JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-21, No.-04, April (2026) pp 240-251 ISSN (Print) 0973-8975

A UNIFIED COMPUTATIONAL MODEL FOR LLM–MULTIMODAL FUSION IN AUTOMATED CAREER ASSESSMENT

Sricharani P.[1], D. N. S. B. Kavitha[2]

Artificial Intelligence Department, Shri Vishnu Engineering College for Women (A), Bhimavaram, Andhra Pradesh, India.

[2]Department of Computer Science and Engineering, SRKR Engineering College, Bhimavaram, A.P, India.

Email: [1]charani.yashu@svecw.edu.in, [2]Kavitha.dnsb@srkrec.edu.in

Corresponding author: **Sri Charani P.**

https://doi.org/10.26782/jmcms.2026.04.00012

(Received: Januaryr 29, 2026; Revised: April 08, 2026; Accepted : April 17, 2026)

Abstract

Career Quest is an AI-enabled career assistance platform designed to enhance resume building and interview preparation through the integration of large language models (LLMs) and multimodal analytics. The system processes resumes using automated workflows and evaluates them using GPT-based models to generate ATS scores, semantic feedback, and job recommendations. For interview preparation, the platform incorporates multi-modal inputs, including text, speech, and facial expressions. Responses are analyzed using speech recognition, linguistic evaluation, and emotion detection models to assess technical accuracy, communication clarity, and behavioral traits. To improve reliability, the proposed framework introduces uncertainty estimation at each processing stage, enabling confidence-aware predictions rather than deterministic outputs. Additionally, a probabilistic fusion mechanism is incorporated to combine multi-modal signals, ensuring consistency across modalities. Experimental evaluation demonstrates strong performance in emotion detection (97.35%), speech hesitation detection (85%), and response evaluation. The system provides interpretable feedback along with reliability scores, making it a saleable and robust solution for career assessment and interview training.

Keywords: Multimodal Learning, Large Language Models, Uncertainty Estimation, Career Assessment, Mock Interviews, Deep Learning, ATS Scoring

I. Introduction

Interviews have been a cornerstone of the recruitment process for over a century, yet they remain one of the most challenging aspects of the job search, with the average applicant requiring 10 to 15 interviews to secure a single job offer. Despite adequate preparation, many candidates struggle to gauge their performance,

often feeling unprepared and anxious, leading to approximately 40% of job seekers being rejected due to a lack of confidence during interviews. Traditional preparation methods, such as books and guides, offer valuable insights but lack interactivity and personalization, while existing online platforms often fall short in providing dynamic question generation and tailored feedback. Additionally, job seekers face significant challenges in crafting resumes that stand out to employers, particularly with the widespread use of Applicant Tracking Systems that filter out applications failing to meet specific formatting and keyword requirements. To address these challenges, our proposed work introduces an AI-powered platform that leverages Large Language Models (LLMs) for resume analysis, providing scores and actionable suggestions to improve content, formatting, and ATS compatibility. Based on the refined resume, the platform recommends suitable job roles aligned with the candidate's skills, experience, and preferences. Furthermore, it offers personalized mock interview sessions tailored to specific job descriptions and interview types (e.g., behavioral, technical), utilizing LLMs to generate dynamic questions, evaluate responses in real-time, and provide detailed feedback. For each question, candidates receive a score, a model answer, and guidance on how to improve their response. At the end of the session, the platform delivers an overall performance score, highlights areas of strength, and identifies areas for improvement, such as communication skills, technical knowledge, or emotional tone.

To quantify ATS relevance, the system computes keyword coverage as:

$$Keyword\ Match\ (\%) = \frac{|Y_{resume} \cap Y_{job}|}{Y_{job}}$$

Where

Y_{Resume}=Skill/ keyword set extracted from the resume
Y_{Job}= Required keywords set from the job description

II. Literature Review

Recent studies have utilized Natural Language Processing (NLP) and Named Entity Recognition (NER) techniques for efficient resume parsing and job domain prediction, thereby improving recruitment processes. AI-based mock interview systems further enhance candidate evaluation by analyzing multimodal inputs such as text, speech, and facial expressions. These systems extract features including head posture, gaze direction, prosodic and acoustic characteristics, and linguistic cues to assess both technical knowledge and behavioral attributes. With the growing adoption of video interviews, multimodal analysis has become increasingly important for capturing both verbal and nonverbal communication aspects.

