



Artificial Intelligence Automatic Speech Recognition (ASR) untuk pencarian potongan ayat Al-Qu'ran

Salamun¹, Sukri², Khairul Amin³, Luluk Elvitaria⁴, Liza Trisnawati⁵

¹Universitas Abdurrah, Teknik Informatika, email: salamun@univrab.ac.id

²Universitas Abdurrah, Teknik Informatika, email: sukri@univrab.ac.id

³Universitas Abdurrah, Teknik Informatika, email: amin@gmail.com

⁴Universitas Abdurrah, Teknik Informatika, email: luluk@univrab.ac.id

⁵Universitas Abdurrah, Teknik Informatika, email: liza.trisnawati@univrab.ac.id

[1] Abstrak

Indonesia merupakan negara dengan jumlah umat muslim terbesar di dunia, yang menjadikan pembacaan ayat-ayat Al-Qur'an sering terdengar di berbagai tempat-tempat umum seperti Mesjid, Mushollah, dan di berbagai kegiatan. Pemanfaatan Automatic Speech Recognition (ASR) sebagai pengenalan kata yang bertujuan untuk mengetahui ayat-ayat Al-Qur'an yang di bacakan untuk menambah pengetahuan mengenai ayat-ayat serta informasi pendukung lainnya sebagai salah satu sarana berdakwah dalam menyampaikan pengetahuan mengenai ayat-ayat Al-Qur'an. Automatic Speech Recognitions (ASR) ini dirancang menggunakan bahasa pemrograman Python dan menggunakan framework Django untuk menampilkan informasi mengenai ayat-ayat yang di bacakan dalam bentuk tampilan berbasis web. Penelitian ini bertujuan untuk menciptakan sebuah teknik dan sistem untuk memasukkan perintah suara ke dalam mesin, agar mesin dapat mengerti apa yang manusia ucapkan dan mematuhi apa yang diperintahkannya. Aplikasi ini mengubah data suara menjadi data text menggunakan sistem pengenalan suara yang bekerja secara otomatis dengan pencocokan pola didigitalkan audio kata yang diucapkan terhadap model komputer dari pola bicara untuk menghasilkan keluaran akhir berupa teks yang di simpan di dalam database.

Kata kunci: Automation Speech Recognition, ASR, Al-Qur'an, Python, Speech Recognition, Django

[2] Abstract

Indonesia is a country with the largest number of Muslims in the world, who read verses of the Qur'an often heard in various public places such as mosques, prayer rooms, and at various activities. Utilization of Automatic Speech Recognition (ASR) as word recognition which aims to find out the verses of the Qur'an that are read to increase knowledge about the verses and other supporting information as a means of preaching in conveying knowledge about the verses of Al-Qur'an. Al-Qur'an. Automatic Speech Recognitions (ASR) is designed using the Python programming language and uses the Django framework to display information about the verses that are read in the form of a web-based display. This research aims to create a technique and system for entering voice commands into machines, so that machines can understand what humans say and obey what they are told. This application converts data into text data using a voice recognition system automatically with a digital audio pattern of spoken words from a speech pattern computer model to produce the final output in the form of text that is stored in the database.

Keywords: Automation Speech Recognition, ASR, Al-Qur'an, Python, Speech Recognition, Django

