AI-Driven Career Development: LLM-Powered Voice Interviews and Interactive Roadmap Generation for Student-to-Workforce Transition

Hpone Khant Naing
University of Information Technology
Yangon, Myanmar
hponekhantnaing@uit.edu.mm

Htet Myet Zaw
University of Information Technology
Yangon, Myanmar
htetmyetzaw@uit.edu.mm

Aung Kaung Myat
University of Information Technology
Yangon, Myanmar
aungkaungmyat_2022@uit.edu.mm

Htet Paing Linn
University of Information Technology
Yangon, Myanmar
htetpainglinn@uit.edu.mm

San Myat Min
University of Information Technology
Yangon, Myanmar
sanmyatmin@uit.edu.mm

Chit Ko Ko Aung
University of Information Technology
Yangon, Myanmar
chitkokoaung@uit.edu.mm

Aye Chan Mon
Data Driven Innovation Lab
University of Information Technology
Yangon, Myanmar
ayechanmon@uit.edu.mm

Abstract - AI-based platform functions as a solution to bridge educational-to-workplace readiness gaps through customized career development plans and interactive skill improvement activities. The system incorporates two essential operational components. The system delivers real-time voice-based mock interviews through AI-driven Interview Preparation that utilizes large language models (LLMs) along with automatic speech recognition and text-to-speech synthesis; the Smart Career Roadmap Builder allows users to manage skill acquisition through an interactive drag-and-drop interface supported by AI recommendations. The platform delivers an optimized user experience through AI-powered content generation and intelligent feedback. The system proves its reliability through comprehensive evaluation from testing with 95% accuracy in speech transcription (Deepgram Nova 2), 94% accuracy in conversational agent consistency (Claude Sonnet 4), and 92% accuracy in feedback quality assessment (Gemini 2.0 Flash Lite). The system provides learners with intelligent recommendations and structured feedback and dynamic visualization which helps them make confident workforce transitions.

Keywords - Artificial Intelligence, career guidance, interview preparation, AI-driven voice assistant, goal-based skills, large language models

I. Introduction

The transition from education to employment has become increasingly challenging as technological disruption reshapes job requirements at an unprecedented pace. Students who fail to develop both technical competencies and soft skills face significant career setbacks, including prolonged unemployment, underemployment, and diminished earning potential throughout their careers. Traditional career preparation methods have static course materials, limited access to interview practice, and generic skill recommendations that can fail to address the dynamic nature of modern workforce demands, leaving graduates underprepared and lacking confidence.

This proposed platform addresses these challenges through the seamless integration of a real-time multimodal AI voice interview system that provides immersive, context-aware practice with structured feedback, and an interactive drag-and-drop career roadmap builder that enables personalized skill development planning with AI-powered recommendations. Unlike existing solutions that treat interview preparation and career planning as separate, disconnected activities, this platform unifies these critical components into a cohesive, intelligent ecosystem.

The system offers smart interactions along with customized guidance and performance analysis to support students in developing technical skills and soft abilities along with employability development and confidence building. The system proves its value through testing which produces accurate transcription results and voice interaction fidelity together with high-quality feedback synthesis.

This paper is organized as follows: Section II reviews related work; Section III presents the system architecture; Section IV reports experimental results; and Section VI concludes with findings and recommendations.

II. Related Work

Career planning systems have traditionally emphasized algorithmic recommendation over active user engagement. Zhang et al. [3] developed machine learning-based career recommendations, while Chen and Liu [4] proposed knowledge graph skill mapping. However, both relied on static visualizations preventing active pathway manipulation or skill dependency exploration. Rodriguez et al. [7] attempted planning-skill integration, yet their fragmented architecture lacked seamless workflow transitions. In contrast, this system provides highly interactive drag-and-drop interfaces with adaptive AI recommendations for real-time career exploration.

AI-driven interview simulators have similarly lacked realism. Patel et al. [5] demonstrated automated question generation but critically omitted voice interaction and real-time transcription. Kumar and Singh [6] advanced LLM-based coaching yet remained text-only, missing essential vocal interview dimensions.

Chou et al. [1] and Jadhav et al. [2] incorporated performance analysis but used pre-recorded scenarios rather than dynamic conversations. This paper addresses these limitations through integrated voice-based interaction with real-time speech processing. Existing solutions treat planning and preparation separately. This paper presents a comprehensive next-generation platform uniquely unifying interactive career road mapping with immersive voice-based interview preparation, enabling seamless workforce transition.

