

Mengukur Keterampilan Menyimak Bahasa Indonesia dengan Menggunakan *Artificial Intelligence*

Musaljon

Email: musa_muhammad30@yahoo.com

Universitas Muhammadiyah Bogor Raya

Abstrak

Keterampilan menyimak merupakan salah satu kompetensi dasar dalam penguasaan bahasa, termasuk Bahasa Indonesia. Namun, penilaian keterampilan menyimak secara manual seringkali menghadapi tantangan seperti bias subjektivitas dan waktu penilaian yang lama. Artikel ini bertujuan untuk mengeksplorasi bagaimana teknologi *Artificial Intelligence* (AI) dapat digunakan untuk mengukur keterampilan menyimak Bahasa Indonesia secara efisien dan objektif. Penelitian ini menggunakan pendekatan eksperimen dengan melibatkan sistem AI berbasis pemrosesan bahasa alami (NLP) untuk menilai performa menyimak dari peserta uji. Hasil penelitian menunjukkan bahwa AI mampu memberikan evaluasi yang akurat, efisien, dan sejalan dengan standar penilaian manusia. Selain itu, integrasi AI juga meningkatkan aksesibilitas dan fleksibilitas dalam pengujian keterampilan menyimak. Artikel ini memberikan kontribusi signifikan pada pengembangan teknologi pendidikan berbasis AI, khususnya dalam pembelajaran Bahasa Indonesia.

Kata Kunci: *Artificial Intelligence*, Keterampilan Menyimak, Bahasa Indonesia, Pemrosesan Bahasa Alami, Teknologi Pendidikan

Pendahuluan

Keterampilan menyimak merupakan salah satu kompetensi penting dalam pembelajaran bahasa, termasuk Bahasa Indonesia, yang mendukung kemampuan komunikasi secara menyeluruh. Menyimak tidak hanya melibatkan pemahaman terhadap informasi literal tetapi juga kemampuan untuk menganalisis konteks, menarik kesimpulan, dan memahami makna implisit (Rahmawati, 2021). Dalam konteks pendidikan, penilaian keterampilan menyimak sering kali menjadi tantangan karena melibatkan aspek subjektif dari penilai dan membutuhkan waktu yang signifikan dalam proses evaluasi manual.

Seiring dengan perkembangan teknologi, *Artificial Intelligence* (AI) menawarkan solusi inovatif untuk mengatasi tantangan tersebut. Teknologi berbasis Natural Language Processing (NLP) telah diterapkan untuk mendukung analisis linguistik, mulai dari penerjemahan otomatis hingga penilaian tulisan (Tanaka, H., & Morita, 2020). Namun, penerapan AI untuk menilai keterampilan menyimak dalam konteks Bahasa Indonesia masih jarang dijelajahi secara mendalam.

Literatur Review

Beberapa studi sebelumnya telah mengkaji penerapan teknologi AI dalam penilaian bahasa:

1. **AI dalam Penilaian Bahasa Inggris:** Penelitian oleh Kim et al. (2020) menunjukkan bahwa sistem NLP berbasis AI mampu memberikan penilaian otomatis terhadap tulisan dan kemampuan berbicara dalam Bahasa Inggris dengan tingkat akurasi lebih dari 90%.
2. **AI untuk Bahasa Lokal:** Widodo et al. (2022) menyoroti tantangan dalam melatih model AI untuk bahasa lokal, seperti Bahasa Indonesia, yang memiliki sumber daya data yang terbatas dibandingkan dengan bahasa global.
3. **Keterampilan Menyimak:** Anderson & Zhang (2021) mengembangkan model berbasis AI untuk menilai keterampilan menyimak dalam bahasa Mandarin. Model ini menggunakan analisis semantik untuk mengevaluasi pemahaman terhadap konten audio, tetapi masih menghadapi kendala dalam menilai respons yang lebih kompleks.

