HIDAY: Guidance, Delivered with Compassion

GUIDE - Genuine Understanding of Islamic Doctrine and Ethics

Khadeeja Shah
Department of Data Science
FAST-NUCES, Islamabad
i211653@nu.edu.com

Nibras Aamir
Department of Data Science
FAST-NUCES, Islamabad
i212683@nu.edu.cpk

ABSTRACT

Conversational AI is revolutionizing how humans interact with technology, yet it largely overlooks ethical and cultural contexts. **Hiday**, an Islamic voice assistant, fills this gap by analyzing spoken conversations in Urdu, identifying ethical breaches such as gossip or disrespectful language, and offering corrective feedback in real-time. Leveraging advancements in speech-to-text, NLP, and text-to-speech technologies, Hiday embodies Islamic principles to foster better communication habits. It not only identifies but also educates users about ethical speech, acting as a technological tool for self-improvement and spiritual growth.

KEYWORDS

Islamic ethics, NLP, voice assistant, real-time feedback

ACM Reference Format:
Khadeeja Shah and Nibras Aamir. 2024. HIDAY: Guidance, Delivered with Compassion: GUIDE - Genuine Understanding of Islamic Doctrine and Ethics. In . ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 PROBLEM STATEMENT

1.1 Ethical Lapses in Everyday Conversations

In informal and professional settings, conversational lapses like gossip, neglecting greetings, or using foul language are common. These behaviors harm relationships, erode trust, and contradict the values of respect and morality emphasized in Islam. Despite the availability of conversational AI tools, there is no system tailored to encourage ethical practices or integrate cultural and religious sensitivities into communication.

1.2 The Need for Hiday

There is a pressing need for an intelligent system that not only processes conversations but also evaluates them against a moral framework. **Hiday** addresses this by providing a proactive, voice-based solution to guide individuals in practicing Islamic ethical principles during conversations.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
Conference'17, July 2017, Washington, DC, USA
© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2 INTRODUCTION

2.1 Contextual Motivation

In many cultures, especially Islamic societies, ethical communication is seen as a reflection of character and faith. However, daily conversations often stray from these values due to social pressures or habits. This behavior, although normalized, is discouraged in Islamic teachings, which emphasize avoiding gossip (*ghibat*), maintaining politeness, and fostering positive social interactions.

Existing AI tools like Alexa or Siri prioritize functionality—setting reminders, fetching information—but lack cultural or ethical awareness. Inspired by this gap, **Hiday** reimagines voice assistants as tools for spiritual and moral improvement, combining the capabilities of advanced AI with Islamic teachings.

2.2 Key Innovations of Hiday

Hiday is not just a tool but a guide. Its key innovations include:

- Real-time conversational analysis in **Urdu**, making it accessible to a broader audience.
- Feedback rooted in Islamic ethics, with clear references to principles from the Quran and Hadith.
- An educational dimension that explains why certain conversational behaviors should be avoided.

3 RELATED WORK

3.1 Advances in Conversational AI

Conversational AI has rapidly advanced with the introduction of systems like GPT-based assistants, multilingual speech-to-text engines, and sentiment analysis models. However, these systems focus primarily on utility rather than ethics.

3.2 Gaps in Existing Solutions

Several tools support Urdu speech recognition, including the **Google Speech API** and open-source tools like **CMU Sphinx**, but these are designed for transcription rather than ethical evaluation. Similarly, NLP tools like **BERT** and **Hugging Face Transformers** excel in multilingual text analysis but are not tailored for religious or cultural ethics.

Hiday integrates these technologies into a cohesive system while addressing the unique requirements of real-time ethical guidance.

4 APPROACH

4.1 System Overview

Hiday is composed of three primary modules:

(1) **Speech Processing Module:** Converts spoken Urdu into text.

(2) **NLP Analysis Module:** Analyzes transcriptions to detect ethical violations.
(3) **Feedback Generation Module:** Responds in Urdu with corrective guidance.

These modules are underpinned by advanced tools and algorithms, ensuring efficiency and accuracy.

4.2 Speech Processing

Tools and Techniques:

- **PyAudio:** Captures live audio streams in real-time.
- **Google Speech-to-Text API:** Converts Urdu audio into text with high accuracy (92%).

Challenges:

- **Accent and Dialect Variations:** Addressed using a diverse dataset covering multiple accents.
- **Noisy Environments:** Filters and denoising algorithms improve recognition in group settings.

