

Koreksi Pelafalan Bahasa Indonesia Anak Usia Dini Menggunakan Word Embedding

^{1*}Aldo Aditya Saputra, ²Risky Aswi Ramadhani, ³Julian Sahertian

^{1,2,3}Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: *¹aldoadityasaputra70@gmail.com, ²risky_aswi@unpkediri.ac.id,

³juliansahertian@unpkediri.ac.id

Penulis Korespondens : Risky Aswi Ramadhani

Abstrak—Kemampuan pelafalan merupakan salah satu aspek penting dalam perkembangan bahasa anak usia dini. Namun, masih terdapat anak yang mengalami kesalahan pengucapan kosakata bahasa Indonesia sehingga memerlukan umpan balik yang tepat untuk membantu proses pembelajaran berbicara. Penelitian ini bertujuan mengembangkan sistem koreksi pelafalan bahasa Indonesia menggunakan *Automatic Speech Recognition* (ASR) dan *Word Embedding FastText*. Sistem memanfaatkan *Faster-Whisper* untuk mengubah suara menjadi teks, *FastText* untuk merepresentasikan kata ke dalam bentuk vektor numerik, serta *Cosine Similarity* untuk menentukan tingkat kemiripan antara hasil transkripsi dan kosakata target. Sistem juga dilengkapi fitur identifikasi jenis kesalahan pelafalan, rekomendasi stimulasi pengucapan, dan *Text-to-Speech* (TTS) untuk menghasilkan pelafalan yang benar. Dataset penelitian terdiri dari 49 data audio kosakata benda dalam lingkungan rumah yang direkam dalam format .wav. Hasil pengujian menunjukkan bahwa sistem mampu melakukan koreksi pelafalan dengan akurasi sebesar 89,80%, di mana 44 data berhasil dikoreksi sesuai target dan 5 data mengalami kesalahan koreksi. Hasil penelitian menunjukkan bahwa kombinasi *Faster-Whisper*, *Word Embedding FastText*, dan *Cosine Similarity* dapat mendukung koreksi pelafalan bahasa Indonesia secara otomatis.

Kata Kunci— *Automatic Speech Recognition*, *FastText*, Koreksi Pelafalan, *Text-to-Speech*, *Word Embedding*

Abstract—Pronunciation ability is an important aspect of early childhood language development. However, some children still experience pronunciation errors in Indonesian vocabulary and require appropriate feedback to support their speaking skills. This study aims to develop an Indonesian pronunciation correction system using *Automatic Speech Recognition* (ASR) and *FastText Word Embedding*. The system employs *Faster-Whisper* to convert speech into text, *FastText* to represent words as numerical vectors, and *Cosine Similarity* to measure the similarity between transcription results and target vocabulary. In addition, the system provides pronunciation error identification, speech stimulation recommendations, and *Text-to-Speech* (TTS) features to generate correct pronunciation audio. The dataset consists of 49 audio recordings of household object vocabulary in .wav format. Experimental results show that the proposed system achieved an accuracy of 89.80%, with 44 data samples correctly matched to the target words and 5 samples incorrectly corrected. These findings indicate that the combination of *Faster-Whisper*, *FastText Word Embedding*, and *Cosine Similarity* can effectively support automatic Indonesian pronunciation correction for early childhood learning.

Keywords— *Automatic Speech Recognition*, *FastText*, Pronunciation Correction, *Text-to-Speech*, *Word Embedding*

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Perkembangan teknologi digital pada era modern memberikan pengaruh terhadap pola komunikasi dan perkembangan bahasa anak usia dini. Dalam proses belajar berbicara, masih terdapat anak-anak yang mengalami kesalahan atau pengucapan yang kurang tepat pelafalan pada beberapa kata bahasa Indonesia. Kondisi tersebut dapat memengaruhi kemampuan komunikasi anak apabila tidak diberikan stimulasi pelafalan secara tepat sejak dini. Oleh karena itu, diperlukan suatu media yang dapat membantu proses latihan pengucapan kata secara sederhana dan interaktif. Salah satu teknologi yang dapat dimanfaatkan adalah *Automatic Speech Recognition* (ASR), yaitu teknologi yang mampu mengenali ucapan manusia dan mengubahnya menjadi teks secara otomatis [1][2]. Teknologi ini berpotensi digunakan untuk membantu proses koreksi pelafalan anak usia dini melalui pendekatan berbasis suara.