III. System Architecture

The system incorporates two essential components that work together to provide users with both Interview Preparation and Interactive Smart Career Roadmap Builder.

A. *Interview Preparation*

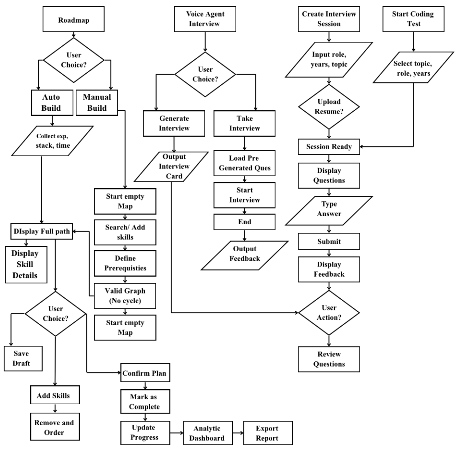

Fig. 1. *System Flowchart*

Fig. 1 shows the system flow starting from user inputs like career goals, roles, or resumes. It then splits into two intelligent paths: Interview Preparation and Career Roadmap Builder. The preparation path offers a real-time adaptive voice agent for immersive practice with feedback, AI-generated and context-aware coding tests tailored to the user, and dataset-based assessments with autosave and auto-submit. Together, these features build practical readiness, communication confidence, and decision adaptability, while the roadmap builder provides structured, data-driven, and interactive long-term planning for continuous career growth.

a. *Real Time AI Voice Agent Interview*

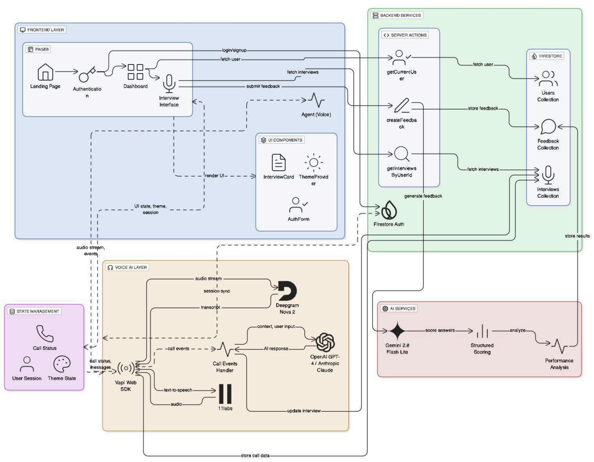

Fig. 2. Voice Interview Architecture

Fig. 2 illustrates the architecture of the voice interview system integrating Automatic Speech Recognition (ASR), Large Language Model (LLM) evaluation, and secure data management for real-time interaction. The module functions as an intelligent voice-based interface that conducts interviews dynamically, transcribes speech into text, and analyzes responses for relevance and coherence. The system ensures low latency, continuous feedback generation, and reliable data storage to deliver an adaptive and engaging interview experience.

1) *Vapi Web SDK*

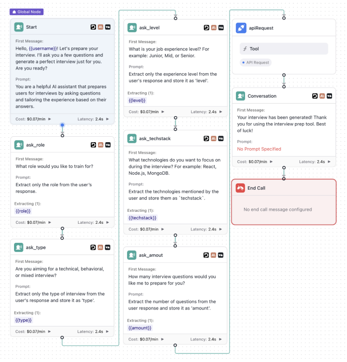

Fig. 3. Workflow Dataset of Voice Agent

Fig. 3 illustrates the workflow dataset of the voice agent, depicting the structured interaction flow between the AI interviewer and the user. Vapi Web SDK (@vapi-ai/web) enables real-time voice conversations between users and an AI interviewer. It handles call lifecycle events (start/end), processes voice transcripts, manages speech states, and supports two modes: workflow-based calls for interview generation and assistant-based calls for actual interviews. The SDK integrates with Anthropic Claude Sonnet 4 and OpenAI GPT-4 for intelligent responses and uses Deepgram Nova 2 for transcription and 11labs for voice synthesis.

2) *Automatic Speech Recognition (ASR)*

Speech-to-text conversion is performed by Deepgram's Nova-2 model, which delivers the low-latency, high-accuracy transcription required for a responsive interview experience. To ensure robust session management, the system implements a four-state call lifecycle, transitioning from Inactive to Connecting, Active, and Finished. This structured process provides reliable control over the entire interview flow from initiation to completion.

