

PERANCANGAN APLIKASI PEMBELAJARAN PELAFALAN BAHASA INGGRIS MENGGUNAKAN TEKNOLOGI PENGOLAHAN SUARA WHISPER

Imam Ahmad Fahrezi¹, Nabila Khatami², Rizal Setiono³, Akbar Azhari Wirajaya⁴, Ahmad Danny Maulana⁵, Joshua Setiawan B. Nahor⁶, Hendra Janu Prasetyo⁷, Mei Lestari⁸

^{1,2,3,4,5,6,7}Program Studi Teknik Informatika, Universitas Indraprasta PGRI
Jalan Raya Tengah No 80, Kelurahan Gedong, Pasar Rebo, Jakarta Timur

¹imamfahrezi24@gmail.com, ²nabilakhatami01@gmail.com, ³rizalsetiono436@gmail.com,
⁴akbarazhari2406@gmail.com, ⁵ahmaddannym@gmail.com, ⁶joshua.setiawan123@gmail.com,
⁷hendrajpras@gmail.com, ⁸mei.lestari@gmail.com

ABSTRAK

Penelitian ini mengembangkan aplikasi pembelajaran pelafalan bahasa Inggris berbasis teknologi pengolahan suara model Whisper. Masalah utama yang dihadapi pembelajar bahasa asing, khususnya bahasa Inggris, adalah pelafalan yang sulit dan seringkali tidak akurat. Tujuan penelitian ini adalah untuk memberikan solusi berupa aplikasi yang dapat mendeteksi dan mengevaluasi pelafalan secara *real-time* menggunakan *speech recognition* dan metrik *Character Error Rate (CER)* dan *Word Error Rate (WER)*. Metode penelitian meliputi perancangan sistem, pengembangan aplikasi menggunakan metode *prototyping*, serta pengujian aplikasi yang melibatkan 20 pengguna. Hasil pengujian menunjukkan aplikasi dapat mendeteksi pelafalan dengan tingkat akurasi yang baik, menghasilkan CER rata-rata antara 8% hingga 16% dan WER rata-rata 15% hingga 25%. Umpan balik dari pengguna menyatakan bahwa aplikasi ini mudah digunakan dan efektif dalam memberikan evaluasi pelafalan yang akurat. Dengan demikian, aplikasi ini berpotensi menjadi alat bantu pembelajaran bahasa Inggris yang efisien dan dapat digunakan secara mandiri oleh pengguna.
Kata Kunci: Pengecekan Pelafalan, *Speech Processing*, *Speech Recognition*.

ABSTRACT

This study developed an English learning application for pronunciation checking based on Whisper's speech processing technology. The main problem faced by foreign language learners, especially in English, is the difficulty in pronouncing accurately. The purpose of this research is to provide a solution in the form of an application that detects and evaluates pronunciation in real-time using speech recognition and the Character Error Rate (CER) metric and Word Error Rate (WER). The research method includes system design, application development using the prototyping method, and application can detect pronunciation with good accuracy, yielding an average CER between 8% to 16% and WER between 15% to 25%. Feedback from users indicates that the application is easy to use and effective in providing accurate pronunciation evaluation. Thus, this application has the potential to be an efficient English learning tool that can be used independently by learners.

Key Word: *Pronunciation Checking, Speech Processing, Speech Recognition.*

PENDAHULUAN

Bahasa Inggris merupakan bahasa internasional yang banyak dipelajari dan digunakan oleh banyak orang, sehingga bahasa Inggris dianggap sangat penting (Ahmad *et al.*, 2019). Namun, belajar bahasa Inggris secara manual, seperti menggunakan kamus atau mengikuti kursus itu memerlukan waktu dan biaya yang cukup besar, sehingga membuat banyak orang enggan untuk mempelajarinya (Jefrizal *et al.*, 2017).

