yang terkait dengan produksi suara [4]. MFCC dapat mengubah *file* gelombang menjadi vektor fitur, yang kemudian dapat digunakan dalam proses *machine learning*.

Dalam beberapa tahun terakhir, algoritma *machine learning* semakin populer dalam bidang penelitian *music information retrieval* (MIR) [5]. *Artificial Neural Network* adalah algoritma *machine learning* yang dapat mengenali pola data tersembunyi. Proses membangun model dikenal sebagai *training*, dan mengevaluasi kinerja model adalah proses *testing*. Algoritma *machine learning* ini umumnya *overfitting* atau *overtraining*, yaitu model bekerja sangat baik pada data *training*, tetapi buruk pada data *testing* atau kasus yang tidak ter-amati. Oleh karena itu, perlu hati-hati saat menggunakan algoritma ini agar komputer dapat mempelajari pola *input* dan menentukan *output* tertentu.

Pada hakikatnya, suara Al-Qur'an adalah sinyal audio. Oleh karena itu, permasalahan mengidentifikasi ayat Al-Qur'an dapat diselesaikan melalui pendekatan yang sama seperti *Music Information Retrieval* (MIR). Setiap muslim

yang membaca Al-Qur'an mungkin memiliki pendekatan yang berbeda untuk membaca ayat-ayat tersebut, seperti tajwid, qiraah, fitur emosional, dan transisi nada. Karena jumlah ayat yang sangat banyak, ayat-ayat Al-Qur'an sulit untuk diidentifikasi dari satu ayat ke ayat lain menggunakan suara pembaca. Oleh karena itu, sangat penting untuk melakukan penelitian tentang klasifikasi surah dan ayat Al-Qur'an.

Dalam penelitian ini, pendekatan akan menggunakan algoritma *Artificial Neural Network* untuk mengklasifikasi surah dan ayat Al-Qur'an. Penelitian oleh [6] melakukan klasifikasi surah terlebih dahulu, kemudian mengklasifikasi ayat, sementara penelitian lain menggunakan model berbasis *Siamese LSTM (Long Short-Term Memory)* [7] di mana kedua penelitian tersebut menggunakan *dataset* yang diambil dari everyayah.com. Al-Qur'an diciptakan dalam bahasa Arab, sehingga pendekatan *Natural Language Processing (NLP)*, seperti *Automatic Speech Recognition (ASR)* [8], [9], [10].

Sebagai tambahan, setelah proses pembelajaran selesai, model akan melakukan *testing* pada data yang belum dilihat sebelumnya untuk memastikan kemampuannya dalam mengidentifikasi ayat Al-Qur'an. Jika hasil *testing* menunjukkan hasil yang baik, maka model dapat digunakan untuk mengidentifikasi ayat Al-Qur'an yang dibacakan oleh qari. Dengan menggunakan ANN, model dapat mempelajari pola-pola yang terkait dengan cara qari membaca Al-Qur'an sehingga dapat menentukan surah dan ayat yang sesuai.

Dengan demikian, tujuan skripsi ini adalah membahas identifikasi ayat-ayat Al-Qur'an melalui suara dengan menggunakan metode *Artificial Neural Network* dan ekstraksi fitur menggunakan algoritma *Mel Frequency Cepstral Coefficient (MFCC)*. Penelitian ini membatasi penggunaan data Al-Qur'an sebanyak 7 surah dan 32 ayat serta 31 *dataset* qari dengan audio *bitrate* lebih besar dari 128 kbps dan *sampling rate* 22050. Skenario percobaan dilakukan untuk memilih model ANN dengan *hyperparameter* terbaik untuk klasifikasi audio dan *validation* model dengan data *test* yang berbeda-beda. Penelitian ini menggunakan metode penelitian kuantitatif dengan pendekatan *machine learning supervised learning* dan evaluasi dilakukan dengan menggunakan metrik-metrik

seperti *f-measure*, *precision*, *recall*, *ROC-AUC*, dan *akurasi* serta melihat *confusion matrix*-nya. Hasil penelitian ini dapat dimanfaatkan untuk mempermudah proses pemahaman dan *validation* ayat Al-Qur'an bagi umat muslim. Ini sangat berguna bagi para qari atau pembaca Al-Qur'an karena membantu memastikan bahwa ayat yang dibacakan sesuai dengan Al-Qur'an yang asli.

