Icon

Description automatically generated

**PROPOSAL TUGAS AKHIR – EF234702**

**Analisis Komparatif Pendekatan Machine Learning dan Transformer Pra-latih untuk Prediksi Kepribadian dari Data Suara**

**Muhammad Aqil Farrukh**

NRP 5025221158

Dosen Pembimbing

Shintami Chusnul Hidayati, S.Kom., M.Sc., Ph.D

NIP 1987202012004

Dosen Ko-pembimbing

Dini Adni Navastara, S.Kom., M.Sc.

NIP 198510172015042001

**Program Studi S-1 Teknik Informatika**

Departemen Teknik Informatika

Fakultas Teknologi Elektro dan Informatika Cerdas

Institut Teknologi Sepuluh Nopember

Surabaya

2025

*Halaman ini sengaja dikosongkan.*

*Halaman ini sengaja dikosongkan*

LEMBAR PENGESAHAN

**<JUDUL PROPOSAL TUGAS AKHIR DITULIS SINGKAT,  
JELAS, DAN MENGGAMBARKAN TEMA POKOK PENELITIAN>**

**PROPOSAL TUGAS AKHIR**

Diajukan untuk memenuhi salah satu syarat

memperoleh gelar Sarjana Komputer pada

Program Studi S-1 Teknik Informatika

Departemen Teknik Informatika

Fakultas Teknologi Elektro dan Informatika Cerdas

Institut Teknologi Sepuluh Nopember

Oleh: **Muhammad Aqil Farrukh**

NRP. 5025221158

Disetujui oleh Tim Penguji Proposal Tugas Akhir:

|  |  |  |
| --- | --- | --- |
| 1. | Shintami Chusnul Hidayati, S.Kom., M.Sc., Ph.D | Pembimbing |
| 2. | Dini Adni Navastara, S.Kom., M.Sc. | Ko-pembimbing |
| 3. | <Nama dan gelar penguji> | Penguji |
| 4. | <Nama dan gelar penguji> | Penguji |

**SURABAYA**

**April, 2025**

*Halaman ini sengaja dikosongkan.*

APPROVAL SHEET

**<FINAL PROJECT TITLE SHOULD BE WRITTEN IN BRIEF,  
YET CLEAR AND REPRESENTING THE MAIN THEME OF THE RESEARCH>**

**FINAL PROJECT PROPOSAL**

Submitted to fulfill one of the requirements

for obtaining a Bachelor of Computer Science degree at

Undergraduate Study Program of Informatics

Department of Informatics

Faculty of Intelligent Electrical and Informatics Technology

Institut Teknologi Sepuluh Nopember

By: **Muhammad Aqil Farrukh**

NRP. 5025221158

Approved by Final Project Proposal Examiner Team:

|  |  |  |
| --- | --- | --- |
| 1. | Shintami Chusnul Hidayati, S.Kom., M.Sc., Ph.D | Advisor |
| 2. | Dini Adni Navastara, S.Kom., M.Sc. | Co-advisor |
| 3. | <Name and title> | Examiner 1 |
| 4. | <Name and title> | Examiner 2 |

**SURABAYA**

**April, 2025**

*Halaman ini sengaja dikosongkan.*

ABSTRAK

**PENGEMBANGAN SISTEM REALITAS VIRTUAL   
UNTUK SIMULASI INTERAKSI MANUSIA-MESIN**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Nama Mahasiswa / NRP** | **:** | | **Muhammad Aqil Farrukh / 5025221158** | |
| **Departemen** | **:** | | **Teknik Informatika FTEIC - ITS** | |
| **Dosen Pembimbing** | **:** | | **Shintami Chusnul Hidayati, S.Kom., M.Sc., Ph.D** | |
| **Dosen Ko-pembimbing** | | **:** | | **Dini Adni Navastara, S.Kom., M.Sc.** | |