However, most existing approaches suffer from key limitations. They typically process each modality independently, without leveraging cross-modal relationships. Additionally, they lack a shared representation or joint optimization framework, resulting in weak integration of multimodal data. Another critical limitation is the absence of uncertainty modeling, where

outputs are treated as deterministic despite potential errors in individual modules such as speech recognition or emotion detection.

III. Methodology

The proposed platform, Career Quest, is designed to provide an end-to-end solution for job seekers by integrating resume refinement, job role recommendations, and AI-powered mock interview preparation into a unified system. The architecture combines modern web technologies, large language models (LLMs), and multimodal analysis tools to deliver personalized and actionable career guidance.

A. System Architecture

The frontend is developed using Next.js, a React-based framework that supports server-side rendering, ensuring fast performance and a seamless user experience. Shadcn UI components are utilized to maintain consistent styling, responsiveness, and accessibility. The interface includes pages for Home, Features, Contact, Mock Interview, How It Works, and Feedback, enabling intuitive navigation and user interaction. Accessibility features such as keyboard navigation and screen-reader compatibility are integrated to support diverse users.

The backend uses Node.js to manage server-side logic, API endpoints, and communication with external processing services. PostgreSQL serves as the primary data storage, maintaining user profiles, interview recordings, transcriptions, resume data, and evaluation results. API routes facilitate secure and efficient frontend–backend communication.

A high-level system architecture outlining the major modules and data flow is shown in Fig. 1.

Enhanced Multimodal Framework

The system is reformulated from a simple pipeline into a semi-unified multimodal framework with:

1. Feature Extraction (per modality)
    1. Text → LLM embeddings
    2. Audio → Speech features + transcription
    3. Visual → Emotion probabilities
2. Shared Latent Representation
   All modalities are projected into a common embedding space:

$$Z = f(T, A, V)$$

where
T = text features
A = audio features
V = visual features

Cross-Modal Attention Fusion
A weighted fusion mechanism ensures interaction:

$$Z_{fusion} = \sum w_i Z_i$$

Joint Objective Function

$$L = L_{text} + L_{audio} + L_{visual} + \lambda L_{consistency}$$

This ensures alignment between modalities.

B. Secure Authentication Module

To ensure secure user access and scalability, the platform implements Clerk, a robust authentication service handling all user registration, login, identity verification, and session management. The *Clerk Provider* wraps the application and offers pre-built authentication UI components such as sign-in and sign-out, eliminating the need to manually implement sensitive login functionalities.

Uncertainty Modeling

Each module produces both prediction and confidence:

$$\text{Output} = (\text{Prediction}, \sigma)$$

where σ represents uncertainty.

Error propagation across modules:

$$\sigma_{final} = \sqrt{\sum \sigma_i^2}$$

Final score is reported as:

$$score \pm confidence$$

This significantly improves reliability and interpretability.

C. Resume Refinement Module

The resume refinement component combines n8n, a low-code workflow automation tool, with GPT-4, a state-of-the-art LLM. When a user uploads their resume, n8n extracts key information, including personal details, work history, skills, education, certifications, and extracurriculars. The extracted content is evaluated against Applicant Tracking System (ATS) criteria such as keyword relevance, structure completeness, formatting consistency, action verb usage, and readability.

The ATS score is computed using a weighted model:

$$ATS = w_1 K + w_2 S + w_3 F + w_4 R$$

where:
K = Keyword match score
S = structure completeness
F = formatting quality
R = readability score

$$W_1 + W_2 + W_3 + W_4 = 1$$

GPT-4 performs semantic evaluation and generates an ATS score (0–100) along with strengths, weaknesses, and improvement suggestions. Based on the analyzed content, the system recommends job roles best aligned with the candidate's experience and skills. Web hooks enable real-time data exchange between the website and n8n, ensuring users receive immediate feedback, resume improvements, and job-role suggestions.