Kemampuan berbicara merupakan salah satu aspek penting dalam perkembangan anak usia dini. Namun, beberapa anak masih mengalami kesulitan dalam mengucapkan kata secara jelas dan tepat. Kesalahan pelafalan yang sering terjadi meliputi penghilangan huruf, penambahan bunyi, maupun substitusi bunyi tertentu. Kondisi tersebut dapat memengaruhi kemampuan komunikasi anak apabila tidak diberikan stimulasi yang tepat sejak dini. Selain itu, proses latihan pelafalan secara manual membutuhkan pendampingan secara terus-menerus sehingga diperlukan media bantu berbasis teknologi yang mampu memberikan umpan balik secara otomatis terhadap hasil pelafalan anak.

Penelitian mengenai *Automatic Speech Recognition* telah banyak dilakukan sebelumnya [3][4]. Sebagian besar penelitian *speech recognition* masih berfokus pada bahasa Inggris, sedangkan pengembangan sistem pengenalan suara untuk bahasa Indonesia masih terbatas [5]. Penelitian lain juga menunjukkan bahwa teknologi pengenalan suara dapat digunakan untuk membantu proses pembelajaran berbasis *audio* dan interaksi manusia dengan komputer [6][7]. Selain itu, penggunaan *Word Embedding FastText* dinilai mampu merepresentasikan hubungan semantik antar kata dengan cukup baik sehingga dapat dimanfaatkan pada proses pencarian kemiripan kata [8]. Teknologi *Text-to-Speech* (TTS) juga telah banyak diterapkan untuk menghasilkan suara sintesis yang dapat digunakan sebagai media bantu pembelajaran pengucapan kata [9].

Meskipun penelitian terkait ASR dan pengolahan suara telah berkembang cukup pesat, sebagian besar penelitian masih berfokus pada proses transkripsi suara umum dan belum banyak diterapkan pada sistem koreksi pelafalan anak usia dini secara otomatis. Selain itu, penelitian sebelumnya umumnya belum menggabungkan proses transkripsi suara, koreksi kata, analisis kesalahan pelafalan, serta pemberian rekomendasi stimulasi dalam satu sistem terintegrasi. Berdasarkan analisis kesenjangan tersebut, penelitian ini mengembangkan sistem koreksi otomatis pelafalan anak usia dini menggunakan *Faster-Whisper* dan *Word Embedding FastText* dengan *cosine similarity* untuk membantu proses stimulasi pelafalan sederhana.

Penelitian ini bertujuan membangun sistem yang mampu melakukan transkripsi suara anak menjadi teks menggunakan *Faster-Whisper*, melakukan koreksi kata menggunakan *Word Embedding FastText*, menganalisis kesalahan pelafalan, serta menghasilkan rekomendasi stimulasi sederhana dan *audio* pelafalan yang benar menggunakan *Text-to-Speech* (TTS). Sistem menerima *input* berupa rekaman suara anak yang mengucapkan kosakata benda di dalam rumah, kemudian melakukan proses pengolahan secara otomatis hingga menghasilkan *output* berupa hasil koreksi dan *audio* pelafalan yang benar.

Dataset penelitian diperoleh dari rekaman suara anak usia dini yang mengucapkan sekitar 50

kosakata benda yang terdapat di dalam rumah. Setiap kata direkam secara terpisah dan disimpan dalam format *audio .wav* agar dapat diproses oleh sistem *Automatic Speech Recognition*. Dengan adanya penelitian ini, diharapkan sistem yang dibangun dapat menjadi media bantu latihan pelafalan sederhana yang lebih interaktif serta membantu memberikan stimulasi pengucapan kata pada anak usia dini berbasis teknologi pengenalan suara.