1. Pendahuluan

Al-Qur'an adalah sumber ajaran dan pedoman hidup umat Islam yang pertama, kitab suci ini menempati posisi sentral dalam segala hal yaitu dalam pengembangan dan perkembangan ilmu pengetahuan dan keislaman. Dewasa ini banyak sekali berkembang berbagai macam metode dalam pembelajaran Al-Qur'an, dan kesemuanya itu memiliki kekurangan dan kelebihan. Salah satu metode yang sering dilakukan dalam pembelajaran Al-Qur'an yaitu metode talaqqi atau yang sering disebut dengan metode sorogan[4]. Penelitian ini bukanlah hal yang baru, terdapat beberapa penelitian sejenis sebagai rujukan bagi peneliti untuk mengembangkan metode Automatic Speech Recognition, diantaranya yaitu penelitian yang dilakukan oleh Nizhny Novgorod State Linguistic University, Nizhny Novgorod, Russia pada tahun 2020 dengan judul Minimum of Information Divergence Criterion for Signals with Tuning to Speaker Voice in Automatic Speech Recognition. Penelitian ini mengembangkan perangkat lunak yang dapat memecahkan masalah pengenalan ucapan otomatis pada tingkat bicara fonetis dasar pemrosesan sinyal[5]. Kemudian penelitian yang dilakukan oleh Shahid Sattari dengan judul Adaptive Windows Multiple Deep Residual Networks For Speech Recognition. Penelitian ini berfokus pada bagaimana cara pembelajaran residual baru, yang mengarah pada pemanfaatan informasi yang lebih baik di lapisan dalam dan memberikan kontrol yang lebih baik pada transfer informasi input[6]. Speech (wicara) adalah modalitas utama komunikasi dengan cara yang alami dan paling efektif untuk bertukar informasi antar manusia. Cara pengungkapan suara itu sendiri dapat diwujudkan dengan sinyal atau bentuk gelombang yang membawa pesan atau informasi. Sinyal suara dikirim, disimpan, dan diproses dengan berbagai cara dalam sistem komunikasi[7]. Automatic Speech Recognition (ASR) ini adalah proses menafsirkan ucapan manusia di komputer. Sistem pengenalan ucapan otomatis yang ideal akan memproses audio ucapan menjadi teks dengan sedikit kesalahan dan mengubah ucapan kata demi kata. Namun karena berbagai faktor, termasuk sebagian besar faktor dalam bicara manusia normal, perubahan kata demi kata bisa dilihat dari gambar sistem Automatic Speech Recognition (ASR) di bawah ini[8]. Sistem Automatic Speech Recognition(ASR) memberikan kemampuan untuk mengkonversi ucapan dalam kata-kata yang dimengerti. Adapun jenis – jenis speech recognition Berdasarkan kemampuan dalam mengenal kata yang diucapkan, terdapat 5 jenis pengenalan kata[9], yaitu: 1.

Kata-kata yang terisolasi : Proses pengidentifikasian kata yang hanya dapat mengenal kata yang diucapkan jika kata tersebut memiliki jeda waktu pengucapan antar kata. 2. Kata-kata yang berhubungan : Proses pengidentifikasian kata yang mirip dengan kata-kata terisolasi, namun membutuhkan jeda waktu pengucapan antar kata yang lebih singkat. 3. Kata-kata yang berkelanjutan: Proses pengidentifikasian kata yang sudah lebih maju karena dapat mengenal kata-kata yang diucapkan secara berkesinambungan dengan jeda waktu yang sangat sedikit atau tanpa jeda waktu. Proses pengenalan suara ini sangat rumit karena membutuhkan metode khusus untuk membedakan kata-kata yang diucapkan tanpa jeda waktu. Pengguna perangkat ini dapat mengucapkan kata-kata secara natural. 4. Kata-kata spontan : Proses pengidentifikasian kata yang dapat mengenal kata-kata yang diucapkan secara spontan tanpa jeda waktu antar kata. Verifikasi atau identifikasi suara : Proses pengidentifikasian kata yang tidak hanya mampu mengenal kata, namun juga mengidentifikasi siapa yang berbicara.

Al-Qur'an adalah sumber ajaran dan pedoman hidup umat Islam yang pertama, kitab suci ini menempati posisi sentral dalam segala hal yaitu dalam pengembangan dan perkembangan ilmu pengetahuan dan keislaman. Dewasa ini banyak sekali berkembang berbagai macam metode dalam pembelajaran Al-Qur'an, dan kesemuanya itu memiliki kekurangan dan kelebihan. Salah satu metode yang sering dilakukan dalam pembelajaran Al-Qur'an yaitu metode talaqqi atau yang sering disebut dengan metode sorogan[4].

Penelitian ini bukanlah hal yang baru, terdapat beberapa penelitian sejenis sebagai rujukan bagi peneliti untuk mengembangkan metode *Automatic Speech Recognition*, diantaranya