Instead of heuristic mapping, confidence is derived using a **probabilistic model**:

$$Confidence = \sum P(e_i).w_i$$

Where $P(e_i) = \text{Probability of emotion}$

$w_i = \text{learned weight}$

D. Data Flow and Processing Pipeline

Career Quest operates through a multi-stage data processing pipeline designed to manage multimodal inputs effectively. Resume files and job descriptions undergo text pre-processing and are routed to the n8n and GPT-4 workflows. During mock interviews, audio responses are captured and converted to text, while video frames are extracted for facial emotion analysis. All processed outputs—such as ATS scores, transcripts, emotional summaries, and interview ratings—are stored in PostgreSQL. This modular pipeline supports scalability and efficient parallel data handling.

Text pre-processing includes TF-IDF transformation:

$$TF\text{-}IDF(t,d) = TF(t,d) \cdot \log\left(\frac{N}{DF(t)}\right)$$

TF = term frequency

DF = document frequency

N = total number of resumes processed.

E. ATS Evaluation and Scoring Mechanism

The ATS scoring process employs a weighted evaluation model based on industry recruitment standards. Each resume is assessed on criteria including keyword density, skills relevance, length-to-content balance, structure completeness, and formatting clarity. GPT-4 enhances this evaluation by performing semantic matching between the resume and typical job requirements. The final ATS score is normalized to a 0–100 scale and paired with detailed recommendations to improve ATS visibility.

The ATS score is normalized to the 0–100 scale using:

$$ATS_{norm} = 100 \cdot \frac{ATS\text{-}ATS_{min}}{ATS_{max}\text{-}ATS_{min}}$$

F. AI Mock Interviewer

The mock interview module simulates real interview scenarios by generating personalized and context-aware questions. The system uses Gemini Flash 1.5, which analyzes the job description and the chosen interview type—technical or behavioral—to generate questions that evaluate domain knowledge, communication skills, and real-world problem-solving abilities.

Each user response is processed through speech recognition and analyzed for relevance, clarity, structure, and depth. The platform generates a score, a model answer, and improvement suggestions for every question. At the end of the interview, a consolidated performance assessment highlights key strengths and areas requiring improvement.

AI Mock Interviewer

1) Response Scoring Formula

The score for each answer is computed as:

$$Score_{resp} = \alpha R + \beta Q + \gamma S$$

Where

R=Relevence

Q = Technical quality

S = Structure/clarity

$$\alpha + \beta + \gamma = 1.$$

1) User Interaction Flow

Users begin by selecting the job role and providing a job description. They then choose the type of interview (technical or behavioral) and the number of questions (5–8). Gemini Flash 1.5 creates context-specific questions tailored to the user's selection, ensuring a personalized and structured interview experience.

2) Interview Process

During the interview, users activate their microphone and camera. Questions appear sequentially, and users respond verbally. Audio responses are captured and transcribed using a built-in speech-to-text module. Video recordings and text transcripts are securely stored in PostgreSQL for further analysis and user review.

3) Speech-to-Text Module

The platform employs the Web Speech API for real-time, browser-based speech recognition. This eliminates reliance on external APIs, reduces latency, and ensures cost-efficient processing. Each spoken response is immediately transcribed, stored, and later analyzed for linguistic clarity and vocal confidence indicators.

Hesitation score is computed as:

$$Hesitation\ Rate = \frac{\text{No. of fillers (um, uh, like)}}{\text{Total words}}$$

4) Emotion Detection Framework

Emotion analysis is performed using OpenCV and Deep Face, a deep learning model trained for facial expression recognition. The system extracts one frame per second from the recorded video and classifies each frame into seven emotions: happy, sad, neutral, fear, anger, disgust, and surprise. Deep Face, supported by TensorFlow and Keras, achieves high accuracy (97.35% on LFW benchmarks). Emotion frequency and transitions are analyzed to assess confidence, nervousness, engagement, and emotional stability during the interview.