**Abstrak**

Realitas Virtual (VR) merupakan teknologi yang memungkinkan pengguna untuk berinteraksi dengan lingkungan virtual yang menyerupai dunia nyata. Penelitian ini bertujuan untuk mengembangkan sistem VR yang dapat digunakan untuk simulasi interaksi manusia-mesin, khususnya dalam konteks pelatihan dan pendidikan. Sistem VR yang dikembangkan menggunakan perangkat keras dan perangkat lunak terkini untuk menciptakan pengalaman yang imersif dan realistis. Dalam penelitian ini, analisis terhadap berbagai aspek interaksi manusia-mesin akan dilakukan, termasuk respons pengguna terhadap berbagai skenario simulasi, tingkat kenyamanan, dan efektivitas pelatihan. Metode pengumpulan data meliputi survei, wawancara, dan pengamatan langsung.

**Kata kunci: *Virtual Reality*, Interaksi Manusia-Mesin, Simulasi, Pelatihan, Pendidikan.**

*Halaman ini sengaja dikosongkan.*

*ABSTRACT*

***DEVELOPMENT OF A VIRTUAL REALITY SYSTEM   
FOR HUMAN-MACHINE INTERACTION SIMULATION***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Full Name / Student ID*** | **:** | | **Muhammad Aqil Farrukh / 5025221158** | |
| ***Department*** | **:** | | ***Informatics* ELECTICS - ITS** | |
| ***Advisor*** | **:** | | **Shintami Chusnul Hidayati, S.Kom., M.Sc., Ph.D** | |
| ***Co-advisor*** | | **:** | | **Dini Adni Navastara, S.Kom., M.Sc.** | |

***Abstract***

*Virtual Reality (VR) technology allows users to interact with a virtual environment that resembles the real world. This research aims to develop a VR system to simulate human-machine interactions, particularly in the context of training and education. The developed VR system utilizes the latest hardware and software to create an immersive and realistic experience. This study examines various aspects of human-machine interaction, including user responses to different simulation scenarios, comfort levels, and the effectiveness of training. Data collection methods include surveys, interviews, and direct observation.*

***Keywords: Virtual Reality, Human-Machine Interaction, Simulation, Training, Education.***

*Halaman ini sengaja dikosongkan.*

DAFTAR ISI

[LEMBAR PENGESAHAN i](#_Toc208028970)

[APPROVAL SHEET iii](#_Toc208028971)

[ABSTRAK v](#_Toc208028972)

[*ABSTRACT* vii](#_Toc208028973)

[DAFTAR ISI ix](#_Toc208028974)

[DAFTAR GAMBAR xi](#_Toc208028975)

[DAFTAR TABEL xiii](#_Toc208028976)

[DAFTAR KODE SUMBER xv](#_Toc208028977)

[BAB 1 PENDAHULUAN 1](#_Toc208028978)

[1.1 Latar Belakang 1](#_Toc208028979)

[1.2 Rumusan Masalah 2](#_Toc208028980)

[1.3 Batasan Masalah 2](#_Toc208028981)

[1.4 Tujuan 3](#_Toc208028982)

[1.5 Manfaat 3](#_Toc208028983)

[1.5.1 Manfaat Teoritis 3](#_Toc208028984)

[1.5.2 Manfaat Praktis 3](#_Toc208028985)

[1.5.3 Manfaat Sosial 3](#_Toc208028986)

[BAB 2 TINJAUAN PUSTAKA 5](#_Toc208028987)

[2.1 Hasil Penelitian Terdahulu 5](#_Toc208028988)

[2.2 Dasar Teori 5](#_Toc208028989)

[BAB 3 METODOLOGI 7](#_Toc208028990)

[3.1 Metode yang Digunakan 7](#_Toc208028991)

[3.2 Bahan dan Peralatan yang Digunakan 7](#_Toc208028992)

[3.3 Urutan Pelaksanaan Penelitian 7](#_Toc208028993)

[DAFTAR PUSTAKA 9](#_Toc208028994)

[LAMPIRAN 11](#_Toc208028995)

*Halaman ini sengaja dikosongkan.*

DAFTAR GAMBAR

[Gambar 2.1 Kotak berwarna biru 5](#_Toc202858853)