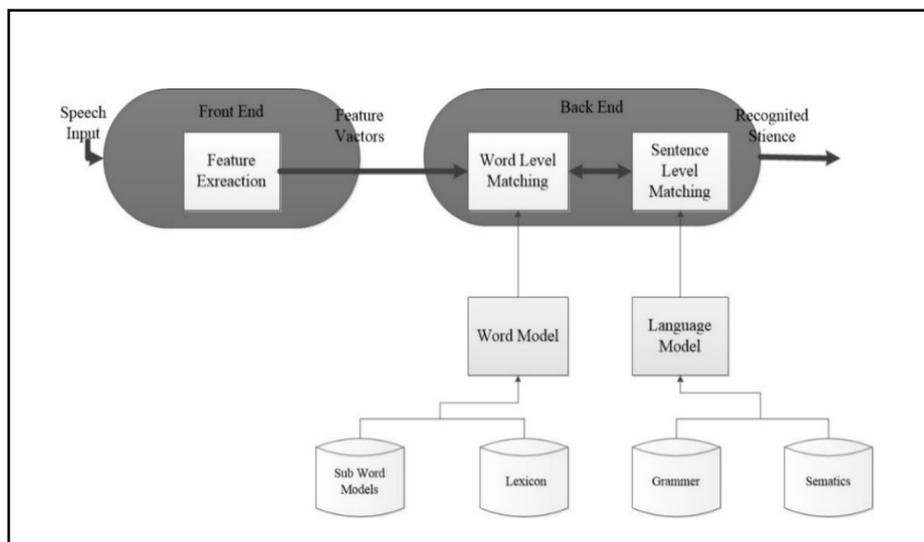
yaitu penelitian yang dilakukan oleh Nizhny Novgorod State Linguistic University, Nizhny Novgorod, Russia pada tahun 2020 dengan judul *Minimum of Information Divergence Criterion for Signals with Tuning to Speaker Voice in Automatic Speech Recognition*. Penelitian ini mengembangkan perangkat lunak yang dapat memecahkan masalah pengenalan ucapan otomatis pada tingkat bicara fonetis dasar pemrosesan sinyal[5].

Kemudian penelitian yang dilakukan oleh Shahid Sattari dengan judul *Adaptive Windows Multiple Deep Residual Networks For Speech Recognition*. Penelitian ini berfokus pada bagaimana cara pembelajaran residual baru, yang mengarah pada pemanfaatan informasi yang lebih baik di lapisan dalam dan memberikan kontrol yang lebih baik pada transfer informasi input[6].

2. Metode Penelitian

Kegiatan penelitian ini dimaksudkan untuk melakukan pengenalan suara yang mengucapkan potongan ayat yang dikeluarkan oleh suara yang dapat di deteksi oleh perangkat. Ayat yang di deteksi adalah potongan ayat yang sudah disimpan dalam basis data bukan keseluruhan ayat dalam satu surat. Potongan ayat tersebut akan di berikan pembobotan agar nantinya dapat di hitung dan di kelompokkan sesuai nama surat nya setelah pengelompokan mendapatkan hasil maka aplikasi akan menampilkan hasil akhir yaitu nama surat dari potongan ayat yang dideteksi dari suara yang ditangkap oleh perangkat.

Speech (wicara) adalah modalitas utama komunikasi dengan cara yang alami dan paling efektif untuk bertukar informasi antar manusia. Cara pengungkapan suara itu sendiri dapat diwujudkan dengan sinyal atau bentuk gelombang yang membawa pesan atau informasi. Sinyal suara dikirim, disimpan, dan diproses dengan berbagai cara dalam sistem komunikasi[7]. Automatic Speech Recognition (ASR) ini adalah proses menafsirkan ucapan manusia di komputer. Sistem pengenalan ucapan otomatis yang ideal akan memproses audio ucapan menjadi teks dengan sedikit kesalahan dan mengubah ucapan kata demi kata. Namun karena berbagai faktor, termasuk sebagian besar faktor dalam bicara manusia normal, perubahan kata demi kata bisa dilihat dari gambar sistem Automatic Speech Recognition (ASR) di bawah ini[8].



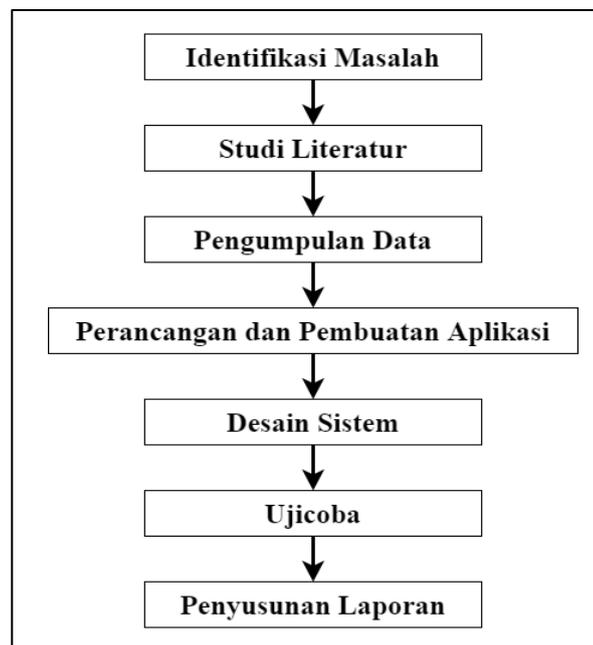
Gambar 1. Block diagram of Automatic Speech Recognition

