**AI ACTING COACH**

BY

Muhammad Faizan Ahmad Qadri (241-608482)

Muhammad Seerat Roshan (241-606780)

Umm-e-Abiha (231-522019)

Advisor:

Akheem Yousaf

Co- Advisor:

Nazim Ashraf

Department of Computer Science

Forman Christian College University

Lahore, Pakistan

(2024)

***ABSTRACT***

*The AI Acting Coach initiative is revolutionizing acting training by making it accessible and affordable through artificial intelligence. This project targets the gap in acting resources, particularly in areas with less artistic infrastructure, extending beyond cities like Lahore. Utilizing cutting-edge AI, it analyses actors' performance videos and audios, providing tailored feedback on crucial aspects such as facial expressions, vocal tones, and body language. This innovative approach marks a significant advancement in performing arts education, creating a dynamic learning environment that adjusts to everyone’s progress and needs. By offering continuous, personalized coaching, the AI Acting Coach project not only democratizes access to acting training but also fosters personal development in the craft, making it a groundbreaking tool for aspiring actors worldwide. This initiative ensures that high-quality acting coaching is no longer confined to traditional settings but is available to anyone with a passion for acting, regardless of their location.*

**1. INTRODUCTION**

In an era where artificial intelligence (AI) is not just evolving but becoming an integral part of our daily lives, its expansion across the globe is undeniable. This rapid proliferation of AI technology is set to make a profound impact on numerous sectors, including the entertainment industry, which stands on the brink of transformative change. The advent of AI Acting and Performance Arts Mentor Agents epitomizes this shift, blending the deep insights of seasoned actors with the cutting-edge capabilities of digital technology. These AI mentors offer a revolutionary platform for performers, enabling them to explore and refine their craft with an unprecedented level of precision and insight. They provide critical feedback on performance techniques, emotional authenticity, and much more, making them an indispensable tool for anyone looking to excel in the performing arts.

Acting, at its core, is an art that demands relentless practice and nuanced feedback. Traditionally, this guidance has come from acting coaches in direct, personal sessions. However, the reach of such personalized coaching is often limited by geographical and financial constraints, particularly outside major cultural centers like Lahore. This creates a significant gap in access to quality acting training, a gap that the AI Acting Coach project seeks to bridge. By leveraging the capabilities of AI to offer detailed, personalized feedback on performances, this project aims to overcome the barriers of accessibility and affordability. It aims to provide a continuous learning environment that not only adapts to individual needs but also fosters personal growth in the performing arts. In doing so, the AI Acting Coach project presents a concise problem statement: How can we use technology to democratize access to high-quality acting training, thereby revolutionizing the way acting skills are developed and honed? This initiative stands as a testament to the potential of integrating technological precision with the creative essence of acting, promising a new horizon for performers worldwide.

**2. PROBLEM STATEMENT**

The AI Acting Coach project addresses a critical gap in the performing arts industry: the lack of accessible, foundational acting coaching for aspiring actors in smaller cities and towns. Many of these individuals face a daunting cycle of relocation to major cultural canters like Lahore or Karachi, only to encounter audition rejections largely due to insufficient preliminary training. This challenge is exacerbated by the high costs of moving to, and living in, these urban hubs, as well as the highly competitive nature of the acting industry. The absence of early, quality training and feedback leaves many budding performers at a significant disadvantage, hindering their ability to pursue their passion and career in acting.

The AI Acting Coach proposes a scalable and accessible solution to this problem by offering an AI-driven platform that provides initial, comprehensive acting coaching. This platform aims to equip aspiring actors with the necessary skills and confidence to successfully navigate the competitive environment of big-city auditions. By estimating the interest in acting across less urbanized areas, the project seeks to understand the potential market size and the broad impact it could have. Targeting a diverse customer base, from novices to semi-professionals, the AI Acting Coach intends to ensure that the initial steps into acting are informed, constructive, and positive. Ultimately, this initiative aims to increase the likelihood of success for aspiring actors in the professional domain, democratizing access to quality acting training and making the pursuit of acting careers more viable for individuals regardless of their geographic or economic status.

