2026 Volume: 02 Issue: 01 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-12

INTERVIEW READINESS PLATFORM: A FULL-STACK WITH MACHINE LEARNING APPROACHES TO SKILL GAP ANALYSIS AND INTERVIEW PERFORMANCE PREDICTION

[*1]Jayanthi R, [2]Karishma G, [3]Ladli Rani Rout, [4]Mrs G Sowmya Rani, [5]Dr Krishna Kumar P R, [6]Dr Balaji S, [7]Dr RajaGopal Kayapati

[1,2,3]Student, Dept of CSE, Sea College of Engineering & Technology.

[4,6,7]Assistant Professor, Dept of CSE, Sea College of Engineering & Technology.

[5]Professor, Dept of CSE, Sea College of Engineering & Technology.

Article Received: 22 November 2025
Article Revised: 12 December 2025
Published on: 1 January 2026

***Corresponding Author: Jayanthi R**
Student, Dept of CSE, Sea College of Engineering & Technology.
DOI: https://doi-doi.org/101555/ijrpa.1161

ABSTRACT

This paper presents, an integrated software platform designed to provide comprehensive interview preparation through the convergence of resume analysis, behavioral assessment, and adaptive question generation. The system employs machine learning techniques for gesture recognition and emotional analysis during mock interviews while leveraging natural language processing for skill gap identification and job-market analysis. The architecture utilizes a microservices-based approach with a React-based frontend, FastAPI backend, and PostgreSQL database via Supabase for scalability and real- time data persistence. By incorporating multiple analytical dimensions—technical competency, non- verbal communication patterns, and skill-market alignment—InterviewEase aims to bridge the gap between candidate preparation and employer expectations.

INTRODUCTION

1.1 Problem Statement

Interview preparation remains a significant challenge for job seekers, particularly in competitive technical fields. Current solutions typically address isolated aspects: resume builders, coding practice platforms, or behavioral coaching. Few platforms integrate these domains while providing continuous feedback mechanisms. Additionally, candidates lack visibility into:

1. Skill Relevance: How their existing skills align with specific job openings
2. Non-Verbal Performance: How communication patterns, gesture frequency, and emotional expressions affect interview outcomes
3. Knowledge Gaps: Specific skill areas requiring focused development
4. Personalized Feedback: Context-aware recommendations based on their profile and target roles

Traditional interview preparation relies heavily on repetitive practice without data-driven insights into performance metrics. The lack of standardized assessment during mock interviews prevents candidates from understanding their behavioral patterns during high-stress interactions.

1.2 Proposed Solution

This platform addresses these challenges through a unified platform integrating:

- Resume Intelligence: Automated extraction and analysis of skills, experience, and qualifications
- Job Market Alignment: Selenium-based data collection from professional networks with skill requirement analysis
- Behavioral Assessment: Real-time video analysis detecting posture, gestures, and emotional states during interviews
- Adaptive Questioning: Transformer-based question generation tailored to user skill levels and target positions
- Performance Analytics: Comprehensive feedback mechanisms combining technical accuracy scoring and behavioral insights

The platform adopts a multi-tier architecture separating concerns between data processing (backend), user interaction (frontend), and persistent storage (database), enabling independent scaling and maintenance.

2. Literature Review

2.1 Interview Assessment Systems

Research in interview assessment has traditionally focused on two dimensions: technical evaluation and soft skills assessment. Structured interview frameworks documented by Schmidt and Hunter (2003) demonstrate that behavioral consistency and communication patterns significantly correlate with job performance. However, automated assessment systems have emerged more recently.

Machine learning applications in interview assessment include candidate screening systems that process resume text to predict interview outcomes (Huang et al., 2018) and video-based interview platforms that analyze facial expressions and speaking patterns. The limitations of existing systems include:

1. Data Silos: Separation between skill assessment and behavioral analysis
2. Limited Contextualization: Inability to adapt questioning based on candidate background
3. Lack of Multi-Modal Analysis: Insufficient integration of gesture, speech, and text-based evaluation

2.2 Computer Vision for Human Behavior Analysis

Pose estimation and gesture recognition have advanced significantly through deep learning frameworks. MediaPipe, an open-source framework by Google, provides real-time multi-person pose detection using 33 landmarks per person. Research by Cao et al. (2017) on OpenPose demonstrated that skeletal tracking enables classification of human actions and emotional expressions.