II. METODE



Gambar 1. Diagram Alur Sistem Koreksi Pelafalan Bahasa Indonesia Anak Usia Dini

Penelitian ini menggunakan metode pengembangan sistem berbasis *Automatic Speech Recognition* (ASR) untuk melakukan koreksi otomatis pelafalan anak usia dini menggunakan suara berbahasa Indonesia. Sistem dikembangkan menggunakan bahasa pemrograman *Python* dengan *framework Streamlit* sebagai antarmuka pengguna. Penelitian ini memanfaatkan model *Faster-Whisper* untuk proses transkripsi suara, *Word Embedding FastText* untuk proses pencarian kemiripan kata, *cosine similarity* untuk menghitung tingkat kemiripan antar kata, serta *Text-to-Speech* (TTS) untuk menghasilkan *output* suara pelafalan yang benar[10][4].

Subjek penelitian berupa rekaman suara anak usia dini yang mengucapkan kosakata benda yang terdapat di dalam rumah. Dataset penelitian terdiri dari sekitar 50 kosakata benda, seperti meja, kursi, pintu, sapu, gelas, dan benda rumah tangga lainnya. Setiap kata direkam secara terpisah dan disimpan dalam format file *audio .wav*. Teknik pengambilan data dilakukan melalui proses perekaman suara secara langsung menggunakan microphone dan perangkat perekam suara digital. Dataset kemudian dipisahkan menjadi *file audio* tunggal untuk setiap kata agar mempermudah proses pengujian sistem.

Perangkat keras yang digunakan dalam penelitian ini meliputi laptop dengan sistem operasi *Windows*, *microphone* untuk proses perekaman suara, dan *speaker* atau *headset* untuk pengujian *output audio*. Perangkat lunak yang digunakan meliputi *Python*, *Streamlit*, *Faster-Whisper*, *FastText*, *Gensim*, dan *gTTS*. *Faster-Whisper* digunakan untuk mengubah suara menjadi teks secara otomatis melalui proses *Automatic Speech Recognition* [11]. Sementara itu, *FastText* digunakan sebagai model *Word Embedding* untuk merepresentasikan kata dalam bentuk vektor [12][13].

A. Input Data

Pada tahap ini sistem menerima satu data suara anak berupa pengucapan kata benda dalam bahasa Indonesia. Setelah diproses oleh model *Automatic Speech Recognition* (ASR), sistem menghasilkan teks hasil pengenalan suara.

Tabel 1. Tabel Dataset/Data Penelitian

No	Kata Benda	Jumlah Audio
1	Baju	1
2	Buku	1
3	Cermin	1
4	Kursi	1
...	...	...
50	Wajan	1

B. Transkripsi Suara

Proses transkripsi dilakukan menggunakan model *Faster-Whisper*. *Audio* masukan dikonversi menjadi teks yang selanjutnya digunakan pada proses koreksi kata.

$$\text{Text} = \text{Whisper}(\text{Audio}) \dots\dots\dots (1)$$

Keterangan:

- a) *Audio*: merupakan *file* suara masukan anak.
- b) *Text*: merupakan hasil transkripsi suara dalam bentuk teks

C. Word Embedding

Langkah selanjutnya adalah mengkonversi kata-kata yang telah ditranskripsikan menjadi representasi vektor menggunakan model *FastText* Bahasa Indonesia. Sebagai contoh, kata “sermin” dan “cermin” akan direpresentasikan menjadi vektor numerik sebagai berikut:

$$\text{Vector}(\text{“sermin”}) = [0.12, -0.45, 0.31, \dots]$$

$$\text{Vector}(\text{“cermin”}) = [0.14, -0.40, 0.29, \dots]$$

Nilai vektor tersebut bukan dihitung secara manual oleh peneliti, tetapi dihasilkan secara otomatis oleh model *Word Embedding FastText* yang telah dilatih sebelumnya pada korpus bahasa Indonesia.

D. Preprocessing

Setelah proses transkripsi selesai, teks pengenalan suara akan diproses menggunakan tahap normalisasi teks. Pada penelitian ini proses normalisasi dilakukan menggunakan fungsi `normalize_text()` yang terdiri dari:

- a) Mengubah seluruh huruf menjadi *lowercase*.
- b) Menghapus spasi berlebih pada awal dan akhir teks.