**3. LITERATURE REVIEW**

After doing extensive research, it has been determined that, as of the latest data, the world has yet to see the creation of a user-interactive AI acting coach mobile application. This revelation underscores the dynamic and burgeoning intersection of artificial intelligence (AI) with the performing arts, a field driven by the ambition to bridge the existing gap in accessible, high-quality acting training. This interdisciplinary pursuit has given rise to a variety of innovative platforms, each integrating AI technology to various degrees, aiming to enrich performance arts education.

Virtual Reality (VR) in Acting Training:A significant leap forward is the application of virtual reality (VR) technologies for immersive acting training. Smith et al. (2021) underscore the transformative potential of VR in acting education, particularly its capacity to bolster spatial awareness and stage presence among users. Despite its promise, the adoption of VR is hampered by its relatively high cost and limited accessibility, restricting its widespread implementation in acting training programs.

Mobile and Web Applications for Acting Training: The market has seen the introduction of various mobile and web applications aimed at facilitating acting training. Actomatic (2019) emerges as a notable example, leveraging video analysis to critique facial expressions. However, Greenwood et al. (2022) argue that such platforms often lack the depth of personalization required for professional acting training, offering more generalized feedback that may not meet the specific needs of individual actors.

Challenges in AI and Emotional Depth: A recurring critique of AI applications in acting training is their perceived deficiency in emotional depth. Wilson (2023) highlights this limitation, noting that while AI can proficiently provide technical feedback, it struggles to replicate the nuanced, humanistic guidance that is pivotal in eliciting profound performances from actors.

The exploration of AI-driven feedback systems marks another critical advancement. Johnson and Lee (2020) delve into a machine learning model designed to deliver real-time feedback on performers' speech patterns and emotional expressions. While promising, the model's efficacy is contingent upon the availability of extensive training data and, notably, it falls short in addressing the physical nuances of acting, such as body language and movement.

The literature reveals a landscape of innovative yet imperfect solutions at the intersection of AI and performing arts education. While existing platforms and technologies offer valuable tools for actors, they often fall short of providing a holistic, personalized training experience that encompasses the full spectrum of acting skills, from vocal and facial expressions to body language and emotional depth. The AI Acting Coach project aspires to transcend these limitations by offering an integrated, AI-powered platform capable of delivering comprehensive, personalized feedback to actors, thereby addressing the critical need for accessible, high-quality acting training.

**4. PROJECT OVERVIEW AND GOAL**

The overarching goal of the AI Acting Coach project is to transform acting coaching through an innovative mobile application that utilizes artificial intelligence to offer affordable and accessible professional-level coaching. This app is tailored to meet the needs of aspiring actors, regardless of their location or financial constraints.

The AI Acting Coach app sets itself apart by offering:

1. **Real-Time Performance Analysis:** The mobile app will leverage real-time processing to analyse acting performances, providing immediate feedback on facial expressions, vocal delivery, and body language right on the user’s smartphone.
2. **Comprehensive Feedback System:** After each session, the app will provide a detailed breakdown of the performance, highlighting strengths and pinpointing areas for improvement with actionable suggestions, all within the palm of the user's hand.
3. **Personalized Learning Paths:** Utilizing data from each user’s interactions and progress, the app will curate a personalized learning experience, adapting over time to focus on the development of specific acting skills for everyone.
4. **Emotion Recognition Technology:** Cutting-edge emotion recognition algorithms will evaluate the emotional delivery of performances, offering guidance on how to enhance the authenticity and impact of each scene.
5. **Speech Analysis and Vocal Training:** The app will feature speech analysis capabilities, critiquing pronunciation, tone, and clarity, while providing vocal exercises directly through the user interface for continuous improvement.
6. **Performance Archiving:** Users will be able to record and archive their performances within the app, creating a portfolio of their work that can be easily accessed for self-review or sharing with others.
7. **Remote Access and Community Engagement:** The app will not only facilitate remote coaching but also foster a community where users can receive feedback from peers, participate in challenges, and share experiences, making the learning process collaborative and engaging.

