INTEGRASI WHISPER DAN GPT-4O SEBAGAI PENDEKATAN DALAM ANALISIS PERKEMBANGAN KOMUNIKASI DAN POLA BAHASA PADA *AUTISM SPECTRUM DISORDER*

Disusun Oleh:

N a m a : Naela Fauzul Muna
NIM : 21523238

PROGRAM STUDI INFORMATIKA – PROGRAM SARJANA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM INDONESIA
2025

HALAMAN PENGESAHAN DOSEN PEMBIMBING

INTEGRASI WHISPER DAN GPT-4O SEBAGAI PENDEKATAN DALAM EVALUASI PERKEMBANGAN KOMUNIKASI DAN POLA BAHASA PADA *AUTISM SPECTRUM DISORDER*

TUGAS AKHIR

Disusun Oleh:

N a m a : Naela Fauzul Muna

NIM : 21523238

Yogyakarta, 16 Januari 2025

Pembimbing,

(Mukhammad Andri Setiawan, S.T., M.Sc., Ph.D.)

HALAMAN PENGESAHAN DOSEN PENGUJI

INTEGRASI WHISPER DAN GPT-4O SEBAGAI PENDEKATAN DALAM EVALUASI PERKEMBANGAN KOMUNIKASI DAN POLA BAHASA PADA *AUTISM SPECTRUM DISORDER*

TUGAS AKHIR

Telah dipertahankan di depan sidang penguji sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer dari Program Studi Informatika – Program Sarjana di Fakultas Teknologi Industri Universitas Islam Indonesia

Yogyakarta, 26-05-2025

Tim Penguji

Mukhammad A. Setiawan, S.T., M.Sc., Ph.D. _____

Anggota 1

Sheila Nurul Huda, S.Kom., M.Cs. _____

Anggota 2

Rahadian Kurniawan, S.Kom., M.Kom. _____

Mengetahui,
Ketua Program Studi Informatika – Program Sarjana
Fakultas Teknologi Industri
Universitas Islam Indonesia

(Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D.)

HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini:

Nama : Naela Fauzul Muna

NIM : 21523238

Tugas akhir dengan judul:

INTEGRASI WHISPER DAN GPT-4O SEBAGAI PENDEKATAN DALAM EVALUASI PERKEMBANGAN KOMUNIKASI DAN POLA BAHASA PADA *AUTISM SPECTRUM DISORDER*

Menyatakan bahwa seluruh komponen dan isi dalam tugas akhir ini adalah hasil karya saya sendiri. Apabila di kemudian hari terbukti ada beberapa bagian dari karya ini adalah bukan hasil karya sendiri, tugas akhir yang diajukan sebagai hasil karya sendiri ini siap ditarik kembali dan siap menanggung risiko dan konsekuensi apapun.

Demikian surat pernyataan ini dibuat, semoga dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 14 Januari 2025

(Naela Fauzul Muna)

HALAMAN PERSEMBAHAN

Penulis mempersembahkan laporan penelitian ini sebagai wujud penghargaan untuk diri sendiri, yang telah berjuang tanpa henti menyelesaikan penelitian ini di tengah berbagai tantangan dan hambatan. Penelitian ini juga dipersembahkan dengan penuh cinta untuk keluarga tersayang, yaitu bapak, ibu, dan mamas yang selalu memberikan doa, dukungan moral, dan materi dengan tulus.

HALAMAN MOTO

Take the risk or lose the chance.

KATA PENGANTAR

Dengan penuh rasa syukur kepada Tuhan Yang Maha Esa, penulis mengucapkan terima kasih atas rahmat dan hidayah-Nya, sehingga laporan tugas akhir yang berjudul ″Integrasi Whisper Dan GPT-4o Sebagai Pendekatan Dalam Analisis Perkembangan Komunikasi dan Pola Bahasa Pada *Autism Spectrum Disorder*″ ini dapat diselesaikan dengan baik.

Laporan ini merupakan salah satu syarat untuk menyelesaikan pendidikan di Program Studi Informatika, Universitas Islam Indonesia. Proses penyusunan laporan ini tidak terlepas dari berbagai kendala dan tantangan, baik yang bersifat teknis maupun non-teknis. Namun, berkat kerja keras, dedikasi, dan bantuan dari berbagai pihak, penulis dapat menyelesaikannya dengan baik. Penulis ingin menyampaikan rasa terima kasih yang mendalam kepada:

1. Allah SWT, Tuhan Yang Maha Esa, atas kasih dan penyertaan-Nya dalam setiap langkah penulis.
2. Keluarga tercinta yang selalu memberikan doa, dukungan moral, dan materi selama penulis menjalani proses penyusunan laporan ini.
3. Bapak Mukhammad Andri Setiawan, S.T., M.Sc., Ph.D., yang telah memberikan arahan, bimbingan, dan motivasi selama proses penelitian dan penulisan laporan ini.
4. Seluruh dosen di Program Studi Informatika, yang telah membekali penulis dengan ilmu dan wawasan yang menjadi fondasi dalam menyelesaikan penelitian ini.
5. Sahabat-sahabat penulis, Tanziila, Siraj, Juan, Kharisma, Asista, Nabila, Zahwa, Ihwan, Nabil, Bintang, Anra, Satriya, Hanuga, Adit, dan Daud, atas kebersamaan, serta semangat yang telah membantu penulis menyelesaikan tugas akhir ini.
6. Seluruh guru dan siswa dari SLB Dian Amanah, SLB Fajar Nugraha, SLB Bina Anggita, dan SLB Fredofios, yang telah berbaik hati memberikan izin dan kerja sama selama pelaksanaan penelitian.

Yogyakarta, 10 Januari 2025

(Naela Fauzul Muna)

SARI

Autism Spectrum Disorder (ASD) adalah gangguan perkembangan yang memengaruhi komunikasi, interaksi sosial, dan perilaku. Penilaian komunikasi pada anak dengan ASD umumnya dilakukan secara manual, sehingga memakan waktu, yang dapat menyebabkan keterlambatan dalam pengembangan program pendidikan dan kurangnya standarisasi akibat evaluasi yang bersifat subjektif. Penelitian ini bertujuan untuk mengembangkan kerangka kerja otomatis (*automated framework)* berbasis Whisper dan GPT-4o guna meningkatkan efisiensi dan akurasi evaluasi kemampuan komunikasi dan pola bahasa pada anak dengan ASD.

Penelitian ini menggunakan pendekatan *Research and Development* (RnD) yang melibatkan anak dengan ASD (kategori verbal ringan dan sedang) serta guru dari empat sekolah autisme di Daerah Istimewa Yogyakarta. Data dikumpulkan melalui wawancara, observasi kelas, rekaman audio, dan instrument penilaian. Whisper digunakan untuk transkripsi otomatis yang terintegrasi dengan GPT-4o untuk diarisasi pembicara dan analisis komunikasi. Kombinasi alat ini menghasilkan pengurangan waktu analisis hingga 89,1% dibandingkan metode manual. Whisper mencapai *Word Error Rate* (WER) rendah untuk anak dengan autisme ringan (rata-rata 5%) dan tingkat kesalahan lebih tinggi untuk autisme sedang (rata-rata 23%). GPT-4o berkontribusi dalam proses dengan akurasi diarisasi pembicara yang tinggi (93,9% untuk autisme ringan dan 89,2% untuk autisme sedang).

Framework ini berhasil mengidentifikasi peningkatan komunikasi secara rinci melalui instrument penilaian, meliputi aspek verbal, pragmatik, semantik, struktur kalimat, dan ekolalia (pengulangan kata) berdasarkan validasi dari psikolog klinis dan guru sekolah luar biasa. Kerangka ini juga memberikan wawasan yang sebelumnya tidak terdeteksi oleh guru, seperti pola perkembangan spesifik pada setiap aspek tersebut.

Kata kunci: *Autism Spectrum Disorder*, Penilaian Komunikasi dan Pola Bahasa, *Automated Framework*, Whisper *Automatic Speech Recognition*, GPT *Large Language Model*.

GLOSARIUM

Ekolalia	pengulangan kata atau frasa yang diucapkan oleh orang lain.
Waterfall	metode pengembangan perangkat lunak.

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PENGESAHAN DOSEN PEMBIMBING ii
HALAMAN PENGESAHAN DOSEN PENGUJI iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR iv
HALAMAN PERSEMBAHAN v
HALAMAN MOTO vi
KATA PENGANTAR vii
SARI viii
GLOSARIUM ix
DAFTAR ISI x
DAFTAR TABEL xii
DAFTAR GAMBAR xiii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah dan Pertanyaan Penelitian 3
1.3 Batasan Masalah 3
1.4 Tujuan Penelitian 3
1.5 Manfaat Penelitian 3
1.6 Metodologi Penelitian 4
1.6.1 Metode Pengumpulan Data 4
1.6.2 Metode Perancangan Sistem 4
1.7 Sistematika Penulisan 5
BAB II KAJIAN PUSTAKA 7
2.1 Landasan Teori 7
2.1.1 Autisme 7
2.1.2 Pengenalan Suara (*Speech Recognition*) 9
2.1.3 Model Pengenalan Suara Whisper 10
2.1.4 Model Bahasa (*Large Language Model*) 13
2.1.5 Model Bahasa GPT 13
2.2 Penelitian Terdahulu 17
BAB III METODOLOGI PENELITIAN 23
3.1 Populasi dan Sampel 23
3.2 Metode Pengumpulan Data 23
3.2.1 Studi Pustaka 23
3.2.2 Studi Lapangan 23
3.3 Metode Perancangan Sistem 24
3.3.1 Analisis Kebutuhan 25
3.3.2 Desain Sistem 25
3.3.3 Penulisan Kode Program 26
3.3.4 Pengujian 26
3.3.5 Pemeliharaan 26
3.4 Kerangka Penelitian 26
BAB IV HASIL DAN PEMBAHASAN 28
4.1 Analisis Kebutuhan 28
4.1.1 Identifikasi Masalah 28
4.1.2 Usulan Pemecahan Masalah 28
4.1.3 Analisis Instrumen Penilaian (*Matrix Based Evaluation*) 28

4.1.4 Analisis Kebutuhan Sistem 30
4.2 Desain Sistem 33
4.2.1 Diagram Alir Sistem (*Flowchart*) 33
4.3 Penulisan Kode Program 34
4.3.1 *Automatic Speech Recognition* dengan Whisper 34
4.3.2 Diarisasi Pembicara dengan GPT-4o 35
4.3.3 Analisis Perkembangan Komunikasi dengan GPT-4o 40
4.3.4 Visualisasi Data Perkembangan Komunikasi 42
4.4 Pengujian 50
4.4.1 Hasil Pengujian ASR Whisper 50
4.4.2 Hasil Pengujian Diarisasi Pembicara 52
4.4.3 Validasi dan Evaluasi *Framework* Perbandingan Analisis Waktu 53
4.5 Pemeliharaan 65
BAB V KESIMPULAN DAN SARAN 65
5.1 Kesimpulan 65
5.2 Saran 66
DAFTAR PUSTAKA 67

DAFTAR TABEL

Tabel 2.1 Detail Model Whisper 11
Tabel 2.2 Perbandingan Alat Ukur Komunikasi ASD dari Penelitian Terdahulu 18

Tabel 4.1 *Matrix-Based Evaluation* Komunikasi & Bahasa Siswa ASD 29
Tabel 4.2 Perbandingan Waktu Analisis Antara Metode Manual dan Metode Otomatis (*Automated Framework*) 53
Tabel 4.3 Perbandingan Analisis Komunikasi Antara Metode Manual 54
Tabel 4.4 Perbandingan Analisis Komunikasi Antara Metode Manual 58

DAFTAR GAMBAR

Gambar 2.1 Arsitektur Whisper (Open AI, 2022). 12
Gambar 2.2 Arsitektur Transformer 15
Gambar 2.3 Proses Kerja *Generative Pre-Trained Transformer* (GPT) (Taskus, 2024). 15

Gambar 3.1 Tahapan Metode *Waterfall* 25
Gambar 3.2 Kerangka Penelitian 27

Gambar 4.1 Diagram Alir Sistem 33
Gambar 4.2 Kode Pengolahan Audio dengan M4A 34
Gambar 4.3 Kode Proses Transkripsi dengan Whisper 35
Gambar 4. 4 Kode Mengonversi JSON ke Model Data 36
Gambar 4.5 Kode Membagi Transkripsi Menjadi Beberapa Bagian 38
Gambar 4.6 *Prompt* pada Proses Diarisasi Pembicara 40
Gambar 4.7 Kode & *Prompt* pada Analisis Perkembangan Komunikasi 42
Gambar 4.8 Kode untuk Memisahkan Kosa Kata 44
Gambar 4.9 Tampilan Kosa Kata & Persentase Perkembangannya 45
Gambar 4.10 Tampilan Daftar Kosa Kata 45
Gambar 4.11 Kode Mengolah Transkripsi dengan Regex 47
Gambar 4.12 Kode untuk Menghitung Respon Relevan & Tidak Relevan 47
Gambar 4.13 Tampilan Perbandingan Respon Relevan & Tidak Relevan 48
Gambar 4.14 Kode untuk Menghitung Perkembangan Respon Relevan & Tidak Relevan 49
Gambar 4.15 Tampilan Tren Perkembangan Respon Relevan & Tidak Relevan 49
Gambar 4.16 Kode Pemrograman Menghitung Frekuensi Ekolalia 50
Gambar 4.17 Tampilan Visualisasi Data Frekuensi Ekolalia 50
Gambar 4.18 Hasil *Word Error Rate* dari Whisper 51
Gambar 4.19 Hasil Akurasi Diarisasi Pembicara oleh GPT-4o 52

BAB I
PENDAHULUAN

1.1 Latar Belakang

Autism Spectrum Disorder (ASD) atau sering disebut autis merupakan disabilitas perkembangan neurologis kompleks yang ditandai oleh gangguan kognisi, komunikasi, perilaku atau keterampilan motorik akibat perkembangan otak yang tidak normal (American Psychiatric Association, 2013). Diagnostik autisme didasarkan pada analisis dan pengamatan perilaku dan ciri khas individu, dengan memperhatikan berbagai aspek seperti komunikasi, interaksi sosial, dan perilaku repetitif (Harrison et al., 2021).

Permasalahan utama yang sering dihadapi individu dengan autisme mencakup kesulitan dalam komunikasi, interaksi sosial, dan aktivitas imajinasi, seperti yang dikemukakan oleh (Mansur, 2016). Pendapat ini didukung oleh penelitian yang dilakukan oleh (Vogindroukas et al., 2022) yang menunjukkan bahwa anak-anak dengan ASD sering menghadapi tantangan dalam berbagai aspek bahasa, termasuk pragmatik, tata bahasa, semantik, sintaks, fonologi, dan morfologi, baik dalam bentuk lisan maupun tulisan.

Autisme dikelompokkan ke dalam beberapa tingkatan yang mencakup autisme ringan, autisme sedang, dan autisme berat. *Childhood Autism Rating Scale* (CARS) membagi autisme menjadi tiga tingkatan berdasarkan tingkat gejala yang ditunjukkan oleh individu: (1) autisme ringan, pada tingkat ini, individu dengan autisme masih dapat menunjukkan sedikit kontak mata, memberikan respons ketika dipanggil namanya, dan kadang-kadang menunjukkan ekspresi wajah serta kemampuan berkomunikasi dua arah, meskipun tidak secara konsisten. (2) autisme sedang, individu dengan tingkat autisme sedang masih menunjukkan sedikit kontak mata dan memberikan respons ketika dipanggil namanya, namun mungkin memiliki sedikit ekspresi wajah dan interaksi sosial yang lebih terbatas daripada individu dengan autisme ringan. (3) autisme berat, pada tingkat ini, individu dengan autisme tidak dapat melakukan kontak mata, berekspresi, dan berinteraksi (Mujiyanti, 2011).

Angka penderita autisme secara global terus meningkat. Menurut data dari *World Health Organization* (WHO) pada Maret 2023, prevalensi autisme mencapai 1:100 rata-rata secara global. Di Amerika Serikat, menurut laporan dari *(Centers for Disease Control and Prevention (CDC), 2024)*, sekitar 1 dari 36 anak didiagnosis dengan autisme. Temuan

tersebut lebih tinggi daripada perkiraan sebelumnya pada tahun 2018 yang menemukan prevalensi 1 dari 44 anak yang didiagnosis autisme.

Dalam penelitian yang dilakukan oleh (Rahmawati, 2024) memaparkan bahwa di Indonesia, menurut Badan Pusat Statistik (BPS) tahun 2020, populasi saat ini mencapai sekitar 270,2 juta jiwa dengan perkiraan sekitar 3,2 juta anak yang menderita autisme. Data dari Pusat Data Statistik Sekolah Luar Biasa menunjukkan bahwa jumlah siswa autisme di Indonesia pada tahun 2019 sebanyak 144.102 siswa, yang mengalami peningkatan dibandingkan dengan tahun 2018 yang tercatat sebanyak 133.826 siswa penderita autisme.

Peningkatan jumlah kasus autisme dan perbedaan dalam fitur bahasa pada individu dengan autisme menunjukkan adanya anomali yang bisa memberikan informasi penting untuk diagnosis. Namun, menurut penelitian yang dilakukan oleh (O'Sullivan et al., 2023) Analisis data komunikasi pada individu dengan autisme selama ini masih dilakukan melalui transkripsi manual oleh profesional terlatih, yang memakan banyak waktu dan biaya. Oleh karena itu, diperlukan pengembangan alat atau teknologi baru yang bisa membantu ahli dalam mengidentifikasi autisme. Salah satu caranya adalah dengan memanfaatkan teknologi *Speech-to-Text* Whisper dan *Large Language Model* GPT dari OpenAI.

Teknologi *Speech-to-text* Whisper adalah sistem model pengenalan suara yang mampu mentranskripsikan audio menjadi teks dengan tingkat akurasi yang tinggi. Menurut penelitian oleh (Ghale et al., 2022), Whisper menunjukkan kinerja yang kuat dan konsisten dalam mengenali suara, bahkan dalam situasi di mana kualitas suara rendah atau terdapat gangguan audio. Di sisi lain, *Large Language Model* (LLM) GPT adalah model bahasa berbasis teknologi *machine learning* yang mampu memahami dan menghasilkan teks mirip bahasa manusia. Dalam penelitian oleh (Yadgar N. Abbas et al., 2023), GPT berhasil mendiagnosis 15 dari 20 kasus gangguan psikiatri yang dirancang oleh psikiater dengan keakuratan 75%, hal ini menunjukkan potensi teknologi GPT dalam bidang diagnostik medis.

Oleh karena itu, integrasi teknologi Whisper dan GPT berpotensi menghasilkan alat atau teknologi yang dapat membantu para ahli dalam mengukur perkembangan komunikasi anak autisme berdasarkan karakteristik bicara dan pola bahasa yang digunakan oleh individu dengan autisme. Penelitian ini diharapkan memberikan kontribusi signifikan dalam pengembangan metode diagnostik autisme yang berbasis teknologi terutama kecerdasan buatan.

1.2 Rumusan Masalah dan Pertanyaan Penelitian

Berdasarkan latar belakang yang telah disampaikan, masalah utama dalam penelitian ini adalah bagaimana membantu guru SLB (Sekolah Luar Biasa) dalam mengukur peningkatan kemampuan komunikasi individu dengan autisme melalui pengembangan sebuah *automated framework* yang mampu mengidentifikasi perkembangan komunikasi berdasarkan karakteristik bicara dan pola bahasa individu dengan autisme, dengan memanfaatkan teknologi pengenalan suara Whisper dan model bahasa GPT-4o.

1.3 Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan yang ditetapkan untuk memperjelas ruang lingkup dan fokus penelitian. Batasan-batasan ini adalah sebagai berikut:

a. Penelitian ini berfokus pada analisis pola bahasa dan komunikasi, tanpa melibatkan aspek lainnya, termasuk perilaku pada individu dengan autisme.
b. Penelitian ini hanya dilakukan pada individu dengan autisme verbal.
c. Penelitian ini menggunakan teknologi Whisper dari OpenAI untuk transkripsi suara ke teks, namun Whisper tidak dapat mendeteksi intonasi dan artikulasi secara spesifik. Fokus utama adalah pada pengenalan kata-kata yang diucapkan dan bukan pada analisis prosodi seperti intonasi dan artikulasi.

1.4 Tujuan Penelitian

Tujuan yang hendak dicapai dari penelitian ini adalah:

a. Membantu guru SLB memantau perkembangan komunikasi individu dengan autisme melalui analisis karakteristik bicara dan pola bahasa yang digunakan oleh individu tersebut.
b. Menganalisis efektivitas teknologi Whisper dalam mengenali dan mentranskripsikan ucapan dari individu dengan autisme.
c. Mengkaji kemampuan model GPT-4o dalam memproses dan memahami pola bahasa yang unik dari individu dengan autisme serta mengaplikasikannya untuk menganalisis perkembangan komunikasi dan pola bahasa.