State of the Art

Teknologi AI berbasis NLP telah mencapai kemajuan signifikan dalam penilaian keterampilan bahasa, termasuk:

- **Analisis Semantik:** Sistem AI dapat memahami hubungan antar kata dan frasa dalam sebuah teks, memungkinkan evaluasi yang lebih mendalam terhadap respons peserta.
- **Pemahaman Konteks:** Model berbasis Transformer, seperti BERT dan GPT, telah menunjukkan kemampuan luar biasa dalam memahami konteks linguistik.
- **Efisiensi Penilaian:** AI memungkinkan penilaian dalam waktu nyata, mengurangi waktu yang dibutuhkan dibandingkan metode manual.

Meskipun demikian, penelitian pada Bahasa Indonesia masih berada pada tahap awal. Tantangan utama meliputi:

1. **Kurangnya Dataset Berkualitas Tinggi:** Bahasa Indonesia memiliki sumber daya yang lebih sedikit dibandingkan bahasa seperti Inggris dan Mandarin.
2. **Variasi Bahasa:** Bahasa Indonesia memiliki ragam variasi formal dan informal yang memengaruhi performa model NLP.
3. **Evaluasi Kreativitas:** AI masih menghadapi kendala dalam menilai respons kreatif yang tidak memiliki jawaban pasti.

Analisis Gap

Berdasarkan literatur yang ada, terdapat beberapa celah yang perlu dijawab:

1. **Konteks Lokal:** Sebagian besar penelitian AI untuk penilaian keterampilan menyimak berfokus pada bahasa global seperti Inggris dan Mandarin. Penelitian untuk Bahasa Indonesia sangat terbatas.
2. **Kompleksitas Respons:** Studi sebelumnya cenderung fokus pada evaluasi fakta sederhana, sementara keterampilan menyimak yang lebih kompleks (seperti interpretasi konteks dan analisis kritis) kurang dieksplorasi.
3. **Aplikasi dalam Pendidikan:** Meskipun AI telah terbukti efisien, implementasinya dalam platform pendidikan di Indonesia belum optimal.

Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengembangkan model AI berbasis NLP yang mampu menilai keterampilan menyimak Bahasa Indonesia.
2. Mengukur tingkat akurasi, efisiensi, dan objektivitas sistem AI dibandingkan dengan metode manual.
3. Mengeksplorasi tantangan dan potensi integrasi AI dalam sistem pendidikan Bahasa Indonesia.

Metode Penelitian

Penelitian ini menggunakan pendekatan eksperimen kuasi dengan desain *pre-test* dan *post-test*. Sistem AI yang digunakan adalah perangkat berbasis NLP yang dirancang untuk menganalisis respons peserta uji terhadap pertanyaan menyimak.

1. Subjek Penelitian

Sebanyak 100 peserta dari mahasiswa prodi Pendidikan bahasa Indonesia Universitas Muhammadiyah Bogor Raya yang dipilih secara acak.

2. Instrumen Penelitian

- Teks audio berbasis narasi dengan berbagai tingkat kompleksitas.
- Sistem penilaian berbasis AI yang mencakup analisis semantik dan sintaksis.

3. Prosedur Penelitian

Penelitian ini dirancang untuk mengevaluasi kemampuan sistem AI dalam mengukur keterampilan menyimak Bahasa Indonesia secara objektif. Prosedur penelitian melibatkan beberapa tahapan, yaitu sebagai berikut:

1. Pengembangan Instrumen Penilaian

Instrumen penilaian menyimak terdiri dari:

- **Teks Audio:** Materi menyimak berupa teks narasi yang telah dirancang dengan berbagai tingkat kesulitan (mudah, sedang, sulit). Teks ini mencakup tema-tema umum seperti budaya, pendidikan, dan sains untuk memastikan relevansi konteks dengan peserta uji. Teks diadaptasi dari sumber-sumber yang valid, seperti buku teks Bahasa Indonesia dan artikel populer.
- **Pertanyaan Evaluasi:** Setiap teks audio disertai dengan 10-15 pertanyaan, baik dalam format pilihan ganda maupun isian singkat. Pertanyaan ini dirancang untuk mengukur pemahaman peserta terhadap informasi faktual, inferensi, dan konteks narasi.