4.3 NLP Analysis

Core Components:

(1) **Regex-Based Rule Detection:** Flags gossip phrases (e.g.,
`kya aap jantay hain ke? (in urdu)`
(2) **Named Entity Recognition (NER):** Identifies entities to flag gossip about specific individuals. Fine-tuned BERT models achieved very high score.
(3) **Sentiment Analysis:** Detects negative tones using Hugging Face transformers.
(4) **Ethical Violation Scoring:** Evaluates the severity of detected issues and prioritizes feedback.

4.4 Feedback Generation

Tools and Techniques:

- **gTTS (Google Text-to-Speech):** Converts ethical feedback into spoken Urdu.
- **Personalized Feedback:** Friendly yet firm tones ensure users are not alienated.

Example Workflow:

`kya tum nay suna Ali nay larai mai kya kya kaha ahmed ko? (in urdu)`

`Gheebat say guraiz krain. (in Urdu)`

5 EVALUATION AND EXPERIMENTS

5.1 Experimental Setup

Datasets:

- Urdu Speech Dataset for training speech-to-text.
- Curated text dataset for ethical NLP tasks (gossip, swearing detection).

Benchmarks:

- **Trigger: greeting**
- Detected: 1
- Responded: 1 (100.00
- Missed Responses: 0 (0.00
- False Positives (Responded without Detection): 0
- **Trigger: abuse**
- Detected: 3
- Responded: 3 (100.00
- Missed Responses: 0 (0.00
- False Positives (Responded without Detection): 0
- **Trigger: gheebat**
- Detected: 6
- Responded: 6 (100.00
- Missed Responses: 0 (0.00
- False Positives (Responded without Detection): 0
- **=== Performance Metrics ===**
- Total Audio Buffers Processed: 28
- Recognized Audio: 16 (57.14
- Unrecognized Audio: 12 (42.86
- Speech Recognition Errors: 12

5.2 Results

User testing showed a 76% positive response rate, with participants reporting improved awareness of ethical communication.

6 SYSTEM DIAGRAMS

Below attached are the diagrams for Hiday to explain every aspect of this model.

6.1 System Architecture Diagram

The system architecture diagram illustrates the overall design and interaction of Hiday's components. It shows the user's device interacting with the Speech-to-Text API, NLP analysis, and Text-to-Speech feedback generation. This modular structure ensures scalability and efficient real-time processing.

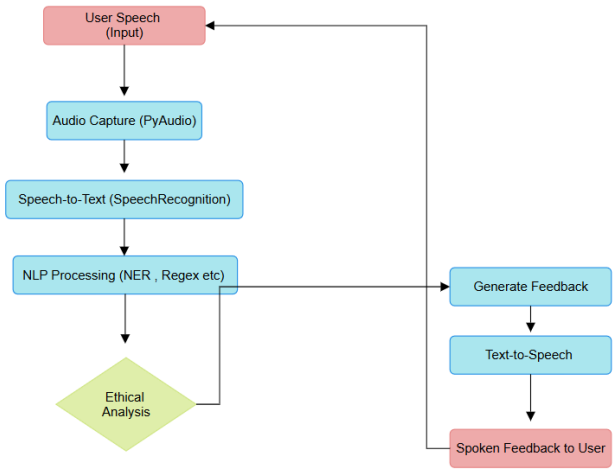

Figure 1: System Architecture Diagram

6.2 Use Case Diagram

The use case diagram maps interactions between the user and Hiday. Primary use cases such as **Provide feedback on language**

and **Generate ethical suggestions** are highlighted. This diagram emphasizes user interactions with the system's functionalities.

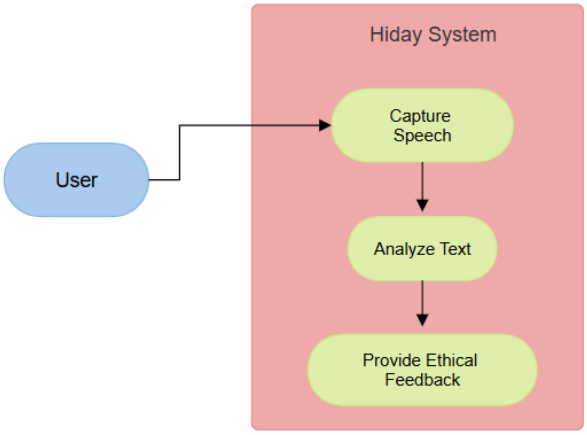

Figure 2: Use Case Diagram

6.3 Activity Diagram

The activity diagram depicts the step-by-step workflow within the system. It starts with audio input and ends with feedback output, showing intermediate steps like audio-to-text conversion, NLP analysis, and feedback generation.