Proses normalisasi pada sistem dapat dituliskan sebagai berikut:

$$[\text{Text}_{\{\text{normalisasi}\}} = \text{lower}(\text{Text})] \dots\dots\dots (2)$$

E. Perhitungan *Cosine Similarity*

$$\text{similarity}(A, B) = \frac{A \cdot B}{|A||B|} \dots\dots\dots (3)$$

Keterangan:

- a) A = vektor kata hasil transkripsi
- b) B = vektor kata target
- c) $A \cdot B$ = hasil perkalian *dot product*
- d) $\|A\|$ dan $\|B\|$ = panjang vektor kata

Semakin tinggi nilai *cosine similarity*, maka semakin besar tingkat kemiripan antara hasil transkripsi dan kosakata target [14].

Sebagai contoh:

- a) Hasil ASR = “sermin”
- b) Kosakata target = [baju, buku, cermin, gelas, kursi, meja]

Tabel 2. Perhitungan *Similarity*

No	Kata Target	Nilai <i>Similarity</i>
1	Baju	0.6003
2	Buku	0.3277
3	Cermin	0.3603
4	Gelas	0.5200
5	Meja	1.0000

F. Koreksi Kata

Tahap koreksi dilakukan dengan memilih kata target yang memiliki nilai *similarity* tertinggi terhadap hasil transkripsi ASR. Proses pemilihan kata terbaik dilakukan menggunakan kondisi berikut:

$$[\text{BestWord} = \max(\text{Similarity})] \dots\dots\dots (4)$$

Pada sistem, proses ini dilakukan secara berulang terhadap seluruh kosakata target hingga diperoleh kata dengan nilai *similarity* terbesar.

G. Analisis Kesalahan Pelafalan

Setelah memperoleh kata hasil koreksi, sistem melakukan analisis kesalahan pelafalan dengan membandingkan hasil ASR dan kata target. Pada simulasi ini:

- a) Hasil ASR = “geles!”
- b) Kata target = “gelas”

Karena terdapat perubahan bunyi huruf “a” menjadi “e”, maka sistem mengklasifikasikan kesalahan sebagai substitusi bunyi.

Tabel 3. Analisis Kesalahan Pelafalan

No	Hasil ASR	Kata Target	Jenis Kesalahan
1	“geles!”	Gelas	Penambahan huruf

H. Rekomendasi Stimulasi

Berdasarkan jenis kesalahan pelafalan yang diperoleh, sistem memberikan rekomendasi stimulasi pengucapan secara otomatis. Penelitian ini rekomendasi diberikan menggunakan fungsi rekomendasi_stimulasi().

Tabel 4. Rekomendasi Stimulasi

No	Jenis Kesalahan	Rekomendasi
1	Penambahan huruf	Latihan fokus pada pelafalan kata target.

I. Text-to-Speech (TTS)

Tahap berikutnya adalah menghasilkan suara pelafalan yang benar menggunakan *Google Text-to-Speech* (gTTS). Sistem akan mengubah kata hasil koreksi menjadi *audio* pelafalan Bahasa Indonesia. Proses *Text-to-Speech* dapat dituliskan sebagai berikut:

$$[\text{Audio} = \text{gTTS}(\text{kata_koreksi})] \dots\dots\dots (5)$$

Sebagai contoh, kata “cermin” akan diubah menjadi suara pelafalan yang benar sehingga dapat diputar kembali dan ditirukan oleh anak.

J. Output Sistem

Tahap terakhir adalah menampilkan seluruh hasil proses pada antarmuka sistem. *Output* yang ditampilkan meliputi:

- a) Hasil transkripsi ASR
- b) Kata hasil koreksi
- c) Nilai *similarity*
- d) Jenis kesalahan pelafalan
- e) Rekomendasi stimulasi
- f) *Audio* pelafalan yang benar

Dengan demikian, sistem ini mampu membantu proses koreksi pengucapan bahasa Indonesia untuk anak usia dini secara otomatis dan interaktif menggunakan metode *Word Embedding* dan *Cosine Similarity*.