Dalam era digital saat ini, pembelajaran bahasa Inggris berkembang pesat dengan

dukungan teknologi (Gusti *et al.*, 2021). Salah satu tantangan utama dalam mempelajari bahasa Inggris adalah menguasai pelafalan (*pronunciation*) kata-kata dengan benar, terutama bagi pemula. *Pronunciation* adalah cara kita mengucapkan atau membunyikan kata-kata. Banyak orang merasa bahwa *pronunciation* adalah aspek tersulit, karena orang Indonesia terbiasa dengan pelafalan bahasa Indonesia, sehingga pengucapan bahasa Inggris terasa aneh (Gd *et al.*, 2021). Aplikasi pembelajaran yang ada seringkali berfokus pada pembelajaran tata bahasa dan hafalan kosakata (*vocabulary*) saja, tetapi kurang memberikan umpan balik yang

spesifik mengenai kesalahan pelafalan (*pronunciation*).

Penggunaan teknologi dalam pembelajaran telah menunjukkan hasil yang positif. Salah satu pendekatan yang banyak digunakan adalah teknologi *Speech recognition*, yang memungkinkan evaluasi dan umpan balik langsung terhadap pelafalan pengguna. *Speech recognition* adalah suatu sistem untuk menerima input kata yang diucapkan dan memungkinkan perangkat mengenali ciri-ciri ucapan yang telah tersimpan dalam basis data (Muhamad M.I. Putra *et al.*, 2020).

Penelitian ini bertujuan untuk mengembangkan aplikasi pembelajaran bahasa Inggris yang memiliki fitur pengecekan pelafalan (*pronunciation*) menggunakan teknologi pengolahan suara berbasis model Whisper untuk menerjemahkan suara yang diucapkan oleh user menjadi teks lalu mengevaluasi pelafalan pengguna menggunakan metrik *Character Error Rate* (CER) untuk kosa-kata dan *Word Error Rate* (WER) untuk kalimat bahasa Inggris. *Word Error Rate* (WER) adalah pendekatan standar untuk mengevaluasi kinerja sistem pengenalan suara kontinu dengan kosakata besar (LVCSR) (Ali & Renals, 2018). Sedangkan CER, meskipun tidak umum seperti WER, tetapi sangat dikenal dan sering digunakan dalam evaluasi ASR (Conneau *et al.*, 2021). Arsitektur Whisper dibangun di sekitar encoder-decoder berbasis transformer, yang menggunakan pembelajaran multitugas untuk melakukan transkripsi, penerjemahan, dan tugas pengenalan ucapan lainnya. Model ini beroperasi pada input audio 16 kHz yang disampel ulang, yang diubah menjadi spectrogram log-mel. Spektrogram ini diproses oleh serangkaian lapisan konvolusional dan penyandian posisional sebelum melewati blok transformer yang memprediksi token teks keluaran (Radford *et al.*, n.d.). Tujuannya adalah untuk memberikan umpan balik yang akurat dan interaktif kepada pengguna mengenai pelafalan mereka, sekaligus meningkatkan pengalaman belajar dengan menyajikan skor pelafalan dalam bentuk persentase.

Penelitian diharapkan dapat memberikan kontribusi signifikan bagi pengembangan metode pembelajaran bahasa Inggris, khususnya dalam meningkatkan kemampuan pelafalan (*pronunciation*) pengguna. Selain itu, aplikasi ini diharapkan dapat menjadi alat bantu yang efektif bagi pengajar dan pelajar dalam proses belajar-mengajar.

METODE PENELITIAN

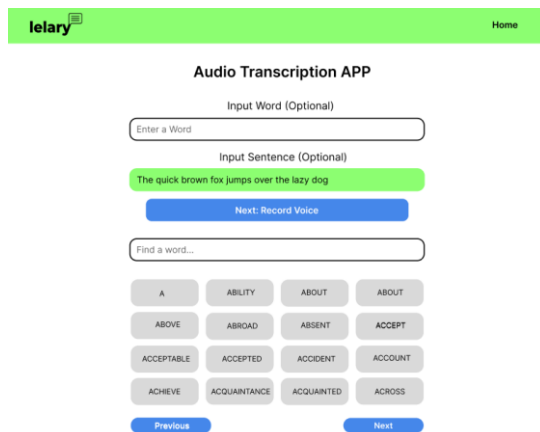
Penelitian ini menggunakan beberapa metode yang meliputi Studi Literatur, Perancangan Sistem, Pengembangan Aplikasi, Pengujian Aplikasi, Analisis Data, serta Evaluasi dan Penyempurnaan. Setiap tahap dilakukan untuk mengembangkan aplikasi pembelajaran bahasa Inggris dengan fitur pengecekan pelafalan (*pronunciation*) yang akurat dan interaktif. Berikut adalah penjelasan dari masing-masing tahapan:

1. Studi Literatur

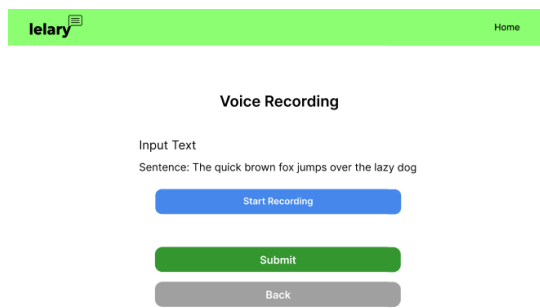
Pada tahap ini dilakukan kajian literatur terkait teknologi pengolahan suara, khususnya penerapan *speech recognition* dalam pembelajaran bahasa asing. Kajian ini mencakup model Whisper dan metrik evaluasi pelafalan seperti *Character Error Rate* (CER) dan *Word Error Rate* (WER) yang digunakan sebagai acuan dalam sistem evaluasi pelafalan (*pronunciation*).

2. Perancangan Sistem

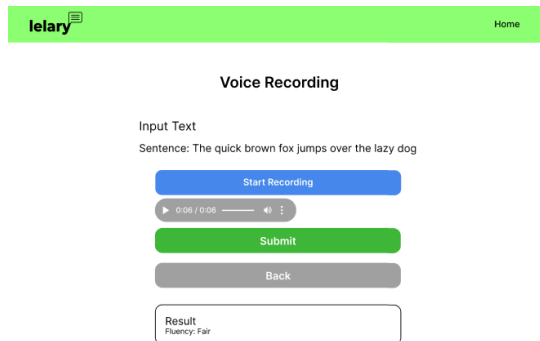
Perancangan sistem aplikasi dilakukan dengan menggunakan model arsitektur perangkat lunak yang mencakup antarmuka pengguna (UI), integrasi teknologi pengolahan suara untuk pengecekan pelafalan (*pronunciation*), serta penyimpanan data kosakata untuk memastikan akurasi dan konsistensi dalam proses pembelajaran. Selain itu, hasil pengolahan suara disajikan secara interaktif kepada pengguna dalam bentuk umpan balik yang mudah dipahami, sehingga dapat membantu meningkatkan keterampilan pelafalan pengguna secara bertahap dan efektif.



Gambar 1. UI Beranda Aplikasi



Gambar 2. UI Untuk Cek Pronunciation



Gambar 3. UI Setelah Cek Pronunciation

3. Pengembangan

Perangkat Lunak Pengembangan aplikasi dilakukan menggunakan metode prototyping, yang memungkinkan pengujian sistem dalam iterasi yang lebih cepat. Alat dan bahasa pemrograman yang digunakan meliputi Python untuk backend dan pengolahan suara, Javascript, HTML untuk frontend dan firebase untuk menyimpan kosakata, fitur utama dari firebase dapat sinkronisasi data secara otomatis tanpa membuat update request atau refresh aplikasi (Wiratno & Hastuti, 2017). serta teknologi Whisper untuk *speech processing*. Model Whisper digunakan untuk

mengubah input suara menjadi teks, dan kemudian dilakukan evaluasi pelafalan (*pronunciation*) menggunakan CER atau WER.

Rumus *Character Error Rate* (CER) didefinisikan sebagai:

$$CER(\%) = \frac{D}{L} \times 100 \quad (1)$$

$$D = Levenshtein\ distance(X, Y) \quad (2)$$

di mana X adalah output yang diprediksi dan Y adalah teks asli yang untuk perbandingan. Jarak D adalah jarak Levenshtein antara X dan Y, dan panjang L adalah panjang skrip kebenaran Y (Jeong *et al.*, 2021).