Sistem *Automatic Speech Recognition*(ASR) memberikan kemampuan untuk mengkonversi ucapan dalam kata-kata yang dimengerti. Adapun jenis – jenis *speech recognition* Berdasarkan kemampuan dalam mengenal kata yang diucapkan, terdapat 5 jenis pengenalan kata[9], yaitu:

1. Kata-kata yang terisolasi : Proses pengidentifikasian kata yang hanya dapat mengenal kata yang diucapkan jika kata tersebut memiliki jeda waktu pengucapan antar kata.
2. Kata-kata yang berhubungan : Proses pengidentifikasian kata yang mirip dengan kata-kata terisolasi, namun membutuhkan jeda waktu pengucapan antar kata yang lebih singkat.
3. Kata-kata yang berkelanjutan: Proses pengidentifikasian kata yang sudah lebih maju karena dapat mengenal kata-kata yang diucapkan secara berkesinambungan dengan jeda waktu yang sangat sedikit atau tanpa jeda waktu. Proses pengenalan suara ini sangat rumit karena membutuhkan metode khusus untuk membedakan kata-kata yang diucapkan tanpa jeda waktu. Pengguna perangkat ini dapat mengucapkan kata-kata secara natural.
4. Kata-kata spontan : Proses pengidentifikasian kata yang dapat mengenal kata-kata yang diucapkan secara spontan tanpa jeda waktu antar kata.

Verifikasi atau identifikasi suara : Proses pengidentifikasian kata yang tidak hanya mampu mengenal kata, namun juga mengidentifikasi siapa yang berbicara.

Kerangka penelitian merefleksikan keterkaitan antara variabel yang diteliti dan menjadi pedoman dalam pemecahan masalah penelitian yang dituangkan dalam bentuk diagram alir dengan instruksi yang berkualitas. Dengan mendeskripsikan langkah-langkah yang dilakukan dalam penelitian berdasarkan kerangka ideologis maka ditentukan tahap akhir penelitian. Oleh karena itu, agar program penelitian dapat berjalan secara sistematis, maka peneliti menjabarkan berbagai langkah yang harus diikuti, seperti pada gambar di bawah ini.

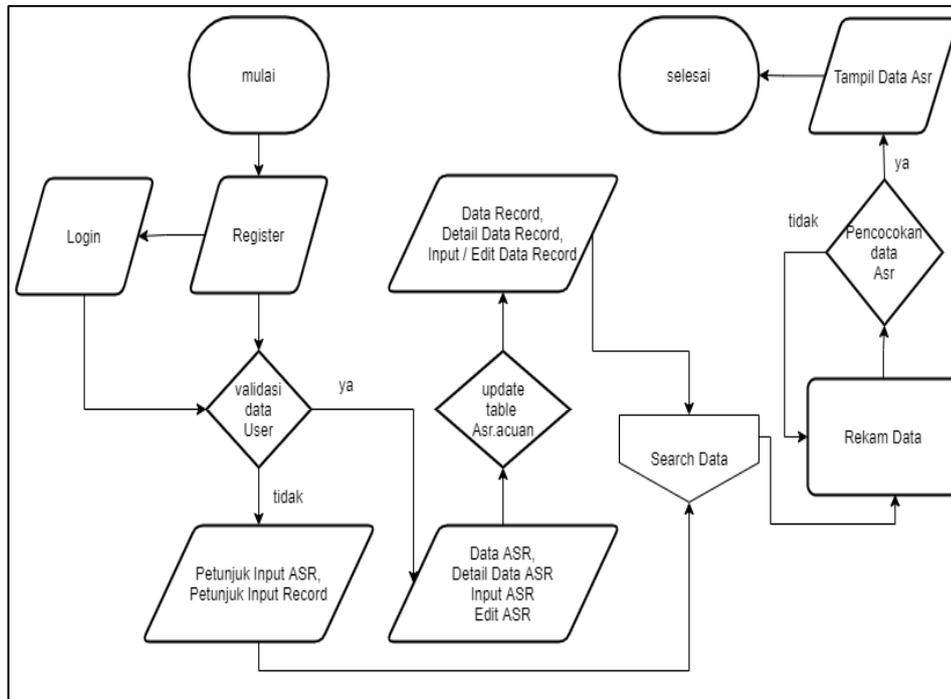


Gambar 2. Kerangka Penelitian

3. Hasil dan Pembahasan

Pada tahapan ini pembuatan flowchart system yang dilakukan setelah tahapan desain dan perencanaan. Analisis sistem yang sedang berjalan membantu untuk menyederhanakan aplikasi yang terperinci, selain itu, menguraikan informasi apa yang dapat dihasilkan oleh sistem yang dirancang. Untuk lebih rinci, selanjutnya akan dijelaskan masing-masing alur dari flowchart