*Halaman ini sengaja dikosongkan.*

DAFTAR TABEL

[Tabel 3.1 Lini Masa Pengerjaan Tugas Akhir 7](#_Toc202858843)

*Halaman ini sengaja dikosongkan.*

DAFTAR KODE SUMBER

*Halaman ini sengaja dikosongkan.*

# PENDAHULUAN

## Latar Belakang

Suara manusia tidak hanya menyampaikan isi pesan, tetapi juga merefleksikan identitas dari penutur. Identitas yang dimaksud berupa usia, aksen budaya, hingga kepribadian dari penutur (Lukac, 2024). Gagasan ini sebenarnya telah dikemukakan sejak awal abad ke-20, dimana penelitian pada tahun 1930-an sudah berusaha mengaitkan ciri-ciri vokal dengan kepribadian individu. Meskipun demikian, temuan di era awal tersebut cenderung belum konsisten, sehingga keterkaitan antara vokal dengan kepribadian manusia tetap menjadi teka-teki ilmiah selama beberapa dekade (Lukac, 2024; Rubio et al., 2024).

Perkembangan teknologi dalam analisis suara dan generalisasi dari analisis *big data* menciptakan kesempatan besar untuk mengestimasi kepribadian manusia berdasarkan kekuatan fitur dalam suara manusia (Rubio et al., 2024). Bidang asesmen kepribadian secara komputasional telah berkembang pesat dimana kita bisa menilai kepribadian seseorang dari sosial media, komunikasi personal, gerakan mata, teks atau tulisan umum, sampai foto wajah seseorang (Lukac, 2024). Dibanding metode sebelumnya, analisis kepribadian melalui suara bisa dimanfaatkan secara efektif untuk beberapa kondisi seperti wawancara kerja atau layanan pelanggan karena kepribadian seseorang sangat berpengaruh (Rubio et al., 2024).

Beberapa penelitian terbaru dalam teknologi analisis suara, sudah menggunakan teori umum yang diterima secara luas seperti *Big Five Personality* dan telah menghasilkan temuan yang lebih jelas seperti fitur-fitur akustik tertentu yang berkorelasi dengan kepribadian manusia. Misalnya, variasi prosodikseperti nada (*pitch*) dan kecepatan bicara bisa menjadi parameter tingkat Ekstraversi seseorang. Eksperimen yang telah dilakukan sebelumnya mengklasifikasikan individu dengan sifat tinggi dan rendah fitur suara. Orang dengan karakteristik vokal yang lantang dan lancar dalam berbicara cenderung dinilai lebih ekstrovert, sedangkan frekuensi suara yang monoton atau rendah diklasifikasikan dengan sifat kurang percaya diri atau dominasi yang rendah (Rubio et al., 2024).

Penelitian yang dilakukan oleh Lukac (2024) mencoba mengestimasi kepribadian seseorang dengan memanfaatkan *convolutional neural network* (CNN) dan transformer pra-latih untuk mengekstrak *embedding* akustik dari suara (seperti intonasi, nada, ritme) dan linguistik (makna teks atau transkrip percakapan). Kedua jenis fitur ini kemudian digabungkan dalam model *gradient boosted trees* untuk memprediksi skor *Big Five* tiap individu. Hasil yang didapatkan menunjukkan bahwa model tersebut mampu memprediksi dengan tingkat akurasi yang cukup baik dengan tingkat kesesuaian (*correlation coefficient*) dengan skor *Big Five* berkisar antara 0.26 hingga 0.39. Setelah dilakukan koreksi statistik untuk menghilangkan “gangguan” dalam data (*disattenuated correlations)* terjadi peningkatan di tingkat kesesuaiannya yaitu 0.39 hingga 0.60. Penemuan ini membuka potensi penggunaan analisis suara sebagai alat untuk mengestimasi kepribadian manusia, dan juga memberikan cara baru untuk memahami hubungan antara suara dan kepribadian (Lukac, 2024).