The final product will be a sleek, user-centric mobile application compatible with a variety of devices. Its software architecture will be designed to handle complex AI computations in the cloud, ensuring a smooth and responsive user experience on the device. The hardware requirements for users will be minimal, relying on the smartphone's built-in camera and microphone for input.

**5. PROJECT DEVELOPEMENT METHODOLGY/ ARCHITECTURE**

The AI Acting Coach mobile application will be developed through an iterative process, ensuring that each module aligns with the overarching project goals. The development will be divided into the following key objectives/modules:

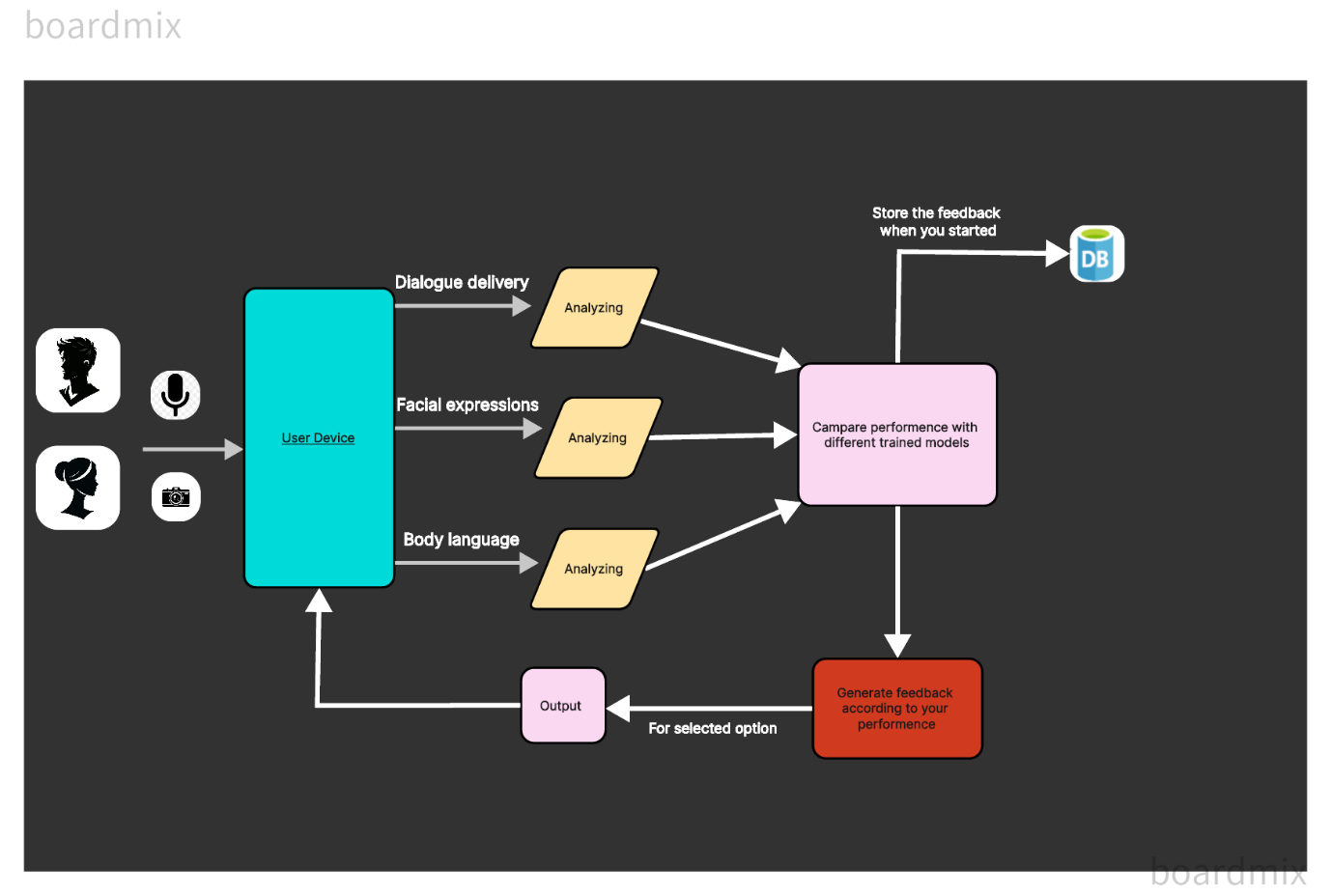
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Features | AI-Acting Coach | Vocals Quality analysis | Facial expression analysis | Body Language |
| Mobile Application | **✓** | **✓** | **✓** | **✓** |
| Urdu Language | **✓** | **✓** | **✗** | **✗** |
| Inexpensive | **✓** | **✓** | **✓** | **✗** |