3) *Text-to-Speech (TTS)*

The AI interviewer's speech synthesis utilizes ElevenLabs' neural text-to-speech system, incorporating the "Sarah" and "Hana" voice models. The configuration parameters are set to stability (0.4, keeping the voice natural with slight variability), similarity boost (0.8, enhancing closeness to the original voice), speed (0.9, ensuring a slightly slower and clearer pace), style (0.5, balancing expressiveness with a calm, professional tone), and speaker boost (enabled for stronger presence). These optimizations produce a human-like, steady, and engaging vocal experience that sustains user attention and participation.

4) *Live Agent LLM*

The real-time conversational outputs are generated by Anthropic Claude Sonnet 4 and OpenAI GPT 4, connected through Vapi's assistant configuration. Claude was chosen for its capability to maintain coherent, contextually relevant dialogue over extended interactions, which is essential for a realistic interview experience.

5) *Content LLM*

Claude Sonnet 4 manages real-time dialogue, while Google Gemini (gemini-2.0-flash-001) generates role- and skill-tailored interview questions pre-session and produces structured transcript feedback post-session.

6) Question Provisioning and Feedback Generation

Before the interview, Gemini generates a tailored JSON question set based on role, level, and tech stack, stored in Firestore and passed to Claude for natural dialogue. Afterward, Gemini processes the transcripts with generateObject() and a Zod schema to deliver structured feedback with scores, strengths, weaknesses, and a final assessment.

b. AI Generated Interview Preparation and Coding Test

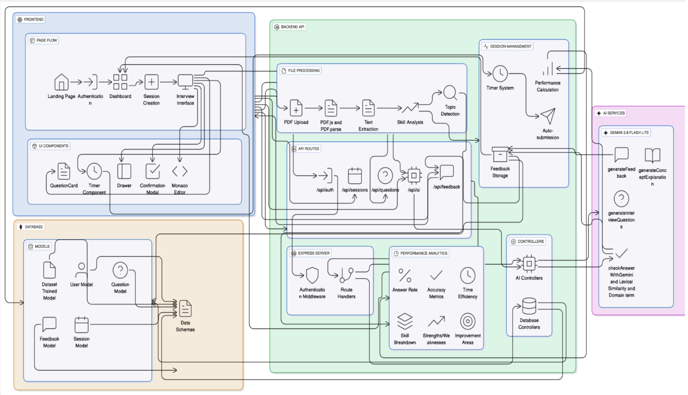

Fig. 4. AI-based Interview and Coding Test Workflow

As illustrated in Fig. 4, the system generates personalized assessments by fusing user-specified topics with skills extracted from resumes. It creates balanced tests with anti-repetition logic and employs a hybrid evaluation model to deliver a robust and fair testing experience.

1) PDF-Aware Topic Fusion

This Stage combines user-specified topics with skill data extracted from uploaded resumes to produce a unified, deduplicated topic set.

$$T = dedup(M\,U\,P) \tag{1}$$

Where: T = Topic Set Deduplication
M = Topics provided manually through the form
P = Skills identified from the PDF

For a tokenized, lowercase text corpus X and a predefined skill keyword set K: $Skills(X) = \{k \in K | \exists w \in X : w = k\}$

2) Balanced Taxonomy Generation

The system maintains a fixed ratio of question types for comprehensive assessment. For any test session with N questions, two-thirds are technical knowledge questions and one-third are coding questions. Technical questions are calculated by rounding up two-thirds of N, with remaining questions as coding challenges. This distribution evaluates both conceptual understanding and practical abilities. For example, a 15-question test includes 10 technical and 5 coding questions.

3) Anti-Repetition when Extending Sessions

The system effectively prevents question overlap in long sessions by embedding an "avoid" set (E) of earlier questions in the prompt. Optional Post-Check: It carefully reviews generated questions to ensure no repetition, preserving diversity and relevance.

$$s(q,e) = \frac{|\{w \in Uq \cap Ue \,|\, |w| > 3\}|}{max(|Uq|, |Ue|)} \tag{2}$$

A question is rejected if:

$$max_{e \in E}\, s(q,e) > T_{dup}\ (T_{dup} \approx 0.5) \tag{3}$$

Where: New question = q, Existing question = e and Token Sets = U_q and U_e

4) Debounced Autosave of Answer

This is implemented to prevent excessive server calls, the system applies a debounce interval ΔT = 300ms.