Dalam domain estimasi kepribadian berdasarkan data suara melalui pendekatan tradisional menggunakan fitur-fitur akustik (misalnya prosodik, spektral, MFCC, jitter, shimmer) yang dipadukan dengan algoritme klasik seperti SVM atau XGBoost telah lama diteliti. Hasil yang didapatkan cukup menjanjikan dan berkembang seiring waktu, namun akurasinya cenderung terbatas pada tingkat korelasi rendah hingga sedang. Berdasarkan studi dari Barchi et al. (2023) dan Rubio et al. (2024), telah digunakan metode klasik untuk penelitian sebelumnya, teteapi model berbasis fitur *handcrafter* (rekayasa fitur manual) semacam ini umumnya hanya mampu menjelaskan sekitar 10%-16% variasi skor kepribadian. Temuan tersebut mengindikasikan bahwa estimasi kepribadian dari suara memang signifikan secara statistik, tetapi performa masih terbatas dan kurang stabil di berbagai konteks atau dataset.

Di sisi lain, pendekatan yang dilakukan dengan transformer pra-latih seperti Wav2Vec2 dan HuBERT untuk prediksi kepribadian membuktikan performanya lebih baik dari kinerja fitur akustik klasik (Barchi et al., 2023). Berdasarkan studi yang telah dilakukan oleh Lukac (2024), kelebihan model berbasis transformer ini adalah kemamuannya menangkap representasi akustik mendalam dan juga semantik langsung dari sinyal suara tanpa memerlukan rekayasa fitur manual. Studi terkini menunjukkan bahwa fitur-fitur yang diekstraksi melalui Wav2Vec2 merupakan paling informatif untuk estimasi kepribadian, jika dikombinasikan dengan embedding transformer dengan fitur akustik tradisional dapat memberikan peningkatan kinerja lebih lanjut (Barchi et al., 2023). Tetapi untuk mengumpulkan dataset suara natural yang mana lebih representatif, memiliki tantangan tersendiri dimana membutuhkan data berukuran besar dan daya komputasi tinggi (Lukac, 2024). Selain itu, belum jelas apakah pendekatan baru ini benar-benar lebih unggul secara konsisten untuk setiap dimensi *Big Five*. Hanya sedikit riset yang secara langsung membandingkan metode klasik dengan transformer secara sistematis. Barchi et al. (2023) bahkan mencatat bahwa sebelum studi mereka, belum ada laporan hasil baseline estimasi kepribadian hanya dari data suara di literatur yang mana menunjukkan bahwa ada gap penelitian. Inilah yang mendorong penelitian di tugas akhir ini, yakni untuk membandingkan pendekatan klasik dan transformer (termasuk menambahkan CNN sebagai baseline) secara komprehensif, guna melihat apakah model pra-latih suara memang memberikan keunggulan yang konsisten dibanding metode konvensional.

Berdasarkan penelitian yang telah dilakukan sebelumnya, analisis kepribadian melalui data suara semakin diakui manfaatnya dalam berbagai aplikasi praktis. Misalnya, dalam konteks rekrutmen dan wawancara kerja, kepribadian kandidat dapat diestimasi secara lebih objektif melalui rekaman suara wawancara dan bisa membantu pengambil keputusan dalam mengurangi penilaian secara subyektif. Di bidang lain seperti interaksi manusia dan komputer, antarmuka cerdas dapat menyesuaikan responsnya berdasarkan kepribadian pengguna dengan deteksi suara, sehingga meningkatkan pengalaman pengguna secara personal. Dengan terus berkembangnya riset di rentang tahun 2020–2025 ini, pendekatan estimasi kepribadian dari data suara diharapkan semakin matang dan siap diimplementasikan secara luas. (Rubio et al., 2024)

## Rumusan Masalah

Penelitian mengenai estimasi kepribadian berbasis suara berkembang pesat dengan adanya kemajuan machine learning dan deep learning. Akan tetapi, masih terdapat kesenjangan dalam menentukan pendekatan mana yang paling efektif antara metode berbasis fitur akustik klasik dan model berbasis Transformer pra-latih. Selain itu, belum banyak penelitian yang secara sistematis membandingkan performa model klasik dengan model pra-latih pada data suara untuk prediksi kepribadian Big Five.