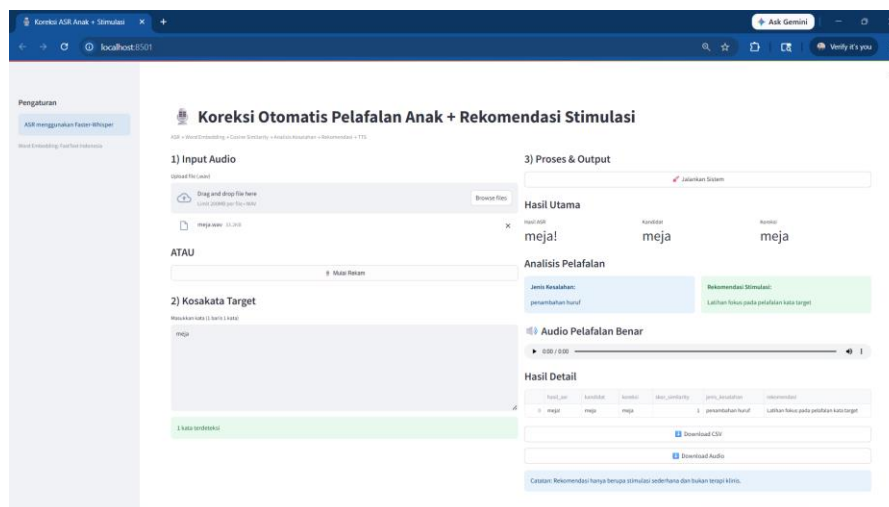
III. HASIL DAN PEMBAHASAN

A. Hasil

Penelitian ini menghasilkan sebuah sistem koreksi pelafalan bahasa Indonesia untuk anak usia dini yang dibangun menggunakan metode *Automatic Speech Recognition* (ASR) berbasis *Faster-Whisper* dan *Word Embedding FastText*. Sistem dikembangkan dalam

bentuk aplikasi berbasis web menggunakan *Streamlit* yang mampu menerima masukan suara anak, melakukan transkripsi otomatis, melakukan koreksi kata berdasarkan nilai *cosine similarity* tertinggi terhadap kosakata target, mengidentifikasi jenis kesalahan pelafalan, serta memberikan rekomendasi stimulasi pengucapan dan *audio* pelafalan yang benar menggunakan *Text-to-Speech* (TTS).

Pengujian sistem dilakukan menggunakan 49 data *audio* yang berisi pengucapan kosakata benda yang terdapat di lingkungan rumah. Setiap data *audio* direkam dalam format .wav dan diuji secara individual untuk mengetahui kemampuan sistem dalam menghasilkan koreksi kata yang sesuai dengan kata target.



Gambar 2. Halaman Utama Sistem

Gambar 2 memperlihatkan tampilan utama dari sistem untuk memperbaiki pelafalan dalam bahasa Indonesia yang telah dikembangkan dalam studi ini. Tampilan sistem ini terdiri dari tiga komponen yang utama, yaitu *input audio*, daftar kosakata sasaran, serta proses dan hasil analisis. Pada bagian *input audio*, pengguna diberi opsi untuk mengunggah *file* suara dengan format .wav atau merekam suara secara langsung menggunakan mikrofon. Kemudian, pengguna perlu memasukkan kosakata sasaran yang akan digunakan sebagai basis untuk koreksi. Setelah proses ditekan, sistem akan melakukan transkripsi suara dengan *Faster-Whisper*, lalu membandingkan hasil transkripsi tersebut dengan daftar kosakata sasaran memakai *Word Embedding FastText* dan *Cosine Similarity*. Hasil dari proses ini ditampilkan dalam bentuk hasil ASR, kandidat kata dengan tingkat kesamaan tertinggi, kata yang telah diperbaiki, analisis jenis kesalahan pelafalan, rekomendasi stimulasi, dan *audio* pelafalan yang benar yang dihasilkan melalui *Google Text-to-Speech* (TTS). Selain itu, sistem ini juga menawarkan fitur untuk mengeksport hasil pengujian ke dalam format CSV dan mengunduh *audio* hasil dari koreksi.