4. Pengujian Aplikasi

Aplikasi diuji dengan melibatkan 20 partisipan yang merupakan pengguna bahasa Inggris sebagai bahasa asing. Pengujian dilakukan secara bertahap, dimulai dari uji coba internal oleh tim pengembang, dilanjutkan dengan uji coba eksternal oleh pengguna. Umpan balik pengguna terkait akurasi pengecekan pelafalan (*pronunciation*) dan kemudahan penggunaan aplikasi dikumpulkan melalui kuesioner, serta data kesalahan pelafalan (*pronunciation*) dianalisis berdasarkan metrik CER dan WER.

5. Analisis Data

Data hasil pengujian, terutama terkait tingkat akurasi pengecekan pelafalan dan persepsi pengguna terhadap aplikasi, dianalisis secara deskriptif. Nilai CER atau WER yang dihasilkan dari pengecekan pelafalan akan dibandingkan dengan standar pelafalan yang benar untuk menilai akurasi sistem. Umpan balik pengguna mengenai kemudahan dan pengalaman penggunaan aplikasi juga dianalisis untuk peningkatan lebih lanjut.

6. Evaluasi dan Penyempurnaan

Berdasarkan hasil analisis, dilakukan evaluasi terhadap fitur-fitur aplikasi yang telah dikembangkan. Jika ditemukan kelemahan, iterasi pengembangan dilanjutkan untuk menyempurnakan sistem sebelum aplikasi diluncurkan secara resmi.

Metode ini diharapkan dapat menghasilkan aplikasi yang mampu memberikan umpan balik yang akurat dan interaktif kepada

pengguna mengenai pelafalan bahasa Inggris, serta meningkatkan efisiensi dalam pembelajaran kosakata secara mandiri.

HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan aplikasi pembelajaran pelafalan bahasa Inggris dengan teknologi pengolahan suara berbasis model Whisper. Hasil penelitian dijabarkan dalam beberapa bagian, mulai dari pengujian sistem, analisis hasil pengecekan pelafalan, hingga umpan balik pengguna terkait efektivitas aplikasi.

1. Hasil Pengujian Sistem

Pengujian sistem dilakukan untuk menilai kinerja aplikasi dalam mendeteksi dan mengevaluasi pelafalan pengguna. Berdasarkan uji coba terhadap 20 partisipan, aplikasi mampu mendeteksi pelafalan kata secara real-time dengan tingkat akurasi yang cukup tinggi. Proses evaluasi pelafalan dilakukan dengan menghitung Character Error Rate (CER) untuk setiap kata yang diucapkan oleh pengguna. Rata-rata CER yang dihasilkan dari pengujian ini berada pada kisaran 8% hingga 15%, yang berarti sebagian besar pengguna hanya membuat sedikit kesalahan dalam pelafalan.

Tabel 1. Hasil Pengujian CER untuk 5 Partisipan

Partisipan	Kosakata	CER
1	Above	15%
2	About	11%
3	Better	8%
4	Detail	13%
5	Downstairs	16%

Dari Tabel 1 di atas, dapat dilihat bahwa rata-rata CER bervariasi antara 8% hingga 16%, menunjukkan bahwa aplikasi cukup baik dalam mengidentifikasi pelafalan pengguna dan memberikan evaluasi berdasarkan standar pelafalan yang benar.

Sedangkan proses evaluasi pelafalan yang berupa kalimat dilakukan dengan menghitung *Word Error Rate* (WER) untuk setiap kata yang diucapkan oleh pengguna. Sebagaimana yang ditunjukkan pada Tabel 2, rata-rata WER

yang dihasilkan dari pengujian ini berada pada kisaran 15% hingga 25%, yang berarti sebagian besar pengguna hanya membuat sedikit kesalahan dalam pelafalan.

Tabel 2. Hasil Pengujian WER untuk 5 Partisipan

Partisipan	Kalimat	WER
1	The quick brown fox jumps over the lazy dog	16%
2	She sells seashells by the seashore.	20%
3	Artificial intelligence is transforming the world	23%
4	I would like to order a cup of coffee, please	25%
5	Technology is advancing at an incredible pace	25%