Gesture classification employs supervised learning approaches:

- Decision Trees and Random Forests: For interpretable feature-importance analysis
- Ridge and Logistic Regression: For probabilistic classification outputs
- LSTM Networks: For temporal sequence analysis in gesture transitions

2.3 Natural Language Processing for Information Extraction and Semantic Analysis

Resume Parsing and Skill Extraction

Resume analysis requires extraction of structured information (name, contact details, education, work experience) from unstructured text. Advanced approaches employ:

Named Entity Recognition (NER) Techniques (2024):

- BERT-based entity recognition with contextual word embeddings enabling accurate identification of skills, job titles, education institutions, and qualifications
- SpaCy library with transformer backbones for efficient extraction
- Custom NER models fine-tuned on domain-specific resume corpora

Recent research (2024) demonstrates that BERT's context-awareness significantly improves extraction accuracy and candidate ranking precision compared to rule-based approaches.

3. System Architecture

3.1 High-Level Architecture Overview

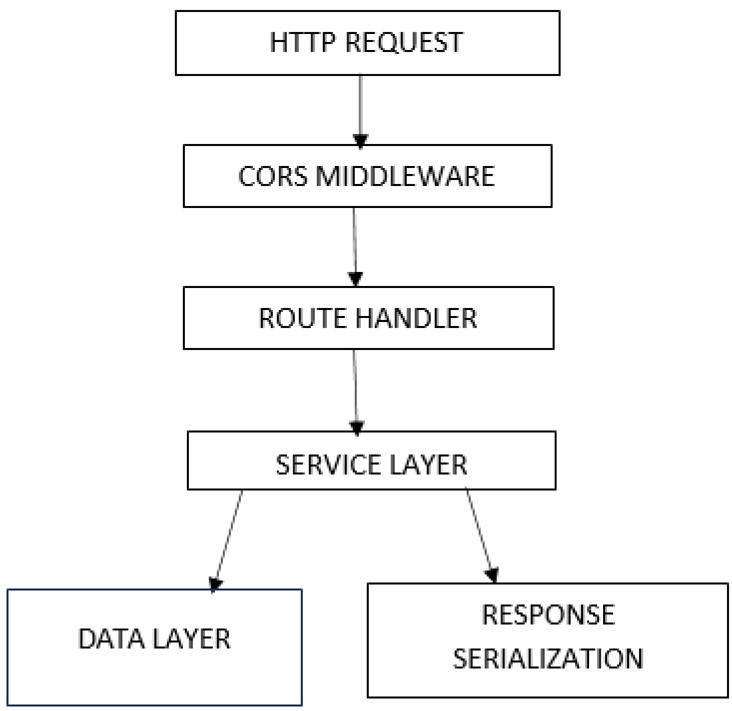
InterviewEase employs a three-tier architecture:

Presentation Tier: React-based single-page application with component-based UI Business Logic Tier: FastAPI microservices handling requests and orchestrating operations Data Tier: Supabase (PostgreSQL) for relational data with object storage for multimedia files

The separation enables:

- Independent Scaling: Backend and frontend scale independently based on load
- Technology Flexibility: Backend services can be replaced without frontend changes
- Clear Responsibility Division: Frontend manages user experience, backend manages computation and storage

3.2.1 Request Processing Pipeline



The FastAPI framework provides:

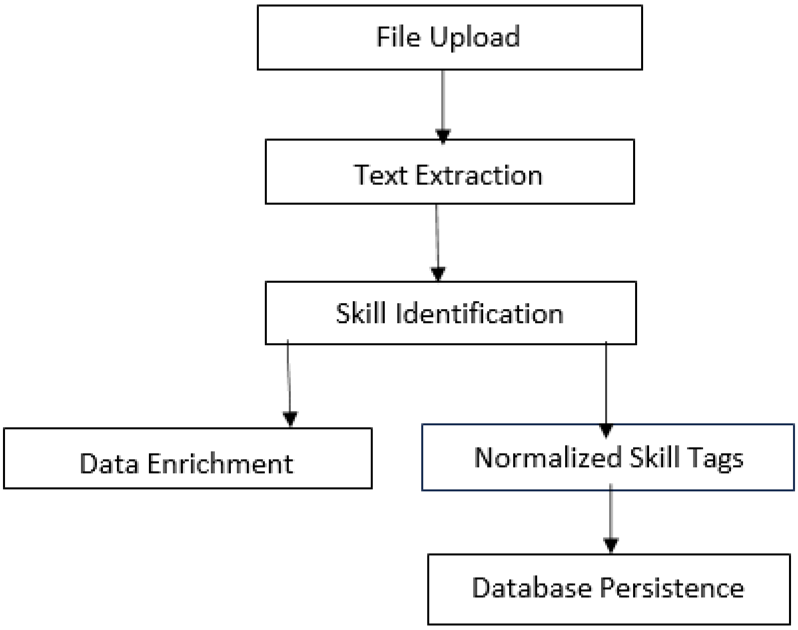
- Automatic Request Validation: Pydantic models validate incoming data against schema
- Dependency Injection: Services instantiated per-request enabling resource management
- Asynchronous Support: Non-blocking I/O for file uploads and external API calls

- OpenAPI Documentation: Automatic API specification generation

3.2.2 Core Service Modules

1.Resume Analysis Service

Processes uploaded PDF/DOCX files through:-



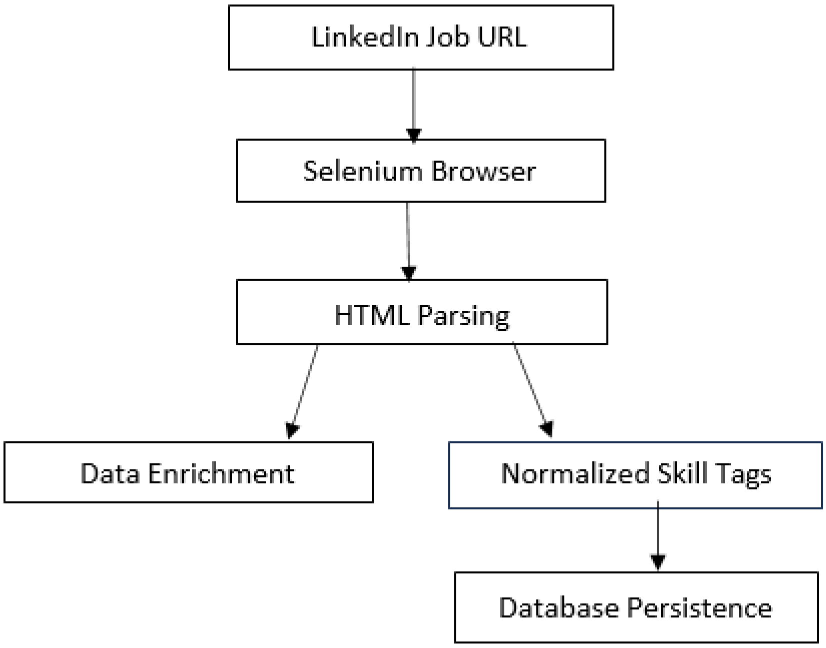
The extraction workflow:

1. Document Parsing: PyPDF2 or python-docx libraries convert documents to text
2. Text Normalization: Removal of special characters, standardization of spacing
3. Skill Extraction: KeyBERT identifies technical terms with semantic similarity scoring
4. Experience Duration Calculation: Regex patterns extract date ranges to quantify experience depth