1.5 Manfaat Penelitian

Penelitian ini memiliki manfaat bagi beberapa pihak, yaitu:

a. Bagi Peneliti

1. Memanfaatkan ilmu yang diperoleh selama masa kuliah untuk mengatasi tantangan dalam kehidupan sehari-hari.
2. Merancang sebuah kerangka kerja otomatis untuk membantu guru-guru Sekolah Luar Biasa (SLB) dalam mengukur perkembangan komunikasi pada individu dengan ASD berdasarkan karakteristik bicara dan pola bahasa yang digunakan.

b. Bagi Guru Sekolah Luar Biasa (SLB)

1. Kerangka kerja otomatis yang dikembangkan ini dapat membantu dan mempermudah guru SLB dalam mengukur perkembangan komunikasi siswa penyandang ASD berdasarkan karakteristik bicara dan pola bahasa.
2. *Matrix-Based Evaluation* yang dikembangkan dapat menjadi standar untuk mengukur perkembangan komunikasi siswa penyandang ASD.

1.6 Metodologi Penelitian

Metodologi penelitian adalah serangkaian langkah yang diikuti oleh peneliti mulai dari tahap perumusan masalah hingga mencapai kesimpulan, yang menciptakan sebuah proses sistematis. Metodologi ini berperan sebagai panduan bagi peneliti selama proses penelitian untuk memastikan bahwa hasil yang diperoleh sesuai dengan tujuan yang ditetapkan.

1.6.1 Metode Pengumpulan Data

Metodologi pengumpulan data adalah serangkaian prosedur sistematis yang digunakan untuk mengumpulkan informasi yang relevan dan valid dalam rangka menjawab pertanyaan penelitian atau mencapai tujuan studi.

a. Studi pustaka

Data dikumpulkan melalui pencarian literatur untuk menemukan informasi yang relevan dengan masalah penelitian. Sumber informasi mencakup buku ilmiah, artikel riset, publikasi ilmiah, serta sumber cetak dan elektronik lainnya.

b. Studi lapangan

Pengumpulan data melalui studi lapangan dilakukan dengan cara sebagai berikut:

1. Wawancara
2. Observasi

1.6.2 Metode Perancangan Sistem

a. Analisis kebutuhan

Dalam tahap ini, peneliti mengevaluasi kebutuhan berdasarkan data yang dikumpulkan dan memastikan bahwa *automated framework* yang dikembangkan memenuhi kebutuhan pengguna.

b. Desain sistem

Pada tahap ini, desain sistem atau perancangan sistem akan dibuat. Perancangan ini berfokus pada desain alur sistem, yang bertujuan untuk menggambarkan proses kerja sistem secara logis dan sistematis.

c. Penulisan Kode

Pada tahap ini, mengimplementasikan rancangan perangkat lunak yang sebelumnya sudah ditentukan. Bahasa pemrograman yang digunakan adalah Swift dengan memanfaatkan *design pattern* MVVM (*Model, View, ViewModel*).

d. Pengujian

Setelah sistem *automated framework* dikembangkan, pengujian dilakukan untuk mengevaluasi performa komponen utama sistem. Pada Whisper, pengujian dilakukan dengan menggunakan metode *Word Error Rate* (WER), yang menghitung jumlah kesalahan dalam transkripsi dibandingkan dengan teks referensi untuk menilai tingkat akurasi transkripsi yang dihasilkan. Fungsi diarisasi pembicara dalam GPT-4o diuji dengan mengukur tingkat akurasi dalam mengidentifikasi dan memisahkan pembicara dalam suatu percakapan. Selain itu, fungsi GPT-4o dalam menganalisis kemajuan komunikasi anak penyandang ASD diuji dengan membandingkan waktu analsis dengan metode manual dan membandingkan hasil analisis dari kedua metode.

1.7 Sistematika Penulisan

a. BAB I - PENDAHULUAN

Bab ini memaparkan latar belakang masalah penelitian, rumusan masalah, batasan masalah, tujuan, manfaat, metode penelitian dan sistematika penulisan.

b. BAB II - KAJIAN PUSTAKA

Bab ini membahas landasan teori yang berkaitan dengan autisme, model pengenalan suara Whisper, model bahasa GPT-4o, dan membandingkan penelitian-penelitian yang telah dilakukan sebelumnya.

c. BAB III - METODOLOGI PENELITIAN

Bab ini menjelaskan mengenai metodologi yang digunakan untuk mencapai tujuan penelitian, meliputi metode pengumpulan data dan metode perancangan sistem.

d. BAB IV - HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dari penerapan model pengenalan suara Whisper dan model bahasa GPT-4o yang digunakan untuk mengukur perkembangan komunikasi dan bahasa anak penyandang autisme. Bab ini juga menyampaikan hasil pengujian serta pengukuran tingkat akurasi *framework* yang dikembangkan.

e. BAB V - KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang diambil dari hasil penelitian dan pembahasan, serta saran untuk penelitian lebih lanjut dan implementasi praktis dari temuan penelitian ini.

BAB II
KAJIAN PUSTAKA

2.1 Landasan Teori

2.1.1 Autisme

Definisi Autisme

Autism Spectrum Disorder adalah gangguan perkembangan saraf yang ditandai dengan kesulitan berkomunikasi dan berinteraksi sosial, serta pola perilaku, minat, atau aktivitas yang terbatas dan berulang. Gangguan ini muncul sejak awal masa perkembangan dan dapat mempengaruhi aspek sosial, pekerjaan, atau kehidupan sehari-hari (Kodak & Bergmann, 2020).

Anak dengan autisme memiliki kemampuan komunikasi yang berbeda dengan anak non-autisme, karena individu dengan ASD cenderung mengalami kesulitan dalam pemahaman kalimat dan kata. Anak-anak dengan autisme sering kali menafsirkan ucapan secara harfiah, sehingga mereka sering tidak memahami makna tersirat. Selain itu, mereka juga menghadapi tantangan dalam memahami dan menggunakan bahasa dalam konteks sosial atau pragmatik.

Hal tersebut diperkuat dengan hasil penelitian yang dilakukan oleh (Boucher, 2003), individu dengan autisme sering mengalami masalah komunikasi yang berkaitan dengan kesulitan dalam interaksi sosial. Meskipun beberapa individu mampu berbicara, bahasa yang digunakan lebih sering bersifat praktis, bukan untuk berinteraksi secara sosial. Pembicaraan mereka cenderung satu arah, tanpa timbal balik, dan sering kali diulang-ulang serta berfokus pada diri sendiri.

Kriteria diagnosis ASD berdasarkan pada *Diagnostic and Statistical Manual of Mental Disorders* edisi kelima (DSM-V) berfokus pada dua domain utama, yaitu gangguan dalam komunikasi dan interaksi sosial, serta pola perilaku yang terbatas dan berulang, termasuk minat yang sempit (Gensler, 2012).

Meskipun gangguan bahasa lisan bukan satu-satunya kriteria untuk diagnosis ASD, kesulitan berbahasa sering kali merupakan gejala yang menyertainya. Penguasaan bahasa fungsional pada usia 2 tahun merupakan indikator penting dari perkembangan yang baik. Kemampuan bahasa lisan juga menjadi faktor kunci yang mempengaruhi hasil perkembangan jangka panjang yang positif bagi anak-anak dengan ASD (Kodak & Bergmann, 2020).

Klasifikasi Autisme

Autisme dalam buku DSM-IV diklasifikasikan sebagai bagian dari Gangguan Pervasive Perkembangan (*Persevasive Development Disorder*). Subtipenya meliputi: (1) Gangguan Autisme, ditandai dengan kesulitan komunikasi, interaksi sosial, dan perilaku yang terbatas atau berulang; (2) Sindrom Asperger, individu memiliki keterampilan verbal yang baik, tetapi tetap mengalami kesulitan sosial dan pola perilaku terbatas, tanpa keterlambatan perkembangan bahasa yang signifikan; (3) *Pervasive Developmental Disorder-Not Otherwise Specified* (PDD-NOS), kategori untuk anak-anak dengan beberapa gejala autisme tetapi tidak memenuhi kriteria lengkap untuk autisme atau sindrom Asperger; (4) Gangguan Rett, gangguan langka pada anak perempuan yang menyebabkan kehilangan keterampilan setelah perkembangan awal normal; (5) *Childhood Disintegrative Disorder* (CDD), ditandai dengan regresi mendadak dalam kemampuan setelah dua tahun perkembangan normal (Diagnostic and Statistical Manual of Mental Disorders (4th ed.), 1994).

Menurut (American Psychiatric Association, 2013), subtipe autisme yang telah dijelaskan sebelumnya digabungkan menjadi satu kategori yang disebut ASD, yang mencakup spektrum gejala dengan berbagai tingkat keparahan. Sedangkan, menurut *Childhood Autism Rating Scale* (CARS) autisme dibagi menjadi tiga tingkatan berdasarkan tingkat gejala yang ditunjukkan oleh individu: (1) autisme ringan, pada tingkat ini, individu dengan autisme masih dapat menunjukkan sedikit kontak mata, memberikan respons ketika dipanggil namanya, dan kadang-kadang menunjukkan ekspresi wajah serta kemampuan berkomunikasi dua arah, meskipun tidak secara konsisten; (2) autisme sedang, individu dengan tingkat autisme sedang masih menunjukkan sedikit kontak mata dan memberikan respons ketika dipanggil namanya, namun mungkin memiliki sedikit ekspresi wajah dan interaksi sosial yang lebih terbatas daripada individu dengan autisme ringan; (3) autisme berat, pada tingkat ini, individu dengan autisme tidak dapat melakukan kontak mata, berekspresi, dan berinteraksi (Mujiyanti, 2011).

Karakteristik Komunikasi Penyandang Autisme

Individu dengan autisme umumnya menghadapi tiga permasalahan utama, yaitu kesulitan dalam komunikasi, interaksi sosial, dan keterbatasan dalam aktivitas imajinasi (Mansur, 2016). Penelitian terbaru yang dilakukan oleh (Vogindroukas et al., 2022) memperkuat pandangan ini dengan menunjukkan adanya heterogenitas yang signifikan dalam tantangan komunikasi pada individu dengan ASD sepanjang hidup mereka. Anak-anak dengan ASD sering kali mengalami hambatan maupun keunggulan dalam berbagai sub-sistem bahasa,

termasuk pragmatik, tata bahasa, semantik, sintaksis, fonologi, dan morfologi. Kesulitan ini bisa muncul dalam komunikasi lisan maupun tulisan, yang menunjukkan bahwa autisme mempengaruhi banyak aspek perkembangan bahasa dan cara berkomunikasi.

Selain itu, beberapa penyandang autisme juga menghadapi kesulitan dalam artikulasi. Walaupun banyak anak autisme dapat berbicara dengan jelas seperti anak-anak lainnya, beberapa memerlukan waktu lebih lama untuk belajar mengucapkan kata-kata dengan benar. Anak-anak dengan autisme yang memiliki kemampuan kognitif baik namun tetap mengalami kesulitan dalam interaksi sosial, sering kali menghadapi tantangan dalam mengucapkan suara tertentu, seperti "r", "l", dan "s". Kesulitan ini dapat berlanjut hingga dewasa, sehingga memengaruhi cara mereka berkomunikasi secara keseluruhan (Tager-Flusberg et al., 2005).

Dalam buku yang ditulis oleh (Tager-Flusberg et al., 2005), mereka juga menjelaskan bahwa individu dengan autisme sering kali memiliki banyak kata dalam ingatan mereka dan menunjukkan performa yang baik dalam tes kosakata. Namun, meskipun mereka memiliki pengetahuan yang baik tentang banyak kata, mereka sering kali mengalami kesulitan dalam menggunakan kata-kata tersebut dengan cara yang biasa digunakan oleh orang lain. Misalnya, mereka cenderung jarang menggunakan kata-kata yang berkaitan dengan pikiran atau perasaan, seperti "mengingat", "berpikir", "tahu", atau "pura-pura", yang merupakan bagian penting dalam komunikasi sehari-hari.

2.1.2 Pengenalan Suara (*Speech Recognition*)

Speech Recognition merupakan salah satu bidang dalam kecerdasan buatan yang berfungsi untuk mengenali ucapan melalui pencocokan sinyal suara. Sinyal suara ini kemudian diubah menjadi representasi numerik dan dibandingkan untuk mengidentifikasi kata-kata yang diucapkan (Sani, 2019).

Sistem pengenalan suara (*speech recognition*) juga dapat didefinisikan sebagai proses mengonversi suara yang diterima menjadi format lain, seperti teks. Oleh karena itu, *speech recognition* sering disebut sebagai *speech-to-text*. Sistem ini dibangun dengan beberapa komponen utama, seperti mikrofon untuk menangkap suara, perangkat lunak pengenalan suara, komputer untuk memproses data suara, dan *soundcard* yang digunakan untuk *input* atau *output* suara (Trima, 2020).

Salah satu kemajuan terbaru dalam teknologi ASR adalah Whisper, yang dikembangkan oleh OpenAI. Teknologi ini menunjukkan kinerja yang luar biasa dalam mengenali ucapan di berbagai kondisi menantang, seperti perbedaan aksen, kebisingan latar belakang, serta dataset

yang sebelumnya tidak dikenal (Radford et al., 2022). Keunggulan utama Whisper terletak pada kemampuannya beroperasi dalam pengaturan *zero-shot*, yang memungkinkan teknologi ini untuk bekerja tanpa perlu pelatihan khusus pada dataset tertentu. Hal ini menjadikan Whisper lebih unggul dibandingkan model ASR lain yang bergantung pada pelatihan berbasis domain tertentu (Radford et al., 2022).

Penelitian terkini berfokus pada pengembangan lebih lanjut dengan mengintegrasikan Whisper dengan LLM untuk meningkatkan kinerja pengenalan ucapan pada domain tertentu. Integrasi ini menggunakan deskripsi kontekstual yang dihasilkan oleh LLM untuk meningkatkan akurasi pengenalan istilah spesifik domain tanpa perlu melakukan perubahan pada arsitektur Whisper. Hasilnya, pendekatan ini berhasil menurunkan Word Error Rate (WER) hingga 15,15%, menunjukkan potensi besar untuk aplikasi ASR di berbagai bidang yang memerlukan pengenalan istilah khusus (Suh et al., 2024).

Dengan kemampuan tersebut, Whisper menjadi teknologi ASR yang ideal untuk penelitian ini, terutama dalam mengenali ucapan guna mengukur kemajuan komunikasi pada anak dengan autisme. Selain itu, teknologi ini dapat dikombinasikan dengan Large Language Models (LLMs) untuk meningkatkan kinerjanya dalam memahami dan memproses pola komunikasi spesifik, sehingga membuatnya semakin efektif untuk tujuan ini.

2.1.3 Model Pengenalan Suara Whisper

Pengertian Whisper

Whisper adalah sebuah model pengenalan suara atau *Automatic Speech Recognition* (ASR) yang dibuka secara publik mulai September 2022. Model ini dikembangkan oleh Open AI yang telah dilatih dengan data multibahasa dan *multitask* yang totalnya sekitar 680.000 jam. Berdasarkan penjelasan dari Open AI, penggunaan dataset yang besar dan beragam menghasilkan peningkatan kestabilan terhadap aksen, kebisingan latar belakang, dan bahasa teknis. Selain itu, Whisper juga dapat mengerti berbagai bahasa dan menerjemahkannya ke dalam bahasa inggris ataupun bahasa lain.

Dalam perkembangannya, model terbaru dari Whisper adalah versi whisper v2-large, yang dapat diakses melalui API OpenAI dengan nama whisper-1. Menurut (Open AI, 2022), Whisper menyediakan beberapa versi model yang disebut *checkpoint*. *Checkpoint* adalah versi model yang sudah dilatih dengan ukuran berbeda-beda, sehingga pengguna bisa memilih sesuai kebutuhan. Misalnya, ada model kecil yang lebih ringan dan cepat, atau model besar yang lebih akurat dalam mengenali bahasa dan konteks yang beragam. Beberapa *checkpoint* hanya untuk

bahasa Inggris, sementara yang lain mendukung berbagai bahasa. Seluruh *checkpoint* atau detail model yang sudah dilatih ini akan diurakan pada Tabel 2.1.

Tabel 2.1 Detail Model Whisper

Ukuran	**Parameter**	**Hanya Bahasa Inggris**	**Multibahasa**
Tiny	39 M	✓	✓
Base	74 M	✓	✓
Small	244 M	✓	✓
Medium	769 M	✓	✓
Large	1550 M	x	✓
Large-v2	1550 M	x	✓

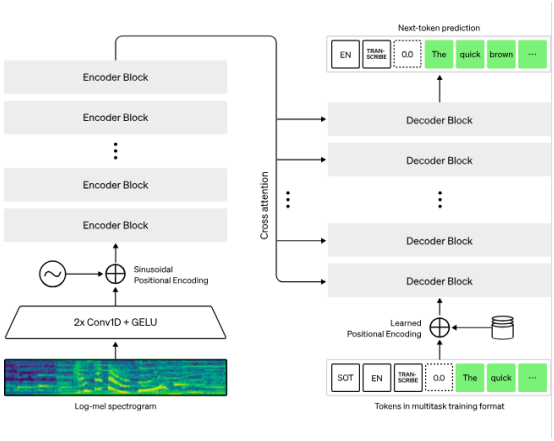
Menurut penelitian oleh (Chen et al., 2024), Kinerja Whisper dalam pengenalan suara ASR dianggap unggul. Whisper ketika diintegrasikan dengan robot, menunjukkan kinerja terbaik. Hal ini dibuktikan dengan *Word Error Rate* (WER) rata-rata Whisper yang sangat rendah, yaitu 1.716%, dan waktu pemrosesan rata-rata yang juga rendah, yaitu 2.639 detik. Ini menandakan bahwa Whisper sangat efektif dalam mengenali dan memproses ucapan dengan akurasi yang tinggi dan dalam waktu yang singkat.

Keunggulan Whisper tidak hanya terletak pada kemampuan multibahasanya, tetapi juga pada kemampuan integrasinya. Whisper dapat dengan mudah diintegrasikan dengan model AI lainnya, seperti yang ditunjukkan dalam penelitian oleh (Vivek et al., 2023), yang mengintegrasikan Whisper dengan GPT-3 untuk menciptakan sebuah *Voice Assistant* yang fleksibel dan mudah diintegrasikan. Hasil penelitian tersebut menunjukkan bahwa *Voice Assistant* yang dihasilkan mampu memahami dan merespons dengan akurat terhadap pertanyaan pengguna, membuatnya menjadi solusi yang kuat dan potensial kedepannya.

Pengertian Whisper

Open AI Whisper dikembangkan menggunakan model *deep learning* yang dilatih dengan data dari website, data yang diambil adalah data multibahasa dan *multitask* yang diawasi. Artinya, Whisper dapat menangani berbagai tugas dalam berbagai bahasa. Saat OpenAI Whisper mendapatkan audio ucapan, Whisper tidak hanya mendengarnya, tetapi juga menganalisisnya. Ini memecah audio menjadi beberapa bagian kecil dan mempelajarinya,

kemudian menguraikan ucapan dengan memprediksi transkripsi yang paling mungkin (Deepgram, 2024).



Gambar 2.1 Arsitektur Whisper (Open AI, 2022).

Pada Gambar 2.1 menjelaskan mengenai arsitektur Whisper berdasarkan yang dipaparkan oleh (Open AI, 2022), arsitektur ini menggunakan pendekatan *end-to-end* yang diimplementasikan sebagai *Transformer encoder-decoder*. Proses kerja pada Whisper dimulai dengan membagi audio yang masuk menjadi potongan-potongan berdurasi 30 detik. Setiap potongan tersebut kemudian diubah menjadi format visual yang disebut log-Mel *spectrogram*. Informasi dari *spectrogram* ini diolah oleh *encoder* yang bertugas memahami konten audio. Selanjutnya, *decoder* mengambil informasi dari *encode*r dan menghasilkan teks yang merupakan terjemahan dari audio. Selain itu, Whisper juga memiliki kemampuan untuk melakukan beberapa tugas khusus, seperti mengidentifikasi bahasa dalam audio, menulis teks dalam berbagai bahasa, dan menerjemahkan ucapan ke bahasa Inggris. Dengan fitur-fitur ini, Whisper mampu mengonversi ucapan menjadi teks yang dapat dibaca sambil menyediakan berbagai fungsi tambahan.