1.2 Rumusan Masalah

Masalah utama yang dicakup dalam skripsi ini adalah mengidentifikasi ayat-ayat dan surah Al-Qur'an melalui suara. Untuk memecahkan masalah tersebut, tiga hal utama harus dilakukan:

1. Melakukan *preprocessing* pada data suara sebelum *training* yaitu dengan mengaplikasikan algoritma MFCC untuk ekstraksi fitur sebagai *input* dari model *Artificial Neural Network*, sehingga model dapat mengidentifikasi ayat-ayat dan surah dalam Al-Qur'an melalui suara.
2. Melakukan analisis studi kasus pada skenario percobaan untuk memilih model *Artificial Neural Network* yang optimal yang dapat mengidentifikasi ayat-ayat Al-Qur'an dengan baik, dengan mempertimbangkan faktor-faktor seperti data MFCC, *hyperparameter*, dan arsitektur.
3. Menguji validitas model dengan data *test* yang terdiri dari 6 macam qari dan kondisi suara yang berbeda-beda, untuk menguji sejauh mana model dapat mengidentifikasi indeks letak surah dan ayat di Al-Qur'an dari suara yang dibacakan.

1.3 Batasan Masalah

Masalah ini dibatasi pada hal-hal berikut:

1. Metode yang digunakan adalah ANN.
2. Ekstraksi fitur yang digunakan adalah MFCC yang sudah di normalisasi dengan metode *Cepstral Mean Variance Normalization* (CMVN)
3. Untuk *dataset* skenario percobaan

- Data Al-Qur'an yang digunakan adalah 7 surah dan 32 ayat, yaitu dari surah Al-Kautsar hingga An-Nas, sehingga klasifikasi ayat yang digunakan hanya 32 kelas
 - Terdapat 31 *dataset* qari dengan waktu yang bervariasi dan audio *bitrate* lebih besar dari 128 kbps.
 - Ukuran total *dataset* adalah 996 audio.
 - *Dataset* seimbang antar kelas.
 - *Dataset* dipecah menjadi tiga bagian: data *train*, data *validation*, dan data *test*.
 - Proporsi *dataset* yang digunakan untuk setiap bagian adalah 80:20 untuk data *train* dan data *test*. Untuk data *train* yang 80 akan dipecah menjadi 70:30 untuk data *train* dan data *validation*
 - Pemecahan *dataset* dilakukan secara acak dan *stratification* dengan menggunakan *library train_test_split* dari *library scikit-learn*.
4. Untuk *preprocessing* skenario percobaan
- Skripsi ini hanya menggunakan *sampling rate* 22050 untuk semua data suara yang digunakan dalam percobaan.
 - Skripsi ini menggunakan konfigurasi bawaan dari MFCC pada *library librosa*, dengan jumlah sampel untuk setiap *frame* STFT (n_fft) adalah 2048, *hop length* (jumlah sampel perpindahan antara setiap *frame* STFT) sebesar $\frac{FFT}{4} = 512$, jumlah *mel filterbank* yang digunakan dalam ekstraksi fitur MFCC (n_mels) sebesar 128, menggunakan *hann window* pada STFT dan *frame STFT* diatur pada titik tengah window.
 - Hanya digunakan 20 koefisien MFCC sebagai fitur *input* dalam model *Artificial Neural Network*.
 - Skripsi ini hanya melakukan *padding* (penambahan) dan *truncate* (pemotongan) pada skenario percobaan dengan menggunakan batasan panjang dari MFCC ($mfcc_len$), yaitu 129.1 *frame* untuk audio selama 3 detik, 453.5 *frame* untuk audio selama 10.5 detik, dan 774.8 *frame* untuk audio selama 18 detik. Sehingga skripsi ini hanya menguji validitas model dengan data *test* yang terdiri dari tiga jenis audio,

yaitu audio yang paling pendek, audio yang paling panjang, dan rata-rata dari audio tersebut [11].