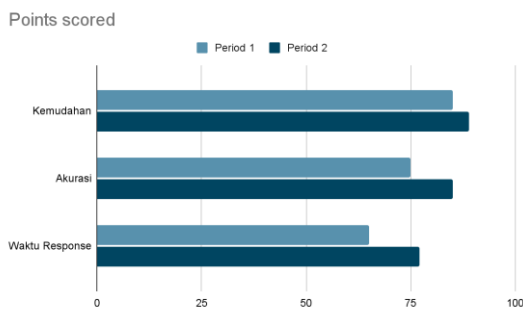
2. Evaluasi Umpan Balik Pengguna

Umpan balik dari partisipan terkait penggunaan aplikasi menunjukkan respons yang positif. Sebagian besar pengguna merasa bahwa aplikasi mudah digunakan dan mampu memberikan umpan balik pelafalan secara akurat. Berikut ini adalah beberapa poin umpan balik yang didapat dari hasil kuesioner:

Kemudahan Penggunaan: 85% dari partisipan merasa bahwa aplikasi memiliki antarmuka yang mudah dipahami dan digunakan, baik bagi pemula maupun pengguna yang sudah mahir.

Akurasi Umpan Balik Pelafalan: 80% dari partisipan menyatakan bahwa hasil evaluasi pelafalan yang diberikan oleh aplikasi akurat dan membantu mereka memahami kesalahan pelafalan yang mereka buat.

Waktu Respon Aplikasi: 75% partisipan merasa bahwa aplikasi mampu memberikan hasil pengecekan pelafalan dengan cepat, tanpa ada keterlambatan yang signifikan.



Gambar 4. Skor Kepuasan Pengguna terhadap Aplikasi

Grafik di atas menunjukkan tingkat kepuasan pengguna yang tinggi terhadap berbagai aspek aplikasi, mulai dari kemudahan penggunaan hingga akurasi evaluasi pelafalan.

3. Pembahasan

Hasil penelitian ini menunjukkan bahwa teknologi pengolahan suara berbasis Whisper efektif dalam mendukung pembelajaran kosakata bahasa Inggris, terutama dalam aspek pengecekan pelafalan. Dengan tingkat akurasi rata-rata CER yang rendah (8%-16%) dan WER (15%-25%), aplikasi ini mampu memberikan umpan balik yang cukup akurat kepada pengguna, sehingga mereka dapat memperbaiki pelafalan mereka dengan cepat dan efisien.

Implementasi fitur pengecekan pelafalan berbasis Whisper memungkinkan aplikasi untuk mengenali variasi suara pengguna dengan baik, bahkan dalam kondisi lingkungan yang tidak sepenuhnya tenang. Hal ini membuktikan bahwa model Whisper mampu mengatasi beberapa tantangan yang biasanya dihadapi oleh teknologi pengenalan suara, seperti gangguan latar belakang dan variasi aksen pengguna.

Umpan balik dari pengguna menunjukkan bahwa aplikasi ini memiliki potensi untuk digunakan sebagai alat bantu pembelajaran bahasa Inggris secara mandiri.

Namun, terdapat beberapa tantangan yang masih perlu diatasi, seperti peningkatan akurasi pengecekan pada kata-kata dengan pelafalan yang mirip dan optimalisasi waktu respons aplikasi. Meskipun waktu respons aplikasi sudah cukup baik, terdapat beberapa kasus di mana sistem membutuhkan waktu

lebih lama untuk memproses pelafalan kata yang lebih kompleks.

4. Implikasi dan Rekomendasi

Berdasarkan hasil penelitian ini, aplikasi pengecekan pelafalan dapat diimplementasikan lebih luas untuk mendukung pembelajaran bahasa Inggris, baik di lingkungan pendidikan formal maupun untuk pembelajaran mandiri. Pengembangan lebih lanjut dapat difokuskan pada peningkatan akurasi pengecekan pelafalan untuk kata-kata dengan pelafalan yang mirip, serta penambahan fitur lain, seperti latihan fonetik atau kosakata tematik.

Implementasi aplikasi ini juga dapat diperluas untuk mendukung pembelajaran bahasa lainnya, dengan menyesuaikan model pengolahan suara dan data latih yang relevan untuk bahasa yang diinginkan. Hal ini dapat meningkatkan cakupan aplikasi dan membantu lebih banyak pengguna dalam belajar bahasa asing.