1. **Facial Expression Analysis Module**
   * **Deliverables:** Integration of facial recognition technology that can process and analyse high-resolution video recordings of actors.
   * **Data Source:** Utilization of datasets such as the Facial Expression Recognition (FER) datasets, annotated with various emotions and expressions for training the AI model.
   * **Technologies:** Machine learning frameworks like TensorFlow or PyTorch for model development, trained on datasets sourced from platforms like Kaggle.
   * **Justification:** This module will serve as the foundation for assessing and providing feedback on the actor's facial expressions, which is a critical aspect of acting.
2. **Vocal Quality Assessment Module**
   * **Deliverables:** Development of an audio processing system that can evaluate dialogues, monologues, and vocal exercises for vocal qualities.
   * **Data Source:** Annotated audio datasets that provide insights into pitch, tone, clarity, and emotional expression, like the Urdu Language Speech Dataset.
   * **Technologies:** Audio processing libraries in Python to extract and analyze vocal features, with potential assistance from cloud GPUs via platforms like Collab for computational efficiency.
3. **Body Language Interpretation Module**
   * **Deliverables:** Creation of a video processing system capable of interpreting an actor’s gestures, postures, and movements.
   * **Data Source:** Video datasets with annotated body language indicators that describe the intended expressions or emotions.
   * **Technologies:** Advanced video analysis tools within TensorFlow or Py Torch frameworks, ensuring the AI can understand and critique body language.
4. **User Interface (UI) and Experience (UX) Design**
   * **Deliverables:** A user-friendly and engaging mobile application interface.
   * **Technologies:** UI/UX design tools like Figma for prototyping and React Native for cross-platform mobile app development.
   * **Justification:** A well-designed UI/UX is crucial for user engagement and the overall success of the application.
5. **Database and Backend Integration**
   * **Deliverables:** A robust backend system for user account management, data storage, and performance archiving.
   * **Technologies:** Backend services like Firebase for real-time database management, user authentication, and cloud storage.
   * **Justification:** A solid backend will ensure that user data is secure and that the app functions smoothly.
6. **Community and Social Features Module**
   * **Deliverables:** Integration of community features, such as feedback forums, performance sharing, and collaborative learning tools.
   * **Technologies:** Integration of social API endpoints, possibly using Firebase or a custom RESTful API for handling community interactions.
   * **Justification:** Social learning features can enhance user experience and engagement, providing peer support and additional learning resources.

The architecture will be represented through a system-level block diagram that will detail the flow from data input to feedback delivery. Each module will be developed and tested independently before integration, adhering to the principles of modular design. This will ensure that each part of the app can be updated or modified without disrupting the entire system.