- Dalam melakukan perhitungan panjang maksimal *frame* dari MFCC, digunakan fungsi sebagai berikut rumus

$$\text{max_len} = \left\lfloor \frac{\text{durasi_dari_audio} \times \text{sampling_rate} - \text{n_fft} + \text{hop_length}}{\text{hop_length}} \right\rfloor$$

5. Untuk arsitektur skenario percobaan

- Dilakukan untuk melakukan pemilihan *hyperparameter* terbaik untuk model *deep learning* yang digunakan untuk klasifikasi audio.
- *Hyperparameter* yang dipilih meliputi *learning rate*, *batch size*, jumlah *hidden layer* dan besarnya *hidden layer* (arsitektur), dan panjang MFCC.
- Lapisan *input* adalah MFCC yang sudah di normalisasi dengan MCVN.
- Model yang digunakan adalah *sequential model* dalam *library keras tensorflow* yang sudah dilakukan *batch normalization* kemudian menggunakan fungsi aktivasi ReLU di tiap layernya dan fungsi aktivasi *softmax* pada *output* lapisan.
- Menggunakan *Adam* sebagai algoritma optimisasi untuk mengoptimalkan model *neural network*.
- Menggunakan *sparse_categorical_crossentropy* sebagai fungsi *loss* untuk masalah klasifikasi *multiclass* dengan label indeks sebagai ayat.
- Menggunakan teknik *early stopping* untuk mencegah *overfitting* pada model yang dibangun. Teknik ini akan menghentikan *training* model jika nilai *val_loss* tidak menurun selama 10 *epoch* berturut-turut. Monitor yang digunakan adalah *val_loss* dan mode yang digunakan adalah *min* (untuk meminimalkan *val_loss*).

6. Variabel yang diperhatikan adalah *f-measure*, *precision*, *recall*, *Receiver Operating Characteristic Area Under the Curve (ROC-AUC)* terakhir, dan *akurasi* dari data *test*.

1.4 Tujuan dan Manfaat Penelitian

Tujuan utama dalam skripsi ini adalah membuat model yang dapat mengidentifikasi ayat-ayat dan surah dalam Al-Qur'an melalui suara. Model tersebut akan mengidentifikasi indeks letak surah dan ayat di Al-Qur'an dari suara yang dibacakan. Adapun tiga uraian tujuan yang akan dilakukan dalam tugas skripsi ini adalah:

1. Melakukan *preprocessing* sebelum *training* yaitu mengaplikasikan algoritma *Mel Frequency Cepstral Coefficient* (MFCC) untuk ekstraksi fitur sebagai *input* dari model *Artificial Neural Network*.
2. Menganalisis studi kasus pada skenario percobaan, yaitu data MFCC, *hyperparameter*, dan arsitektur untuk memilih model *Artificial Neural Network* yang optimal yang dapat mengidentifikasi ayat-ayat Al-Qur'an dengan baik.
3. Menguji validitas model yang sudah di-*training* dengan data *train* dan *validation* dengan data *test*.

1.5 Metode Penelitian

Metodologi penelitian yang digunakan dalam skripsi ini adalah penelitian kuantitatif dengan pendekatan permasalahan *machine learning supervised learning* untuk klasifikasi data audio (pengumpulan data, *Preprocessing*, *training* model, dan evaluasi). Langkah-langkah yang akan dilakukan adalah sebagai berikut:

1. Studi Literatur dan Pengumpulan Data: Tahap ini melibatkan pengumpulan literatur, jurnal, paper, dan bahan bacaan yang berkaitan dengan penelitian menggunakan *Artificial Neural Network* dan *Mel Frequency Cepstral Coefficients* (MFCC). Selain itu, juga dilakukan pengambilan data suara Al-Qur'an dari sumber Everyayah.com.
2. *Preprocessing Data*: Data suara Al-Qur'an diubah menjadi representasi waktu-frekuensi dengan menggunakan metode MFCC (*Mel Frequency*

Cepstral Coefficients). Metode ini dapat mengekstrak ciri-ciri penting dari gelombang suara seperti amplitudo, frekuensi, dan *phase*.

3. Simulasi *training* dan *testing*: Dilakukan simulasi untuk membangun model ANN yang dapat mengidentifikasi surah dan ayat Al-Qur'an sesuai skenario percobaan. Pada tahap ini, model ANN dibangun dengan menggunakan data *train* dan data *validation* yang dihasilkan dari pembagian *dataset* menggunakan fungsi *train_test_split* dari *library scikit-learn*. Selanjutnya, akan dilakukan evaluasi studi kasus dengan menggunakan data *test* pada seluruh model kemudian akan dipilih model yang optimal untuk dianalisis hasilnya untuk mengetahui performa model yang dibangun berdasarkan metrik-metrik yang telah ditentukan sebelumnya, seperti *f-measure*, *precision*, *recall*, *ROC-AUC*, dan *akurasi*, dan dapat melihat *confusion matrix*-nya dari data *test*.

