[NED University of Engineering and Technology]  
[16-March-2025]

**Voice-to-Text with Sentiment & Intent Analysis**

[AI-Powered Voice Transcription and Sentiment Analysis System]

Contents

[Introduction 3](#_Toc193026801)

[Objective 3](#_Toc193026802)

[Scope… 3](#_Toc193026803)

[Methodology 3](#_Toc193026804)

[Research and Data Collection 3](#_Toc193026805)

[Development 3](#_Toc193026806)

[Testing and Evaluation 3](#_Toc193026807)

[Documentation 4](#_Toc193026808)

[Tools and Technologies 4](#_Toc193026809)

[Expected Outcomes 4](#_Toc193026810)

[Challenges and Risks 4](#_Toc193026811)

[Conclusion 4](#_Toc193026812)

# Introduction

With the growing need for automated speech processing in industries such as customer support, podcast analytics, and mental health monitoring, this project aims to develop an AI-powered voice transcription system. The system will convert spoken language into text and analyze the sentiment and intent of the conversation, providing valuable insights for various applications.

# Objective

* Develop a speech-to-text system using OpenAI Whisper for high-accuracy transcription.
* Implement sentiment analysis using RoBERTa to classify speech as positive, negative, or neutral.
* Extend the system to detect intent and urgency in conversations.

# Scope

The project will cover:

* Speech Processing: Convert audio files into text using Whisper.
* Sentiment Analysis: Identify positive, neutral, or negative sentiment from transcribed text.
* Intent Analysis: Detect urgency or intent (e.g., complaint, request, inquiry).

# Methodology

## Research and Data Collection

* Investigate best practices for speech-to-text and sentiment analysis.
* Explore *Hugging Face* models to optimize accuracy for real-world applications.

## Development

* Set up the backend using FastAPI and integrate OpenAI Whisper for transcription.
* Implement RoBERTa for sentiment analysis and intent classification.

## Testing and Evaluation

* Assess transcription accuracy using benchmark datasets.
* Evaluate sentiment and intent detection with real-world conversations.

## Documentation

* Document API endpoints and provide usage guidelines.

# Tools and Technologies

|  |  |
| --- | --- |
| Component | Technology |
| Speech to Text | OpenAI Whisper |
| Sentiment Analysis | RoBERTa NLP model |

# Expected Outcomes

* A functional prototype that accurately transcribes speech and analyzes sentiment and intent.

# Challenges and Risks

* Speech Quality Variability:

Background noise and accents may impact transcription accuracy.

* Model Limitations:

Sentiment models may misclassify nuanced emotions.

# Conclusion

This project leverages advanced NLP and speech recognition technologies to develop a practical AI-driven transcription and analysis tool. With applications across various industries, the system serves as a valuable demonstration of machine learning’s impact on real-world scenarios.