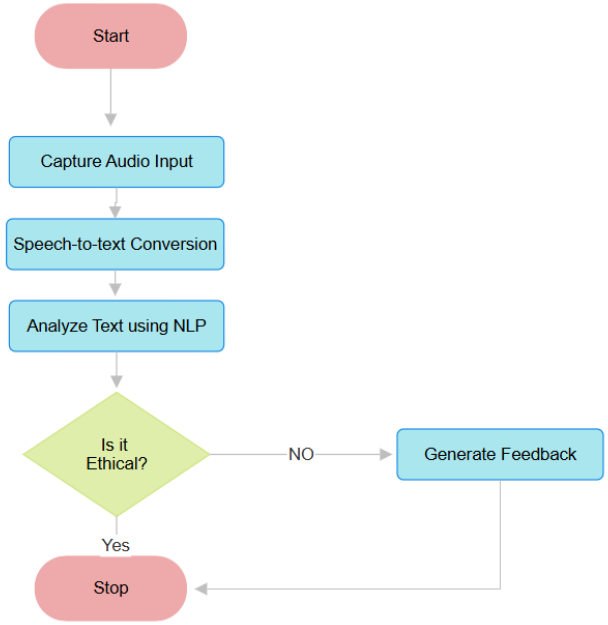

Figure 3: Activity Diagram

6.4 Domain Model

The domain model highlights the key entities in the system and their relationships. Components such as **Audio input**, **Text analysis**, and **Feedback generation** are shown with their attributes and interconnections.

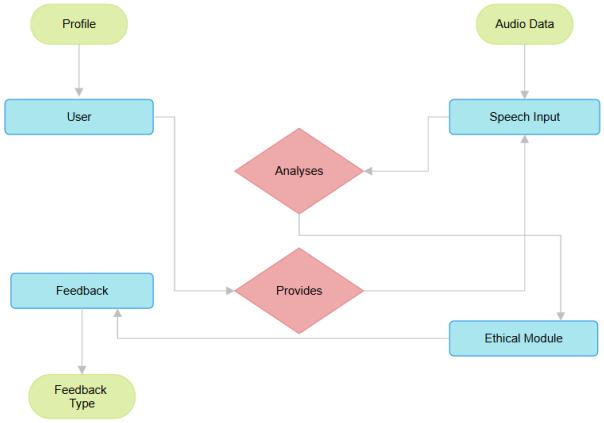

Figure 4: Domain Model

6.5 Sequence Diagram

The sequence diagram describes the temporal interaction between components during a user request. It includes the process of capturing audio, analyzing text, and providing feedback.

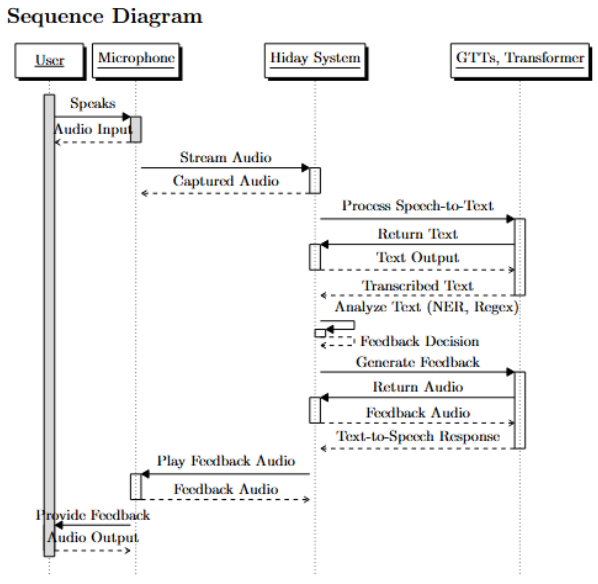

Figure 5: Sequence Diagram

6.6 Deployment Diagram

The deployment diagram outlines the physical arrangement of components. It shows the user device interacting with the Hiday server, transformers, and feedback mechanisms.