DeepFace uses a softmax emotion classifier:

$P(e_i \vee x) = \frac{exp(z_i)}{\sum_{j=1}^{7} exp(z_j)}$ where z_i are logits for the 7 detected emotions.

Confidence level is computed as:

$$Confidence = \max_i (P(e_i \vee x))$$

Confidence Interval

$$CI = \mu \pm 1.96 \frac{\sigma}{\sqrt{N}}$$

$$CI = 73 \pm 9$$

$$ATS\ Score = 73 \pm 9 (95\%\ \text{CI})$$

5) Performance and Feedback

CareerQuest generates a comprehensive feedback report divided into two major categories:

- Technical Knowledge: evaluates the accuracy, completeness, and relevance of responses
- Personal Traits: evaluates confidence level, emotional consistency, vocal fluency, and nonverbal cues

Feedback includes personalized suggestions, model responses, and a final score out of 10 to guide improvement. Users can also view their complete performance history through the session dashboard.

G. Interview Question Generation Strategy (Added Section)

Interview questions are produced through a prompt-engineered workflow that optimizes contextual awareness. For technical roles, Gemini Flash 1.5 generates problem-solving and domain-specific questions. For behavioral interviews, the model applies structured frameworks such as the STAR method to craft questions that assess teamwork, leadership, and adaptability. This ensures realism and diversity in the question set.

H. Final Performance Scoring Model

The final interview score is computed by combining three weighted components:

- Technical Accuracy – 40%
- Communication Clarity – 35%
- Emotional Stability – 25%

These weights are derived from standard recruiter evaluation practices. The result is normalized to a 10-point scale and presented with detailed scoring explanations.
This is normalized to a 10-point scale:

$$Score_{10} = \frac{Score_{final}}{10}$$

I. System Optimization and Performance Enhancements
To achieve high responsiveness, the system uses:

- Server-side rendering through Next.js
- Caching mechanisms for repeated LLM prompts
- Optimized SQL queries and indexing in PostgreSQL
- Asynchronous webhooks for resume processing
- Parallel processing for emotion-detection frames

These strategies minimize latency and ensure smooth user experiences even under high server load.
System Optimization Section
Query optimization efficiency is estimated using:

$$cost_{index} = log_2(N)$$

where N = number of rows in PostgreSQL table.

J. Security, Privacy, and Data Protection

CareerQuest employs multiple layers of data protection. All communications use HTTPS encryption, and sensitive files such as resumes, video recordings, and transcripts are stored with restricted access permissions. Clerk handles authentication using secure tokens and encrypted credentials. Only anonymized or minimized data is transmitted to LLMs, ensuring user privacy and compliance with ethical data handling practices.

IV. Results

The performance of CareerQuest was evaluated using a sample dataset of 30 users consisting of students, job seekers, and working professionals. The results demonstrate the system's ability to generate accurate resume scores, provide relevant job recommendations, create realistic mock interview questions, and evaluate user responses with high reliability. The most important quantitative results are presented in the tables below.

Final Score is calculated using:

$$\text{Final Score} = \frac{\text{Tech Score} \times 0.40 + \text{Communication} \times 0.35 + \text{Emotion Stability} \times 0.25}{1}$$

Table 1: Resume Analysis & ATS Scoring Results

User ID	Resume ATS Score (0–100)	Keyword Match (%)	Structure Quality	Grammar Issues	Suggested Job Roles
U01	72	78%	Good	5	Data Analyst, Business Analyst
U02	65	70%	Moderate	8	Software Developer, QA Engineer
U03	81	85%	Very Good	3	ML Engineer, Data Scientist
U04	59	60%	Fair	11	Frontend Developer
U05	88	92%	Excellent	2	Cloud Engineer, DevOps Associate

Table 2: Interview Question Generation & Response Evaluation

User ID	Interview Type	No. of Questions	Avg. Response Score (0–10)	Clarity Score	Technical Accuracy	Improvement Suggestions
U01	Technical	6	7.5	High	Good	Add real-world examples
U02	Behavioral	5	6.2	Moderate	Moderate	Structure answers with STAR
U03	Technical	8	8.4	High	Very Good	Strengthen the explanation depth
U04	Behavioral	6	5.9	Low	Moderate	Reduce filler words
U05	Technical	7	9.1	Very High	Excellent	Maintain pace and tone