Tabel 5. Hasil Pengujian Sistem

No	Keterangan	Nilai
1	Jumlah data uji	49
2	Data terkoreksi benar	44
3	Data terkoreksi salah	5
4	Akurasi sistem	89,80%

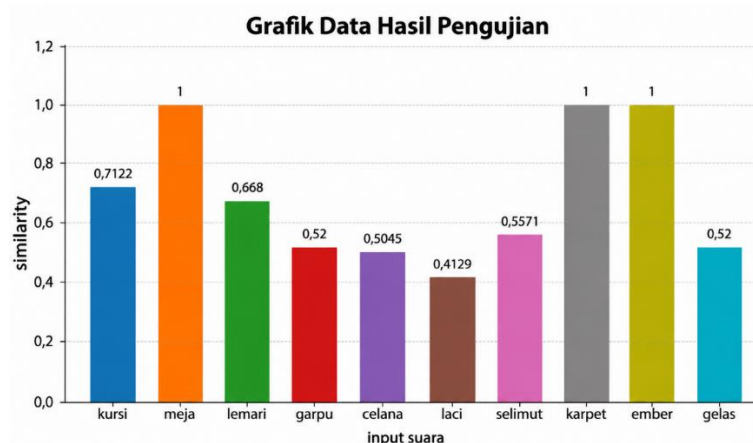
Berdasarkan hasil pengujian pada Tabel 5, diperoleh sebanyak 44 data berhasil dikoreksi sesuai dengan kata target dan 5 data mengalami kesalahan koreksi. Akurasi sistem dihitung menggunakan perbandingan jumlah data yang berhasil dikoreksi dengan jumlah keseluruhan data uji sehingga diperoleh nilai akurasi sebesar 89,80%.

Selain menghasilkan kata koreksi, sistem juga mampu mengidentifikasi jenis kesalahan pelafalan yang dilakukan anak, seperti penghilangan huruf, penambahan huruf, dan substitusi bunyi. Hasil analisis tersebut kemudian digunakan untuk menghasilkan rekomendasi stimulasi pengucapan yang sesuai dengan jenis kesalahan yang terdeteksi.

Berikut beberapa data hasil pengujian:

Tabel 6. Contoh Data Hasil Pengujian

Input Suara	Hasil ASR	Hasil Koreksi	<i>Similarity</i>	Jenis Kesalahan
“kursi”	korsi!	kursi	0.7122	penambahan huruf
“meja”	meja!	meja	1	penambahan huruf
“lemari”	lamari	lemari	0.668	substitusi bunyi
“garpu”	Karpu	Garpu	0.52	substitusi bunyi
“celana”	Jalana	Celana	0.5045	substitusi bunyi
“laci”	Latih!	Laci	0.4129	penambahan huruf

Gambar 3. Grafik Hasil Pengujian *Similarity*

Berdasarkan Gambar 3, nilai *cosine similarity* yang dihasilkan sistem berada pada rentang 0,4129–1,0000. Nilai tersebut menunjukkan tingkat kedekatan antara kata hasil transkripsi ASR dengan kosakata target yang direpresentasikan menggunakan *Word Embedding FastText*. Tiga kata, yaitu meja, karpet, dan ember, memperoleh nilai *similarity* sebesar 1,0000, sedangkan nilai terendah diperoleh pada kata laci sebesar 0,4129. Perbedaan nilai *similarity* menunjukkan bahwa setiap kata memiliki tingkat kedekatan vektor yang berbeda terhadap kosakata target. Semakin tinggi nilai *similarity* yang diperoleh, semakin besar kemungkinan sistem memilih kata target yang sesuai sebagai hasil koreksi. Temuan ini menunjukkan bahwa pendekatan *Word Embedding* dan *Cosine Similarity* dapat digunakan untuk membantu proses koreksi pelafalan bahasa Indonesia pada anak usia dini melalui pencocokan kata berdasarkan kedekatan representasi vektor.

B. Pembahasan

Hasil penelitian menunjukkan bahwa integrasi *Faster-Whisper* dan *Word Embedding FastText* mampu digunakan untuk membangun sistem koreksi pelafalan bahasa Indonesia bagi anak usia dini. *Faster-Whisper* berperan dalam mengubah suara menjadi teks, sedangkan *Word Embedding FastText* digunakan untuk mencari kata yang memiliki tingkat kemiripan tertinggi dengan hasil transkripsi melalui perhitungan *cosine similarity*.