Keunggulan Whisper

Whisper memiliki beberapa keunggulan dibandingkan TTS (*Speecht-To-Text*) yang sudah ada menurut (Latentview, 2024) yaitu:

a. Fleksibilitas dalam Bahasa dan Tugas, Whisper bisa mengenali bahasa dari audio yang diberikan dan secara otomatis beralih tugas sesuai kebutuhan.
b. Akurasi transkripsi, Whisper sangat akurat dalam menyalin ucapan menjadi teks, yang sangat berguna untuk keperluan penelitian, proses hukum, dan membantu orang yang membutuhkan aksesibilitas.

c. Adaptasi Domain, Whisper bisa menyesuaikan diri dengan berbagai jenis penggunaan, seperti saat mendengarkan kuliah, wawancara, atau bahkan saat merekam hal-hal yang teknis.

Kekurangan Whisper

Whisper masih dalam tahap pengembangan dan memiliki beberapa keterbatasan yang perlu diperhatikan oleh pengembang. Pertama, ukuran fail unggahan dibatasi hingga maksimal 25MB. Oleh karena itu, Whisper merekomendasikan kepada pengembang untuk mengompres fail audio agar sesuai dengan batasan ukuran tersebut. Kedua, sistem ini dapat mengalami halusinasi, yaitu situasi ketika sistem ASR menghasilkan teks kata-kata atau frasa yang sebenarnya tidak terdapat dalam audio yang diucapkan. Hal ini terjadi karena faktor internal (misalnya data pelatihan, arsitektur model, dan *overfitting*) serta faktor eksternal (misalnya audio yang kompleks) (Gladia, 2024).

2.1.4 Model Bahasa (*Large Language Model*)

Large Language Model (LLM) atau Model Bahasa Besar adalah algoritma pembelajaran mendalam (*Deep Learning*) yang mampu memahami, meringkas, menerjemahkan, memprediksi, dan menghasilkan teks serta konten lainnya berdasarkan data yang sangat besar. Model ini merupakan salah satu penerapan transformator yang paling berhasil. Fungsinya tidak hanya terbatas pada pembelajaran bahasa manusia untuk kecerdasan buatan, tetapi juga mencakup kemampuan seperti menulis kode perangkat lunak dan masih banyak lagi (Laraswati, 2023).

LLM seperti *Generative Pre-trained Transformer* (GPT) telah membuat kemajuan signifikan dalam beberapa tahun terakhir. Model-model ini dilatih pada sejumlah besar data teks dan mampu menghasilkan teks yang menyerupai bahasa manusia, menjawab pertanyaan, serta menyelesaikan tugas-tugas lain yang berhubungan dengan bahasa dengan tingkat akurasi yang tinggi (Kasneci et al., 2023). Maka dari itu, model LLM sangat cocok untuk menganalisis perkembangan anak autisme setelah mendapatkan hasil transkripsi dari Whisper**.** Dengan analisis berbasis teks dari transkripsi ini, LLM dapat mengidentifikasi pola-pola kemajuan dalam komunikasi dan interaksi bahasa anak autisme.

2.1.5 Model Bahasa GPT

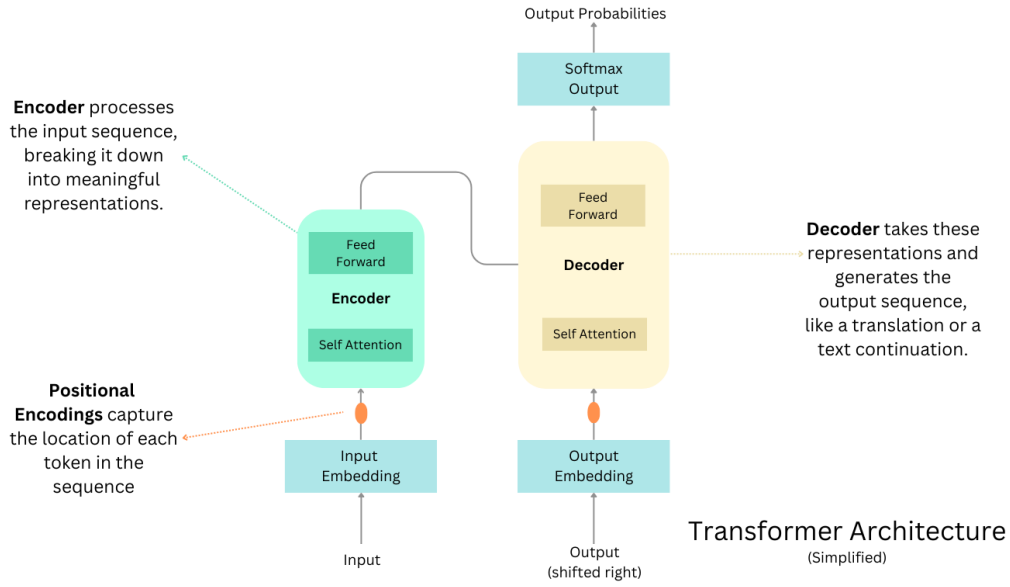
Pengertian GPT

Generative Pre-Trained Transformer (GPT) adalah model bahasa yang dikembangkan oleh Open AI. Sesuai dengan namanya, GPT merupakan AI berbasis Transformer yang dapat membuat konten baru (*generative*) berdasarkan data pelatihannya (*pre-trained*). GPT dirancang untuk tugas *Natural Language Processing* (NLP) yang melibatkan dua aspek penting: (1) menganalisis dan menentukan makna kalimat (pemahaman bahasa alami), dan (2) menghasilkan kalimat baru berdasarkan masukan (pembuatan bahasa alami) (Cheng et al., 2023).

Menurut (IPSpecialist, 2024), dalam pengembangan model bahasa GPT Open AI telah merilis empat versi GPT yaitu, (1) GPT-1, model ini dirilis pada tahun 2018. GPT-1 memiliki 117 juta parameter dan dilatih pada kumpulan data teks yang besar. Meskipun ukurannya kecil dibandingkan versi yang lebih baru, GPT-1 berfungsi dengan baik dalam tugas pemrosesan bahasa alami, seperti menyusun frasa, menjawab pertanyaan, dan menghasilkan rangkaian tulisan yang kohesif. (2) GPT-2, dirilis pada tahun 2019. GPT-2 memiliki 1,5 miliar parameter dan telah menunjukkan peningkatan koherensi, memori konteks, dan kemampuan menulis inovatif, memungkinkannya menyusun cerita, merespons pertanyaan, dan bahkan membuat artikel teknis dengan presisi. (3) GPT-3, dirilis pada tahun 2020. GPT-3 memiliki 175 miliar parameter, kemampuan pembelajaran *few-shot* dan *zero-shot* GPT-3 merupakan salah satu fitur yang paling menonjol, memungkinkannya menyelesaikan soal dengan sedikit atau tanpa contoh, dan menunjukkan tingkat generalisasi yang sangat baik. (4) GPT-4, diperkenalkan oleh OpenAI pada 14 Maret 2023, namun pelatihan GPT-4 sebenarnya sudah selesai pada pertengahan tahun 2022. GPT-4 adalah generasi keempat dalam rangkaian model GPT dan dirancang dengan fokus pada keselamatan, multimodalitas, dan kinerja. Model ini dapat menerima masukan teks dan gambar, memungkinkannya menghasilkan keluaran teks berdasarkan informasi visual.

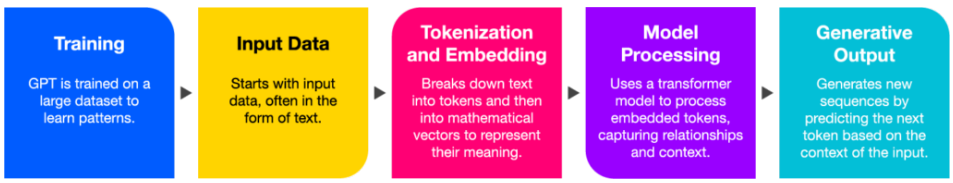
Cara Kerja GPT

Natural Language Processing (NLP) dan *machine learning* adalah cara kerja GPT untuk merespons pertanyaan dari pengguna seperti halnya manusia (Radford et al., 2018). Respons yang dihasilkan oleh GPT berdasarkan pola-pola yang sudah dipelajari dari data saat dilatih, ditambah dengan kemampuan untuk membuat jawaban yang kreatif dan masuk akal berdasarkan struktur otak tiruannya.



Gambar 2.2 Arsitektur Transformer

Dalam pengembangannya, GPT menggunakan arsitektur Transformer yang merupakan jenis dari *Recurrent Neural Network* (RNN) yang dirancang khusus untuk menangani data berurutan, seperti teks. Berdasarkan Gambar 2.2, arsitektur transformer ini terdiri dari dua komponen utama: (1)*Encoder*, bertugas memproses teks masukan dan menghasilkan representasi kontinu dari urutan masukan tersebut. Proses ini dilakukan melalui serangkaian mekanisme perhatian mandiri (*self-attention*) yang memungkinkan model untuk menimbang pentingnya berbagai bagian dari urutan masukan terhadap satu sama lain. (2)*Decoder*, bertanggung jawab untuk menghasilkan teks keluaran berdasarkan urutan masukan dan representasi kontinu yang dihasilkan oleh *encoder*. *Decoder* melakukan ini dengan memprediksi token selanjutnya dalam urutan, berdasarkan token-token sebelumnya. Jadi, secara sederhana, *encoder* membaca dan memahami teks masukan, dan *decoder* menggunakan pemahaman tersebut untuk menghasilkan teks keluaran.



Gambar 2.3 Proses Kerja *Generative Pre-Trained Transformer* (GPT) (Taskus, 2024).

Gambar 2.3 menjelaskan tentang proses kerja GPT dalam beberapa tahap utama menurut (Taskus, 2024) yaitu: (1) *training*, GPT dilatih menggunakan dataset besar untuk mempelajari pola-pola tertentu. Di tahap ini, GPT memahami berbagai struktur dalam data yang diberikan, yang memungkinkan model untuk mengembangkan kemampuan dasar dalam memproses bahasa. (2) *input data*, proses dimulai dengan *input data* yang biasanya dalam bentuk teks. Di tahap ini, teks yang ditulis oleh pengguna akan dimasukkan ke dalam sistem. (3) *tokenization and embedding*, teks yang dimasukkan dipotong-potong menjadi bagian kecil seperti kata atau frasa. Setiap potongan ini diubah menjadi angka-angka yang merepresentasikan makna dan informasi lainnya. Langkah ini penting agar sistem bisa mengerti dan dapat bekerja dengan teks tersebut. (4) *model processing*, GPT menggunakan arsitektur Transformer untuk memahami hubungan antara potongan-potongan teks tersebut. Sistem ini mencoba mengerti konteks dan makna keseluruhan dari teks yang dimasukkan. (5) *generative output*, setelah memahami konteks, GPT akan membuat teks baru berdasarkan informasi yang telah dipelajari sebagai respon.

Keunggulan GPT

GPT memberikan keuntungan yang signifikan dalam membantu mencari informasi, kecepatannya dan keakuratannya sangat baik. Berikut adalah beberapa keunggulan GPT.

a. Efisiensi dan kecepatan, GPT meningkatkan efisiensi dan mempercepat proses dengan mengotomasi tugas-tugas rutin, seperti menjawab pertanyaan yang sering diajukan. Hal ini membuat pekerjaan kita lebih efisien dan meminimalkan kesalahan. Menurut (Taskus, 2024), GPT sangat efisien berkat kecepatannya dalam memproses dan menghasilkan teks dengan cepat. Keunggulan ini sangat terlihat dalam pemrosesan bahasa berskala besar yang membutuhkan kecepatan tinggi.
b. Pemahaman kontekstual yang baik, menurut (Frieda Rong, 2021) model GPT dilatih dengan jumlah data teks yang sangat besar, yang memungkinkan model untuk belajar pola dan hubungan dalam bahasa. Pelatihan ini memungkinkan model untuk mengenali dan merespons petunjuk spesifik konteks, seperti urutan kata, sintaks, dan semantik.
c. Kemampuan multibahasa, hal ini didukung oleh data pelatihan yang ekstensif, yang berasal dari kumpulan besar data internet dalam berbagai bahasa. Selain itu, GPT memanfaatkan teknik yang disebut pembelajaran *Zero-Shot*, di mana GPT mampu menganalisis teks dalam berbagai bahasa tanpa perlu pelatihan khusus untuk setiap bahasa tersebut. Penelitian yang dilakukan oleh (Rathje et al., 2024) menunjukkan bahwa GPT

lebih unggul dari metode analisis berbasis kamus dalam mengidentifikasi komponen psikologis dari 47.925 tweet dan judul berita dalam berbagai bahasa.

Kekurangan GPT

Model GPT memiliki serangkaian kekurangan yang cukup signifikan, kekurangan-kekurangan ini mencakup berbagai aspek, yaitu:

a. Halusinasi, Model GPT terutama GPT-3, dikenal memiliki fenomena yang disebut halusinasi. Hal ini merujuk pada kecenderungan model untuk menghasilkan informasi yang salah atau menyesatkan. Halusinasi ini bisa terjadi karena beberapa alasan, seperti data pelatihan model tidak lengkap atau keterbatasan model dalam memahami dan memproses bahasa (Lacy, 2024).
b. Respon yang bias, bias dalam model GPT merujuk pada kecenderungan model untuk menghasilkan respon yang tidak seimbang berdasarkan asumsi atau stereotip. Menurut (Hendrick, 2024), respon yang bias dalam model GPT biasanya disebabkan oleh beberapa faktor. Pertama, kecenderungan ini timbul karena adanya bias yang sudah ada dalam data yang digunakan untuk melatih model tersebut. Kedua, kurangnya variasi dalam bahasa yang tercakup dalam data pelatihan juga berkontribusi terhadap respons yang bias. Ketiga, model-model ini sering kali gagal memahami konteks dari pertanyaan pengguna, yang dapat menyebabkan jawaban yang tidak tepat atau tidak sesuai.
c. Kesulitan dalam memahami emosi, model GPT tidak memiliki pengalaman pribadi dan emosi sendiri meskipun dirancang untuk memproses dan menghasilkan teks yang mirip dengan manusia berdasarkan pola dalam data pelatihan. Model GPT dapat mengenali dan memahami konsep emosi serta bahasa yang terkait dengan emosi, tetapi model ini kesulitan untuk memprediksi intensitas emosi dan respon penanganan akurat (Tak & Gratch, 2023).

2.2 Penelitian Terdahulu

Banyak peneliti telah berupaya menemukan metode terbaik untuk mendeteksi dan mengukur perkembangan komunikasi pada penyandang autisme melalui teknologi *Artificial Intelligence* (AI). Salah satu pendekatan menggunakan model BERT dan ChatGPT-3 (*text-davinci-003*) untuk mendeteksi autisme dengan menganalisis dialog teks antara orang tua dan anak melalui analisis sentimen. Sistem yang diusulkan menunjukkan potensi untuk deteksi ASD yang terjangkau dan mudah diakses melalui analisis teks. Namun, terdapat kelemahan

pada ChatGPT-3, di mana model ini unggul dalam memahami teks tetapi kurang efisien dalam menangani respons bahasa alami yang melibatkan agregasi data, seperti melakukan perhitungan atau menganalisis data untuk tugas seperti *cosine similiarity* (Mukherjee et al., 2023).

Pendekatan lain yaitu pengembangan *LENA Autism Screening*, yang menggunakan rekaman audio dan algoritma statistik untuk membedakan vokalisasi anak dari suara lingkungan, memungkinkan deteksi autisme secara dini. Pendekatan ini mampu dalam mengidentifikasi tanda-tanda ASD dari ucapan anak. Namun, masih ada ruang untuk perbaikan, seperti kebutuhan akan pengumpulan data yang lebih komprehensif dan standar dari anak-anak yang telah didiagnosis dengan ASD, serta pendekaran ini memerlukan alat dengan biaya yang signifikan (Richards et al., 2010).

Penelitian yang dilakukan oleh (O'Sullivan et al., 2023) menerapkan model Whisper dari OpenAI untuk pengenalan suara otomatis (ASR), yang menunjukkan kemampuannya menganalisis fitur linguistik yang relevan untuk diagnosis ASD dan memberikan keluaran yang selaras dengan *Vineland Adaptive Behavior Scales-II (VABS-II)*. Namun, studi ini memiliki keterbatasan, seperti kebutuhan diarisisasi pembicara manual untuk membedakan antara anak dan terapis atau orang tua dalam percakapan, serta cakupan fitur linguistik yang diuji masih terbatas, seperti jumlah unit komunikasi (*communication units/c-units*) per menit, pengulangan ucapan (ekolalia), dan kompleksitas morfologi (O'Sullivan et al., 2023).

Tabel 2.2 Perbandingan Alat Ukur Komunikasi ASD dari Penelitian Terdahulu

No.	Alat Diagnosis	Deskripsi	Kelebihan	Gap dan Keterbatasan
1.	Kecerdasan Buatan *(Natural Language Text Using BERT and ChatGPT Models)*	Menurut penelitian oleh Mukherjee, et al. 2023, model BERT dan ChatGPT digunakan untuk untuk mendeteksi gejala ASD melalui	1. BERT menghasilkan akurasi yang tinggi, yaitu 83%. 2. BERT dan ChatGPT adalah model bahasa besar (LLM)	Menurut (Mukherjee et al., 2023), sistem berbasis transformer dan LLM seperti ChatGPT dapat meningkatkan kinerja diagnosis autisme, namun memiliki

		analisis teks dialog orang tua. BERT dan ChatGPT digunakan untuk menganalisis sentimen setiap pernyataan dari orang tua untuk mendeteksi gejala ASD. Data dikumpulkan dari berbagai grup autisme di media sosial dan sumber internet lainnya.	yang dilatih menggunakan teknik pembelajaran mesin canggih, termasuk pembelajaran penguatan dari umpan balik manusia (RLHF), yang memungkinkan mereka untuk menangani tugas-tugas pemrosesan bahasa alami dengan sangat baik 3. Kedua model ini mampu menganalisis sentimen dari setiap pernyataan orang tua dengan efektif, sehingga dapat mendeteksi gejala autisme dengan lebih tepat. (Mukherjee et al., 2023).	beberapa keterbatasan yaitu: 1. Perhitungan Bahasa Alami**:** ChatGPT (*text-davinci-003*) tidak selalu akurat dalam perhitungan yang kompleks atau dalam menghitung jumlah dari beberapa nilai. 2. Kebutuhan komponen yang berfungsi baik: untuk mendapatkan hasil yang akurat, semua komponen sistem harus berfungsi dengan baik. Jika ada masalah atau kerusakan pada salah satu bagian, hal ini bisa mempengaruhi kinerja keseluruhan sistem.

2.	Pengenalan Suara (*LENA Autism Screening)*	*LENA Autism Screening* adalah alat skrining otomatis yang dikembangkan oleh LENA *Foundation* untuk mendeteksi risiko ASD pada anak-anak. Alat ini menggunakan teknologi perekaman audio dan analisis vokalisasi untuk mengidentifikasi karakteristik vokal yang tidak biasa pada anak dengan ASD. Sistem ini terdiri dari perekam audio digital yang ditempatkan pada pakaian khusus yang dikenakan oleh anak, serta perangkat lunak yang menganalisis data rekaman menggunakan algoritma pengenalan suara dan model statistik untuk membedakan vokalisasi anak dari	1. Kemampuan skrining skala besar, dengan desain yang relatif sederhana, tidak mengganggu, dan melibatkan sedikit interaksi, data rekaman audio dapat diperoleh dalam jumlah yang cukup untuk melakukan skrining. 2. Penggunaan algoritma statistik yang valid dan andal, algoritma statistik otomatis yang digunakan oleh sistem ini telah terbukti valid dan andal dalam membedakan vokalisasi anak dari suara orang dewasa dan suara	1. Kemungkinan besar sampel ASD tidak cukup merepresentasikan seluruh populasi anak-anak dengan ASD karena data diambil tidak hanya dari anak yang telah terdiagnosa ASD. 2. Tidak ada anak dengan ASD dalam penelitian ini yang dievaluasi langsung oleh profesional terlatih dan bersertifikat dengan menggunakan alat diagnosis standar. Artinya, diagnosis ASD untuk anak-anak dalam penelitian ini hanya berdasarkan laporan orang tua, bukan dari pemeriksaan langsung oleh ahli. 3. Kemungkinan membutuhkan

		suara lingkungan lainnya (Richards et al., 2010).	lingkungan sekitar. 3. Pengurangan waktu dan sumber daya dalam proses skrining 4. Membantu deteksi dini autism 5. Objektivitas dalam penilaian, sifat otomatis dari skrining ini memberikan tingkat objektivitas yang sulit dicapai dengan penilaian manusia. (Richards et al., 2010)	biaya yang cukup besar untuk membeli alatnya. (Richards et al., 2010)
3.	Pengenalan Ucapan Otomatis (*Automatic Speech Recognition*) dengan *Whisper Model From Open AI*	Menurut penelitian yang dilakukan oleh (O'Sullivan et al., 2023), penggunaan model Whisper dari OpenAI untuk *Automatic Speech Recognition* (ASR) pada individu dengan	1. Menunjukkan kemampuannya menganalisis fitur linguistik yang relevan untuk diagnosis ASD dan memberikan keluaran yang selaras dengan	1. Performanya yang lebih rendah untuk individu autisme dengan IQ di bawah 70. 2. Diarisasi pembicara masih dilakukan secara manual 3. Cakupan fitur linguistik yang diuji

		autisme dapat menggantikan transkripsi manual yang mahal dan memakan waktu. Hasil penelitian menunjukkan bahwa Whisper dapat sebanding dengan solusi komersial dalam mengidentifikasi fitur bahasa yang penting untuk diagnosis ASD.	*Vineland Adaptive Behavior Scales-II (VABS-II)* 2. Model ini gratis untuk digunakan. 3. Whisper telah sebanding dengan solusi komersial.	masih terbatas, seperti jumlah unit komunikasi (*communication units/c-units*) per menit, pengulangan ucapan (ekolalia), dan kompleksitas morfologi.