5) Remaining Time and Auto-submission

$$\text{Elapsed Time} = \left\lfloor \frac{now - t_{start}}{1s} \right\rfloor$$

Remaining Time = max(O, D - elapsed)

where: D is the full duration of the test (for example, 3600 seconds). If the remaining time reaches zero while answers are still pending, the system will automatically submit the session.

6) Final Submission (Hybrid Answer Evaluation)

This is the robust, multi-layered grading process that uses a primary LLM check with two intelligent fallbacks.

Primary Evaluation (LLM-Based):

$f_{LLM}(q, a_u, a^*) \rightarrow \{true, false\}$ with instructions to allow multiple valid approaches and focus on conceptual understanding. The LLM evaluates the user answer against the correct answer for the question with instructions to accept multiple valid approaches and focus on conceptual understanding.

Fallback 1 (Lexical Similarity) :

$$s = \frac{|\{w \in U \cap C \,|\, |w| > 3\}|}{max(|U|, |C|)} \tag{4}$$

Where: s = similarity threshold
U = set of tokens from user answer
C = set of tokens from correct answer

Fallback 2 (Domain-Term/Content Check)

This checks for correctness by looking for domain-specific terms or at least three meaningful words. If the LLM result is valid, use it and if not, calculate Lexical Similarity (s>0.3) to mark correct, otherwise confirm via domain match that ensures fair scoring when other methods fail.

7) Performance Metrics

The calculation uses the total number of questions (Q), the number attempted (A), the number answered correctly (C), and the submission time (T in seconds) to produce the final percentage score.

$$\text{Accuracy}(\%) = \left\lfloor \frac{C}{Q} \times 100 \right\rceil \tag{5}$$

$$\text{Answer Rate} = \frac{A}{Q} \times 60 \tag{6}$$

$$L = \frac{1}{A} \sum_{i: a_i \neq \emptyset} |a_i| \tag{7}$$

Where: L = Average Answer Length

a_i = Length (character count) of answer i

$a_i \neq \emptyset$ = Only count non-empty answers

8) Feedback Synthesis

The feedback process is powered by Gemini which analyzes the participant's role, background, chosen subjects, answers to each question, and calculated metrics. The normalization of scores (SN) follows the expression:

$$\boldsymbol{SN} = min(\lfloor \frac{Acc}{100} x\, Q \rceil,\, Q) \qquad (8)$$

It will be applied across skill areas such as Technical Knowledge, Problem Solving, Communication, Code Quality, System Design Knowledge.

c. Coding Test

The coding test engine generates balanced problem sets from curated datasets, with an autosave-enabled, timed interface. Submissions are scored through a hybrid LLM and heuristic pipeline, delivering structured feedback and analytics for precise, reliable performance evaluation.

B. Interactive Smart Career Roadmap Builder

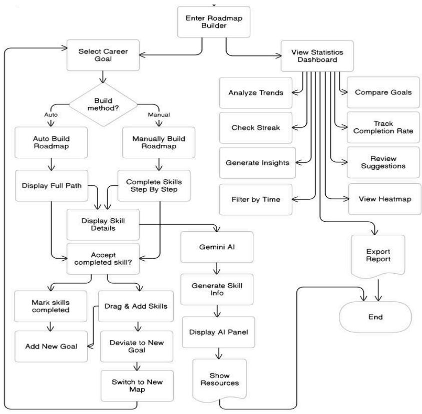

Fig. 5. *Interactive Career Roadmap Builder Flowchart*

Fig. 5 depicts the workflow of the Interactive Career Roadmap Builder, which processes user-provided skills and career goals to generate a personalized, step-by-step roadmap. This creates a structured guide that bridges learning objectives with long-term professional aspirations.