Nilai akurasi sebesar 89,80% menunjukkan bahwa metode yang digunakan mampu memberikan hasil koreksi yang cukup baik pada sebagian besar data uji. Kesalahan yang masih muncul umumnya terjadi ketika hasil transkripsi ASR berbeda jauh dari kata target sehingga nilai *similarity* yang dihasilkan menjadi rendah. Kondisi tersebut menyebabkan sistem memilih kata yang kurang sesuai meskipun masih berada pada kosakata yang tersedia.

Dibandingkan dengan penelitian sebelumnya yang berfokus pada pengenalan ucapan bahasa Indonesia atau klasifikasi suku kata menggunakan model *deep learning*, penelitian ini menawarkan pendekatan yang berbeda karena tidak hanya melakukan pengenalan suara, tetapi juga menyediakan mekanisme koreksi pelafalan berdasarkan kemiripan kata serta rekomendasi stimulasi pengucapan. Dengan demikian, penelitian ini mengisi celah penelitian (*research gap*) pada pengembangan sistem pendukung pelatihan pelafalan bahasa Indonesia untuk anak usia dini yang bersifat interaktif dan mudah digunakan.

Kontribusi utama penelitian ini terletak pada pemanfaatan *Word Embedding FastText* untuk membantu proses koreksi pelafalan setelah tahap pengenalan suara. Pendekatan tersebut memungkinkan sistem memberikan umpan balik yang lebih informatif dibandingkan sistem ASR konvensional yang hanya menghasilkan teks hasil transkripsi tanpa melakukan koreksi kata.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem koreksi pelafalan bahasa Indonesia untuk anak usia dini menggunakan *Faster-Whisper* sebagai *Automatic Speech Recognition* (ASR), *Word Embedding FastText* sebagai representasi kata, *cosine similarity* sebagai metode pengukuran kemiripan, dan *Text-to-Speech* sebagai media umpan balik pelafalan yang benar.

Hasil pengujian terhadap 49 data *audio* menunjukkan bahwa sistem mampu mengoreksi 44 data dengan benar dan menghasilkan tingkat akurasi sebesar 89,80%. Selain melakukan koreksi kata, sistem juga mampu mengidentifikasi jenis kesalahan pelafalan dan memberikan rekomendasi stimulasi pengucapan sesuai dengan kesalahan yang terdeteksi.

Penelitian ini berhasil mengisi celah penelitian yang masih terbatas pada pengembangan sistem koreksi pelafalan bahasa Indonesia untuk anak usia dini. Berbeda dengan penelitian sebelumnya yang umumnya berfokus pada pengenalan suara atau klasifikasi suku kata, penelitian ini mengintegrasikan proses pengenalan suara, koreksi pelafalan, analisis kesalahan, dan rekomendasi stimulasi dalam satu sistem yang terintegrasi. Untuk penelitian selanjutnya, jumlah kosakata dan data *audio* dapat diperluas serta dilakukan pengembangan metode koreksi yang mempertimbangkan aspek fonetik agar tingkat akurasi sistem dapat ditingkatkan.