2. Persiapan Sistem AI

Sistem AI yang digunakan dalam penelitian ini adalah model berbasis Natural Language Processing (NLP) yang telah dilatih menggunakan dataset Bahasa Indonesia. Dataset mencakup:

- **Korpus Bahasa Indonesia:** Berisi berbagai teks naratif, dialog, dan percakapan formal serta informal.
- **Dataset Penilaian Menyimak:** Dataset khusus yang memetakan respons menyimak peserta terhadap teks audio tertentu.

Sistem AI dilengkapi dengan kemampuan analisis semantik (untuk memahami makna), analisis sintaksis (untuk memahami struktur), dan evaluasi kontekstual (untuk memahami hubungan antar informasi).

3. Tahap Pelaksanaan Penelitian

Pelaksanaan penelitian dilakukan dalam tiga sesi:

- **Sesi 1: Pengujian Awal (Pre-test)**
Peserta mendengarkan teks audio melalui perangkat elektronik, seperti komputer atau ponsel, yang telah dilengkapi dengan aplikasi pengujian. Mereka diminta menjawab pertanyaan secara tertulis atau lisan setelah mendengarkan setiap teks. Jawaban direkam untuk dianalisis secara manual dan menggunakan AI.
- **Sesi 2: Pelatihan Sistem AI**
Sebelum tahap pengujian utama, sistem AI diuji dengan data uji awal untuk memvalidasi keakuratannya. Hasil ini dibandingkan dengan penilaian manual untuk memastikan bahwa AI dapat memahami nuansa Bahasa Indonesia dengan baik.
- **Sesi 3: Pengujian Utama (Post-test)**
Peserta diuji kembali dengan teks audio berbeda namun setara dalam tingkat kesulitan. Proses yang sama dilakukan untuk pengumpulan respons peserta. Semua jawaban langsung dianalisis oleh sistem AI.

4. Proses Analisis Data

Analisis data melibatkan beberapa tahap berikut:

- **Analisis Respons oleh Sistem AI:** AI mengevaluasi respons peserta berdasarkan kriteria seperti:
 - *Kesesuaian makna:* Apakah jawaban peserta sesuai dengan inti informasi dalam teks audio.
 - *Ketepatan struktur:* Apakah struktur bahasa dalam respons sesuai dengan kaidah Bahasa Indonesia.
 - *Kelengkapan informasi:* Apakah peserta menangkap semua elemen penting dari teks yang diperdengarkan.
- **Validasi Hasil AI:** Skor yang diberikan oleh AI dibandingkan dengan skor manual dari tim ahli bahasa untuk menilai tingkat akurasi sistem.

5. Pengumpulan Feedback Peserta

Setelah sesi pengujian, peserta diminta mengisi kuesioner untuk memberikan umpan balik terkait pengalaman mereka menggunakan sistem AI dalam penilaian keterampilan menyimak. Hal ini untuk mengevaluasi aspek *usability* dan *user experience* dari teknologi yang digunakan.

6. Pengolahan dan Penyajian Data

Hasil penelitian diproses menggunakan metode statistik deskriptif dan inferensial. Analisis korelasi digunakan untuk menilai kesesuaian antara skor AI dan skor manual, sementara analisis waktu digunakan untuk membandingkan efisiensi AI dengan metode konvensional.

Hasil dan Pembahasan

Penelitian ini menghasilkan sejumlah temuan yang relevan terkait penggunaan *Artificial Intelligence* (AI) dalam mengukur keterampilan menyimak Bahasa Indonesia. Temuan ini mencakup analisis akurasi penilaian, efisiensi waktu, serta aspek konsistensi dan objektivitas.

1. Akurasi Penilaian AI vs Penilaian Manual

Untuk mengukur akurasi sistem AI, hasil penilaian dari AI dibandingkan dengan penilaian manual oleh dua ahli bahasa yang bertindak sebagai *gold standard*. Berikut data perbandingannya:

Kategori	Skor Rata-rata (Manual)	Skor Rata-rata (AI)	Korelasi (r)
Pemahaman Faktual	85,4%	84,8%	0,93
Inferensi	78,2%	77,5%	0,91
Konteks Narasi	80,5%	79,7%	0,92

Interpretasi Hasil:

Korelasi yang tinggi ($r > 0,90$) menunjukkan bahwa sistem AI memiliki tingkat akurasi yang hampir setara dengan penilaian manual dalam semua kategori keterampilan menyimak. Hal ini menunjukkan kemampuan AI untuk memahami dan mengevaluasi nuansa Bahasa Indonesia dengan baik.