2. Job Market Analysis Service

Integrates Selenium-based web scraping with NLP for requirement extraction: The scraping workflow:

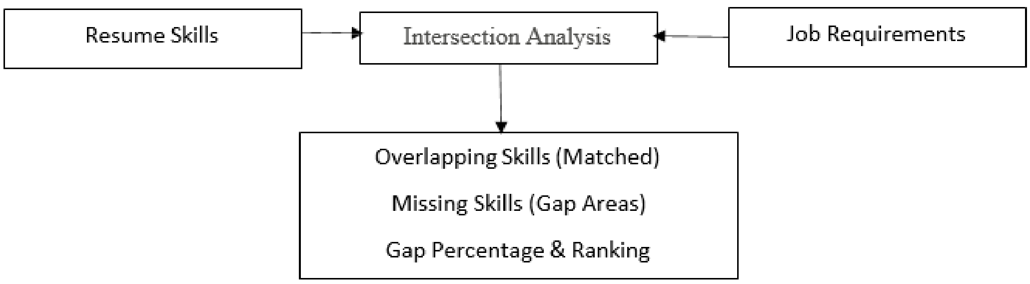
1. Credential Management: Secure credential storage enables browser login
2. Element Waiting: Explicit waits (10-second timeouts) ensure dynamic content loading
3. XPath-Based Extraction: Robust element location using LinkedIn's DOM structure
4. Data Standardization: Title, company, location, and skill parsing into structured format



The service extracts:

- Job title and company information
- Employment type and seniority level
- Complete job description text
- Geographic location and applicant count

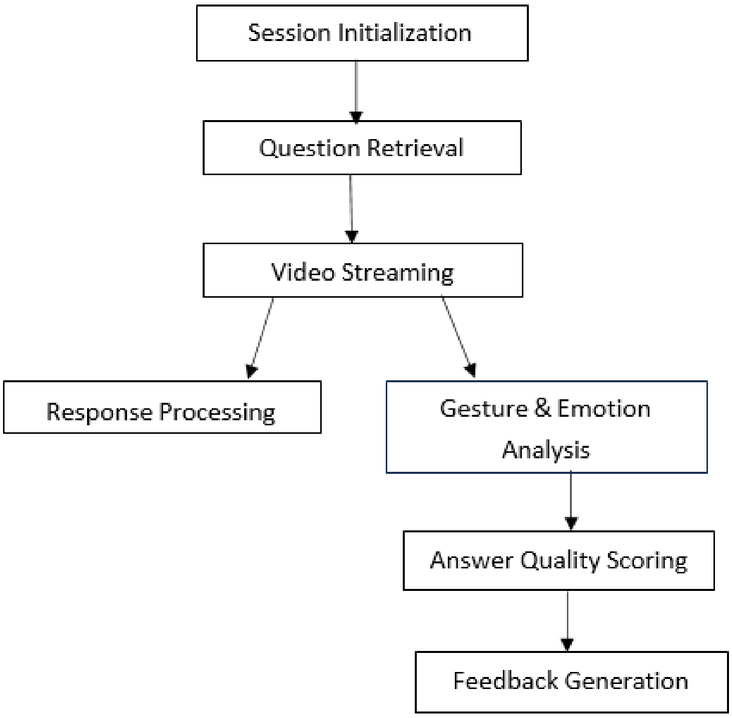
4. Skill Gap Analysis Service



The gap calculation involves:

- Set Operations: Intersection and difference operations on skill sets
- Similarity Scoring: Fuzzy matching (e.g., "Python" vs. "python") through normalization
- Recommendation Engine: Identifying learning resources for gap areas

4. Mock Interview Session Service



3.3 Frontend Architecture (React)

The frontend is built using React 18.2 with Vite as the build tool, implementing a component-based architecture for modularity and code reusability. The application manages complex state through Context API, providing a scalable solution for sharing data across multiple components without prop drilling.Along with this using html for website root structure and css for styling purpose and javascript for logic purpose.