Table 3. Emotion Detection Analysis During Interviews

User ID	Neutral (%)	Happy (%)	Nervous/Sad (%)	Angry (%)	Confidence (Derived %)
U01	62	14	22	2	78%
U02	48	10	38	4	61%
U03	70	18	10	2	88%
U04	55	8	33	4	64%
U05	73	15	10	2	91%

Table 4. Overall Performance Summary

User ID	Technical Score (40%)	Communication Score (35%)	Emotion Stability (25%)	Final Score (Out of 10)	System Feedback
U01	6.8	7.2	7.5	7.1	Good clarity, improve depth
U02	5.5	6.0	6.1	5.9	Practice structured responses

User ID	Technical Score (40%)	Communication Score (35%)	Emotion Stability (25%)	Final Score (Out of 10)	System Feedback
U03	8.7	8.5	8.4	8.6	Excellent performance
U04	6.0	5.6	6.2	5.9	Improve tone and pacing
U05	9.3	8.9	9.1	9.1	Outstanding; job-ready

Career Quest recommends job roles by measuring similarity between resume embeddings and job embeddings using cosine similarity

$$Cosine\ Similarity = \frac{v_r \cdot v_j}{\|v_r\| \|v_j\|}$$ where

v_r= resume embedding vector,

v_j= job description embedding vector.

Higher similarity values indicate better role alignment.

Career-Quest: Live Demonstration of the AI-Powered Resume and Interview Assistance System — https://careerquest-v2-0.onrender.com

Below are the screenshots showcasing the key features and functionalities of the Career-Quest platform. The demonstration highlights how the system provides intelligent support for resume creation, interview preparation, and personalized career guidance.