2. Efisiensi Waktu Penilaian

Efisiensi waktu dianalisis berdasarkan rata-rata waktu yang diperlukan untuk menilai setiap respons peserta:

Metode	Waktu per Respons	Total Waktu (100 Respons)
Penilaian Manual	45 detik	4500 detik (75 menit)
Penilaian dengan AI	3 detik	300 detik (5 menit)

Interpretasi Hasil:

AI terbukti 15 kali lebih cepat dibandingkan penilaian manual. Hal ini menjadikan AI sebagai solusi yang sangat efisien, terutama untuk skala besar seperti ujian nasional atau sertifikasi bahasa.

3. Objektivitas dan Konsistensi Penilaian

Objektivitas diukur dengan menghitung deviasi standar skor antara dua penilai manual dan AI.

Metode	Deviasi Standar Skor
Penilaian Manual	$\pm 3,5$
Penilaian dengan AI	$\pm 0,8$

Interpretasi Hasil:

Sistem AI memberikan skor yang lebih konsisten dibandingkan penilai manusia. Ini menunjukkan bahwa AI mampu mengurangi bias subjektivitas dalam penilaian.

4. Feedback Peserta

Sebanyak 85% peserta menyatakan bahwa mereka merasa nyaman dengan proses penilaian berbasis AI. Beberapa poin penting dari kuesioner:

- **Responsif terhadap teks audio:** 78% peserta merasa bahwa sistem AI mampu memahami jawaban mereka secara akurat.
- **Kemudahan penggunaan:** 82% peserta menyatakan bahwa platform berbasis AI mudah digunakan.
- **Kekhawatiran terhadap teknologi:** 15% peserta khawatir AI mungkin tidak mampu menilai respons yang lebih kompleks atau kreatif.

5. Analisis Kualitatif dari Jawaban Peserta

Sistem AI menunjukkan performa yang sangat baik dalam menilai respons faktual dan inferensi langsung. Namun, pada pertanyaan yang melibatkan opini atau interpretasi kreatif, AI kadang memberikan skor lebih rendah dibandingkan penilaian manual. Hal ini menandakan bahwa AI masih memerlukan pengembangan lebih lanjut dalam menangani variasi respons yang tidak langsung.

Pembahasan

Kekuatan Teknologi AI

Penelitian ini menegaskan potensi AI sebagai alat yang efektif untuk mengukur keterampilan menyimak Bahasa Indonesia. Dengan efisiensi waktu dan tingkat akurasi yang tinggi, AI dapat diandalkan untuk menggantikan metode manual, terutama dalam skenario skala besar.

Keterbatasan Teknologi AI

Namun, sistem AI saat ini masih menghadapi tantangan dalam menangani pertanyaan yang

membutuhkan interpretasi kreatif. Hal ini menunjukkan bahwa, meskipun AI dapat meningkatkan efisiensi, peran manusia tetap diperlukan untuk beberapa aspek tertentu.

Implikasi untuk Pengembangan Pendidikan

Integrasi AI dalam penilaian keterampilan menyimak dapat menjadi inovasi signifikan dalam sistem pendidikan. Selain mempercepat proses evaluasi, teknologi ini juga dapat membantu institusi pendidikan menyediakan umpan balik yang lebih objektif kepada siswa.

Rekomendasi Pengembangan

1. **Dataset yang Lebih Beragam:** Melatih sistem AI dengan dataset yang mencakup variasi respons kreatif dan interpretatif.
2. **Kolaborasi Peneliti:** Mengundang kolaborasi antara ahli bahasa dan pengembang AI untuk menyempurnakan algoritma NLP.
3. **Integrasi ke Platform Pendidikan:** Mengimplementasikan teknologi ini pada platform ujian daring untuk meningkatkan skala penggunaannya.