1.6 Sistematika Penulisan

Sistematika penulisan pada skripsi ini terdiri dari lima bab dan di dalam setiap bab terdiri dari beberapa subbab.

BAB 1 : PENDAHULUAN

Pada bab ini akan dipaparkan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, serta sistematika penulisan skripsi.

BAB 2 : LANDASAN TEORI

Bab ini berisi penjelasan teori-teori yang berkaitan dengan topik yang akan dibahas.

BAB 3 : KLASIFIKASI AYAT AL-QUR'AN DENGAN *MEL FREQUENCY CEPSTRAL COEFFICIENT* DAN *ARTIFICIAL NEURAL NETWORK*

Pada bab ini akan dijelaskan mengenai teoritis dan analisis menggunakan metodologi dalam klasifikasi ayat Al-Qur'an dengan menggunakan *Mel Frequency Cepstral Coefficient* (MFCC) dan *Artificial Neural Network* (ANN). Untuk klasifikasi audio menggunakan *supervised learning* untuk data audio, Metodologi yang terdiri dari enam langkah akan dilakukan. *Pengumpulan Data*

yaitu mengumpulkan dan menyimpan data audio yang relevan *Ekstraksi Fitur* yaitu mengubah data audio menjadi vektor fitur yang merepresentasikan karakteristik suara sesuai skenario percobaan yaitu *Mel-Frequency Cepstral Coefficients* (MFCC). *Preprocessing* yaitu membersihkan dan menormalkan data fitur untuk mengurangi *noise* dan meningkatkan kualitas data, sesuai skenario percobaan yaitu dengan melakukan CMVN, memotong dan menambahkan audio sesuai panjang maksimal MFCC. *Pemilihan Model* yaitu memilih arsitektur model yang sesuai dengan tujuan skenario klasifikasi, yaitu menggunakan model ANN. *Training Model* yaitu melatih model dengan data fitur dan label kelas yang tersedia menggunakan algoritma optimasi dan fungsi *loss* sesuai skenario percobaan. *Evaluasi Model* yaitu menguji kinerja model dengan data fitur dan label kelas yang belum pernah dilihat sebelumnya menggunakan metrik evaluasi sesuai skenario percobaan yaitu *akurasi, precision, recall, ROC-AUC* terakhir, dan *f-measure*. Jika hasil evaluasi model belum memuaskan, kita dapat melakukan langkah tambahan yaitu *Optimasi Model* yaitu meningkatkan kinerja model dengan menyesuaikan parameter atau *hyperparameter* dengan menggunakan *sweep Weight and Biases* (WandB)

BAB 4 : STUDI KASUS DAN ANALISIS KLASIFIKASI AYAT AL-QUR'AN DENGAN MEL FREQUENCY CEPSTRAL COEFFICIENT DAN ARTIFICIAL NEURAL NETWORK

Bab 4 akan membahas studi kasus dan analisis dari klasifikasi ayat Al-Qur'an menggunakan MFCC dan ANN. Terdapat beberapa subbab yang akan dijelaskan di dalamnya, seperti analisis hasil klasifikasi ayat Al-Qur'an, evaluasi performa model klasifikasi, analisis *confusion matrix*, dan analisis kesalahan klasifikasi. Selanjutnya, pada subbab diskusi hasil, akan dibahas mengenai pembahasan hasil klasifikasi, perbandingan hasil dengan penelitian terkait, serta interpretasi hasil dan kesimpulan yang dapat diambil dari penelitian ini.

BAB 5 : PENUTUP

Bab 5 sebagai penutup, akan membahas kesimpulan dari penelitian dan saran-saran untuk penelitian selanjutnya. Pada bab ini, akan dipaparkan

kesimpulan sebagai jawaban dari rumusan masalah yang diajukan, serta saran yang berisi rekomendasi untuk penelitian selanjutnya pada topik yang serupa.