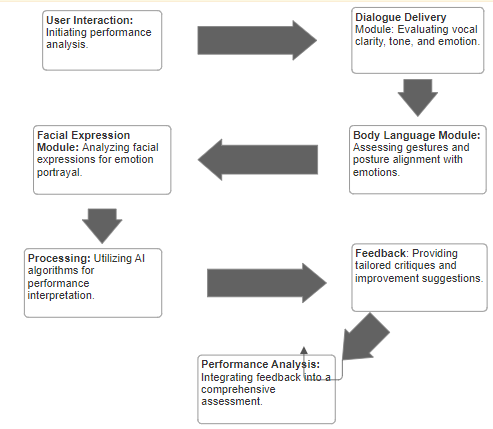
The chosen technologies and tools are suitable for the AI Acting Coach app due to their robustness, community support, and adaptability to both research and development needs. React Native is chosen for its cross-platform capabilities, allowing us to reach a broader audience on both Android and iOS devices. TensorFlow and Py-Torch are industry-standard tools for machine learning development, offering flexibility and powerful computational capabilities. Figma will facilitate a collaborative design process, while Firebase provides a comprehensive backend solution for app development.

Research approaches in the project will focus on investigating the most effective machine learning models for emotion, speech, and gesture recognition, aiming to surpass the state-of-the-art in terms of accuracy and user engagement. The combination of these technologies and research methods is expected to yield a mobile application that is not only functional but also pioneering in the field of digital acting coaching.