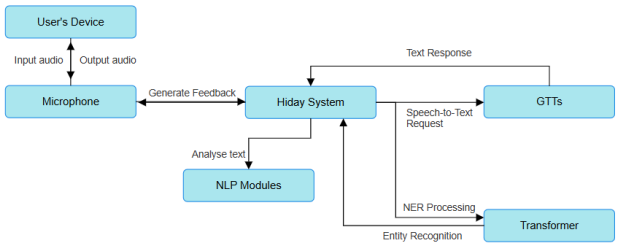

Figure 6: Deployment Diagram

[a4paper,12pt]article amsmath graphicx

EXPLANATION OF MODEL OUTPUTS

The running outputs of the model in the screenshots can be summarized as follows:

1. Initialization

- The system calibrates for ambient noise, setting an energy threshold to **934.9103721716871**.
- The Named Entity Recognition (NER) model is loaded.

2. Speech Recognition

- The assistant begins listening for input. When valid audio input is detected, it transcribes the spoken text into written form in **Urdu**.
- If the input audio is unclear or contains excessive noise, the system outputs:

  `Speech Recognition could not understand audio.`

3. Response Generation

- Upon successful transcription, the assistant generates a response based on the input. Example responses include:

  `Gheebat na krain (in urdu)`

- If the assistant is currently speaking, it logs:

  `Assistant is currently speaking. Ignoring input.`

  Additional input during this time is ignored until the current response is completed.

4. Control Handling

- The system manages multiple iterations of input and output. Each cycle involves listening, transcribing, and responding, with proper handling of overlapping inputs.

6.7 Hiday output

Below are the screenshots of the running outputs of the model.

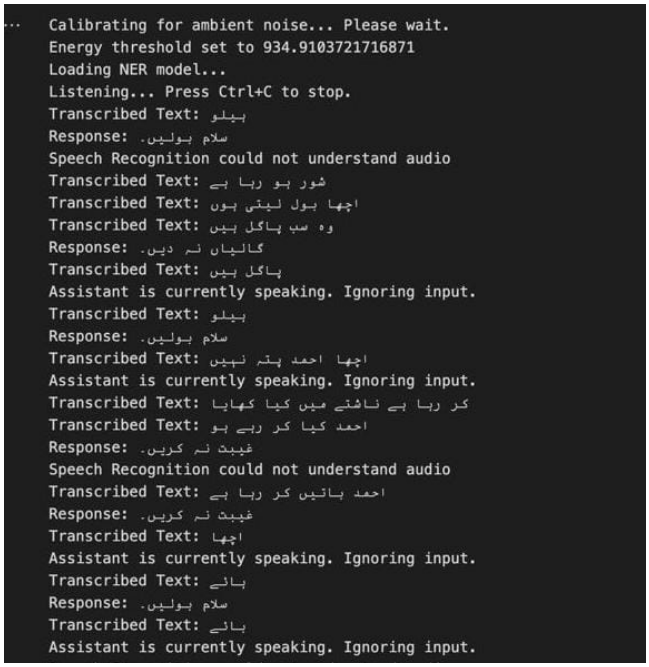

Figure 7: image 1

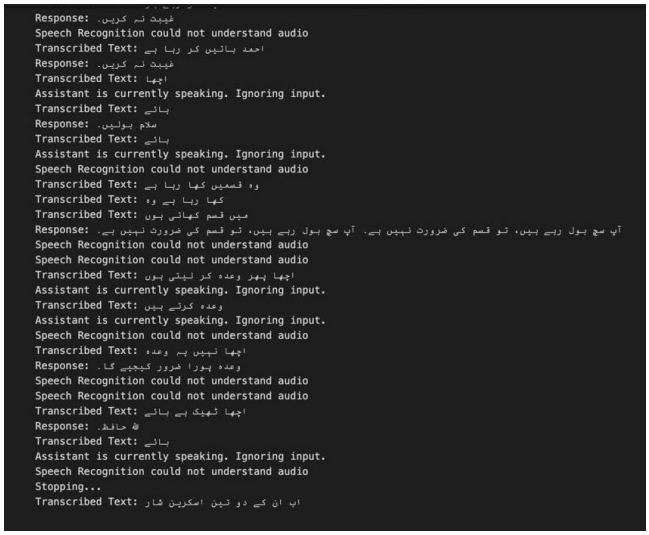

Figure 8: image 1

Conclusion

The outputs showcase a functional speech-to-text and response system with clear handling of edge cases, such as noisy input.