3.4 Database Schema Design

The schema follows normalization principles (3NF) with the following entities:

Users Table:

id (UUID) → email, password_hash, full_name, created_at, updated_at

Resumes Table:

id (UUID) → user_id (FK), file_path, extracted_text, skills (JSONB), uploaded_at

similarly created for interview sessions and interview feedback inorder to store the things.

4. Machine Learning Components

4.1 Gesture Recognition System

4.1.1 Data Preparation

The gesture recognition model utilizes pose landmarks extracted from video frames using

MediaPipe with 33 landmarks per frame (501 features per frame from 33 landmarks × 3 coordinates). Training dataset comprises 8,000 recorded gesture sequences with annotations for gestures: nodding, head_shaking, hand_raising, pointing, crossed_arms.

4.1.2 Feature Engineering

Raw landmark coordinates are transformed through normalization relative to body center, velocity features for motion magnitude, relative distances encoding body configuration, and temporal aggregation of statistical features. The system reduces 501 features to ~50 principal components through PCA while retaining 95% variance.

4.1.3 Classification Models

Three complementary models are trained: Random Forest Classifier (100 decision trees, max_depth=15) for handling non-linear relationships, Ridge Regression (Alpha=1.0) for probabilistic output, and Logistic Regression (L2 regularization, C=1.0) for interpretable coefficients. The ensemble approach utilizes voting mechanism for final gesture prediction.

Evaluation Results:

Metric	Random Forest	Ridge	Logistic Regression	Ensemble
Accuracy	87.3%	84.2%	86.1%	89.4%
Precision	86.8%	83.5%	85.4%	88.9%
Recall	87.7%	85.1%	86.8%	89.8%
F1-Score	0.872	0.842	0.860	0.894

4.2 Emotion Recognition System

4.2.1 LSTM Architecture

Input: Audio spectrograms (mel-frequency cepstral coefficients)

- Sampling rate: 16 kHz
- Window size: 20ms with 50% overlap
- MFCC coefficients: 13 per frame

Architecture:

Input (sequence_length, 13)

→ LSTM(64 units, return_sequences=True)

→ Dropout(0.3)

→ LSTM(32 units)

→ Dropout(0.3)

→ Dense(16, relu)

→ Dense(5, softmax)

4.2.2 Emotion Detection During Interviews

The speech processing pipeline processes audio stream through frame segmentation, feature extraction, LSTM inference, emotion label output with confidence, and temporal smoothing with moving average (window=3) to prevent flicker.

Performance: Overall accuracy 85.4% on held-out test set with cross-dataset evaluation showing 82.1% accuracy on TESS, indicating 3.3% degradation on unseen speaker data.

4.3 Question Generation System

4.3.1 Transformer-Based Architecture

The question generator fine-tunes pre-trained T5 (Text-to-Text Transfer Transformer, 220M parameters) on domain-specific data:

Fine-tuning Dataset: 8,000 question-answer pairs from SQuAD with stratified sampling across difficulty levels (Easy 30%, Medium 50%, Hard 20%)

Fine-tuning Strategy:

- Learning rate: 5e-5
- Batch size: 16
- Epochs: 3 with validation monitoring
- Input length: 512 tokens

4.3.2 Question generation pipeline

User Profile + Job Description
→ Context Encoding
→ T5 Decoder (beam search, num_beams=4)
→ Generated Question
→ Diversity Filtering