V. DAFTAR PUSTAKA

- [1] H. Wijaya, "Teknologi Pengenalan Suara tentang Metode , Bahasa dan Tantangan : Systematic Literature Review," vol. 7, no. 2, 2024, doi: 10.32877/bt.v7i2.1888.
- [2] H. Azis, E. P. Indreswari, and R. Wisudawanto, "PEMANFAATAN TEKNOLOGI AUTOMATIC SPEECH RECOGNITION DALAM MENCIPTAKAN PEMBELAJARAN INKLUSIF: IMPLEMENTASI, EFEKTIFITAS DAN TANTANGAN," vol. 5, no. 1, pp. 27–33, 2025, [Online]. Available: <https://doi.org/10.36728/ganesha.v5i1.4171>
- [3] N. Asrin, A. Soraya, S. M. R. Nasution, and N. Hakimah, "Transformasi Pembelajaran Bahasa Indonesia di Sekolah Dasar Melalui Pendekatan Deep Learning," EUNOIA: Jurnal Pendidikan Bahasa Indonesia, vol. 5, no. 2, pp. 145–156, 2025, [Online]. Available: <http://dx.doi.org/10.30821/eunoia.v5i2.4766>
- [4] R. H. Alfando, "KLASIFIKASI TEKS BERITA BERBAHASA INDONESIA MENGGUNAKAN MACHINE LEARNING DAN DEEP LEARNING: STUDI LITERATUR," vol. 7, no. 1, pp. 681–686, 2023, [Online]. Available: <https://doi.org/10.36040/jati.v7i1.6486>
- [5] D. S. Kartawijaya and T. W. Sen, "SINTECH Journal | 95 Analysis Of Neural Network Architectures For Syllable-Based Voice Recognition In Indonesian," vol. 8, no. 2, pp. 95–103, 2025, [Online]. Available: <https://doi.org/10.31598>

- [6] N. A. Mutmaina, “Sistem Pembelajaran Multimedia dengan Teknologi Voice Recognition untuk Berkebutuhan Khusus,” vol. 3, no. 2, pp. 470–473, 2025, [Online]. Available: <https://doi.org/10.61722/jssr.v3i2.4342>
- [7] A. Atikah Putri, S. Agustian, and R. Abdillah, “Penerapan Metode Logistic Regression Untuk Klasifikasi Sentimen Pada Dataset Twitter Terbatas,” *Jurnal Sistem Informasi*, vol. 7, no. 1, pp. 95–107, 2025, [Online]. Available: <https://doi.org/10.31849/zn.v7i1.24804>
- [8] B. E. S. Dewi, “Pengukuran Kemiripan Kalimat Bahasa Indonesia Menggunakan Representasi Word Embedding Fasttext,” *Jurnal Teknologi Informasi dan Digital*, vol. 3, no. 1, pp. 20–29, 2025, [Online]. Available: <https://doi.org/10.65624/tridi.v3i1.99>
- [9] S. Sarif and A. AR, “Efektivitas Artificial Intelligence Text to Speech dalam Meningkatkan Keterampilan Membaca,” *Jurnal Naskhi Jurnal Kajian Pendidikan dan Bahasa Arab*, vol. 6, no. 1, pp. 1–8, 2024, doi: 10.47435/naskhi.v6i1.2697.
- [10] M. R. Waskito, A. D. Rahajoe, and A. L. Nurlaili, “Implementasi Metode Collaborative Filtering Menggunakan Algoritma Cosine Similarity Dan Jaccard Similarity Pada Sistem E-Commerce,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3S1, pp. 4307–4316, 2024, doi: 10.23960/jitet.v12i3s1.5315.
- [11] I. A. N. Fathony et al., “Optimisasi Whisper Speech-to-Text Bahasa Indonesia dengan Hybrid Cloud dan Multi-Engine,” *Elkom: Jurnal Elektronika dan Komputer*, vol. 18, no. 1, pp. 60–72, 2025, [Online]. Available: <https://doi.org/10.51903/5n2d3s08>
- [12] M. D. Rahman, A. Djunaidy, and F. Mahananto, “Penerapan Weighted Word Embedding pada Pengklasifikasian Teks Berbasis Recurrent Neural Network untuk Layanan Pengaduan Perusahaan Transportasi,” *Jurnal Sains dan Seni ITS*, vol. 10, no. 1, 2021, doi: 10.12962/j23373520.v10i1.56145.
- [13] M. Mustasaruddin, E. Budianita, M. Fikry, and F. Yanto, “Klasifikasi Sentiment Review Aplikasi MyPertamina Menggunakan Word Embedding FastText dan SVM (Support Vector Machine),” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 3, p. 526, 2023, doi: 10.30865/json.v4i3.5695.
- [14] D. Iskandar and A. Kurniawati, “Analisis Perbandingan Teknik Word2vec dan Doc2vec dalam Mengukur Kemiripan Dokumen Menggunakan Cosine Similarity,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 1, pp. 133–144, 2025, doi: 10.25126/jtiik.2025129143.