Berdasarkan yang telah diuraikan pada Tabel 2.2, dapat diambil kesimpulan bahwa penelitian sebelumnya telah mencoba berbagai pendekatan berbasis teknologi, seperti kecerdasan buatan, pengenalan suara otomatis, dan analisis vokalisasi, untuk mendeteksi dan mengukur perkembangan komunikasi pada individu dengan ASD.

Untuk mengisi keterbatasan penelitian sebelumnya, peneliti mengusulkan *automated framework* yang menggabungkan model Whisper dari OpenAI dengan GPT-4o. Whisper akan menyediakan transkripsi otomatis dan akurat dari ucapan, sementara GPT-4o akan memungkinkan diarisisasi pembicara secara otomatis, menghilangkan kebutuhan manual dan mengatasi keterbatasan utama dalam skala dan efisiensi. *Framework* ini juga akan memperluas analisis untuk mencakup fitur linguistik yang kurang tereksplorasi, seperti kompleksitas kalimat, ekolalia, dan pragmatik, yang menawarkan pemahaman yang lebih komprehensif tentang pola komunikasi yang terkait dengan ASD, yang sangat penting untuk mengidentifikasi keterlambatan perkembangan dan tantangan secara lebih mendalam. Untuk menjembatani kesenjangan lebih lanjut, peneliti mengimplementasikan *Matrix-Based Evaluation*, yang dikembangkan dari studi literatur dan telah divalidasi melalui umpan balik dari guru di sekolah khusus autisme.

BAB III
METODOLOGI PENELITIAN

3.1 Populasi dan Sampel

Penelitian ini dilaksanakan pada bulan April hingga Desember 2024 untuk memperoleh data yang beragam dari berbagai tingkat ASD. Penelitian dilakukan di beberapa sekolah luar biasa yang fokus pada pendidikan anak penyandang autisme di Daerah Istimewa Yogyakarta, yaitu SLB Bina Anggita, SLB Dian Amanah, SLB Fredofios, dan SLB Fajar Nugraha.

Populasi penelitian mencakup dua kelompok utama, yaitu anak-anak yang telah didiagnosis dengan ASD dan memiliki kemampuan komunikasi verbal, serta guru-guru yang secara aktif terlibat dalam proses pendidikan anak-anak tersebut di sekolah luar biasa. Sampel penelitian terdiri dari 2 anak dengan autisme ringan beserta guru pendampingnya, 2 anak dengan autisme sedang beserta guru pendampingnya, serta 6 guru SLB yang diwawancarai untuk memberikan wawasan tambahan terkait pola komunikasi dan proses pembelajaran anak-anak tersebut.

3.2 Metode Pengumpulan Data

3.2.1 Studi Pustaka

Studi pustaka adalah salah satu cara dalam mengumpulkan data dalam penelitian, yaitu data dikumpulkan melalui pencarian literatur untuk menemukan informasi yang relevan dengan masalah penelitian. Dengan melakukan studi literatur, peneliti dapat mengidentifikasi kekurangan dalam pengetahuan yang telah ada, membentuk landasan teori, dan menggali pemahaman yang lebih mendalam mengenai subjek penelitian. Sumber-sumber informasi yang digunakan dalam penelitian ini mencakup buku ilmiah, artikel riset, publikasi ilmiah, serta website kredibel.

3.2.2 Studi Lapangan

a. Wawancara

Wawancara adalah metode pengumpulan data yang dilakukan melalui percakapan langsung antara pewawancara dan narasumber. Tujuannya yaitu untuk mendapatkan informasi mendalam yang berkaitan dengan topik penelitian. Pada penelitian ini penulis menggunakan teknik wawancara semi-terstruktur yaitu pewawancara memiliki kerangka pertanyaan dasar tetapi diperbolehkan untuk mengeksplorasi lebih dalam dengan

mengajukan pertanyaan tambahan berdasarkan respons narasumber. Wawancara pada penelitian ini dilakukan di Sekolah Luar Biasa (SLB) Bina Anggita Yogyakarta, SLB Dian Amanah, SLB Fredofios, dan SLB Fajar Nugraha, peneliti mengadakan wawancara dengan guru-guru dari sekolah tersebut untuk mendapatkan informasi karakteristik komunikasi dan pola bahasa yang biasa digunakan individu penyandang autisme.

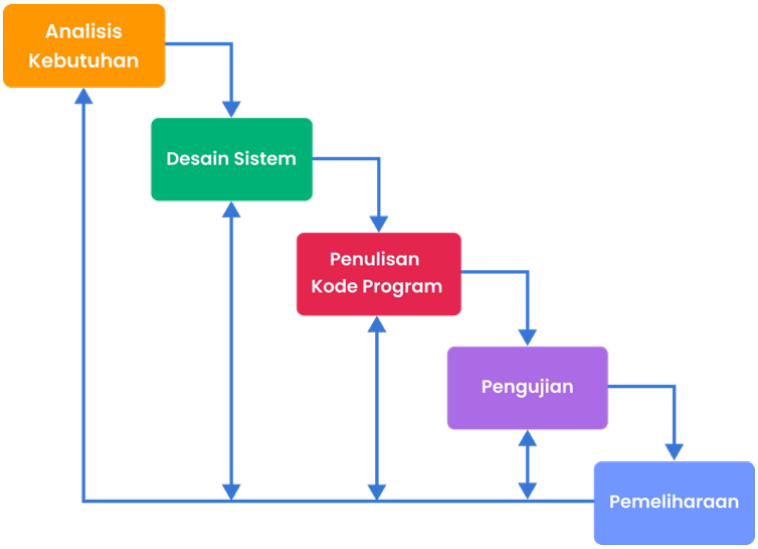
b. Observasi

Observasi adalah metode pengumpulan data dalam penelitian yang melibatkan pengamatan langsung atas masalah yang sedang diteliti tanpa intervensi dari peneliti. Pada penelitian ini, observasi mencakup beberapa aspek di antaranya: (1) karakteristik vokal, yaitu termasuk intonasi, volume, dan kecepatan merespon; (2) struktur bahasa, meliputi sintaksis, kemampuan gramatikal, dan penggunaan kata; (3) isi komunikasi, yang mencakup observasi tentang kesesuaian respons dan kecenderungan perulangan atau ekolalia; dan (4) merekam percakapan antara guru SLB dan anak penyandang ASD untuk dianalisis kemajuan komunikasinya.

3.3 Metode Perancangan Sistem

Pada penelitian ini, dalam perancangan sistem atau *automated framework* mengimplementasikan metode *waterfall*. Metode *waterfall* adalah pendekatan dalam pengembangan perangkat lunak yang awalnya diperkenalkan oleh Royce pada tahun 1970-an, (Wahid, 2020). Metode ini mirip dengan air terjun yang mengalir dari atas ke bawah, yaitu setiap tahap harus selesai sebelum pindah ke tahap selanjutnya.

Metode *waterfall* memiliki beberapa tahapan seperti yang bisa kita lihat dari Gambar 3.1 Tahapan Metode *Waterfall* yaitu, (1) Analisis Kebutuhan, (2) Desain Sistem, (3) Penulisan Kode Program, (4) Pengujian, dan (5) Pemeliharaan. Semua tahap tersebut saling berhubungan satu sama lain membentuk analisis dan desain yang terstruktur.



Gambar 3.1 Tahapan Metode *Waterfall*

3.3.1 Analisis Kebutuhan

Pada tahap ini, peneliti mengevaluasi kebutuhan berdasarkan data yang dikumpulkan dan memastikan bahwa sistem yang dikembangkan memenuhi kebutuhan pengguna. Analisis kebutuhan meliputi identifikasi masalah, usulan pemecahan masalah, *analisis matrix-based evaluation,* dan analisis kebutuhan sistem.

a. Identifikasi masalah, yaitu mengidentifikasi masalah yang didapat dalam aktivitas pengumpulan data sebelumnya melalui studi pustaka dan studi lapangan.
b. Usulan pemecahan masalah, mengusulkan solusi atas masalah yang telah diidentifikasi sebelumnya.
c. Analisis *matrix-based evaluation*, yaitu dengan meninjau instrumen penilaian yang digunakan untuk mengevaluasi perkembangan komunikasi individu autisme berdasarkan data yang telah diperoleh.
d. Analisis kebutuhan sistem, yaitu menganalisis spesifikasi sistem berdasarkan hasil identifikasi di atas.

3.3.2 Desain Sistem

Pada tahap ini, peneliti akan melakukan desain sistem dengan membuat diagram alir aplikasi yang terstruktur. Diagram alir aplikasi ini dirancang untuk memetakan setiap langkah proses pengguna dalam aplikasi, dari awal masuk hingga selesai tugas. Diagram alir ini memastikan bahwa setiap fungsi dirancang sesuai dengan kebutuhan pengguna.

3.3.3 Penulisan Kode Program

Selama tahap pengkodean sistem ini, peneliti mengimplementasikan desain yang telah disusun menggunakan bahasa pemrograman Swift, bahasa yang digunakan khusus untuk pengembangan aplikasi berbasis iOS. Dalam proses ini, peneliti memanfaatkan dependensi dari OpenAIKit untuk membantu pengintegrasian aplikasi dengan Model Whisper dan Model GPT. Peneliti juga memanfaatkan beberapa library pihak ketiga untuk meningkatkan efektivitas dan efisiensi sistem.

Selain itu, untuk meningkatkan pemisahan tanggung jawab dan memudahkan pemeliharaan kode, peneliti mengadopsi design pattern MVVM (*Model-View-ViewModel*). Dengan menggunakan MVVM, aplikasi ini akan memisahkan logika presentasi (*View*), logika bisnis (*Model*), dan pengolahan data (*ViewModel*).

3.3.4 Pengujian

Setelah sistem *automated framework* dikembangkan, pengujian dilakukan untuk mengevaluasi performa komponen utama sistem. Pada Whisper, pengujian dilakukan dengan menggunakan metode *Word Error Rate* (WER), yang menghitung jumlah kesalahan dalam transkripsi dibandingkan dengan teks referensi untuk menilai tingkat akurasi transkripsi yang dihasilkan. Fungsi diarisasi pembicara dalam GPT-4o diuji dengan mengukur tingkat akurasi dalam mengidentifikasi dan memisahkan pembicara dalam suatu percakapan. Selain itu, fungsi GPT-4o dalam menganalisis kemajuan komunikasi anak penyandang ASD diuji dengan membandingkan waktu analsis dengan metode manual dan membandingkan hasil analisis dari kedua metode.

3.3.5 Pemeliharaan

Bagian pemeliharaan pada metode pengembangan *waterfall* bertujuan untuk memastikan bahwa aplikasi terus bekerja dengan baik setelah diluncurkan. Pada penelitian ini, tahap pemeliharaan meliputi pengoperasian sistem yang telah dikembangkan dan perbaikan beberapa kesalahan yang tidak ditemukan pada tahapan sebelumnya.

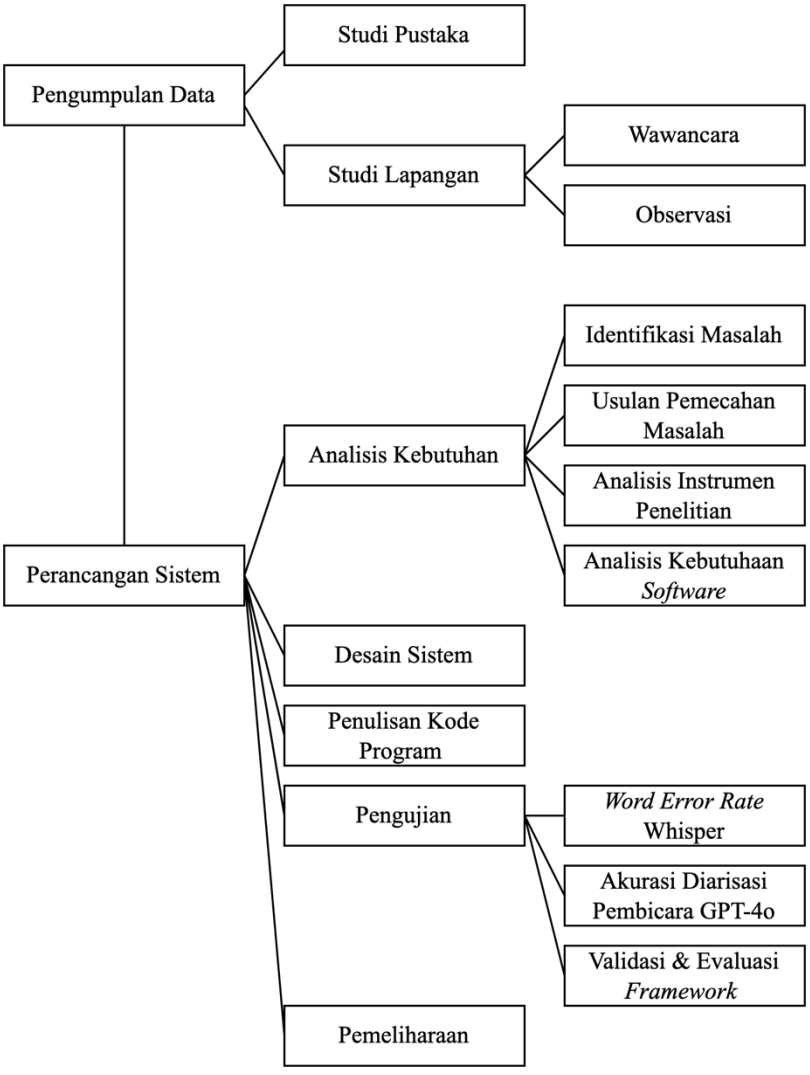
3.4 Kerangka Penelitian

Kerangka penelitian yang dijelaskan pada Gambar 3.2 mencakup beberapa tahapan, dimulai dari pengumpulan data yang dilakukan melalui studi pustaka, wawancara, dan

observasi. Setelah data terkumpul, dilakukan analisis kebutuhan untuk mengidentifikasi masalah, merumuskan solusi, serta menentukan kebutuhan sistem yang diperlukan.

Tahap berikutnya adalah perancangan sistem, yang meliputi proses desain, penulisan kode, dan pengujian sistem menggunakan metode *waterfall*. Pada tahap pengujian, dilakukan evaluasi terhadap kinerja sistem, termasuk pengukuran *Word Error Rate* (WER) untuk modul Whisper, evaluasi akurasi untuk diarisasi pembicara, serta evaluasi dan validasi *framework* menggunakan metode perbandingan, yaitu membandingkan waktu proses dan kualitas hasil analisis.

Tahap akhir dari kerangka penelitian ini adalah pemeliharaan sistem. Pemeliharaan dilakukan untuk memastikan bahwa sistem tetap berfungsi secara optimal setelah diimplementasikan, serta mengatasi masalah yang mungkin muncul di masa mendatang.



Gambar 3.2 Kerangka Penelitian

BAB IV
HASIL DAN PEMBAHASAN

4.1 Analisis Kebutuhan

4.1.1 Identifikasi Masalah

Dalam penelitian ini, identifikasi masalah meliputi tantangan utama yang dihadapi dalam memantau perkembangan komunikasi individu dengan ASD. Secara lebih spesifik, masalah-masalah yang diidentifikasi melalui studi pustaka dan studi lapangan mencakup.

a. Keterbatasan dalam pendekatan manual, analisis perkembangan komunikasi pada individu dengan autisme umumnya dilakukan secara manual, yang memakan waktu dan biaya karena membutuhkan transkripsi dan interpretasi oleh tenaga ahli.
b. Ketidakpastian dalam metode deteksi, saat ini guru-guru di SLB umumnya melakukan deteksi perkembangan komunikasi anak penyandang autisme berdasarkan *feeling* atau intuisi, dengan cara mengingat apakah anak tersebut kemarin sudah mampu atau belum dalam keterampilan komunikasi tertentu. Metode ini tidak didukung oleh matriks atau alat ukur yang tetap, sehingga berpotensi menyebabkan ketidakakuratan dalam pemantauan perkembangan dan evaluasi kemampuan komunikasi anak.

4.1.2 Usulan Pemecahan Masalah

Untuk mengatasi masalah yang telah diidentifikasi, peneliti mengusulkan beberapa langkah pemecahan yang difokuskan pada efisiensi dalam mengukur perkembangan komunikasi anak penyandang autisme. Usulan pemecahan masalah dalam penelitian ini adalah:

a. Membantu guru SLB dalam membuat transkripsi dan interpretasi perkembangan komunikasi dan bahasa anak autisme, dengan mengembangkan sistem berbasis teknologi yang mengintegrasikan model pengenalan suara Whisper dan LLM GPT dari OpenAI.
b. Peneliti akan mengidentifikasi instrumen penelitian atau *matrix-based evaluation* sebagai solusi untuk mengatasi ketidakpastian dalam metode deteksi yang selama ini hanya bergantung pada *feeling* atau ingatan guru mengenai kemampuan komunikasi anak ASD di masa sebelumnya.

4.1.3 Analisis Instrumen Penilaian (*Matrix Based Evaluation*)

Instrumen penilaian atau Matrix Based Evaluation dikembangkan berdasarkan tinjauan literatur dari penelitian oleh Tager-Flusberg et al. (2005) dan Vogindroukas et al. (2022), yang

kemudian divalidasi melalui wawancara dengan guru sekolah luar biasa (SLB) dan observasi langsung di kelas. Matriks ini berfokus pada lima aspek utama komunikasi: kemampuan verbal, pragmatik, semantik, struktur kalimat, dan ekolalia. Matriks ini digunakan untuk mengevaluasi kemajuan komunikasi setiap anak dengan menganalisis kemampuan verbal, pragmatik, semantik, struktur kalimat, dan frekuensi ekolalia mereka. Setiap aspek dinilai berdasarkan respons anak dalam situasi nyata, dengan data analisis diperoleh dari interaksi antara guru dan siswa.

Pembobotan instrumen penilaian diurutkan berdasarkan prioritas, dengan kemampuan verbal dan pragmatik sebagai aspek sekuensial yang dinilai secara berurutan karena peran utamanya dalam fondasi komunikasi. Sebaliknya, semantik, struktur kalimat, dan ekolalia dianggap bersifat paralel, artinya ketiga aspek ini dinilai secara simultan setelah kemampuan verbal dan pragmatik terpenuhi. Semantik, struktur kalimat dan ekolalia saling berkaitan dan memberikan wawasan tambahan tentang perkembangan bahasa anak, tetapi tidak harus dipenuhi secara berurutan.

Matriks ini telah divalidasi menggunakan penilaian ahli dari enam guru berpengalaman di sekolah khusus autisme. Proses validasi melibatkan tinjauan menyeluruh terhadap indikator matriks untuk memastikan kesesuaiannya dengan praktik nyata di kelas dan tantangan komunikasi yang biasanya dihadapi oleh anak-anak dengan ASD. Hasil akhir dari matriks, termasuk deskripsi dan indikator yang telah ditetapkan untuk setiap aspek, diuraikan dalam Tabel 4.1.

Tabel 4.1 *Matrix-Based Evaluation* Komunikasi & Bahasa Siswa ASD

Aspek Penilaian	**Deskripsi**	**Indikator**
Verbal	Kemampuan berbicara dan mengeluarkan suara.	Mengeluarkan suara, mengucapkan kata bermakna.
Semantik	Pemahaman makna kata dan instruksi.	Memahami instruksi sederhana, kosa kata sehari-hari, membuat dan memahami kalimat pertanyaan, menambah kosa kata baru.
Pragmatik	Kemampuan interaksi sosial.	Menyapa, melakukan percakapan dua arah, mengucapkan terimakasih, tolong, dll.
Struktur Kalimat	Penyusunan kalimat dan tata bahasa.	Membuat kalimat sederhana, menambahkan imbuhan, menggunakan konjungsi.

Ekolalia	Frekuensi mengulang kata atau frasa yang didengar.	Jumlah dan frekuensi ekolalia, mengulang topik yang sedang dibicarakan atau mengulangi pertanyaan orang.