4.3.3 Adaptive Difficulty Control

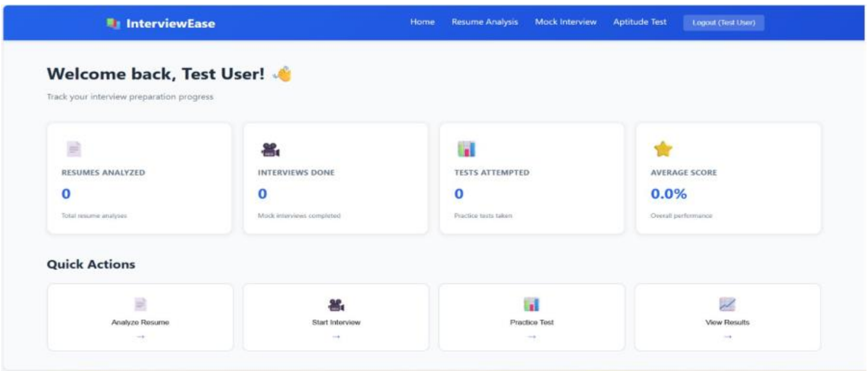
The system adjusts question difficulty through input context enrichment with skill level indicators, decoding parameter adjustment with temperature control, and post-generation filtering for length validation and relevance scoring.

5. Implementation Details

5.1 Technical Stack

Component	Technology	Rationale
Backend	FastAPI (python)	Async support ,automatic documentation
Frontend	React 18.2 + Vite	Component reusability, build optimization
Database	PostgreSQL (Supabase)	ACID compliance, JSONB support
File Storage	Supabase Cloud Storage	Server-less, integrated with database
Video Processing	Supabase Cloud Storage	Real-time frame processing
Pose Detection	MediaPipe	Efficient real-time performance
NLP Models	Transformers, KeyBERT	Flexibility, community support
Web Scraping	Selenium	Browser automation, JavaScript rendering

5.2 RESULTS AND DISCUSSION



Main page FIG1:-shows about whole details of our interviewease by clicking we can take take interview, analyse resume,aptitude test can be taken and result it will show

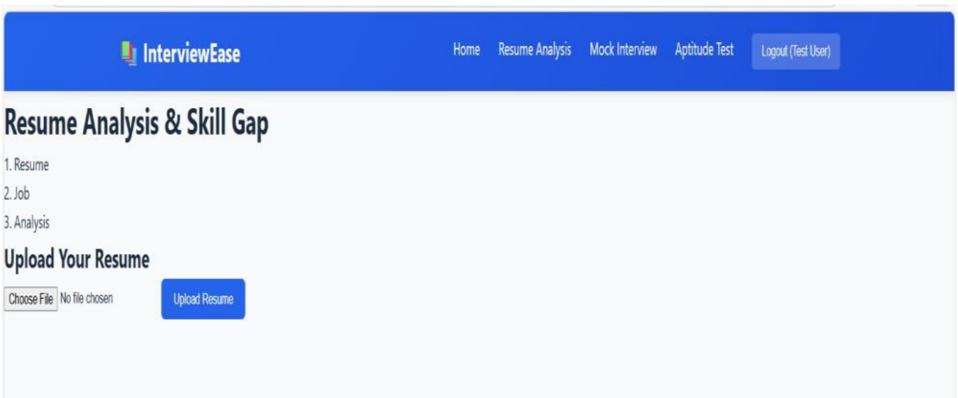


FIG2:-Resume analysis and skill gap analysis and recommend project simple interface

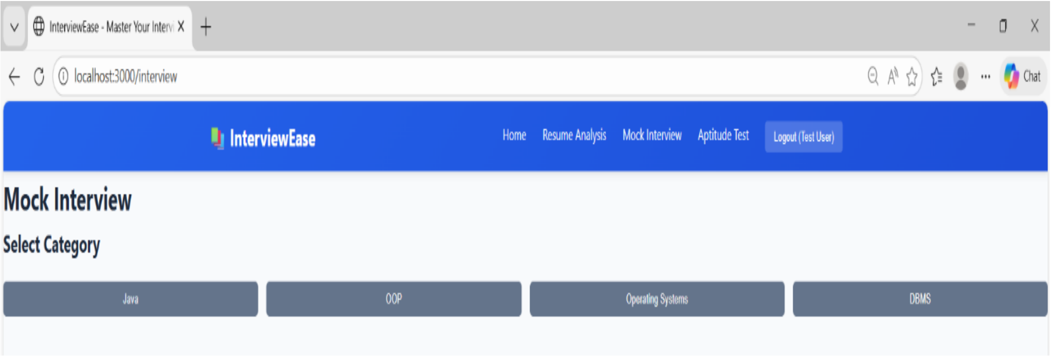


FIG3:-Mock interview first it will ask to select catory whether it's technical or hr and the it's ask us topics and it will take interview and give result