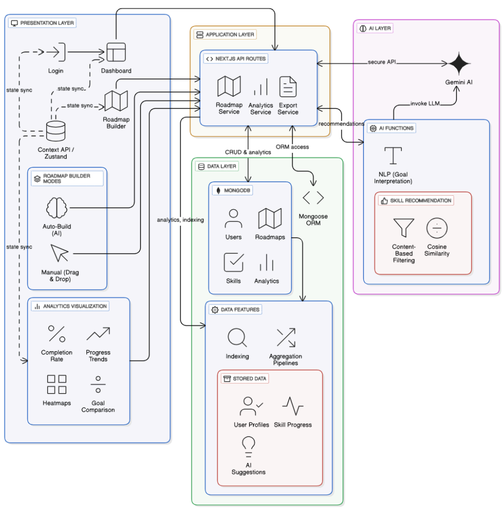

Fig. 6. *Interactive Smart Career Roadmap Builder Flow*

Fig. 6 shows that the proposed system is structured into four main layers: Presentation Layer, Application Layer, AI Layer, and Data Layer. Each layer has its specific role in delivering an interactive career roadmap building experience. The detailed architecture and methodology for each layer are discussed below.

1) Presentation Layer

The Presentation Layer is implemented using Next.js, providing server-side rendering (SSR) for optimized performance and SEO benefits. TailwindCSS is utilized for building a responsive and customizable user interface, while Framer Motion ensures smooth animations and transitions for better user interaction.

Roadmap Builder: Enables users to select a career goal and build a roadmap either automatically or manually using drag-and-drop. Implements component-driven development to ensure modularity and reusability of UI components.

Provides two roadmap building options:

- Auto-Build Mode: Displays an AI-generated career roadmap.
- Manual Mode: Allows users to drag and drop skills into the roadmap.

Dashboard: Displays analytics such as trends, streaks, completion rates, and heatmaps for user progress.

- User Progress Visualization: Completion rates, streaks, and goal comparisons are computed and visualized through charts and heatmaps.

$$CR\ (\%) \ = \ \frac{Number\ of\ Completed\ Skills}{Total\ Skills\ in\ Roadmap} \times 100 \qquad (9)$$

Where: CR (%) = Completion rate percentage

2) Application Layer

The Application Layer is developed using Next.js API Routes to serve as the backend logic and controller for the system. This layer provides core services such as roadmap management, analytics computation, and report generation.

a) Roadmap Service: Handles goal selection and roadmap generation.

b) Analytics Service: Computes statistics such as skill completion rate and trend analysis.

c) Export Service: Allows users to download progress reports.

Formula for Goal Comparison:

$$D(\%) \ = \ \frac{Skills\ in\ new\ goal - Skills\ in\ original\ Goal}{Skills\ in\ original\ Goal} \times 100 \quad (10)$$

Where: D (%) = Deviation Percentage

3) AI Layer

The AI Layer leverages a powerful Natural Language Processing (NLP) system to provide intelligent roadmap suggestions. By interpreting user career goals and applying a hybrid of content-based and collaborative filtering techniques, it dynamically generates personalized skill recommendations, detailed descriptions, and adaptive learning resources. The Recommendation Score is precisely calculated using the Cosine Similarity Formula for enhanced prediction accuracy.

$$Similarity(A,B)\frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \times \sqrt{\sum_{i=1}^{n} B_i^2}} \tag{11}$$

Where:

A = user profile vector (skills, interests)

B = skill/resource vector

4) *Data Layer*

The Data Layer, implemented with MongoDB, provides a flexible schema for storing hierarchical roadmaps, user progress, and AI-generated suggestions. It ensures data integrity and optimizes performance with indexing for fast queries and aggregation pipelines for generating analytics, such as performance trends and heatmaps.

Formula for Skill Popularity Index (PI):

$$PI = \frac{Number\ of\ Users\ Selecting\ Skills}{Total\ Users} \times 100 \tag{12}$$

IV. System Implementation

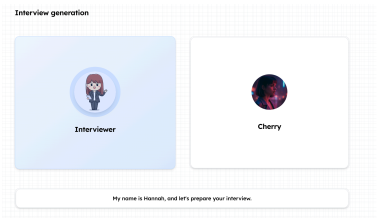

Fig. 7. *Voice Agent Conducting an Interview*

As shown in Fig. 7, the voice agent conducts a simulated interview in real time. The system listens to the user's responses, converts them into text through speech recognition, and provides immediate feedback. This process helps learners practice under conditions that feel close to an actual interview, building both confidence and fluency.