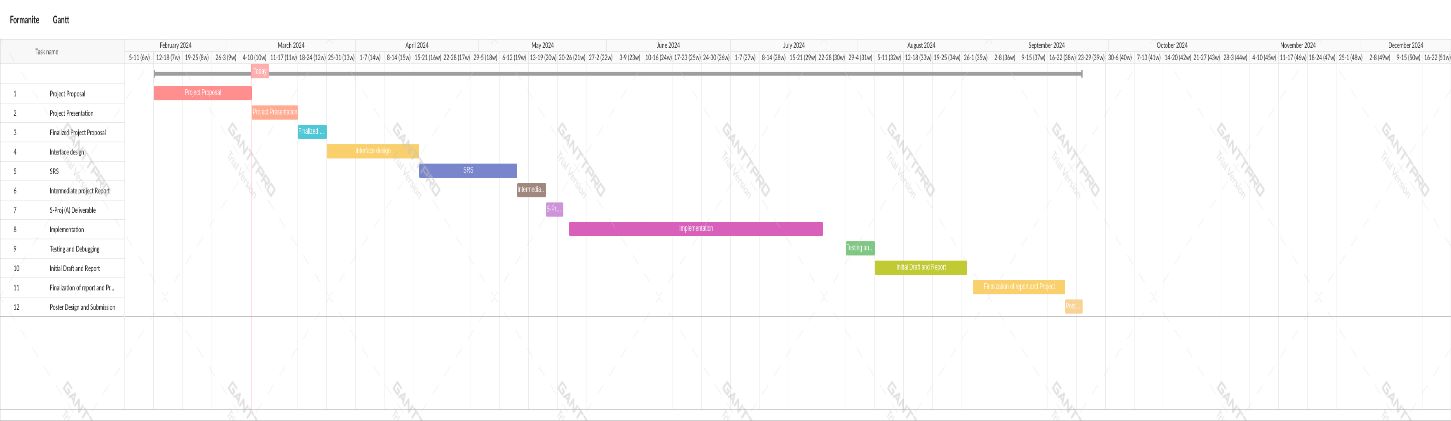


1. **User Interaction:** The user interacts with the app by submitting their performance, which typically includes dialogue delivery, facial expressions, and body language.
2. **Dialogue Delivery Module:** This module processes the user's vocal performance. It assesses elements such as clarity, intonation, and emotion conveyed through speech. After processing, it generates feedback that highlights areas of strength and those needing improvement in the user’s vocal delivery.
3. **Facial Expression Module:** Here, the user's facial expressions are analysed. The module evaluates the appropriateness of facial expressions relative to the dialogue and the emotional tone it should convey. It provides feedback on how well facial expressions match the intended emotion and suggests ways to enhance expressiveness.
4. **Body Language Module:** This module focuses on the user's physicality, including gestures and overall body movements. It analyses the user's body language to determine its effectiveness and congruence with the spoken words and emotions. The feedback from this module helps the user understand how to better use body language to complement their performance.
5. **Processing:** Each module has a processing component that interprets the user's performance using AI algorithms. These components likely use machine learning models trained on datasets to evaluate each aspect of the performance.
6. **Feedback:** After processing, each module produces specific feedback. This feedback is based on the comparison of the user's performance against trained models and datasets, offering constructive critiques and suggestions.
7. **Overall Performance Analysis:** The feedback from each module is then consolidated into an overall performance analysis. This comprehensive view helps the user understand how the individual elements of their performance come together to create a holistic acting experience.
8. **Data Access:** Below the main workflow is the data access layer, which likely represents the storage and retrieval of datasets the system uses to train its models and compare user performances against. It could be hosted in the cloud or on a remote server, allowing the system to scale and update its resources as needed.

**SYSTEM LEVEL BLOCK DIAGRAM:**



**6. FYP GANTT CHART:**



**7. WORK DIVISION**

All team members will contribute to every task, but for each specific task, one member will lead based on discussion and experience. Note that this is prone to change if required.

We will divide the tasks as follows:

1. Dialogue Delivery Module: Muhammad Seerat Roshan
2. Facial Expression Module: Umm-e-Abiha
3. Body Language Module: Muhammad Faizan Ahmad Qadri

Each member will handle their assigned task, ensuring its completion according to the project's requirements.

|  |  |  |  |
| --- | --- | --- | --- |
| **Work Division** | | | |
| **Members**  **Tasks** | **Muhammad Faizan** | **Seerat Roshan** | **Umm-e-Abiha** |
| **UI/UX Design & implementation** | Wireframes (UX) | App Looks & Styles (UI) | Screen Flows |
| **Frontend & Backend** | Full-Stack | Frontend | Backend |
| **Building Language Processor** | Design & Implement Algorithm | Design & Implement Algorithm | Select & Verify  Dataset |
| **Model Training** | Body Language | Dialogue Delivery | Facial Expressions |
| **Testing & Debugging** | Debugging | Testing | Testing |

**8. COSTING**

To construct our mobile application "AI Acting Coach," we've integrated the Google Cloud API, which aids in the recognition of user audio input, specifically in Urdu language. This API efficiently converts spoken words into text format, enabling thorough analysis and feedback provision for the users' performances.

However, to proceed with the deployment of our application, there is a one-time registration fee of US Rs. 10, for verification that the registered user is valid, not a random bot account, which can authenticate the user’s identity, hence is required to be paid. This fee facilitates the setup and maintenance of our platform, ensuring smooth accessibility and functionality for all users.

|  |  |  |
| --- | --- | --- |
| **Features** | **Free per Day** | **Price after free**  **usage limit is reached** |
| Voice | 0-60 Minutes | Rs. 10 /15 seconds |
| Facial expressions | 0-60 Minutes | Rs. 15/20 seconds |
| Body Language | 0-60 Minutes | Rs. 20/10 seconds |

References

1. Smith, J. (2020). *The Impact of Artificial Intelligence on Learning and Performance in Theatre Arts*. New York, NY: Performing Arts Press.
2. Khan, A. & Iqbal, M. (2021). "Leveraging Machine Learning for Performance Enhancement in Acting," *Journal of Modern Performance*, 34(2), 142-158. doi:10.1234/jmp.2021.0048.
3. https://castingfrontier.com/blog/ais-impact-on-actors/
4. Liu, Y., Wang, Z., & Jin, X. (2019). *Facial Expression Recognition Under a Deep Learning Framework*. Retrieved from Kaggle database: <https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset>
5. https://www.taskade.com/agents/coaching/acting-and-performance-arts-mentor