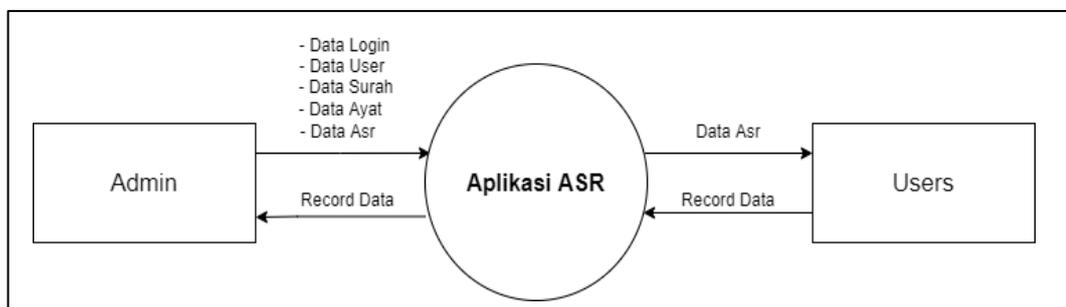
system yang berjalan pada aplikasi. Pada flowchart aplikasi ini menjelaskan aliran sistem dari aplikasi Automatic Speech Recognition (ASR) seperti Gambar 3 berikut:



Gambar 3. Flowchart System

3.1. Context Diagram

Dalam hasil penelitian dan perancangan aplikasi automatic speech recognitions (ASR), didapatkan beberapa informasi tentang proses-proses yang terjadi pada saat pengiriman data ke dalam database yang dikelompokkan berdasarkan hak akses. Hak akses tersebut di gambarkan pada gambar context diagram dibawah ini.

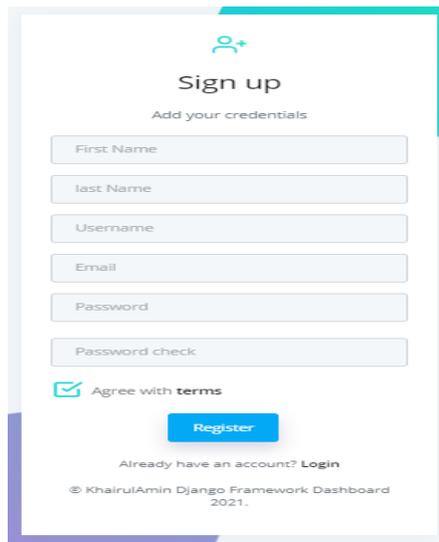


Gambar 4. Context Diagram Aplikasi

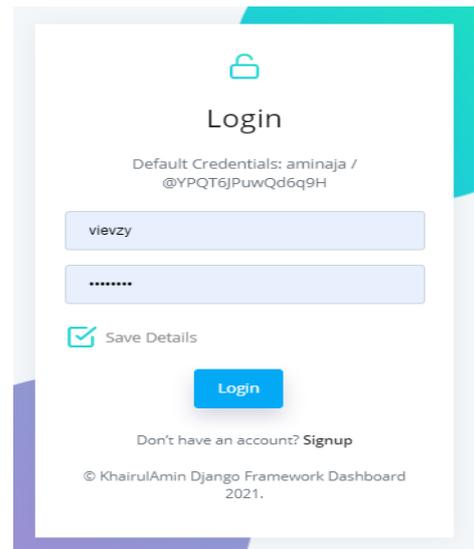
3.2. Pengujian Aplikasi

Admin merupakan hak akses untuk melakukan beberapa perintah di dalam aplikasi diantaranya menambahkan data Asr secara manual, melakukan record untuk memperbaharui data Asr, melihat data dari database Surah, Ayat, dan Asr. Tetapi tidak bisa melakukan penambahan data ke dalam

database Surah dan ayat, serta tidak bisa mengedit atau menghapus data User. Tetapi admin juga dapat melakukan keseluruhan yang dilakukan oleh Users. Register, merupakan halaman inputan data untuk mendapatkan ID agar bisa login kedalam sistem sebagai admin. Login merupakan halaman untuk masuk kedalam aplikasi dengan mengisi username dan password yang di dapat dari id User.