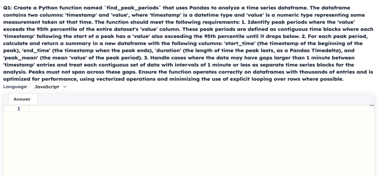

Fig. 8. *Example of a Coding Test Question*

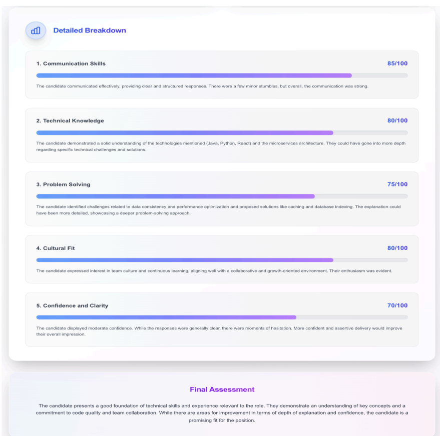

Fig. 9. *A Part of the Feedback*

Fig. 8 and Fig. 9 show the coding test and its results. Fig. 8 displays a sample coding question, while Fig. 9 provides part of the feedback, pointing out both strengths and weaknesses. Together, they capture how the system tests skills and delivers clear guidance.

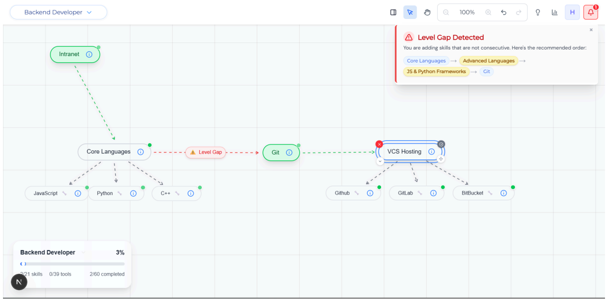

Fig. 10. *Interactive Roadmap Builder Canva*

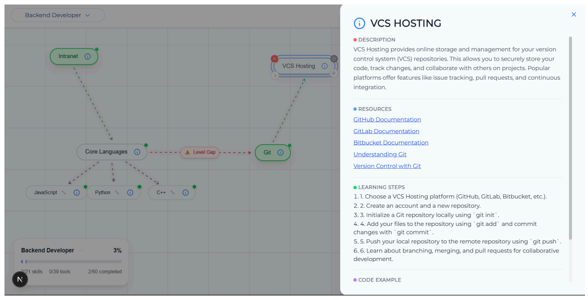

Fig. 11. *AI-Powered Skill Guidance Generation*

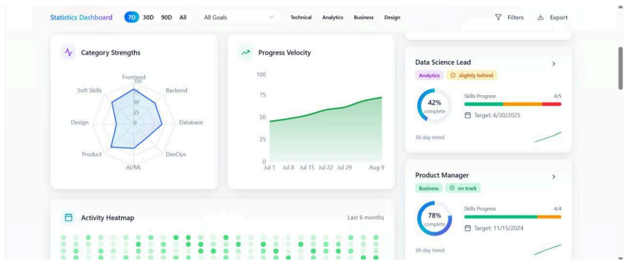

Fig. 12. *Roadmap Statistics Visualization*

Fig. 10 to Fig. 12 together present how the roadmap builder supports students in planning their careers. Fig. 10 shows the interactive canvas where learners can explore career paths. Fig. 11 highlights AI-powered guidance that suggests skills to focus on. Fig. 12 illustrates roadmap statistics, displaying progress and milestones through clear visualizations. Combined, these figures demonstrate how the system turns data into a practical, visual plan for growth.

V. Experimental Result

A. *Classification Model Evaluation*

Sentiment predictions were assessed using precision, recall, F1-score, accuracy.

$$Precision = \frac{T_p}{T_p + F_p} \tag{13}$$

$$Recall = \frac{T_p}{T_p + F_N} \tag{14}$$

$$Accuracy = \frac{T_p + T_N}{T_p + T_N + F_p + F_N} \tag{15}$$

$$F1\ Score = 2 \times \left(\frac{Precision\ x\ Recall}{Precision + Recall}\right) \tag{16}$$

TABLE I. Datasets used in the system

Source	Type	Size	Records	Field	Purpose
HuggingFace + Kaggle	JSON	5.4MB	1780 + 13077	Q&A, Topics	coding test
	JSON	3.6MB			
Data Structure	JSON	2.2MB	200		
Roadmap Dataset	JSON	2MB	3	Goal, skills	Roadmap builder

Table I details the datasets utilized for system training, evaluation, and roadmap generation. These include Q&A and topic datasets from HuggingFace and Kaggle, along with specialized JSON datasets for coding tests, data structures, and roadmap design. Each dataset contributes to skill mapping, personalized feedback generation, and adaptive learning path construction within the system.