4.1.4 Analisis Kebutuhan Sistem

Analisis kebutuhan sistem adalah proses untuk mengidentifikasi dan mendokumentasikan kebutuhan serta persyaratan yang diperlukan dalam pengembangan sistem, sehingga sistem yang dibangun sesuai dengan kebutuhan pengguna dan mampu berfungsi secara efektif dan efisien dalam menyelesaikan masalah yang diidentifikasi. Tahapan dalam analisis kebutuhan sistem meliputi:

Kebutuhan Fungsional

Kebutuhan fungsional menjelaskan fitur-fitur utama yang harus dimiliki sistem untuk mencapai tujuan yang telah ditetapkan. Adapun kebutuhan fungsional dari sistem yang akan dikembangkan yaitu:

a. Sistem mampu mendeteksi percakapan antara anak penyandang autisme dan guru di SLB.
b. Sistem dapat menganalisis perkembangan komunikasi dan pola bahasa anak penyandang anak penyandang autisme dari waktu ke waktu, termasuk perubahasan dalam verbal, semantik, pragmatik, struktur kalimat, dan ekolalia.
c. Sistem menyediakan visualisasi data yang dapat menunjukkan secara jelas perkembangan komunikasi dan pola bahasa anak penyandang autisme, sehingga memudahkan guru SLB dalam memahami kemajuan anak.

Kebutuhan Pengguna

Kebutuhan pengguna mencakup fitur dan kemudahan yang diharapkan pengguna akhir dari sistem agar mereka dapat menggunakan dan memanfaatkan sistem secara optimal. Dalam penelitian ini, pengguna utama adalah guru di SLB yang berinteraksi dengan anak-anak penyandang autisme. Berikut adalah kebutuhan pengguna yang harus dipenuhi oleh sistem:

a. Kemudahan pengoperasian, sistem dirancang untuk digunakan oleh guru SLB, termasuk mereka yang sebagian telah berusia lanjut dan mungkin pemula dalam penggunaan teknologi. Oleh karena itu, sistem harus memiliki antarmuka yang intuitif dan ramah pengguna dengan navigasi yang sederhana.

b. Akses data perkembangan anak, sistem harus memungkinkan guru-guru SLB untuk mengakses data perkembangan komunikasi dan pola bahasa setiap anak secara rinci. Dengan demikian, guru dapat memantau kemajuan anak dari waktu ke waktu. Data ini perlu disajikan melalui antarmuka yang intuitif dan mudah dipahami.
c. Visualisasi laporan yang informatif, pengguna membutuhkan laporan visual yang ringkas dan mudah dipahami untuk mendukung proses evaluasi perkembangan komunikasi dan pola bahasa anak autisme. Laporan ini harus menampilkan data seperti perubahan dalam aspek verbal, semantik, pragmatik, struktur kalimat, dan ekolalia anak dalam format visual yang jelas.

Kebutuhan Teknis

Kebutuhan teknis berfokus pada aspek infrastruktur dan perangkat teknologi yang dibutuhkan agar sistem dapat berfungsi secara optimal dan sesuai dengan kebutuhan pengguna. Dalam penelitian ini, sistem yang akan dikembangkan harus memenuhi kebutuhan teknis berikut:

a. Teknologi pengenalan suara (Whisper)

Teknologi pengenalan suara yang akan digunakan adalah Whisper, model ASR terbaru dari OpenAI yang dirilis pada September 2022. Whisper dikembangkan menggunakan data multibahasa dan multitugas sebanyak 680.000 jam, yang membuatnya mampu mengenali beragam aksen, mengatasi kebisingan latar belakang, dan memahami istilah teknis.

Berdasarkan penjelasan OpenAI, volume dan keragaman data pelatihan ini berkontribusi pada peningkatan kestabilan dan akurasi dalam beragam kondisi. Whisper menyediakan berbagai ukuran model dengan jumlah parameter yang berbeda-beda untuk menyesuaikan kebutuhan pengguna. Terdapat model berukuran *tiny* dengan 39 juta parameter, *base* dengan 74 juta parameter, small dengan 244 juta parameter, dan *medium* dengan 769 juta parameter. Semua ukuran ini tersedia dalam versi *English-only* dan multibahasa. Untuk aplikasi yang memerlukan akurasi lebih tinggi, Whisper juga menyediakan model *large* dan *large*-v2 dengan 1,55 miliar parameter, namun keduanya hanya tersedia dalam versi multibahasa.

Pada penelitian ini, model Whisper versi v2-*large* akan digunakan melalui OpenAI API dengan nama *whisper-1*, yang memberikan performa optimal untuk kebutuhan pengenalan suara yang akurat dalam berbagai konteks.

b. Teknologi model bahasa (GPT-4o)

Dalam penelitian ini, GPT (*Generative Pre-Trained Transformer*) dari OpenAI digunakan untuk menganalisis hasil transkripsi percakapan antara guru SLB dan anak autisme. GPT adalah model berbasis Transformer yang dirancang untuk tugas Pemrosesan Bahasa Alami (NLP), mencakup kemampuan memahami konteks dan menghasilkan teks baru. Salah satu keunggulan utama GPT adalah efisiensi dan kecepatannya dalam melakukan analisis, di mana model ini dapat mengotomatisasi berbagai tugas seperti memberikan respon atau jawaban, yang membuat analisis lebih cepat dan mengurangi risiko kesalahan.

GPT menyediakan beberapa versi, masing-masing memiliki kemampuan dan kapasitas berbeda untuk memenuhi berbagai kebutuhan analisis. Versi-versi ini termasuk GPT-4o, yang dikembangkan untuk meningkatkan ketepatan dan konsistensi dalam tugas-tugas analitis yang lebih kompleks dan melibatkan pemahaman konteks yang mendalam.

Pada penelitian ini, versi model yang akan digunakan adalah GPT-4o untuk menghasilkan analisis yang lebih mendalam dan komprehensif terhadap transkripsi percakapan antara guru SLB dan anak autisme.

Batasan Sistem

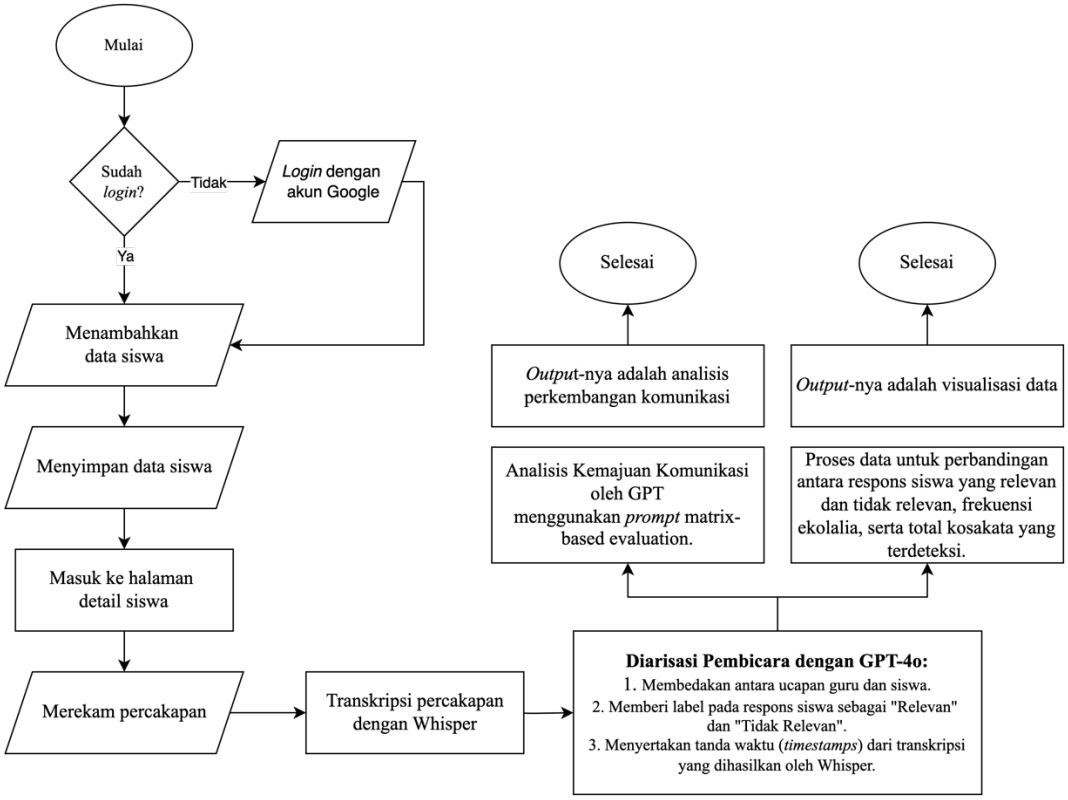
Terdapat beberapa batasan pada sistem yang akan dikembangkan dalam penelitian ini yaitu:

a. Keterbatasan akurasi transkripsi audio, anak penyandang autisme mungkin memiliki intonasi dan artikulasi yang tidak biasa, sehingga Whisper seringkali kesulitan mendeteksinya.
b. Ukuran fail dan durasi rekaman, Model Whisper memiliki batasan durasi rekaman dengan ukuran maksimum fail hanya sebesar 25MB. Karena itu, fail yang berdurasi panjang perlu di-*chunking* agar sesuai dengan batasan ini.
c. Pemahaman konteks yang terbatas, Model GPT-4o mungkin mengalami kesulitan atau bahkan gagal memahami konteks komunikasi anak penyandang autisme secara mendalam, terutama dalam hal emosi dan ekspresi.

4.2 Desain Sistem

4.2.1 Diagram Alir Sistem (*Flowchart*)

Diagram alir atau *flowchart* adalah representasi visual yang menggambarkan urutan langkah-langkah dalam suatu proses kerja sistem. Pada sistem yang dikembangkan dalam penelitian ini, diagram alir digunakan untuk menggambarkan proses yang dilalui oleh pengguna, yaitu guru SLB, dalam menggunakan aplikasi. Diagram ini ini juga mencakup bagaimana sistem memberikan respon terhadap setiap aksi yang dilakukan oleh guru. *Flowchart* yang menggambarkan alur tersebut akan ditampilkan pada Gambar 4.1.



Gambar 4.1 Diagram Alir Sistem

Flowchart yang ditampilkan pada Gambar 4.1 menunjukkan urutan langkah yang dilalui oleh guru dalam aplikasi, di mana sistem akan menangani proses transkripsi dan analisis data. Proses dimulai ketika guru membuka aplikasi, yang kemudian akan memeriksa status *login* guru. Jika guru belum *login*, mereka akan diminta untuk *login* menggunakan akun Google. Namun, jika guru sudah *login* sebelumnya, mereka bisa langsung melanjutkan ke langkah berikutnya tanpa perlu *login* ulang.

Setelah berhasil *login*, guru dapat menambahkan data siswa, seperti informasi dasar mengenai siswa tersebut, ke dalam aplikasi. Setelah data siswa dimasukkan, aplikasi akan

menyimpan data tersebut. Kemudian, guru akan diarahkan ke halaman detail siswa, yang berisi informasi lebih lanjut mengenai siswa.

Pada halaman detail siswa, guru akan merekam percakapan dengan siswa. Setelah percakapan selesai, audio tersebut kemudian ditranskripsi menggunakan Whisper, menghasilkan keluaran berupa teks. Setelah proses transkripsi, diarisisasi pembicara dilakukan menggunakan GPT untuk membedakan antara ucapan siswa dan guru serta memberi label pada respons siswa sebagai "Relevan" atau "Tidak Relevan." Tanda waktu (*timestamps*) juga ditambahkan pada setiap segmen percakapan.

Data terstruktur tersebut kemudian dianalisis untuk menilai kemajuan komunikasi menggunakan evaluasi berbasis matriks, seperti yang dijelaskan dalam Tabel 4.1. Hasil analisis disajikan dalam bentuk ringkasan teks maupun visualisasi data.

4.3 Penulisan Kode Program

4.3.1 *Automatic Speech Recognition* dengan Whisper

Pada tahap ini, peneliti membagi proses penulisan kode pemrograman menjadi dua, yaitu:

Langkah 1: Pengolahan Audio dengan Format M4A

Peneliti menggunakan Format audio MPEG4 AAC digunakan karena dikenal efisien dalam pengelolaan ukuran fail. Format ini menyediakan kualitas audio tinggi dengan ukuran fail yang lebih kecil dibandingkan dengan format lainnya, seperti WAV. Kode pemrograman ditampilkan pada Gambar 4.2.

```
audioRecorder = try AVAudioRecorder(url: captureURL,
   settings: [
      AVFormatIDKey: Int(kAudioFormatMPEG4AAC),
      AVSampleRateKey: 12000,
      AVNumberOfChannelsKey: 1,
      AVEncoderAudioQualityKey: AVAudioQuality.high.rawValue ])
```

Gambar 4.2 Kode Pengolahan Audio dengan M4A

Langkah 2: Proses Transkripsi dengan Whisper

Setelah audio dipersiapkan, langkah berikutnya adalah mengirim fail tersebut ke model Whisper untuk transkripsi. Whisper mendeteksi suara dalam audio dan mengonversinya menjadi teks. Kode pemrograman ditampilkan pada Gambar 4.3.

```
public func generateAudioTransciptions(audioData: Data, fileName: String =
"recording.m4a") async throws -> String {
    var request = URLRequest(url: URL(string:
       "https://api.openai.com/v1/audio/transcriptions")!)

    let boundary: String = UUID().uuidString
        request.timeoutInterval = 30
        request.httpMethod = "POST"
        request.setValue("Bearer \(apiKey)", forHTTPHeaderField:
          "Authorization")
        request.setValue("multipart/form-data; boundary=\(boundary)",
          forHTTPHeaderField: "Content-Type")

        let bodyBuilder = MultipartFormDataBodyBuilder(boundary: boundary,
entries: [
            .file(paramName: "file", fileName: fileName, fileData:
                audioData, contentType: "audio/mpeg"),
            .string(paramName: "model", value: "whisper-1"),
            .string(paramName: "result", value: "raw"),
            .string(paramName: "language", value: "id"),
            .string(paramName: "response_format", value: "verbose_json"),
            .string(paramName: "temperature", value: "0")
        ])
        request.httpBody = bodyBuilder.build()
        let (data, resp) = try await urlSession.data(for: request)
        guard let text = String(data: data, encoding: .utf8) else {
            throw "Invalid format"
        }
        return text
    }
```

Gambar 4.3 Kode Proses Transkripsi dengan Whisper

4.3.2 Diarisasi Pembicara dengan GPT-4o

Setelah transkripsi dari Whisper diperoleh, langkah berikutnya adalah melakukan diarisasi pembicara untuk membedakan siapa yang berbicara dalam percakapan (misalnya, guru atau siswa). Berikut adalah rincian langkah-langkah yang dilakukan dalam memproses percakapan:

Langkah 1: Mengonversi JSON ke Model Data

Transkripsi JSON yang dihasilkan oleh Whisper, yang bersifat verbose, diubah menjadi model data yang lebih ringkas. Proses ini mengurangi ukuran data dan memungkinkan pemrosesan yang lebih cepat oleh GPT-4o, serta menghindari batasan panjang input:

```
let transcriptionResult =
try JSONDecoder().decode(TranscriptionResult.self, from: transcriptionData)

#Model Data
struct Segment: Decodable {
     var start: Double
     var end: Double
     let text: String
      }
struct TranscriptionResult: Decodable {
     let segments: [Segment]
}
```

Gambar 4. 4 Kode Mengonversi JSON ke Model Data

Langkah 2: Membagi Transkripsi ke dalam Beberapa Bagian

Transkripsi dibagi menjadi bagian-bagian sepanjang 60 detik untuk menciptakan segmen yang lebih mudah dikelola, sehingga memungkinkan GPT-4o memproses setiap bagian dengan lebih efektif dan efisien. Untuk mencegah potensi kalimat terpotong di batas waktu antar-chunk, peneliti memanfaatkan pendekatan overlap, di mana sebagian kecil durasi (misalnya, 5 detik) dari akhir satu chunk tumpang tindih dengan awal chunk berikutnya. Overlap ini memastikan bahwa kalimat atau ucapan yang berada di batas waktu tetap utuh dalam setidaknya satu chunk, meningkatkan akurasi diarisasi pembicara oleh GPT-4o dengan menjaga konteks percakapan.

Namun, pendekatan overlap dapat menyebabkan duplikasi data karena segmen transkripsi dalam periode overlap (misalnya, detik 55–60) muncul di dua chunk sekaligus. Untuk mengatasi masalah ini, peneliti menerapkan strategi pasca-pemrosesan berupa penggabungan segmen berdasarkan timestamp. Pendekatan ini melibatkan pemeriksaan waktu mulai (start) dan akhir (end) setiap segmen untuk memastikan bahwa hanya segmen dengan rentang waktu unik yang dipertahankan dalam hasil akhir, sehingga duplikasi dapat dihilangkan. Selain itu, peneliti menambahkan metadata pada setiap segmen untuk mencatat asal chunk-nya (misalnya, chunk 1 atau chunk 2). Metadata ini mempermudah verifikasi segmen yang berasal dari periode

overlap selama pasca-pemrosesan, memastikan penggabungan data dilakukan secara tepat dan efisien. Hasil dari setiap *chunk* kemudian dikirim ke GPT-4o untuk dilakukan diarisasi pembicara. Kode pemrograman ditampilkan pada Gambar 4.5.

```
let chunkedTranscription: [[Segment]]
if transcriptionResult.segments.isEmpty {
   chunkedTranscription = []
} else {
   let totalDuration = transcriptionResult.segments.last!.end -
       transcriptionResult.segments.first!.start
   if totalDuration < 60.0 {
       chunkedTranscription = [transcriptionResult.segments]
   } else {
       chunkedTranscription = chunkTranscription(segments:
       transcriptionResult.segments, chunkSizeInSeconds: 60.0)}
}
#Fungsi untuk chuck transkripsi
func chunkTranscription(segments: [Segment], chunkSizeInSeconds: Double,
overlapInSeconds: Double = 5.0) -> [[Segment]] {
    var chunkedTranscription: [[Segment]] = []
    var currentChunk: [Segment] = []
    var currentTime = segments.first?.start ?? 0
    let chunkSizeWithOverlap = chunkSizeInSeconds - overlapInSeconds

    for segment in segments {
        if currentChunk.isEmpty || (segment.end - currentTime) <
chunkSizeWithOverlap {
            var adjustedSegment = segment
            adjustedSegment.start += currentTime
            adjustedSegment.end += currentTime
            currentChunk.append(adjustedSegment)
            currentTime = adjustedSegment.end
        } else {
            chunkedTranscription.append(currentChunk)
            currentTime = segment.start - overlapInSeconds
            currentChunk = segments.filter { $0.start >= currentTime &&
$0.end <= (currentTime + chunkSizeInSeconds) }
        }
    }
    if !currentChunk.isEmpty {
```

```
        chunkedTranscription.append(currentChunk)
    }
    return chunkedTranscription
}
```

Gambar 4.5 Kode Membagi Transkripsi Menjadi Beberapa Bagian

Langkah 3: Mengirim ke GPT-4o untuk Diarisasi Pembicara

GPT-4o membantu proses diarisasi dengan mengidentifikasi pola percakapan dan membedakan antara pembicara. Peneliti menerapkan sebuah *prompt* yang mencakup beberapa komponen: persona, tugas, instruksi, contoh, instruksi format keluaran, dan permintaan pengguna. Format *prompt* ini secara signifikan meningkatkan kinerja model berbasis GPT (He et al., 2024).

Pendekatan ini memungkinkan GPT-4o untuk membedakan pola berbicara antara guru dan siswa, serta memberi anotasi pada respons siswa yang relevan maupun tidak relevan. Selain itu, *timestamp* ditambahkan ke setiap segmen percakapan untuk memudahkan identifikasi waktu. *Prompt* yang digunakan untuk pemrosesan diarisasi dengan GPT-4o ditampilkan pada Gambar 4.6.

```
Role:
You are a transcription processor responsible for processing dialogue
transcripts. These transcripts contain segments of speech from multiple
speakers (e.g., interactions between a teacher and a student or group
discussions). Your main task is to:
  - Separate and label speakers.
  - Transcribe exactly what is spoken.
  - Mark relevance for student responses.
  - Ensure timestamp format is correct.

Instructions:
Separate and Label Speakers:
  - Identify and label each speaker as either Guru (Teacher) or Siswa
      (Student).
  - In cases of multiple speakers in one segment (e.g., teacher-teacher-
      student interaction), clearly separate and label each speaker’s
      turn.
  - Always label the speaker for each segment, even if there is overlap in
      speaking.
```

Transcribe Exactly:

- Even if the speech is unclear, fragmented, or distorted (e.g., due to autism), transcribe exactly what was spoken without any correction or interpretation.
- Preserve the exact wording of what was said, including hesitations, incomplete sentences, or unclear speech.