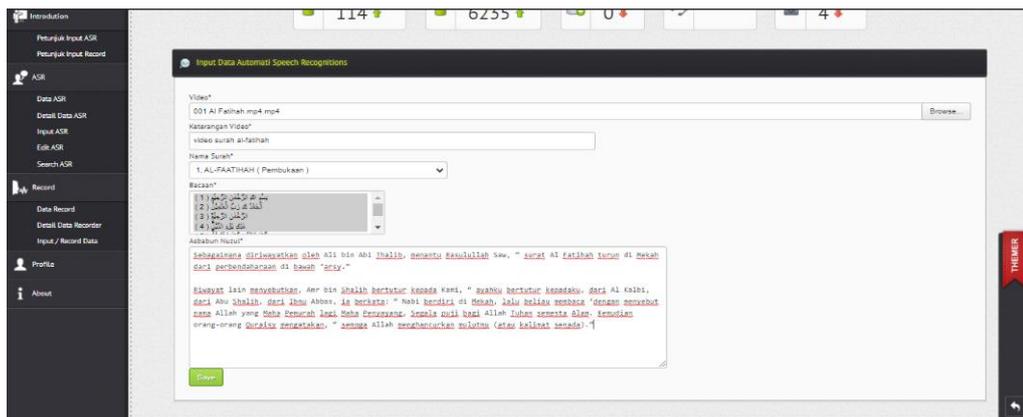


Gambar 5. Halaman Register



Gambar 6. Halaman Login

Input ASR, diminta untuk mengisi keseluruhan field yang tersedia sesuai dengan validasi tiap-tiap field.

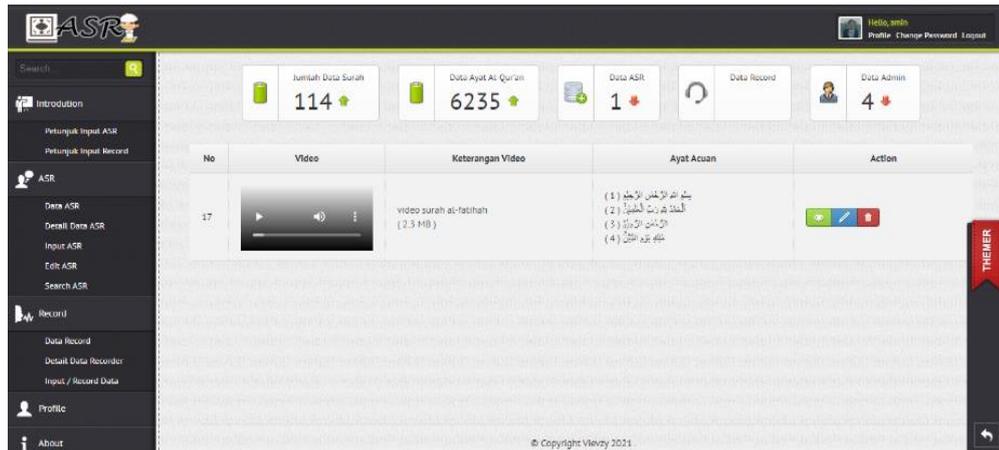


Gambar 7. Input ASR

Berdasarkan gambar 7 di atas terdapat 5 bagian untuk inputan data serta terdapat 1 tombol. Berikut adalah deskripsi yang terdapat pada tampilan Input ASR :

1. Video, merupakan upload file berupa video menggunakan validasi dengan extensi tertentu.
2. Keterangan Video, merupakan inputan sebagai informasi tambahan mengenai video yang di upload admin.

3. Nama Surah, merupakan inputan field dengan tampilan combobox dengan memilih salah satu dari pilihan yang tersedia.
4. Bacaan, merupakan tampilan inputan dengan multi select. Yaitu bisa memilih beberapa pilihan dari pilihan yang tersedia dengan cara menekan dan menahan tombol ctrl+pilih.
5. Asbabun Nuzul, merupakan tampilan inputan textarea dengan minimum data yang di input sebanyak 25 karakter.
6. Tombol Save, merupakan tombol eksekusi untuk mengirimkan data kemudian memvalidasi data dari aplikasi yang telah di inputkan kedalam database.



Gambar 8. Data ASR

Berdasarkan gambar 8 di atas terdapat 1 tampilan data ASR dalam bentuk table serta terdapat 3 tombol. Berikut adalah deskripsi yang terdapat pada tampilan Data ASR :

1. No, merupakan nomor id dari data Asr yang di tampilkan.
2. Video, pada field ini menampilkan Thumbnail dari file video serta dapat memutar video dengan ukuran Thumbnail.
3. Keterangan Video, berisikan informasi mengenai file video yang di tampilkan.
4. Ayat Acuan, berisikan informasi mengenai ayat-ayat yang akan di bacakan sebagai acuan dalam pencarian data.
5. Action :
 - a. Tombol Detail, untuk pindah ke halaman detail berdasarkan dengan nomor id yang dikirimkan.
 - b. Tombol Edit, untuk pindah ke halaman Edit Data berdasarkan dengan nomor id yang dikirimkan.
 - c. Tombol Delete, digunakan untuk menghapus data.