SIMPULAN DAN SARAN

Aplikasi pembelajaran bahasa Inggris berbasis teknologi Whisper berhasil memberikan umpan balik pelafalan dengan akurasi tinggi, ditunjukkan oleh rata-rata Character Error Rate (CER) sebesar 8-16% dan Word Error Rate (WER) sebesar 15-25%. Nilai ini menunjukkan tingkat kesalahan yang relatif rendah, meskipun saat ini aplikasi hanya menampilkan hasil evaluasi tanpa memberikan masukan spesifik. Hasil ini tetap dapat menjadi tolok ukur bagi pengguna dalam menilai kemampuan pelafalan mereka secara mandiri, sehingga aplikasi dinilai efektif sebagai alat evaluasi dalam pembelajaran bahasa Inggris.

Untuk meningkatkan manfaat aplikasi, disarankan agar pengembangan selanjutnya mencakup fitur pemberian umpan balik yang lebih mendetail mengenai kesalahan pelafalan, seperti menunjukkan bagian kata atau fonem yang tidak tepat. Selain itu, aplikasi ini berpotensi diperluas untuk pembelajaran bahasa asing lainnya dengan melakukan penyesuaian teknologi pengolahan suara dan penggunaan data latih yang sesuai. Penelitian lanjutan juga dapat berfokus pada peningkatan akurasi sistem guna menurunkan

nilai CER dan WER, serta memperkaya fitur evaluasi agar mencakup aspek seperti intonasi, ritme, dan tekanan (stress) dalam pengucapan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada dosen pembimbing atas arahan dan bimbingannya. Terima kasih juga kepada rekan-rekan atas kontribusi yang diberikan dalam penelitian ini. Semoga penelitian ini bermanfaat dan berkontribusi pada pengembangan ilmu, khususnya dalam pembelajaran bahasa Inggris berbasis teknologi.

DAFTAR PUSTAKA

- Ahmad, F. I., Afirianto, T., & Akbar, M. A. (2019). *Perancangan Game Pembelajaran Pengucapan Bahasa Inggris Berbasis Pengenalan Suara* (Vol. 3, Issue 9). <http://j-ptiik.ub.ac.id>
- Ali, A., & Renals, S. (2018). *Word Error Rate Estimation for Speech Recognition: e-WER*. <https://github.com/qcri/e-wer>
- Conneau, A., Baevski, A., Collobert, R., Mohamed, A., & Auli, M. (2021). Unsupervised cross-lingual representation learning for speech recognition. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 1*, 346–350. <https://doi.org/10.21437/Interspeech.2021-329>
- Gd, I. P., Sudiarmika, A., Gde, I. A., Putra, S., Perdana, K. S., Hari, K., Dewi, S., Komputer, S., & Bali, S. (2021). *Prosiding Seminar Nasional CORISINDO 2021 Multidisiplin Ilmu*.
- Gusti, I., Agung, A., & Susanthi, D. (2021). Kendala dalam Belajar Bahasa Inggris dan Cara Mengatasinya. *Linguistic Community Service Journal* |, 1(2), 2021. <https://doi.org/10.22225/licosjournal.v1i2.2658>
- Jefrizal, Jaroji, & Agus Tedyyana. (2017). *Aplikasi English Teacher Sebagai Alat Bantu Belajar English Conversation Berbasis Android dengan Menerapkan Voice Recognition*.
- Jeong, J., Raton Mondol, S. I. M. M., Kim, Y. W., & Lee, S. (2021). An effective learning method for automatic speech recognition in Korean ci patients' speech. *Electronics (Switzerland), 10*(7). <https://doi.org/10.3390/electronics10070807>
- Muhamad M.I. Putra, Sherwin R. U. A. Sompie, & Sary Paturusi 3. (2020). *Implementasi Speech Recognition pada Aplikasi Pembelajaran Bahasa Inggris untuk Anak*.
- Radford, A., Kim, J. W., Xu, T., Brockman, G., Mcleavy, C., & Sutskever, I. (n.d.). *Robust Speech Recognition via Large-Scale Weak Supervision*.
- Wiratno, A. R., & Hastuti, K. (2017). Implementation of Firebase Realtime Database to track BRT Trans Semarang. *Scientific Journal of Informatics, 4*(2), 2407–7658. <http://journal.unnes.ac.id/nju/index.php/sji>