Berdasarkan latar belakang tersebut, rumusan masalah penelitian ini adalah sebagai berikut:

1. Bagaimana performa model berbasis fitur akustik klasik dengan metode machine learning (misalnya SVM, XGBoost) dalam melakukan prediksi kepribadian dari data suara?
2. Bagaimana performa model Transformer pra-latih (misalnya Wav2Vec2, HuBERT) dibandingkan dengan metode klasik dalam tugas prediksi kepribadian berbasis suara?
3. Bagaimana performa model berbasis CNN (misalnya CNN akustik) dibandingkan dengan metode klasik dan Transformer pra-latih dalam tugas prediksi kepribadian berbasis suara?
4. Bagaimana perbedaan hasil antara pendekatan feature extraction (frozen embedding) dan fine-tuning pada model Transformer pra-latih untuk estimasi kepribadian?
5. Apakah kombinasi fitur akustik klasik dengan embedding Transformer pra-latih dapat meningkatkan akurasi prediksi kepribadian dibanding penggunaan salah satu pendekatan saja?

## Batasan Masalah

Untuk menjaga fokus dan keterkelolaan penelitian, maka batasan masalah yang ditetapkan adalah sebagai berikut:

1. Penelitian hanya membahas prediksi kepribadian *Big Five* (*Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism*).
2. Dataset yang digunakan adalah dataset publik berlabel Big Five; apabila dataset berbahasa Indonesia tidak tersedia, digunakan dataset berbahasa Inggris (audio-only) dengan dokumentasi pra-proses yang jelas.
3. Data yang dianalisis hanya berupa suara/audio. Data visual atau multimodal lain tidak termasuk dalam lingkup penelitian ini.
4. Model yang dibandingkan mencakup machine learning klasik (SVM, XGBoost), CNN akustik, dan Transformer pra-latih (misalnya Wav2Vec2, HuBERT), serta variasinya (frozen vs fine-tuning).
5. Evaluasi dilakukan menggunakan metrik kuantitatif (Pearson correlation, RMSE, MAE) dengan metode cross-validation, tanpa melibatkan uji persepsi manusia.
6. Split evaluasi bersifat speaker-independent (tidak ada pembicara yang sama antara train/valid/test) untuk mencegah data leakage.

## Tujuan

Tujuan penelitian ini adalah untuk melakukan analisis komparatif pendekatan machine learning klasik dan Transformer pra-latih dalam estimasi kepribadian berbasis suara. Tujuan spesifik dari penelitian ini adalah:

1. Mengevaluasi performa model berbasis fitur akustik klasik dengan metode machine learning (SVM, XGBoost).
2. Menganalisis efektivitas model Transformer pra-latih (Wav2Vec2, HuBERT) untuk prediksi kepribadian berbasis suara.
3. Mengevaluasi efektivitas CNN dalam memprediksi kepribadian berbasis suara, serta membandingkannya dengan metode klasik dan Transformer pra-latih.
4. Membandingkan hasil antara feature extraction (frozen) dan fine-tuning pada model Transformer pra-latih.
5. Mengkaji potensi fusion antara fitur akustik klasik dan embedding Transformer dalam meningkatkan akurasi prediksi.
6. Mengidentifikasi model terbaik yang dapat digunakan sebagai acuan untuk penelitian lebih lanjut dalam bidang personality computing berbasis suara.

## Manfaat

Penelitian ini diharapkan dapat memberikan manfaat baik secara teoritis, praktis, maupun sosial.

### Manfaat Teoritis

1. Menambah wawasan dan pengetahuan dalam bidang computational personality assessment berbasis suara.
2. Memberikan kontribusi pada literatur mengenai perbandingan performa model klasik dan Transformer pra-latih dalam analisis kepribadian.
3. Mengidentifikasi faktor teknis (fitur, model, strategi training) yang memengaruhi akurasi prediksi kepribadian berbasis audio.