No	Surah	Bacaan Ayat	Data Record	Action
19	1, AL-FAATIHAH (Pembukaan)	بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ الْحَمْدُ لِلّٰهِ رَبِّ الْعَالَمِیْنَ الرَّحْمٰنِ الرَّحِیْمِ		  
18	1, AL-FAATIHAH (Pembukaan)	بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ الْحَمْدُ لِلّٰهِ رَبِّ الْعَالَمِیْنَ الرَّحْمٰنِ الرَّحِیْمِ مَلِكِ یَوْمِ الدِّیْنِ اِیُّهَا نَبِیُّنَا مُحَمَّدٌ وَآلِهٖ وَسَلَّمَ كُنْتُمْ الْمُرْسَلُونَ مِرْطَابُ الْبَیِّنِ الْمَعْتَدِ عَلَيْهِمْ ۚ عَجْرُ الْمُطَّوَّبِ عَلَيْهِمْ وَلَا الْمَشْهُورِ	بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ الْحَمْدُ لِلّٰهِ رَبِّ الْعَالَمِیْنَ الرَّحْمٰنِ الرَّحِیْمِ مَلِكِ یَوْمِ الدِّیْنِ اِیُّهَا نَبِیُّنَا مُحَمَّدٌ وَآلِهٖ وَسَلَّمَ وَارِثَةِ سَلْمِیْنَ	  

Gambar 9. Data Record

Berdasarkan gambar 9 di atas terdapat 1 table dan terdapat 3 tombol. Berikut adalah deskripsi yang terdapat pada tampilan Data Asr:

1. No, merupakan nomor id dari data Record yang di tampilkan.
2. Surah, brisikan informasi mengenali nama surah dalam Al-Qur'an dalam teks bahasa indonesia beserta arti surah.
3. Bacaan Ayat, berisikan informasi mengenai ayat-ayat yang akan di bacakan.
4. Data Record, merupakan informasi dari hasil data record dari halaman input record.



Gambar 10. Detail Data Record

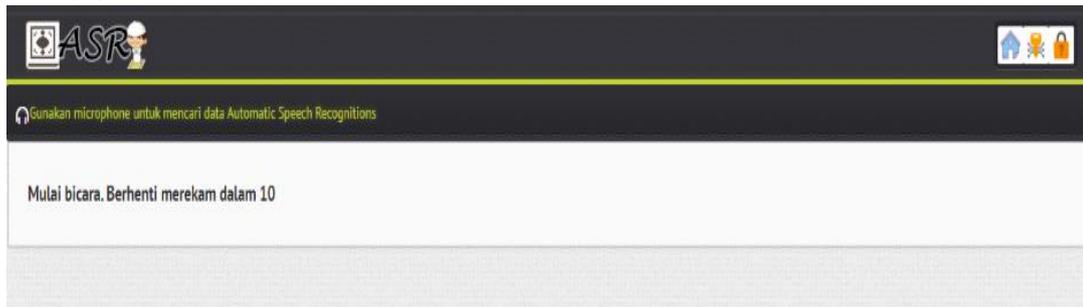
Dari gambar 10 diatas terdapat Detail Data Record, menampilkan informasi lebih detail mengenai data record yang di input berdasarkan id secara terperinci.



Gambar 11. Input / Record Data

Berdasarkan gambar 11 di atas terdapat 4 tombol serta bacaan ayat-ayat Al-Qur'an. Berikut adalah deskripsi yang terdapat pada tampilan Input / Recorder data:

1. Recorder, tombol ini berfungsi untuk mengaktifkan microphone kemudian menampung data audio yang di inputkan melalui record.
2. Pause / Resume, tombol ini berfungsi sebagai tombol untuk berhenti sejenak dan kemudian dapat melanjutkan proses recorder apabila dalam proses sebelumnya terdapat kendala.
3. Stop, tombol ini berfungsi untuk memberhentikan proses recorder kemudian data tersebut di kirimkan untuk di lakukan pembacaan data.
4. Textarea Recorder, bagian ini di gunakan untuk menampilkan hasil dari pembacaan data.
5. Recorder Save, merupakan tombol eksekusi untuk mengirimkan data aplikasi kedalam database.