Mark Relevance:

- For each student's line, determine if their response is relevant to the conversation.
- If relevant, mark it as (Relevan).
- If not relevant, mark it as (Tidak Relevan).
- Teacher's lines do not need a relevance marking.

Timestamp Format:

- Ensure each piece of dialogue has the correct timestamp in the format: [start_time - end_time].
- If the timestamp exceeds 59 seconds, convert the timestamp into a minutes:seconds format (e.g., [1:00 - 1:05]).
- Maintain the integrity of the timestamps as provided in the original transcription.
- Do not alter or change the timestamps, except when converting to minutes if it exceeds 59 seconds.

Context:

You are tasked with processing a transcription that includes segments of speech from multiple speakers, often containing incomplete or unclear speech, especially from students with autism. The transcript includes both teacher and student dialogue, and you must label each speaker correctly, transcribe exactly as heard, mark relevance for student responses, and ensure timestamps are accurate.

Constraints:

- Do not modify or interpret unclear speech.
- Only label Siswa's lines with relevance, marking them as (Relevan) or (Tidak Relevan).
- Maintain the original timestamps and convert them to the minutes:seconds format if they exceed 59 seconds.
- Do not add any additional words or explanations that are not part of the example format.

```
 [1:00 - 1:05] Guru: "Roti apa, Akmal?"
 [1:07 - 1:09] Guru: "Akmal?"
 [1:12 - 1:04] Siswa: "Roti" (Tidak Relevan)

Do not add any additional words or explanations that are not part of the
example format.
```

Gambar 4.6 *Prompt* pada Proses Diarisasi Pembicara

4.3.3 Analisis Perkembangan Komunikasi dengan GPT-4o

Setelah transkripsi diproses menggunakan GPT-4o dan menghasilkan versi baru dengan format diarisasi pembicara, langkah selanjutnya adalah menganalisis perkembangan komunikasi anak dengan ASD dibandingkan dengan transkripsi sebelumnya. Analisis ini dilakukan dengan bantuan GPT-4o, sehingga memungkinkan perbandingan dan evaluasi yang mendetail terkait perkembangan komunikasi dari waktu ke waktu. Analisis ini dilakukan berdasarkan *matrix-based evaluation* sebagaimana dijelaskan dalam Tabel 4.1, yang disertakan ke dalam *prompt* untuk memandu proses evaluasi. Berikut adalah kode untuk proses tersebut:

```
let combinedText = """
Rekaman pertama: \(transcriptions[1].transcriptionText)
Rekaman kedua: \(transcriptions[0].transcriptionText)"""
let client = OpenAIClient(apiKey: "")
let analysis = try await client.promptChatGPT(
prompt: """

Role : Anda adalah seorang ahli perkembangan anak yang mengkhususkan diri
dalam mengevaluasi kemajuan komunikasi dan pola bahasa pada anak autisme.

Instructions : Buatlah analisis secara mendalam dan detail dengan bahasa
yang mudah dipahami dari dua rekaman interaksi antara seorang anak autisme
dan guru, dengan menilai lima aspek berikut:

1. Verbal: Evaluasi kemampuan anak dalam mengeluarkan suara dan mengucapkan
kata bermakna. Perhatikan apakah terdapat peningkatan atau penurunan dalam
penggunaan suara atau kata yang bermakna. Analisis juga apakah anak mampu
mengucapkan kata dengan benar misal anak mengucapkan ovipal namun
seharusnya ovipar menunjukkan anak masih kesulitan mengucapkan huruf R
```

2. **Pragmatik:** Nilai kemampuan anak dalam interaksi sosial. Tinjau apakah anak mulai menyapa, merespon saat dipanggil, melakukan percakapan dua arah, atau mengucapkan kata-kata seperti "terima kasih" dan "tolong." Identifikasi perubahan spesifik yang terjadi.

3. **Semantik:** Tinjau pemahaman anak terhadap makna kata dan instruksi. Pertimbangkan apakah anak memahami instruksi sederhana, menguasai kosa kata sehari-hari, mampu membuat dan memahami kalimat pertanyaan, serta menambah kosa kata baru.

4. **Struktur Kalimat:** Analisis tata bahasa dan penyusunan kalimat anak. Evaluasi apakah anak mampu membuat kalimat sederhana, menambahkan imbuhan, serta menggunakan konjungsi untuk membentuk kalimat yang lebih kompleks.

5. **Ekolalia:** Pantau frekuensi anak dalam mengulang kata atau frasa yang didengar. Hitung jumlah dan frekuensi ekolalia, dan perhatikan apakah anak mengulang topik yang sedang dibicarakan atau mengulangi pertanyaan dari orang lain.

Untuk setiap aspek, jelaskan perubahan yang terjadi pada kemampuan anak, baik peningkatan maupun penurunan, atau jika tidak ada perubahan sama sekali.

Context : Setiap rekaman berisi sesi di mana anak berinteraksi dengan guru SLB untuk mengukur progress kemampuan komunikasi dan pola bahasa. Tujuan perbandingan ini adalah untuk memberikan gambaran perkembangan dalam aspek verbal, pragmatik, semantik, struktur kalimat, dan frekuensi serta jenis ekolalia.

Constraints :

- Hindari penggunaan istilah teknis yang kompleks dan gunakan bahasa yang mudah dipahami. Batasi analisis hingga maksimal 500 kata.
- Jika tidak siswa tidak ada kemajuan, maka tuliskan saja tidak ada kemajuan bahkan penurunan

Examples :

- Verbal: "Pada Rekaman 1, anak hanya mengeluarkan suara dasar, sementara di Rekaman 2, ia mulai mengucapkan kata bermakna seperti 'mama' atau 'makan.' atau hanya bubbling dan 'eeee'

```
   - Pragmatik: "Di Rekaman 2, anak mulai merespons saat dipanggil dan
     berusaha mengucapkan kata seperti 'terima kasih' dan 'tolong'."
   - Semantik: "Anak menunjukkan pemahaman yang lebih baik terhadap
     instruksi sederhana di Rekaman 2 dibandingkan Rekaman 1. dan kosa
     katanya juga bertambah awalnya hanya 5 jenis kosa kata sekarang bisa
     10 jenis kosa kata"
   - Struktur Kalimat: "Pada Rekaman 2, anak mulai menyusun kalimat
     sederhana, seperti 'Mau main bola.' sudah bisa menambahkan imbuhan
     dan konjungsi juga"
   - Ekolalia: "Frekuensi echolalia berkurang di Rekaman 2, Di rekaman
     satu masih dua kali muncul.""
   - Kesimpulan: "Komunikasi anak mengalami peningkatan jika dibandingkan
     dengan rekaman sebelumnya. Setelah dilihat dari anaisis verbal,
     pragmatik, semantik, struktur kalimat, dan ekolalia."

berikan hasil analisis dengan struktur sama persis examples, jangan
pisahkan rekaman 1 lalu rekaman 2, di dalam verbal harus tercantum rekaman
1 dan 2, begitu juga yang lainnya.

ini adalah transkripsinya : \(combinedText)
""")
```

Gambar 4.7 Kode & *Prompt* pada Analisis Perkembangan Komunikasi

Kode yang dijelaskan pada Gambar 4.7 bertujuan untuk menganalisis perkembangan komunikasi anak dengan ASD menggunakan dua transkripsi interaksi. Proses dimulai dengan menggabungkan kedua transkripsi menjadi satu teks yang dikirimkan sebagai input ke GPT-4o melalui *prompt* yang dirancang khusus untuk mengevaluasi lima aspek komunikasi: verbal, pragmatik, semantik, struktur kalimat, dan ekolalia. Hasil analisis yang diterima diproses untuk memisahkan setiap aspek serta kesimpulannya, kemudian diperbarui ke antarmuka pengguna dan disimpan ke dalam basis data untuk dokumentasi.

4.3.4 Visualisasi Data Perkembangan Komunikasi

Setelah menganalisis perkembangan komunikasi pada siswa dengan ASD peneliti mengolah hasil analisis dan transkripsi melalui beberapa langkah. Pertama, data komunikasi siswa dikategorikan berdasarkan relevansi terhadap konteks pembicaraan dengan memberi label "Relevan" dan "Tidak Relevan". Selanjutnya, setiap kosakata yang diucapkan siswa dipisahkan dan dihitung jumlah penggunaannya untuk mendapatkan gambaran pola bahasa

yang digunakan. Peneliti juga mendeteksi adanya pengulangan kata atau frasa (ekolalia) sebagai bagian dari analisis komunikasi. Hasil pengolahan data ini kemudian divisualisasikan untuk memberikan gambaran yang lebih jelas mengenai perkembangan komunikasi siswa dan mendukung interpretasi hasil secara sistematis.

Visualisasi Data Kosa Kata Terdeteksi

Visualisasi data mengenai kosakata yang terdeteksi pada anak diawali dengan pengumpulan data berupa transkripsi percakapan yang telah diolah menjadi diarisasi pembicara. Transkripsi tersebut kemudian diproses lebih lanjut menggunakan logika pemrograman, seperti yang ditunjukkan pada Gambar 4.8, untuk mengurai dan mendeteksi setiap kosakata yang digunakan oleh anak.

```
func calculateUniqueVocabulary(from transcriptions: [Transcription]) -> Int{
    var uniqueWords: Set<String> = []
    let excludedWords: Set<String> = ["guru", "siswa", "relevan", "tidak
relevan"]

    for transcription in transcriptions {
    let lines = transcription.transcriptionText.lowercased().components
   (separatedBy: .newlines)

     for line in lines {
        let trimmedLine = line.trimmingCharacters(in: .whitespacesAnd
          Newlines)
          if trimmedLine.contains("anak autisme") || trimmedLine.contains
          ("siswa") {
          if let messageStart = trimmedLine.firstIndex(of: ":") {
            let message = trimmedLine[trimmedLine.index(after:
             messageStart)...] .trimmingCharacters(in: .whitespacesAndNew
             lines)

         let words = message.components(separatedBy: .whitespacesAndNew
          lines).filter { !$0.isEmpty }

         for word in words {
         if word.range(of: #"^\d+(\.\d+)?$"#, options: .regularExpression)
             != nil || word.contains(":") {
```

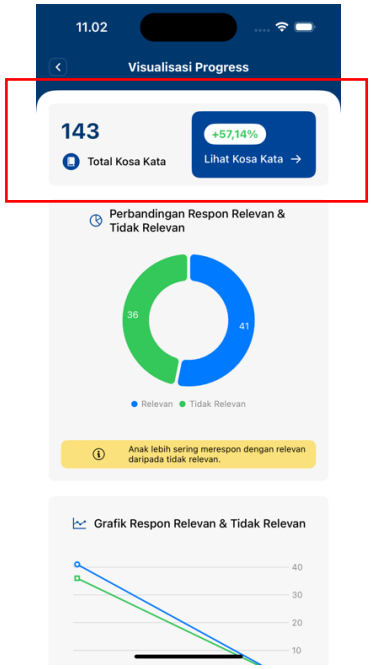
```
            continue
            }
        let cleanedWord = word.trimmingCharacters(in: .punctuationCharacters
            )
        if !excludedWords.contains(cleanedWord) {
            uniqueWords.insert(cleanedWord)
                                }
                            }
                        }
                    }
                }
            }
```

Gambar 4.8 Kode untuk Memisahkan Kosa Kata

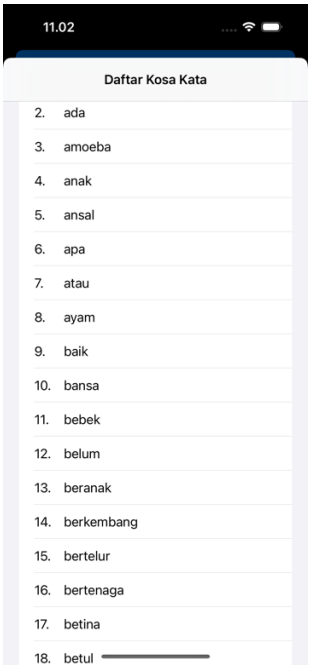
Fungsi *calculateUniqueVocabulary* dirancang untuk menghitung jumlah kosakata unik dari data transkripsi yang diberikan. Prosesnya diawali dengan mendeklarasikan kumpulan (*set*) kosong untuk menyimpan kata-kata unik serta daftar kata yang dikecualikan, seperti "guru", "siswa", "relevan", dan "tidak relevan". Fungsi ini kemudian memproses setiap transkripsi dengan memecah teks transkripsi menjadi baris-baris terpisah. Pada setiap baris, jika ditemukan kata-kata yang terkait dengan "anak autisme" atau "siswa", fungsi akan mencari bagian teks setelah tanda titik dua (:) yang mengindikasikan pesan utama.

Setiap kata dalam teks tersebut dipecah menjadi elemen individu, dibersihkan dari spasi kosong, angka, format waktu, serta tanda baca. Fungsi juga memastikan kata-kata dari daftar pengecualian tidak dimasukkan ke dalam kumpulan kosakata unik. Setelah semua kata diproses, kumpulan kosakata unik ini diubah menjadi *array* menggunakan fungsi *Array(uniqueWords)* untuk mempermudah penyimpanan.

Selanjutnya, array kosakata unik tersebut disimpan ke dalam penyimpanan lokal menggunakan *UserDefaults* dengan kunci "*uniqueVocabularyWords*". Penyimpanan ini memungkinkan data kosakata unik untuk diakses kembali di masa mendatang tanpa perlu memproses ulang data transkripsi. Dengan langkah ini, fungsi tidak hanya menghitung jumlah kosakata unik tetapi juga menyimpan hasilnya secara efisien untuk ditampilkan pada aplikasi seperti yang ditunjukkan pada Gambar 4.9 dan Gambar 4.10.



Gambar 4.9 Tampilan Kosa Kata & Persentase Perkembangannya



Gambar 4.10 Tampilan Daftar Kosa Kata

Visualisasi Data Respon Relevan & Tidak Relevan

Visualisasi data respon relevan dan tidak relevan diolah dari hasil transkripsi yang telah ditandai sebagai relevan atau tidak relevan selama proses diarisasi pembicara. Selanjutnya,

transkripsi dipisahkan untuk mengambil data relevansi menggunakan fungsi parseUtterances yang ditujukkan pada Gambar 4.11. Fungsi ini bertugas mengolah data transkripsi menjadi format yang lebih terstruktur. Proses dimulai dengan memecah teks transkripsi menjadi baris-baris terpisah. Setiap baris kemudian dianalisis menggunakan teknik pencocokan pola (*regex*) untuk mengekstrak informasi penting, seperti waktu mulai dan selesai dialog, identitas pembicara (Guru atau Siswa), teks dialog, dan relevansi percakapan siswa.

```
func parseUtterances(from formattedText: String, transcriptionid: UUID) ->
[Utterance] {
        var utterances = [Utterance]()
        let lines = formattedText.split(separator: "\n")

        let pattern = #"^\[(\d{1,2}:\d{2}) - (\d{1,2}:\d{2})\]
          (Guru|Siswa):\s+"([^"]+)"(?: \((Relevan|Tidak Relevan)\))?$"#
        let regex = try? NSRegularExpression(pattern: pattern, options: [])

        for line in lines {
            if let match = regex?.firstMatch(in: String(line), options: [],
             range: NSRange(location: 0, length: line.utf16.count)) {
                if let startTimeRange = Range(match.range(at: 1), in: line),
                   let endTimeRange = Range(match.range(at: 2), in: line),
                   let speakerRange = Range(match.range(at: 3), in: line),
                   let textRange = Range(match.range(at: 4), in: line) {

                    let startTime = Double(line[startTimeRange]) ?? 0.0
                    let endTime = Double(line[endTimeRange]) ?? 0.0
                    let speaker = String(line[speakerRange])
                    let text = String(line[textRange])
                    var relevance = false
                    if speaker == "Siswa",
                       let relevanceRange = Range(match.range(at: 5), in:
                       line) {
                        relevance = line[relevanceRange] == "Relevan"}

                    let utterance = Utterance(
                        id: UUID(),
                        transcriptionid: transcriptionid,
                        speaker: speaker,
                        starttime: startTime,
```

```
                        endtime: endTime,
                        text: text,
                        relevance: relevance )
                    utterances.append(utterance)}
            } else { print("No regex match for line: \(line)")}}
        return utterances
      }
```

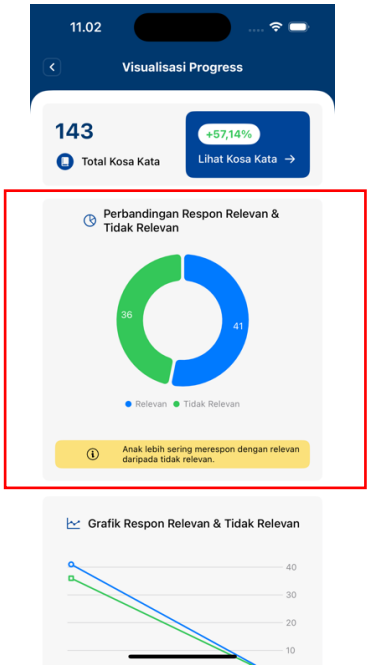
Gambar 4.11 Kode Mengolah Transkripsi dengan Regex

Setelah transkripsi diproses dengan pencocokan pola, data relevansi respon siswa diolah lebih lanjut menggunakan dua fungsi utama. Fungsi *calculateRelevanceCounts* yang ditampilkan pada Gambar 4.12 digunakan untuk menghitung jumlah dialog relevan dan tidak relevan yang diucapkan siswa dengan memfilter data berdasarkan atribut *relevance* dan *speaker*. Hasil perhitungan ini disimpan dalam *UserDefaults* untuk memudahkan akses tanpa perlu menghitung ulang yang kemudian di tampilkan hasilnya di sistem seperti pada Gambar 4.13.

```
func calculateRelevanceCounts() {
        self.relevantCount = utterances.filter { $0.relevance && $0.speaker
          == "Siswa" }.count
        self.notRelevantCount = utterances.filter { !$0.relevance &&
          $0.speaker == "Siswa" }.count

        UserDefaults.standard.set(self.relevantCount, forKey:
          "relevantCount")
        UserDefaults.standard.set(self.notRelevantCount, forKey:
          "notRelevantCount")
    }
```

Gambar 4.12 Kode untuk Menghitung Respon Relevan & Tidak Relevan



Gambar 4.13 Tampilan Perbandingan Respon Relevan & Tidak Relevan

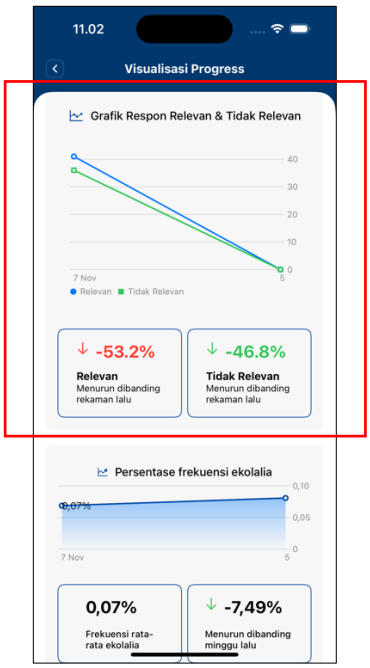
Selain memvisualisasikan perbandingan respon relevan dan tidak relevan, sistem juga menampilkan pola peningkatan atau penurunan jumlah respon relevan dan tidak relevan dari waktu ke waktu. Untuk mendukung analisis ini, fungsi *calculateProgressData* yang ditunjukkan pada Gambar 4.14 digunakan untuk mengelompokkan transkripsi berdasarkan ID (*transcriptionId*) dan menghitung jumlah dialog relevan serta tidak relevan dalam setiap grup. Data hasil pengelompokan ini, yang mencakup tanggal, ID transkripsi, serta jumlah respon relevan dan tidak relevan, disimpan dalam daftar *progressData*. Kemudian hasil perhitungan di visualisasikan di sistem seperti yang ditampilkan pada Gambar 4.15.

```
func calculateProgressData() -> [(date: Date, transcriptionId: UUID,
relevantCount: Int, notRelevantCount: Int)] {
        let groupedUtterances = Dictionary(grouping: utterances) {
          $0.transcriptionid }

        for (transcriptionId, utterances) in groupedUtterances {
            if let transcription = transcriptions.first(where: { $0.id ==
             transcriptionId }) {
                let relevantCount = utterances.filter { $0.relevance &&
                    $0.speaker == "Siswa" }.count
                let notRelevantCount = utterances.filter { !$0.relevance &&
                    $0.speaker == "Siswa" }.count
```

```
                let date = Date(timeIntervalSince1970: TimeInterval
                    (transcription.createdAt) / 1000)
                progressData.append((date, transcriptionId, relevantCount,
                    notRelevantCount))
            }
        }
        return progressData
    }
```