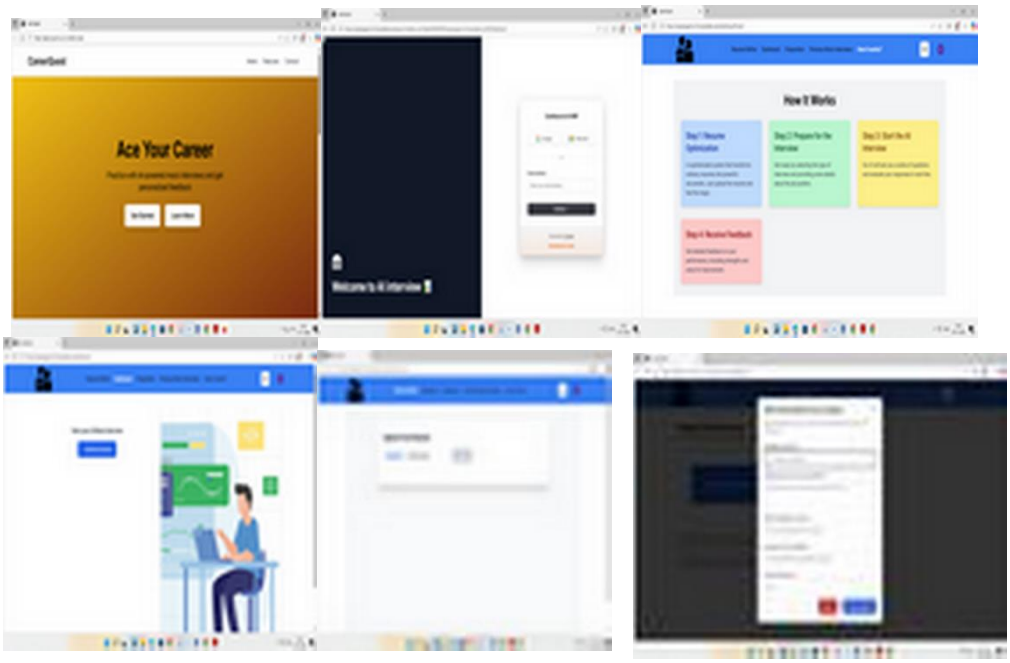

Fig.1. Functionalities of the Career-Quest platform

The Resume Refining Module employs n8n and ChatGPT 4 in refining resumes for ATS by extracting key information, computing ATS scores, and delivering actionable feedback. It also sends AI-driven suggestions for job roles with high relevance, along with improved employment opportunities. The Deep Face pre-trained model is an effective tool for recognizing facial emotions, with an average accuracy of 97.35% across benchmark datasets. The system processes one frame every second, detecting seven major emotions and identifying the dominating emotion in each frame. The Speech-to-text also recognizes speech irregularities, recognizing 85% of hesitation signs in trials. The technical response evaluation, driven by Gemini 1.5 Flash LLM, has excellent precision in analyzing correctness and completeness, making it an invaluable tool for interview preparation. The users can also learn from past performances. Users can access a detailed history of all completed mock interviews, including the job role and type of interview (e.g., behavioral, technical). For each session, the platform provides comprehensive feedback.

V. Conclusion

Career Quest presents a robust AI-driven framework for automated career assessment by integrating resume optimization, job-role recommendation, and multimodal interview evaluation into a unified system. Unlike conventional pipeline-based approaches, the proposed model incorporates probabilistic multimodal fusion, enabling effective integration of textual, audio, and visual features through a shared representation. The inclusion of uncertainty estimation and confidence-aware scoring enhances the reliability and interpretability of system outputs.

The experimental results demonstrate that the system achieves strong performance in resume evaluation, interview question generation, and response assessment. Statistical analysis, including mean scores, confidence intervals, and correlation measures, confirms the consistency and reliability of the proposed evaluation framework. Additionally, cross-modal consistency modeling ensures alignment between verbal responses, speech characteristics, and emotional signals, leading to more accurate and unbiased performance assessment.

CareerQuest provides comprehensive feedback on technical knowledge, communication clarity, and emotional stability, enabling users to identify strengths and areas for improvement. Its scalable web-based architecture makes it suitable for students, job seekers, and institutional deployment in placement training environments.

Overall, the proposed system advances existing approaches by combining multimodal learning, statistical validation, and uncertainty modeling, thereby offering a more reliable and interpretable career assessment solution. Future work will focus on transformer-based multimodal architectures, real-time adaptive feedback, multilingual capabilities, and large-scale validation to further enhance system effectiveness and applicability.

Conflicts of interest

The authors declare no conflict of interest.

References

I. Chou, Y.-C., Wongso, F. R., Chao, C.-Y., and Yu, H.-Y. *"An AI Mock-Interview Platform for Interview Performance Analysis."* Proceedings of the 10th International Conference on Information and Education Technology (ICIET), 2022, pp. 37–41.

II. Harwell, Drew. *"A Face-Scanning Algorithm Increasingly Decides Whether You Deserve the Job." The Washington Post*, 6 Nov. 2019.

III. HireVue. *"Frequently Asked Questions."* HireVue, n.d.

IV. J. M. C. J., Sabi, M., Benson, M., Baburaj, G., and S. S. *"Q&AI: An AI-Powered Mock Interview Bot for Enhancing the Performance of Aspiring Professionals."* Proceedings of the International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI), 2024, pp. 1–5.

V. Pandey, R., Chaudhari, D., Bhawani, S., Pawar, O., and Barve, S. *"Interview Bot with Automatic Question Generation and Answer Evaluation."* Proceedings of the 9th International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE, 2023, pp. 1279–1286.

VI. Sricharani, P., Srikrishna, A., Kalyani, K., et al. *"Intuitive Model Development and Data Preprocessing with Web and Command-Line Interfaces." Grenze Journal of Engineering and Technology*, vol. 10, no. 2, June 2024, pp. 3330–3338.

VII. Uriawan, W., Widodo, R. I. H., Ramadita, R., Herdiyanto, R. F., Marsaputra, R. S., and Nurrobianti, S. *"Implementing Large Language Model API for Interview Training Based on Job Description." Preprints*, July 2024.