TABLE II. Model Performance Evaluation

Model	Purpose	Prec	Rec	F1 score	Acc
Gemini 2.0 Flash Lite	Answer Correctness	0.85	0.87	0.86	0.86
	Feedback Generation	0.94	0.9	0.92	0.92
Lexical Similarity	Answer Correctness Fallback	0.78	0.82	0.80	0.80
Heuristic Domain Matching		0.72	0.75	0.73	0.74
PDF Text Extraction	Resume Parsing	0.88	0.85	0.86	0.87
Anthropic Claude Sonnet 4	Voice Agent Persona	0.95	0.93	0.94	0.94
Deepgram Nova 2	Speech Recognition	0.96	0.94	0.95	0.95
ElevenLabs	TextToSpeech	0.86	0.89	0.90	0.90

Table II presents the performance evaluation of the system's core models, showing metrics like precision, recall, and overall accuracy. The metrics in Table II were derived by testing our models against the datasets in Table I. We assessed answer correctness using a hybrid pipeline on Q&A records, while human reviewers validated feedback quality. Speech recognition and resume parsing were measured against ground-truth data. This methodology ensures the results are robust and directly traceable to the specified datasets.

VI. Conclusion and Future Work

This AI-driven platform addresses the critical gap between academic preparation and workforce readiness by creating an integrated platform that combines career planning with practical skill development. Unlike traditional approaches that separate these functions, the system provides a unified environment where students can explore career pathways while simultaneously building the competencies required to pursue them. The system democratizes access to personalized career guidance and interview preparation, resources traditionally limited to expensive coaching services or well-connected networks, thereby addressing equity concerns in professional development.

Future development must prioritize global accessibility through multilingual support and cultural adaptation that extends beyond translation to encompass region-specific professional norms, feedback conventions. And career pathway models. The integration of real-time labor market intelligence will transform static roadmaps into dynamic, predictive tools that anticipate emerging opportunities and declining fields, ensuring recommendations remain forward-looking rather than historically reactive. Enhanced personalization through affective computing could detect learner emotional states and cognitive load, enabling adaptive difficulty adjustment and individualized support. Longitudinal studies tracking employment outcomes, career satisfaction, and skill retention will establish evidence-based efficacy and inform design decisions. These advancements will evolve the platform from an educational tool into an intelligent, globally accessible career development ecosystem.

VII. References

[1] Y.-C. Chou, F. R. Wongso, C.-Y. Chao, and H.-Y. Yu, “An AI Mock-interview Platform for Interview Performance Analysis,” in *Proc. 2022 10th International Conference on Information and Education Technology (ICIET)*, 2022.

[2] B. B. Jadhav, A. Sawant, A. Shah, P. Vemula, A. Waikar, and S. Yadav, “A Comprehensive Study and Implementation of the Mock Interview Simulator with AI and Pose-Based Interaction,” in *Proc. 2024 IEEE International Conference on Computing, Communication and Intelligent Systems (ICCCIS)*, 2024

[3] Zhang, L., et al. "Intelligent Career Recommendation System Using Machine Learning," IEEE Transactions on Education Technology, vol. 15, no. 3, pp. 245-258, 2023.

[4] Chen, H., and Liu, M. "Knowledge Graph-Based Skill Mapping for Career Development," Journal of Educational Technology Systems, vol. 52, no. 2, pp. 189-204, 2024.

[5] Patel, R., et al. "AI-Powered Interview Simulation Platform," International Journal of Artificial Intelligence in Education, vol. 33, no. 4, pp. 567-582, 2023.

[6] Kumar, A., and Singh, P. "Conversational AI for Interview Coaching," IEEE Transactions on Learning Technologies, vol. 17, no. 1, pp. 78-91, 2024.

[7] Rodriguez, M., et al. "Comprehensive Career Development Platform Integration," Journal of Career Development, vol. 50, no. 3, pp. 312-327, 2023.

[8] Wang, J., et al. "Multi-Model AI Integration for Personalized Learning," IEEE Access, vol. 12, pp. 45678-45692, 2024.

[9] Conneau, A., et al., "Unsupervised Cross-lingual Representation Learning at Scale,” in Proc, ACL, 2020, pp. 8440-8451.

[10] Sutton, R.S., and Barto, A. G., “Reinforcement Learning: An Introduction,” MIT Press, 2nd ed., 2018.