Gambar 4.14 Kode untuk Menghitung Perkembangan Respon Relevan & Tidak Relevan



Gambar 4.15 Tampilan Tren Perkembangan Respon Relevan & Tidak Relevan

Visualisasi Data Frekuensi Ekolalia

Visualisasi data frekuensi ekolalia dikembangkan berdasarkan teks transkripsi yang dianalisis menggunakan fungsi *calculateEcholaliaFrequency*. Fungsi ini menghitung frekuensi ekolalia dalam teks dengan menganalisis pengulangan kata. Teks yang diinput terlebih dahulu dikonversi menjadi huruf kecil untuk memastikan konsistensi, kemudian dipisahkan berdasarkan spasi atau baris untuk menghasilkan daftar kata. Frekuensi kemunculan setiap kata dihitung menggunakan struktur data kamus (*dictionary*), dan kata-kata yang muncul lebih dari satu kali diidentifikasi sebagai kata berulang. Frekuensi ekolalia kemudian dihitung sebagai rasio antara jumlah kata berulang dan total kata dalam teks, dengan hasil yang dikembalikan dalam bentuk bilangan desimal. Proses pemrograman yang digunakan untuk analisis ini

ditampilkan pada Gambar 4.16 dan visualisasi data frekuensi ekolalia ditunjukkan pada Gambar 4.17.

```
func calculateEcholaliaFrequency(text: String) -> Double {
        let words = text.lowercased().components(separatedBy:
          .whitespacesAndNewlines)
        let totalWords = words.count
        var wordCounts: [String: Int] = [:]

        for word in words {
            wordCounts[word, default: 0] += 1
        }
        let repeatedWords = wordCounts.filter { $0.value > 1 }.count
        return Double(repeatedWords) / Double(totalWords)
    }
```

Gambar 4.16 Kode Pemrograman Menghitung Frekuensi Ekolalia

Gambar 4.17 Tampilan Visualisasi Data Frekuensi Ekolalia

4.4 Pengujian

4.4.1 Hasil Pengujian ASR Whisper

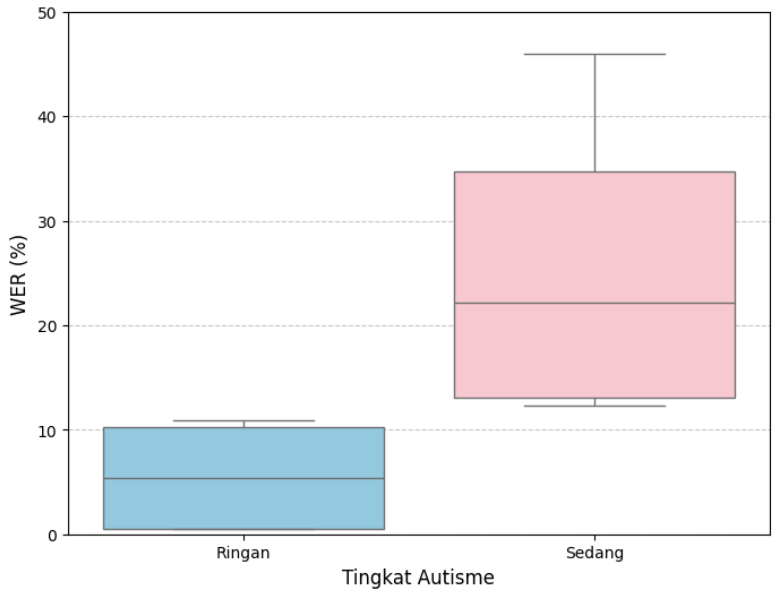
Untuk mengevaluasi kinerja model transkripsi Whisper ASD dalam mentranskripsikan percakapan antara anak dengan ASD dan guru, peneliti menggunakan *Word Error Rate* (WER) sebagai metrik pengukurannya. Hasil WER dibandingkan antara anak dengan tingkat autisme

ringan dan sedang, yang divisualisasikan pada Gambar 4.2. *Word Error Rate* (WER) dihitung menggunakan rumus 4.1.

$$\text{WER} = \frac{S + D + I}{\text{N}}.100\% \qquad (4.1)$$

Dengan keterangan:

a. S: Jumlah substitusi (kata yang salah ditranskripsikan)
b. D: Jumlah kata yang dihapus
c. I: Jumlah kata yang disisipkan
d. N: Jumlah total kata dalam referensi (*ground truth*)



Gambar 4.18 Hasil *Word Error Rate* dari Whisper

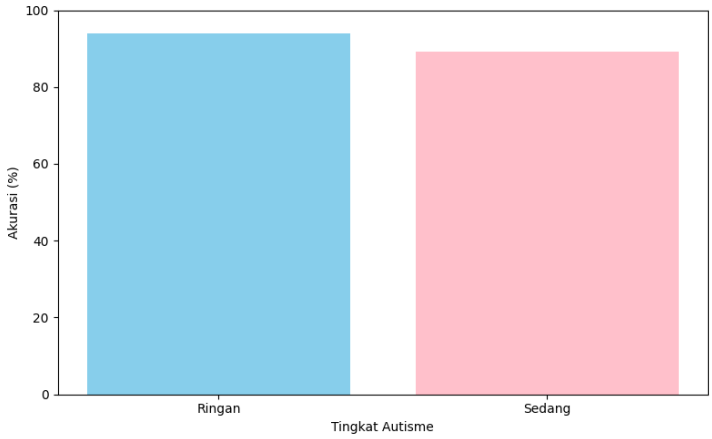
Hasil penelitian yang ditampilkan pada Gambar 4.18 menunjukkan bahwa rata-rata *Word Error Rate* (WER) untuk anak dengan autisme ringan relatif rendah, berkisar antara 0,56% hingga 10%, dengan nilai median sekitar 5%. Distribusi nilai WER yang sempit mencerminkan konsistensi hasil transkripsi, menyoroti kemampuan model untuk menangkap ucapan yang terstruktur dan jelas dengan akurat. Sebaliknya, rata-rata WER untuk anak dengan autisme sedang jauh lebih tinggi, yaitu berkisar antara 13% hingga 46%, dengan nilai median sekitar 23%. Sebaran nilai WER yang luas pada kelompok ini menunjukkan tingkat kompleksitas yang lebih tinggi dalam pola komunikasi mereka.

Temuan penelitian ini mengonfirmasi bahwa kemampuan komunikasi verbal anak dengan autisme secara signifikan memengaruhi kinerja sistem ASR. Pada anak dengan autisme ringan, di mana ucapan cenderung lebih jelas dan terstruktur, sistem ASR menunjukkan kinerja yang optimal, sebagaimana dibuktikan dengan nilai WER yang rendah dan stabil. Sebaliknya, anak dengan autisme sedang menghadirkan tantangan komunikasi yang lebih besar, yang mengakibatkan nilai WER yang lebih tinggi. Hasil ini sejalan dengan temuan (O'Sullivan et al., 2023) yang menunjukkan bahwa sistem ASR bekerja lebih efektif dalam memproses pola ucapan yang jelas dan stabil.

4.4.2 Hasil Pengujian Diarisasi Pembicara

Penelitian ini mengevaluasi akurasi *speaker diarization* menggunakan model GPT-4o pada dua tingkat autisme: ringan dan sedang. Akurasi dihitung menggunakan rumus berikut:

$$\text{Akurasi} = \frac{Jumlah\ Prediksi\ yang\ Benar}{Total\ Prediksi}.100\% \qquad (4.2)$$



Gambar 4.19 Hasil Akurasi Diarisasi Pembicara oleh GPT-4o

Gambar 4.19 mengilustrasikan hasil akurasi diarisasi pembicara pada dua kategori autisme. Untuk autisme ringan, sistem mencapai akurasi sebesar 93,9%, yang menunjukkan kinerja tinggi dalam mengidentifikasi pembicara pada kategori ini. Sebaliknya, untuk autisme sedang, akurasi sistem sedikit lebih rendah, yaitu sebesar 89,2%, mencerminkan tantangan dalam memproses pola komunikasi kompleks yang khas pada kelompok ini. Perbandingan ini menunjukkan adanya kesenjangan kinerja sebesar 4,7% antara kedua kategori, yang

mengindikasikan bahwa sistem lebih efektif dalam menangani pola komunikasi yang terstruktur dan jelas.

Temuan ini menunjukkan bahwa kinerja sistem diarisasi pembicara dipengaruhi oleh kompleksitas komunikasi verbal pada individu dengan autisme. Pada anak dengan autisme sedang, sistem hanya mencapai akurasi sebesar 89,2%. Penurunan kinerja ini disebabkan oleh tantangan seperti *echolalia*, interaksi dua arah yang terbatas, dan kebutuhan guru untuk sering mengulang pertanyaan. Faktor-faktor ini menciptakan ambiguitas yang menyulitkan sistem untuk membedakan pembicara secara efektif. Sebaliknya, sistem menunjukkan akurasi yang lebih tinggi (93,9%) pada anak dengan autisme ringan. Ucapan mereka cenderung lebih terstruktur dan melibatkan komunikasi dua arah yang konsisten, yang membantu model dalam membedakan pembicara dengan lebih akurat.

Meskipun akurasi bervariasi tergantung pada kompleksitas komunikasi, GPT-4o terbukti menjadi solusi yang menjanjikan untuk diarisasi pembicara otomatis. Namun, sistem ini masih memerlukan keterlibatan guru, terutama untuk anak dengan autisme sedang, seperti memberikan koreksi dan penyesuaian terhadap hasil diarisasi. Dengan mengatasi keterbatasan yang diidentifikasi dalam penelitian sebelumnya oleh (O'Sullivan et al., 2023), di mana diarisasi pembicara dilakukan secara manual, otomatisasi ini meningkatkan penelitian tentang autisme dengan menyediakan alat yang andal dan dapat diskalakan untuk mengidentifikasi dan membedakan pembicara.

4.4.3 Validasi dan Evaluasi *Framework* Perbandingan Analisis Waktu

Waktu analisis antara metode manual dan kerangka kerja otomatis yang mengintegrasikan Whisper dan GPT-4o dibandingkan. Hasil perbandingan tersebut disajikan dalam Tabel 4.2.

Tabel 4.2 Perbandingan Waktu Analisis Antara Metode Manual dan Metode Otomatis (*Automated Framework*)

Tahap Analisis	**Analisis Manual (Menit)**	***Automated Framework* (Menit)**	**Pengurangan Waktu**
Transkripsi	Tidak Dilakukan	0,2	-
Diarisasi Pembicara	Tidak Dilakukan	0,7	-
Analisis Komunikasi	3	0,3	90%

Penulisan Laporan	8	0	100%
Total Waktu	11	1,2	89,1%

Tabel 4.2 menyajikan perbandingan antara *automated framework* yang diusulkan dengan metode manual. Pada metode manual, proses transkripsi dan diarisasi pembicara tidak dilakukan. Sistem otomatis mampu menyelesaikan transkripsi dalam waktu 0,2 menit dan diarisasi pembicara dalam 0,7 menit.

Untuk analisis komunikasi, proses manual membutuhkan waktu 3 menit, sementara kerangka kerja otomatis menguranginya menjadi 0,3 menit, menghasilkan pengurangan waktu sebesar 90%. Demikian pula, penulisan laporan yang dalam metode manual membutuhkan waktu rata-rata 8 menit dihilangkan sepenuhnya (0 menit) dengan menghasilkan hasil terstruktur yang langsung siap digunakan, sehingga mencapai pengurangan waktu sebesar 100%.

Automated Framework berhasil mengurangi total waktu pemrosesan dari 11 menit menjadi 1,2 menit, dengan pencapaian pengurangan waktu sebesar 89,1%. Peningkatan yang signifikan ini menunjukkan kemampuan sistem untuk menyederhanakan tugas-tugas yang sebelumnya memakan waktu, termasuk transkripsi, diarisasi pembicara, analisis komunikasi, dan pembuatan laporan. Dengan mengotomatisasi proses-proses tersebut, *automated framework* ini mampu mengatasi keterbatasan metode manual yang sering kali memakan waktu dan membutuhkan banyak sumber daya.

Perbandingan Hasil Analisis oleh Guru SLB

Peneliti telah membandingkan hasil analisis antara metode manual yang dilakukan oleh guru SLB dan *automated framework* yang mengintegrasikan Whisper dan GPT-4o. Hasil perbandingan tersebut disajikan dalam Tabel 4.1 dan Tabel 4.4.

Tabel 4.3 Perbandingan Analisis Komunikasi Antara Metode Manual dan Otomatis pada Autisme Ringan

Aspek	**Analisis Manual (Anak 1)**	***Automated Framework* (Anak 1)**	**Analisis Manual (Anak 2)**	***Automated Framework* (Anak 2)**
Verbal	-	Pada Rekaman 1, anak mengucapkan kata seperti "Vegetatif" dan	-	Pada Rekaman 1, anak menunjukkan kemampuan verbal

		"Generatif," namun terdapat kesalahan pengucapan seperti "Ovipal" menjadi "Ovipar." Pada Rekaman 2, pengucapan membaik untuk istilah seperti "Ovipar," dan anak mulai menyusun jawaban panjang seperti "Telur, embrio, ayam dewasa."		yang cukup konsisten. Anak mampu mengucapkan kata bermakna seperti "baik," "sudah," "roti," dan "uang." Namun, terdapat pengulangan kata "pakai" yang tidak relevan, menunjukkan anak masih memiliki kesulitan memahami konteks. Anak juga mampu mengucapkan angka seperti "dua" dan "lima ratus," meskipun sering tidak relevan dengan pertanyaan guru. Pada Rekaman 2, kemampuan verbal anak cenderung stagnan. Anak hanya mengeluarkan suara dasar dan kata sederhana seperti "seratus," namun tidak ada peningkatan dalam pengucapan kata bermakna atau kalimat yang relevan. Anak juga tidak menunjukkan perbaikan dalam artikulasi atau

				penyampaian kata-kata baru.
Pragmatik	-	Rekaman 1 menunjukkan respons pasif dengan kalimat pendek. Pada Rekaman 2, anak lebih responsif, seperti menjelaskan siklus hidup ayam, dan memahami konteks sosial.	-	Rekaman 1 menunjukkan anak merespons instruksi sederhana. Pada Rekaman 2, anak lebih responsif dengan jawaban seperti "Pakai helm" dan mulai memahami interaksi sosial lebih kompleks.
Semantik	-	Anak memahami instruksi sederhana di kedua rekaman. Pada Rekaman 2, ada peningkatan pemahaman konsep seperti "perkembangan generatif" dan kemampuan mengidentifikasi reproduksi hewan.	-	Pada Rekaman 1, anak memahami instruksi sederhana tetapi kadang salah menjawab. Pada Rekaman 2, pemahaman meningkat dengan jawaban seperti "Empat kali dua sama dengan delapan."
Struktur Kalimat	-	Rekaman 1 menunjukkan penggunaan kalimat sederhana seperti "Bebek, ayam." Pada Rekaman 2, anak mulai menggunakan kalimat lebih kompleks, seperti "Reproduksi generatif melalui pembuahan."	-	Rekaman 1 menunjukkan kalimat sederhana seperti "Baik." Pada Rekaman 2, anak mulai menggunakan kalimat lebih kompleks, seperti "Habis menggosok gigi terus pakai handuk."
Ekolalia	-	Pada Rekaman 1, anak sering mengulang frasa	-	Pada Rekaman 1, anak sering mengulang kata

		seperti "Reproduksi tunas." Pada Rekaman 2, pengulangan berkurang dan menjadi lebih relevan, seperti "Ayam bertelur tiga minggu jadi embrio."		seperti "Tidur." Pada Rekaman 2, frekuensi pengulangan menurun, dan anak mulai memberikan jawaban lebih relevan.
Emosi	-	-	-	-
Kesimpulan	Kosakata anak bertambah.	Komunikasi anak menunjukkan peningkatan signifikan antara Rekaman 1 dan Rekaman 2. Aspek verbal, pragmatik, semantik, struktur kalimat, dan ekolalia semuanya menunjukkan perkembangan positif. Anak semakin mampu memahami dan merespons percakapan dengan relevan, mencerminkan kemajuan dalam keterampilan komunikasi dan pola bahasa.	Komunikasi dua arah sudah cukup baik, dengan peningkatan kecepatan respons terhadap pertanyaan. Tantangan utama terletak pada saat anak kehilangan fokus, meskipun anak mulai menunjukkan konsistensi dalam merespons sesuai konteks.	Komunikasi anak menunjukkan peningkatan antara Rekaman 1 dan Rekaman 2. Dalam aspek verbal, baik pengucapan maupun panjang respons mengalami peningkatan. Dalam aspek pragmatik, interaksi sosial anak menjadi lebih responsif dan terlibat dalam percakapan. Dalam aspek semantik, pemahaman terhadap instruksi sederhana dan penggunaan kosakata terlihat lebih kokoh. Struktur kalimat juga meningkat, dengan anak mulai membentuk kalimat yang lebih lengkap. Frekuensi ekolalia berkurang, menunjukkan

				kemajuan dalam komunikasi yang lebih terfokus.

Tabel 4.4 Perbandingan Analisis Komunikasi Antara Metode Manual dan Metode Otomatis pada Autisme Sedang

Aspek	**Analisis Manual (Anak 1)**	***Automated Framework* (Anak 1)**	**Analisis Manual (Anak 2)**	***Automated Framework* (Anak 2)**
Verbal	-	Pada Rekaman 1, anak menunjukkan kemampuan verbal yang cukup baik, seperti mengucapkan kata-kata bermakna "Naila," "Baik," "Om jagi," dan "Coklat." Namun, terdapat beberapa pengulangan yang tidak relevan, seperti "Senek" dan "Tadi diantar om jagi." Di Rekaman 2, anak lebih berinisiatif dalam berbicara dan menambahkan kata-kata baru seperti "tempe," "mobil kuning," dan "pencil." Selain itu, pengucapan lebih konsisten dibandingkan rekaman sebelumnya, meskipun masih ada beberapa kesalahan dalam pelafalan.	-	Pada Rekaman 1, anak mulai menunjukkan kemampuan pragmatik dengan merespons pertanyaan guru, seperti ketika ditanya "apa kabar" dan menjawab "baik." Namun, anak seringkali gagal memahami konteks percakapan, terbukti dari pengulangan kata-kata yang tidak relevan, seperti "pakai" atau "uang," tanpa melanjutkan ke percakapan dua arah. Pada Rekaman 2, tidak ada peningkatan signifikan dalam aspek pragmatik. Anak tidak tampak

				merespons panggilan guru dengan konsistensi, seperti ketika guru berulang kali memanggil "Akmal" tanpa respons yang memadai. Anak juga belum menunjukkan inisiatif untuk memulai atau mempertahankan percakapan.
Pragmatik	-	Di Rekaman 1, anak hanya merespons pertanyaan guru dan cenderung pasif dalam berbicara. Siswa belum menampilkan banyak inisiatif dalam percakapan. Di Rekaman 2, ada peningkatan yang signifikan di mana anak dapat menjawab pertanyaan dengan lebih tepat dan kadang meniru sapaan dan respons yang diberikan oleh guru, seperti "Iya, pinter" dan "Pinter, tos dulu." Anak juga sudah bisa mengekspresikan keinginan seperti "mau belajar."	-	Pada Rekaman 1, anak mulai menunjukkan kemampuan pragmatik dengan merespons pertanyaan guru, seperti ketika ditanya "apa kabar" dan menjawab "baik." Namun, anak seringkali gagal memahami konteks percakapan, terbukti dari pengulangan kata-kata yang tidak relevan, seperti "pakai" atau "uang," tanpa melanjutkan ke percakapan dua arah. Pada Rekaman 2, tidak ada peningkatan

				signifikan dalam aspek pragmatik. Anak tidak tampak merespons panggilan guru dengan konsistensi, seperti ketika guru berulang kali memanggil "Akmal" tanpa respons yang memadai. Anak juga belum menunjukkan inisiatif untuk memulai atau mempertahankan percakapan.
Semantik	-	Pada Rekaman 1, anak menunjukkan pemahaman dasar terhadap instruksi sederhana dan dapat mengenali nama-nama dasar seperti "rambut," "mata," dan "kaki." Namun, saat diminta menjelaskan fungsi atau deskripsi, jawabannya kadang-kadang tidak relevan atau salah. Di Rekaman 2, anak menunjukkan pemahaman yang lebih baik terhadap instruksi dan lebih banyak menambah kosa kata baru seperti "mobil kuning," "kuping," "celana," dan	-	Pada Rekaman 1, anak memahami beberapa kata dan instruksi sederhana, seperti "roti apa" dan "uang berapa," meskipun jawaban sering tidak relevan. Kosakata anak masih terbatas pada kata-kata dasar seperti "dua," "ratus," dan "rupiah," tanpa penambahan kosakata baru yang bermakna. Pada Rekaman 2, pemahaman semantik anak tidak menunjukkan peningkatan. Anak