Kesimpulan dan Rekomendasi

Hasil analisis menunjukkan bahwa skor yang dihasilkan oleh sistem AI memiliki korelasi tinggi ($r = 0,92$) dengan skor yang diberikan oleh penilai manusia. Hal ini menunjukkan bahwa AI mampu menilai keterampilan menyimak dengan tingkat akurasi yang hampir setara dengan manusia. Penilaian berbasis AI memerlukan rata-rata waktu 3 detik per respons, dibandingkan dengan 45 detik yang diperlukan oleh penilai manusia. Sistem AI menghilangkan potensi bias subjektivitas, sehingga memberikan hasil yang lebih konsisten.

Penelitian ini menunjukkan bahwa AI berbasis NLP memiliki potensi besar dalam mengukur keterampilan menyimak Bahasa Indonesia. Sistem ini tidak hanya akurat dan efisien tetapi juga mengurangi bias subjektivitas dalam penilaian.

Sebagai langkah lanjutan, disarankan untuk mengembangkan dataset Bahasa Indonesia yang lebih komprehensif untuk melatih sistem AI. Selain itu, integrasi teknologi ini ke dalam platform pendidikan nasional dapat menjadi inovasi signifikan untuk meningkatkan kualitas pembelajaran Bahasa Indonesia.

Referensi

- Anderson, J., & Zhang, Y. (2021). AI-based Listening Assessment in Mandarin: Challenges and Opportunities. *Journal of Computational Linguistics and Language Education*, 12(4), 45–63.
- Kim, S., Park, J., & L. (2020). Enhancing Language Assessment with AI: A Focus on Writing and Speaking Skills. *International Journal of Language Technology*, 8(2), 67–89.
- Rahmawati, N. (2021). Penilaian Keterampilan Menyimak dalam Pembelajaran Bahasa Indonesia. *Jurnal Pendidikan Bahasa*, 5(2), 123–135.

- Tanaka, H., & Morita, K. (2020). AI Applications in Language Assessment: A Review. *International Journal of AI in Education*, 30(1), 1–20.
- Widodo, A., et al. (2022). NLP for Indonesian Languages: Opportunities and Challenges. *Indonesian Computational Linguistics Journal*, 10(3), 215–230.

KUESIONER

Kuesioner ini bertujuan untuk mengumpulkan data mengenai pengalaman peserta dalam menggunakan sistem AI untuk penilaian keterampilan menyimak Bahasa Indonesia. Hasil dari kuesioner ini akan digunakan untuk meningkatkan pengembangan sistem AI di masa depan.

Bagian 1: Identitas Responden

1. Nama (opsional): _____
2. Usia: _____
3. Jenis Kelamin:
 - Laki-laki
 - Perempuan
 - Lainnya: _____
4. Tingkat Pendidikan:
 - SMA/ sederajat
 - Diploma
 - Sarjana
 - Pascasarjana
 - Lainnya: _____

Bagian 2: Pengalaman dengan Sistem AI

Skala Likert 1–5

(1: Sangat Tidak Setuju, 5: Sangat Setuju)

1. Sistem AI mudah digunakan.
 - 1
 - 2
 - 3
 - 4
 - 5
2. Penilaian oleh AI terasa akurat dan sesuai dengan kemampuan saya.
 - 1
 - 2
 - 3
 - 4
 - 5

3. Sistem AI memberikan umpan balik yang jelas dan mudah dipahami.
- 1
 - 2
 - 3
 - 4
 - 5
4. Proses penilaian menggunakan AI lebih cepat dibandingkan metode manual.
- 1
 - 2
 - 3
 - 4
 - 5
5. Saya merasa percaya diri dengan hasil penilaian yang diberikan oleh sistem AI.
- 1
 - 2
 - 3
 - 4
 - 5

Bagian 3: Saran dan Masukan

1. Apa hal yang paling Anda sukai dari sistem AI ini?

2. Apa tantangan atau kesulitan yang Anda hadapi saat menggunakan sistem ini?

3. Apakah Anda memiliki saran untuk meningkatkan sistem AI ini?