Gambar 12. Search_asr

Pada gambar 12 di atas terdapat fitur Search_asr, pada halaman ini akan langsung merekam dan membaca data yang di inputkan melalui suara yang diambil dari microphone selama 15 detik. Kemudian menampilkan informasi berkenaan dengan pencocokan data yang di inputkan seperti surah yang di baca, ayat-ayat dalam bentuk format video, serta informasi lainnya berupa asbabun nuzul

4. Kesimpulan

Dari penelitian serta pembahasan pada Aplikasi Automatic Speech Recognitions (ASR) ini, maka penulis dapat mengambil beberapa kesimpulan yaitu:

1. Aplikasi Automatic Speech Recognitions (ASR) ini dirancang menggunakan bahasa pemrograman Python dan menggunakan framework Django untuk menampilkan informasi mengenai ayat-ayat yang di bacakan qori dalam bentuk tampilan berbasis web.
2. Aplikasi ini mengubah data suara menjadi data menggunakan sistem pengenalan suara yang bekerja secara otomatis dengan pencocokan pola digitalisasi audio, kata yang diucapkan terhadap model komputer dari pola bicara untuk menghasilkan keluaran akhir berupa teks yang di simpan di dalam database.
3. Pengujian dilakukan pada surah Al-Fatihah ayat 1 sampai 3, peneliti memilih ayat ini karena surah Al-Fatihah merupakan surah pembuka di dalam Al-Qur'an.
4. Pengujian juga dilakukan pada surah An-Nas ayat 1 sampai 3, peneliti memilih surah ini karena surah An-Nas merupakan surah penutup di dalam Al-Qur'an.

Daftar Pustaka

- [1] N. Anggraini, A. Kurniawan, L. K. Wardhani, and N. Hakiem, "Speech recognition application for the speech impaired using the android-based google cloud speech API," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 16, no. 6, pp. 2733–2739, 2018, doi: 10.12928/TELKOMNIKA.v16i6.9638.
- [2] F. Gozali and R. S. Suharto, "Pemanfaatan Fitur Google Voice Recognition Pada Smartphone Untuk Pengendalian Peralatan Rumah Tangga," *JETri J. Ilm. Tek. Elektro*, vol. 16, no. 2, p. 165, 2019, doi: 10.25105/jetri.v16i2.3620.
- [3] P. W. N. Banamtuan, H. Djahi, and A. A. Maggang, "Pemanfaatan Speech Recognition Pad Smartphone Android Sebagai Sistem Pengontrolan Pintu Berbasis Mikrokontroler," *J. Media Elektro*, vol. 8, no. 1, pp. 72–78, 2019, doi: 10.35508/jme.v8i1.1421.
- [4] Althaf Husein, "Al-Qur'an Di Era Gadget: Studi Deskriptif Aplikasi Qur'an Kemenag," *J. Online Stud. Al-Qur'an*, vol. 16, no. 1, pp. 55–68, 2020, doi: 10.21009/jsq.016.1.04.
- [5] V. Savchenko, "Minimum of Information Divergence Criterion for Signals with Tuning to Speaker Voice in Automatic Speech Recognition," *Radioelectron. Commun. Syst.*, vol. 63, pp. 42–54, Jan. 2020, doi: 10.3103/S0735272720010045.
- [6] T. Zoughi, M. Homayoonpoor, and M. Deypir, "Adaptive Windows Multiple Deep Residual networks for Speech Recognition," *Expert Syst. Appl.*, vol. 139, p. 112840, Jul. 2019, doi: 10.1016/j.eswa.2019.112840.
- [7] A. Joshua, S. Priyadharsini, S. Aravinth, P. Jayaraman, K. Balachandar, and D. Meganathan, "A review on recent trends and development in speech recognition system," *J. Adv. Res. Dyn. Control Syst.*, vol. 12, no. 1 Special Issue, pp. 521–528, 2020, doi: 10.5373/JARDCS/V12SP1/20201099.
- [8] W. Brundage, "Automatic Speech Recognition and Its Application.," *Int. Res. J. Eng. Technol.*, vol. 3, no. 5, pp. 247–252, 2016.
- [9] Y. R. Dewi, F. Satria, M. Rahayu, V. R. V, and A. Uno, "Implementasi Voice Recognition pada Sistem Pengawasan Anak-Anak Dalam Berkata Kasar Melalui Smartphone Dengan Koneksi WiFi," *Ind. Res. Work. Natl. Semin.*, vol. 11, no. 1, pp. 20–25, 2020, doi: 10.35313/irwns.v11i1.1962.