		"tempe." Siswa juga bisa mengidentifikasi dua objek dalam konteks, seperti jumlah mata dan hidung.		belum mampu memahami instruksi sederhana seperti "lihat" dengan konsisten, dan sering mengulang kata yang sama tanpa memahami maknanya. Tidak ada indikasi penambahan kosa kata baru dibandingkan Rekaman 1.
Struktur Kalimat	-	Di Rekaman 1, anak mulai mampu membentuk kalimat sederhana, tetapi sering kali hanya menggunakan satu atau dua kata saat merespons. Dalam Rekaman 2, terlihat sedikit peningkatan, karena mulai menggunakan lebih banyak kata, contohnya "Naik mobil kuning" dan "Mata untuk melihat." Terdapat usaha menggunakan lebih banyak konjungsi atau frasa untuk memperluas jawaban, tetapi masih membutuhkan dukungan dalam membentuk kalimat yang lebih kompleks.	-	Pada Rekaman 1, anak hanya mampu menyusun kata tunggal atau frasa sederhana, seperti "dua ratus rupiah," tanpa membentuk kalimat yang lengkap atau kompleks. Tidak ada penggunaan imbuhan atau konjungsi dalam ucapan anak. Pada Rekaman 2, kemampuan struktur kalimat anak tetap stagnan. Anak hanya mengulang kata-kata yang diajarkan oleh guru, seperti "seratus," tanpa kemampuan menyusun kalimat baru. Tidak ada

				perkembangan dalam membentuk kalimat sederhana maupun kompleks.
Ekolalia	-	Frekuensi ekolalia tetap ada di kedua rekaman, namun lebih berkurang di Rekaman 2. Di Rekaman 1, anak beberapa kali mengulangi kata atau frasa tanpa menambah makna yang signifikan. Di Rekaman 2, meskipun tetap mengulang kata atau frasa yang sudah dikatakan oleh guru, konteks penggunaan mulai lebih relevan dan tampaknya digunakan untuk memahami pertanyaan yang diberikan.	-	Pada Rekaman 1, ekolalia cukup sering terjadi. Anak mengulang kata-kata guru seperti "pakai," "uang," dan "dua ribu" tanpa relevansi terhadap percakapan. Ekolalia ini menunjukkan bahwa anak belum memahami konteks percakapan secara mendalam. Pada Rekaman 2, frekuensi ekolalia tetap tinggi. Anak mengulang kata-kata seperti "seratus" dan "lihat" berulang kali tanpa makna yang relevan. Hal ini menunjukkan tidak ada penurunan dalam pola ekolalia dibandingkan Rekaman 1.
Emosi	Anak sering menunjukkan kemarahan dan kesulitan	-	Anak kadang datang ke sekolah dengan	-

	mengatur emosi.		suasana hati buruk, enggan mengikuti kegiatan, dan lebih sering berteriak atau bermain.	
Kesimpulan	Anak masih kesulitan menyusun kalimat dan sering berbicara dengan kalimat yang patah-patah.	Komunikasi anak menunjukkan kemajuan dari Rekaman 1 ke Rekaman 2. Terdapat peningkatan kemampuan dalam semua aspek yang dianalisis, dengan peningkatan lebih jelas dalam pragmatik, semantik, dan struktur kalimat. Frekuensi ekolalia juga menurun, menunjukkan anak mulai lebih aktif dalam memahami dan membentuk komunikasi yang lebih bermakna.	Anak belum menunjukkan kemampuan komunikasi yang baik dan masih berbicara dengan kata yang tidak jelas.	Kemampuan komunikasi anak tidak menunjukkan peningkatan signifikan antara Rekaman 1 dan Rekaman 2. Dalam aspek verbal, pragmatik, semantik, struktur kalimat, dan ekolalia, anak cenderung stagnan dengan pengulangan pola yang sama. Evaluasi lebih lanjut diperlukan untuk memahami faktor-faktor yang memengaruhi perkembangan komunikasi anak.

Tabel 4.3 adalah rangkuman membandingkan analisis manual dan kerangka kerja otomatis untuk autisme ringan, di mana sistem otomatis mampu mengungkap banyak peningkatan yang tidak terlihat melalui analisis manual. Misalnya, anak menunjukkan perbaikan pengucapan, seperti pada kata "Ovipar," serta mampu membentuk kalimat yang lebih panjang dan jelas. Anak juga terlihat lebih aktif dalam berinteraksi, menjawab pertanyaan dengan konteks yang tepat, serta memahami konsep yang lebih kompleks seperti

"perkembangan generatif." Selain itu, analisis otomatis mencatat penurunan frekuensi pengulangan kata (ekolalia) dengan pengulangan yang lebih relevan secara kontekstual.

Untuk autisme sedang, seperti dijelaskan dalam Tabel 4.4, sistem otomatis mampu membedakan perkembangan setiap anak. Pada satu anak, terdeteksi peningkatan dalam kosakata dan respons sosial, sementara pada anak lainnya, perkembangan cenderung stagnan. Sistem juga mengidentifikasi perbedaan antara pengulangan kata yang relevan dan tidak relevan, yang tidak dicatat dalam metode manual.

Perbandingan antara metode manual dan analisis otomatis menunjukkan bahwa *automated framework* yang mengintegrasikan Whisper dan GPT-4o memberikan hasil yang lebih mendalam dan terperinci dibandingkan metode manual. Analisis otomatis tidak hanya mempercepat proses, tetapi juga mengungkap temuan yang tidak terdeteksi sebelumnya oleh guru, memberikan nilai tambah yang signifikan dalam mengevaluasi perkembangan komunikasi anak dengan ASD.

Temuan ini mengatasi keterbatasan penelitian sebelumnya, seperti yang diungkapkan oleh (O'Sullivan et al., 2023), yang mencatat kurangnya cakupan analisis fitur linguistik, seperti unit komunikasi per menit, ekolalia, dan kompleksitas morfologi. Kerangka kerja dalam penelitian ini mampu mengatasi keterbatasan tersebut dengan menawarkan analisis yang lebih terperinci, termasuk aspek verbal, pragmatik, semantik, struktur kalimat, dan ekolalia. Misalnya, sistem mampu mengidentifikasi stagnasi pada beberapa anak sekaligus mencatat peningkatan signifikan dalam kosakata dan respons sosial pada anak lainnya. Temuan ini memberikan wawasan yang lebih dalam tentang perkembangan individu, sehingga dapat menjadi dasar untuk merancang intervensi terapi yang lebih tepat dan sesuai dengan tujuan terapi di dunia nyata.

Validasi oleh Psikolog Klinis

Untuk memastikan keakuratan automated framework, hasil analisis komunikasi divalidasi oleh dua psikolog klinis yang berpengalaman lebih dari lima tahun di bidang psikologi anak. Validasi dilakukan dengan mendengarkan rekaman percakapan anak dengan autisme ringan dan sedang yang diambil pada November hingga Desember 2024. Psikolog memberikan hasil penilaian komunikasi anak dan membandingkannya dengan hasil sistem.

Pada anak dengan autisme ringan, psikolog klinis menilai kemampuan komunikasi anak mengalami peningkatan, namun masih memerlukan pelatihan kejelasan bicara dan stimulasi untuk lebih meningkatkan kemampuan komunikasinya.

Sedangkan untuk anak dengan autisme sedang, psikolog mencatat anak kurang ekspresif dan memiliki rentang komunikasi pendek. Hasil analisis framework sesuai dengan penilaian psikolog, terutama dalam mengenali pola pengulangan dan penggunaan kalimat sederhana. Namun, framework tidak dapat menangkap ekspresi wajah, tingkat konsentrasi, kejelasan bicara, atau emosi anak, yang menjadi batasan dibandingkan penilaian manual psikolog.

4.5 Pemeliharaan

Pada penelitian ini, tahap pemeliharaan meliputi pengoperasian sistem yang telah dikembangkan dan perbaikan beberapa kesalahan yang tidak ditemukan pada tahapan sebelumnya.

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini berhasil mengembangkan kerangka kerja otomatis yang mengintegrasikan model Whisper dari OpenAI dan GPT-4o untuk membantu guru SLB dalam menilai kemampuan komunikasi anak-anak dengan ASD. Kerangka kerja ini mampu mengotomatisasi proses transkripsi, diarisasi pembicara, dan analisis fitur linguistik, seperti kemampuan verbal, pragmatik, semantik, struktur kalimat, dan ekolalia, sehingga menawarkan solusi yang lebih efisien dibandingkan metode manual.

Berdasarkan pengujian, keenam guru SLB menyatakan bahwa kerangka kerja ini sangat membantu, dengan dua manfaat utama: pengurangan waktu analisis hingga 89,1% dan penyediaan analisis yang lebih objektif serta mendalam, yang seringkali mengungkap aspek komunikasi yang terlewatkan dalam penilaian manual. Kerangka kerja ini juga memberikan wawasan mendetail tentang perkembangan komunikasi dan pola bahasa anak-anak dengan ASD. Pada anak dengan ASD ringan, kerangka kerja ini berhasil mendeteksi peningkatan pelafalan, kompleksitas kalimat, dan penurunan ekolalia yang lebih konteksual. Sementara itu, pada anak dengan ASD sedang, kerangka kerja ini mengidentifikasi stagnasi perkembangan kemampuan verbal dan semantik, serta mengungkap tantangan spesifik dalam komunikasi.

Validasi oleh dua psikolog klinis menunjukkan bahwa automated framework dapat mendeteksi kemampuan komunikasi anak sesuai dengan penilaian manual. Namun, untuk anak dengan autisme ringan, framework tidak dapat mengenali kebutuhan pelatihan kejelasan bicara dan stimulasi tambahan untuk meningkatkan komunikasi. Untuk anak dengan autisme sedang,

framework tidak mampu mendeteksi ekspresi wajah, tingkat konsentrasi, kejelasan bicara, atau emosi anak, yang menjadi batasan dibandingkan penilaian manual psikolog.

Dengan mengatasi berbagai keterbatasan dalam penelitian sebelumnya, seperti minimnya analisis fitur linguistik dan ketergantungan pada diarisasi pembicara manual, penelitian ini berkontribusi pada pengembangan solusi yang lebih efisien, mudah diakses, dan hemat biaya untuk penilaian kemampuan komunikasi pada anak-anak dengan ASD.

5.2 Saran

Meskipun penelitian ini memberikan kontribusi yang signifikan, kerangka kerja yang dikembangkan masih memiliki beberapa keterbatasan yang perlu diperhatikan untuk penelitian selanjutnya.

a. Kerangka kerja otomatis yang dikembangkan belum mampu menganalisis intonasi, kelancaran berbicara, dan isyarat emosional, yang sangat penting untuk mengevaluasi perkembangan komunikasi dan pola bahasa anak dengan ASD. Oleh karena itu, penelitian di masa depan disarankan untuk mengintegrasikan teknologi tambahan yang mampu mendeteksi aspek-aspek tersebut agar analisis menjadi lebih komprehensif.
b. Pada penelitian ini, ditemukan beberapa kesalahan dalam proses diarisasi pembicara, yang berdampak pada ketepatan analisis berikutnya. Penelitian mendatang perlu memastikan akurasi fitur diarisasi pembicara dengan mengoptimalkan algoritma atau metode yang digunakan, sehingga hasil analisis dapat lebih andal dan konsisten.
c. Data yang digunakan dalam penelitian ini diambil dari empat SLB di Daerah Istimewa Yogyakarta, sehingga cakupan populasi masih terbatas. Penelitian berikutnya perlu memperluas pengambilan data dari wilayah yang lebih luas dengan populasi yang lebih beragam agar hasil penelitian memiliki generalisasi yang lebih baik.

Dengan mengatasi keterbatasan-keterbatasan tersebut, diharapkan kerangka kerja ini dapat terus berkembang menjadi solusi yang lebih efektif, efisien, dan mampu memberikan manfaat yang lebih luas dalam mendukung evaluasi komunikasi anak dengan ASD.

DAFTAR PUSTAKA

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.).

Boucher, J. (2003). Language development in autism. *International Journal of Pediatric Otorhinolaryngology*, *67*(SUPPL. 1). https://doi.org/10.1016/j.ijporl.2003.08.016

Centers for Disease Control and Prevention (CDC). (2024, May 16). *Data and Statistics on Autism Spectrum Disorder*. https://www.cdc.gov/autism/data-research/index.html?utm_source=chatgpt.com

Chen, X., Luo, K., Gee, T., & Nejati, M. (2024). *Does ChatGPT and Whisper Make Humanoid Robots More Relatable?* http://arxiv.org/abs/2402.07095

Cheng, S. W., Chang, C. W., Chang, W. J., Wang, H. W., Liang, C. S., Kishimoto, T., Chang, J. P. C., Kuo, J. S., & Su, K. P. (2023). The now and future of ChatGPT and GPT in psychiatry. In *Psychiatry and Clinical Neurosciences* (Vol. 77, Issue 11, pp. 592–596). John Wiley and Sons Inc. https://doi.org/10.1111/pcn.13588

Deepgram. (2024, June 25). *OpenAI Whisper*. https://deepgram.com/ai-glossary/openai-whisper

Frieda Rong. (2021). *Extrapolating to Unnatural Language Processing with GPT-3's In-context Learning: The Good, the Bad, and the Mysterious*. The Stanford AI Lab Blog. https://ai.stanford.edu/blog/in-context-learning/

Gensler, D. (2012). Autism Spectrum Disorder in DSM-V : Differential Diagnosis and Boundary Conditions . *Journal of Infant, Child, and Adolescent Psychotherapy*, *11*(2), 86–95. https://doi.org/10.1080/15289168.2012.676339

Ghale, A., K, J., & Verma C, D. (2022). Instant Transcription and Translation Tool using OpenAI?s Whisper ASR Model. *International Journal of Science and Research (IJSR)*, *11*(12), 185–188. https://doi.org/10.21275/SR221203164929

Gladia. (2024, March). *What is OpenAI Whisper?* https://www.gladia.io/blog/what-is-openai-whisper

Harrison, J. E., Weber, S., Jakob, R., & Chute, C. G. (2021). ICD-11: an international classification of diseases for the twenty-first century. In *BMC Medical Informatics and Decision Making* (Vol. 21). BioMed Central Ltd. https://doi.org/10.1186/s12911-021-01534-6

He, J., Rungta, M., Koleczek, D., Sekhon, A., Wang, F. X., & Hasan, S. (2024). *Does Prompt Formatting Have Any Impact on LLM Performance?* http://arxiv.org/abs/2411.10541

Hendrick, R. (2024). *Why do GPT models produce harmful responses?* https://www.deepchecks.com/question/why-do-gpt-models-produce-harmful-responses/?utm_source=chatgpt.com

IPSpecialist. (2024). *The History of ChatGPT: From GPT-1 to GPT-4*. https://ipspecialist.net/the-history-of-chatgpt-from-gpt-1-to-gpt-4/

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. In *Learning and Individual Differences* (Vol. 103). Elsevier Ltd. https://doi.org/10.1016/j.lindif.2023.102274

Kodak, T., & Bergmann, S. (2020). Autism Spectrum Disorder: Characteristics, Associated Behaviors, and Early Intervention. In *Pediatric Clinics of North America* (Vol. 67, Issue 3, pp. 525–535). W.B. Saunders. https://doi.org/10.1016/j.pcl.2020.02.007

Lacy, L. (2024, April). *Hallucinations: Why AI Makes Stuff Up, and What's Being Done About It*. https://www.cnet.com/tech/hallucinations-why-ai-makes-stuff-up-and-whats-being-done-about-it/

Laraswati, B. D. (2023, March 16). *Large Language Models, Cikal Bakal Lahirnya Chat GPT*. https://blog.algorit.ma/large-language-models/

Latentview. (2024, February 1). *Retrieved from Audio Analytics with OpenAI: How Whisper Transforms Audio into Insights*. https://www.latentview.com/blog/audio-analytics-with-openai-how-whisper-transforms-audio-into-insights/

Mansur. (2016). Hambatan Komunikasi Anak Autis. *Al-Munzir*, *9*.

Mujiyanti, D. M. (2011). Tingkat Pengetahuan Ibu dan Pola Konsumsi Pada Anak Autis Di Kota Bogor. *Institut Pertanian Bogor*.

Mukherjee, P., Gokul, R. S., Sadhukhan, S., Godse, M., & Chakraborty, B. (2023). Detection of Autism Spectrum Disorder (ASD) from Natural Language Text using BERT and ChatGPT Models. *IJACSA) International Journal of Advanced Computer Science and Applications*, *14*(10). https://doi.org/10.14569/IJACSA.2023.0141041

Open AI. (2022). *Introducing Whisper*. https://openai.com/index/whisper/

O'Sullivan, J., Bogaarts, G., Kosek, M., Ullmann, R., Schoenenberger, P., Chatham, C., Nobbs, D., Murtagh, L., Lindemann, M., Parish-Morris, J., Liberman, M., Aponte, E., Dorn, J., & Lipsmeier, F. (2023). Automatic Speech Recognition for ASD Using the Open-Source Whisper Model from OpenAI. *International Society for Autism Research*.

Radford, A., Karthink, N., Tim, T., & IIya, I. (2018). *Improving Language Understanding by Generative Pre-Training*. https://gluebenchmark.com/leaderboard

Radford, A., Kim, J. W., Xu, T., Brockman, G., Mcleavey, C., & Sutskever, I. (2022). *Robust Speech Recognition via Large-Scale Weak Supervision*. https://github.com/openai/

Rahmawati, S. (2024). Optimalisasi Fokus: "Strategi Pembelajaran untuk Meningkatkan Konsentrasi pada Anak dengan Gangguan Spektrum Autisme (GSA)." In *Jurnal Kependidikan* (Vol. 13, Issue 2). https://jurnaldidaktika.org

Rathje, S., Mirea, D. M., Sucholutsky, I., Marjieh, R., Robertson, C. E., & Van Bavel, J. J. (2024). GPT is an effective tool for multilingual psychological text analysis. In *Proceedings of the National Academy of Sciences* (Vol. 121, Issue 34).

Richards, J. A., Xu, D., & Gilkerson, J. (2010). *Development and Performance of the LENA Automatic Autism Screen*.

Sani, M. I. (2019, May 8). *Speech Recognition*. Binus University. https://mti.binus.ac.id/2019/05/08/speech-recognition/

Suh, J., Na, I., & Jung, W. (2024). *Improving Domain-Specific ASR with LLM-Generated Contextual Descriptions*. https://doi.org/10.21437/Interspeech.2024-377

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and Communication in Autism. In *Handbook of Autism and Pervasive Developmental Disorders: Diagnosis, Development, Neurobiology, and Behavior* (pp. 335–354).

Tak, A. N., & Gratch, J. (2023). *Is GPT a Computational Model of Emotion? Detailed Analysis*.

Taskus. (2024, February 28). *What are Generative Pre-trained Transformers (GPT)?* https://www.taskus.com/insights/what-is-gpt/

Vivek, V. S., Raj, P. P., Santhosh, K., & Ganesan, Dr. V. (2023). Voice Assistant based on GPT-3 and Whisper. *International Journal for Research in Applied Science and Engineering Technology*, *11*(3), 827–831. https://doi.org/10.22214/ijraset.2023.49536

Vogindroukas, I., Stankova, M., Chelas, E. N., & Proedrou, A. (2022a). Language and Speech Characteristics in Autism. In *Neuropsychiatric Disease and Treatment* (Vol. 18, pp. 2367–2377). Dove Medical Press Ltd. https://doi.org/10.2147/NDT.S331987

Vogindroukas, I., Stankova, M., Chelas, E. N., & Proedrou, A. (2022b). Language and Speech Characteristics in Autism. In *Neuropsychiatric Disease and Treatment* (Vol. 18, pp. 2367–2377). Dove Medical Press Ltd. https://doi.org/10.2147/NDT.S331987

Wahid, A. (2020). Analisis Metode Waterfall Untuk Pengembangan Sistem Informasi. *Jurnal Ilmu-Ilmu Informatika Dan Manajemen STMIK*. https://www.researchgate.net/publication/346397070

Yadgar N. Abbas, Hemn A. Hassan, Dyari Qadir Hamad, Sabah Jalal Hasan, Diyar A. Omer, Suhaib H. Kakamad, Fahmi H. Kakamad, Hussein M. HamaSalih, Marwan N. Hassan, Hawbash M. Rahim, Rawezh Q. Salih, Berun A. Abdalla, Shvan H. Mohammed, & Yousif M. Mahmood. (2023). Role of ChatGPT and Google Bard in the Diagnosis of Psychiatric Disorders: A Comparative Study. *Barw Medical Journal*. https://doi.org/10.58742/4vd6h741