### Manfaat Praktis

1. Memberikan acuan model AI yang efektif untuk prediksi kepribadian berbasis suara.
2. Memberikan rekomendasi pendekatan (klasik vs Transformer pra-latih) yang lebih tepat untuk digunakan pada aplikasi nyata.
3. Menyediakan kerangka eksperimen yang dapat direplikasi oleh peneliti lain atau pengembang sistem AI.

### Manfaat Sosial

1. Mendukung pengembangan aplikasi AI yang dapat memahami karakteristik individu secara lebih personal melalui suara.
2. Membuka peluang pemanfaatan teknologi ini di bidang rekrutmen, pendidikan, dan layanan pelanggan dengan tetap memperhatikan etika dan privasi.
3. Mendorong inovasi riset interdisipliner di Indonesia dalam bidang psikologi komputasional dan kecerdasan buatan.

*Halaman ini sengaja dikosongkan.*

# TINJAUAN PUSTAKA

## Hasil Penelitian Terdahulu

Lorem ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum ipsum sebagaimana tampak pada Gambar ‎2.1. Keterangan gambar (*figure caption*) ditulis dalam bentuk kalimat biasa. Jadi yang dikapitalkan hanya huruf pertama pada kata pertama di keterangan tersebut.

Gambar ‎2.1 Kotak berwarna biru

## Dasar Teori

Berdasarkan penelitian sebelumnya (Soldati et al., 2007), eksplorasi di VR mampu menghasilkan data yang beragam.

*Halaman ini sengaja dikosongkan.*

# METODOLOGI

## Metode yang Digunakan

## Bahan dan Peralatan yang Digunakan

## Urutan Pelaksanaan Penelitian

Penelitian Tugas Akhir ini akan dilaksanakan selama enam bulan dari Maret sampai dengan September 2024. Lini masa pengerjaan Tugas Akhir bisa dilihat pada Tabel ‎3.1. Judul tabel perlu ditulis dalam format *Title Case*, yang berarti setiap kata diawali huruf kapital, kecuali untuk kata depan seperti ’di’, ’ke’, ’dari’, ’yang’, ’untuk’, ’kepada’, dan sebagainya.

Tabel ‎3.1 Lini Masa Pengerjaan Tugas Akhir

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No | Aktivitas | MAR | APR | MEI | JUN |  | JUL | AGU | SEP |
| 1 | Studi literatur |  |  |  |  |  |  |  |  |
| 2 | Empati konteks permasalahan |  |  |  |  |  |  |  |  |
| 3 | Definisi spesifikasi kebutuhan |  |  |  |  |  |  |  |  |
| 4 | Ideasi solusi |  |  |  |  |  |  |  |  |
| 5 | Pembuatan purwarupa solusi |  |  |  |  |  |  |  |  |
| 6 | Evaluasi |  |  |  |  |  |  |  |  |
| 7 | Penulisan laporan Tugas Akhir |  |  |  |  |  |  |  |  |

*Halaman ini sengaja dikosongkan.*

DAFTAR PUSTAKA

Barchi, R., Pepino, L., Gauder, L., Estienne, L., Meza, M., Riera, P., Ferrer, L. (2023) Apparent personality prediction from speech using expert features and wav2vec 2.0. Proc. SMM23, Workshop on Speech, Music and Mind 2023, 21-25, doi: <https://doi.org/10.21437/SMM.2023-5>

Lukac, M. (2024). Speech-based personality prediction using deep learning with acoustic and linguistic embeddings. Scientific Reports 14, 30149. <https://doi.org/10.1038/s41598-024-81047-0>

Pearsell, S., & Pape, D. (2023). The effects of different voice qualities on the perceived personality of a speaker. *Frontiers in Communication*, *7*, 909427. <https://doi.org/10.3389/fcomm.2022.909427>

Rubio, V. J., Aguado, D., Toledano, D. T., & Fernández-Gallego, M. P. (2024). Feasibility of big data analytics to assess personality based on voice analysis. *Sensors*, *24*(22), 7151. <https://doi.org/10.3390/s24227151>

*Halaman ini sengaja dikosongkan.*

LAMPIRAN