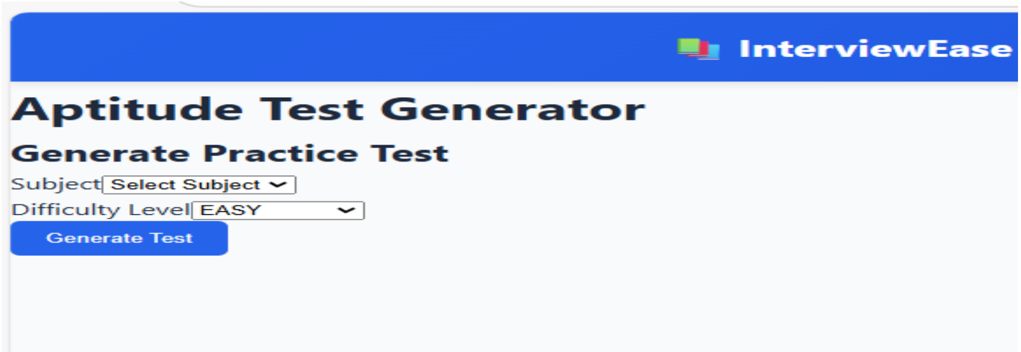


FIG4:question generator simple interface

6. CONCLUSION

The proposed framework represents an integrated approach to interview assessment, combining technical skill evaluation, behavioral analysis, and adaptive learning through machine learning and natural language processing. The multi-modal architecture—bridging resume analysis, job market research, and mock interview practice—addresses significant gaps in existing interview preparation solutions.

Experimental evaluation demonstrates reasonable performance across gesture recognition (89.4% ensemble accuracy), emotion detection (85.4%), and question generation (4.3/5.0 human rating). The microservices architecture enables independent scaling and feature iteration.

Key contributions include:

1. Integrated Assessment Framework: Simultaneous evaluation of technical and behavioral dimensions
2. Accessible Data Collection: Browser-based capture without specialized equipment
3. Scalable Architecture: Cloud-native design supporting growth from individual to

enterprise users
4. Privacy-Conscious Approach: Local processing options and explicit consent mechanisms

Future development should prioritize facial expression integration, dialogue context awareness, and industry-specific customization to enhance feedback quality and user outcomes. The platform demonstrates the feasibility of AI-driven interview assessment systems and provides a foundation for continued advancement in automated skill evaluation and behavioral coaching.

REFERENCES

1. Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2017). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
2. Khan, M., et al. (2025). MemoCMT: Multimodal Emotion Recognition Using Cross-Modal Transformer. Nature Scientific Reports.
3. Lian, H., et al. (2023). A Survey of Deep Learning-Based Multimodal Emotion Recognition. IEEE Transactions on Affective Computing, 14(3), 1234-1256.
4. Schmidt, F. L., & Hunter, J. E. (1998). The Validity and Utility of Selection Methods in Personnel Psychology. Psychological Bulletin, 124(2), 262-274.
5. Senarathne, P., et al. (2021). Automate Traditional Interviewing Process Using Natural Language Processing. In Proceedings of International Conference on Applied Machine Learning.
6. Wu, Y., et al. (2025). Multi-modal Emotion Recognition in Conversation Based on Prompt Learning. Nature Scientific Reports.
7. Xu, L., et al. (2022). ViTPose: Simple Vision Transformer Baselines for Human Pose Estimation. In Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
8. Xu, L., et al. (2023). ViTPose++: Vision Transformer Foundation Model for Generic Body Pose Estimation. In Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
9. Xu, L., et al. (2025). PoseSynVIT: Lightweight and Scalable Vision Transformers for Human Pose Estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence.