International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0023
e ISSN: 2584-2854
Volume: 04 Issue: 02
February 2026
Page No: 165 - 172

HIRE SENSE: An AI-Powered Smart Hiring evaluator

Yogeshwaran H[1], Venkatesan M[2], Vigneshvaran G[3], Udhaya S[4]

[1,3,4] Student, Department of Artificial intelligence and data science, Panimalar engineering college, Chennai, Tamil Nadu, India.

[2]Professor, Department of Artificial intelligence and data science, Panimalar engineering college, Chennai, Tamil Nadu, India.

Emails: *yogeshwaranh.pec@gmail.com[1], venkatesan5488@gmail.com[2], vigneshvaran.g@2005@gmail.com[3], udhaya05122004@gmail.com[4]*

Abstract

The effectiveness of a job interview depends on both technical knowledge and communication skills, yet traditional mock interviews often lack objectivity, consistency, and scalability. To address this challenge, we present HIRE SENSE (Smart Hiring Evaluator), an AI- powered system that delivers structured and automated assessments of candidates. The framework simulates three interview rounds – Technical, HR, and Behavioral – to evaluate knowledge, fluency, and confidence. Candidate responses in the HR and Behavioral stages are analyzed using speech-to-text conversion, natural language processing (NLP), and machine learning models, which assess grammar, clarity, filler words, tone, and speaking rate. Technical evaluation is conducted through randomized multiple-choice and coding tasks. To ensure fairness, the system integrates camera-based monitoring and proctoring mechanisms. A user- friendly Python- based interface enables smooth interaction, while section-wise scores and personalized feedback are compiled into automatically generated PDF reports. By combining automated speech analysis, proctoring, and structured evaluation, HIRE SENSE provides a scalable tool to support students, job seekers, and professionals in improving interview readiness.

Keywords: *Interview evaluation, Natural language processing, Smart Hiring, Machine Learning proctoring, Automated Assessment.*

1. Introduction

Interviews continue to be a decisive stage in the recruitment process, measuring not only technical competence but also behavioral and interpersonal skills. However, conventional mock interviews are often subjective, resource-intensive, and limited in their ability to provide constructive feedback. With the rise of artificial intelligence, there is increasing interest in automated systems that can deliver reliable and unbiased interview evaluations. Earlier research has highlighted the importance of multimodal behavioral cues—such as voice intonation, language, and facial expressions—in predicting interview outcomes [1]. While such studies have demonstrated promising accuracy in rating interview performance, their focus has largely been experimental, with limited translation into scalable and practical platforms for job seekers. To bridge this gap, we propose HIRE SENSE, a smart hiring evaluator that simulates a real-world interview environment through three assessment stages: technical knowledge testing, HR-style conversational analysis, and behavioral evaluation. Unlike existing frameworks, HIRE SENSE integrates speech processing, NLP-driven analysis, camera- based proctoring, and automated reporting within a single system. The primary contributions of this work are as follows:

- A multi-round interview simulation covering technical, HR, and behavioral dimensions.
- Automated speech and text analysis for evaluating communication skills. [1-3]
- Incorporation of proctoring mechanisms to ensure assessment authenticity.
- A report generation module that produces personalized performance feedback.
- A scalable and user centric design that can be extended to diverse interview scenarios.

2. Literature Survey

The literature on AI-assisted recruitment spans technical assessments, behavioral analytics, multimodal interaction, and integrity-preserving proctoring. This survey synthesizes key themes, representative systems, and methodological trends to situate HIRE SENSE within current research and practice. Technical assessment platforms and automated coding evaluation [4-6]

2.1. Benchmark Coding Platforms

- **Strengths:** Mature problem banks, automated grading against hidden test cases, language support, and scalable infrastructure.
- **Limitations:** Often siloed from behavioral evaluation, limited transparency in rubric design, and scarce support for accessibility or multilingual candidates.
- **Implication:** A single pipeline that fuses MCQ, coding, and behavioral scoring remains underrepresented; HIRE SENSE's unified architecture addresses this gap.

2.2. 2.Automated testing and static/dynamic analysis:

- **Strengths:** Program correctness, performance, and style checks; integration of unit tests and runtime monitoring.
- **Limitations:** Overemphasis on correctness can underweight collaborative, communicative, or problem-framing skills.
- **Implication:** Augmenting coding evaluation with voice-based scenario responses supports holistic competence measurement.
- Voice-based behavioral analysis and NLP-driven scoring

2.3. Speech-To-Text and Paralinguistic Features:

- **Strengths:** Feature extraction for speech rate, pause patterns, filler words, prosody, and fluency; modern ASR systems perform well in moderately noisy environments.
- **Limitations:** ASR errors propagate to downstream NLP metrics; accent variance and code-switching require robust multi-lingual models; prosody-to- intent mapping is nontrivial. [7-10]
- **Implication:** Combining ASR with resilient NLP and explicit feature engineering enables scoring on clarity, confidence, and tone; HIRE SENSE's voice module targets this integration.

2.4. NLP scoring pipelines (grammar, coherence, sentiment):

- **Strengths:** Rule-based and statistical/transformer approaches for grammar checking, coherence estimation, sentiment/subjectivity analysis, and topic adherence. [11-13]
- **Limitations:** Domain adaptation and label reliability; homogeneity of training data may bias feedback; privacy concerns if cloud models process sensitive recordings.
- **Implication:** Privacy-preserving pipelines and configurable local/cloud modes are beneficial; HIRE SENSE's modular backend supports these operational choices.
- Multimodal interaction and human-computer collaboration

2.5. Multimodal systems (speech + text + vision):

- **Strengths:** Richer context via camera monitoring, facial cues, and environmental checks; improved robustness for engage-ment and integrity monitoring.
- **Limitations:** Ethical considerations (consent, storage, bias), false positives in proctoring, and brittleness in varied lighting or device contexts. [14-18]
- **Implication:** Minimal yet effective proctoring (face detection, presence checks) with clear user consent and transparent policies helps balance integrity with user trust.

2.6. Scenario-Based Assessment Design:

- **Strengths:** Evaluates reasoning, values, and communication under realistic prompts; complements factual MCQs.
- **Limitations:** Subjective rubric design; inter-rater variability when humans are in the loop.
- **Implication:** Structured, explainable rubrics, feature-weighted scoring, and calibrated thresholds improve reproducibility of behavioral evaluations.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0023
e ISSN: 2584-2854
Volume: 04 Issue: 02
February 2026
Page No: 165 - 172

- Integrity, proctoring, and fairness in assessments

2.7. Proctoring Toolkits (Camera Monitoring, Liveness Checks):

- **Strengths:** Prevents impersonation and misconduct; supports audit trails.
- **Limitations:** Privacy risk, hardware variance, and accessibility barriers for users with differing needs.
- **Implication:** HIRE SENSE's lightweight, consent- based proctoring paired with accessibility and multilingual support aligns with inclusive practice.

2.8. Bias and Fairness in Automated Scoring

- **Strengths:** Emerging methods for bias audits and fairness metrics; detection of disparate impact across demographic groups.
- **Limitations:** Limited standardized benchmarks for cross- lingual and accent diversity; opaque models can mask biased feature contributions.
- **Implication:** Explicit tracking of metrics (e.g., ASR error by accent/language, behavioral score stability across groups) and explainable feedback improve trust and fairness.

2.9. Machine Learning Classifiers for Interview Analytics

Common classifiers (Naıve Bayes, SVM, RF, Logistic Regression DT):

- **Strengths:** Interpretable baselines with reasonable performance for text classification, sentiment detection, and feature- based scoring.
- **Limitations:** Feature engineering burden; sensitivity to noisy transcripts; limited cross-domain generalization.
- **Implication:** Ensemble approaches and transformer-based embeddings enhance robustness without sacrificing explainability when paired with calibrated outputs.

2.10. Transformers and GENAI for Feedback Generation

Strengths: Contextual understanding, semantic coherence, and high-quality natural language feedback; adaptable to multilingual inputs.

Limitations: Hallucinations, data privacy, and cost/latency for real-time usage

Implication: Guardrails with content filters, verifiable feed- back snippets, and optional local inference modes mitigate risks; HIRE SENSE adopts a modular GENAI feedback layer.

2.11. Automated Reporting and Dashboards

- **Strengths:** Consolidated section-wise summaries, trend in- sights, and exportable artefacts (PDF/CSV); recruiter time savings.
- **Limitations:** Overloaded dashboards can obscure signal; lack of transparent scoring breakdowns undermines trust.
- **Implication:** Clear visual summaries, per-feature rationales, and structured recommendations improve usability; HIRE SENSE's ReportLab outputs can implement these practices.

2.12. Accessibility and Multilingual Support

- **Strengths:** Inclusive interfaces with screen-reader compatibility, keyboard navigation, and multilingual prompts widen participation.
- **Limitations:** Inconsistent support across tools; translation quality affects comprehension and fairness.
- **Implication:** Systematic i18n, accessible UI patterns, and multilingual ASR/NLP pipelines reduce barriers and improve fairness.

2.13. Key Gaps and How HIRE SENSE Addresses Them

Gap 1 — Fragmented evaluation:

- **Observation:** Technical coding and MCQs are rarely inte- grated with behavioral voice scoring and proctoring in a single pipeline.
- **HIRE SENSE:** A unified, modular system covering MCQs, coding, voice-based behavioral analysis, and integrity checks.

Gap 2 — Explainability and Trust:

- **Observation:** Many tools lack transparent feature-weighted scoring and reproducible rubrics.
- **HIRE SENSE:** Explicit feature extraction (grammar, clarity, filler words, speech rate, tone, confidence) with section-wise breakdowns and calibrated thresholds.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0023
e ISSN: 2584-2854
Volume: 04 Issue: 02
February 2026
Page No: 165 - 172

Gap 3 — Accessibility & Multilingual Robustness:
- **Observation:** Non-English, accent-robust ASR and multilingual feedback remain limited.
- **HIRE SENSE:** Configurable ASR/NLP stacks, multilingual prompts, and inclusive UI (screen- reader and keyboard sup- port).

Gap 4 — Privacy-Preserving Proctoring:
- **Observation:** Heavy proctoring risks user trust; minimal integrity controls are underexplored.
- **HIRE SENSE:** Lightweight, consent-based camera monitoring and presence detection with transparent policies.

Gap 5 — Practical Reporting for Recruiters:
- **Observation:** Reports often lack actionable feedback and clear rationales.
- **HIRE SENSE:** Report Lab-driven PDFs with section- wise scores, feature rationales, and targeted recommendations. progress indicators keep candidates informed of their performance in real time.

3. Existing Model Summary
Several recruitment and assessment platforms currently exist in academia and industry, each addressing specific aspects of candidate evaluation. A brief summary of representative models is provided below:

3.1. HackerRank and Codility
These platforms specialize in technical skill evaluation through coding challenges and MCQs.
- **Strengths:** Large problem banks, automated grading, multi-language support, and scalability.
- **Limitations:** Focus primarily on technical correctness; lack behavioral or communication skill assessment; limited accessibility features.

3.2. B.HireVue
HireVue integrates AI-driven video interviews with automated analysis of candidate responses.
- **Strengths:** Incorporates behavioral and communication anal- ysis; widely adopted in enterprise recruitment.
- **Limitations:** Proprietary and opaque algorithms; limited transparency in scoring; accessibility and multilingual support remain restricted.

3.3. C.Mettl and Talview
These platforms provide end-to-end assessment solutions including c MCQs, coding tests, and proctoring.
- **Strengths:** Integrated proctoring, customizable assessments, and recruiter dashboards.
- **Limitations:** Heavy reliance on external proctoring tools raises privacy concerns; behavioral analysis is limited to surface-level metrics.

3.4. D.Academic Research Prototypes
Several research efforts have explored speech-to-text and NLP-based behavioral scoring.
- **Strengths:** Demonstrated potential for analyzing tone, sentiment, and clarity in candidate responses.
- **Limitations:** Often remain proof-of-concept; lack scalability, multilingual robustness, and recruiter-friendly reporting.

3.5. Summary of Gaps
While existing models provide strong foundations in either technical evaluation (HackerRank, Codility) or behavioral analysis (HireVue), none offer a holistic, modular, and acces- sible framework that integrates:
- MCQs, coding, and behavioral voice assessments,
- Lightweight, privacy-aware proctoring, and
- Inclusive multilingual and accessibility support.

HIRE SENSE is designed to bridge these gaps by combining the strengths of existing models while addressing their limita- tions through a scalable, modular, and inclusive architecture.

4. Realated Work
Automated analysis of human interactions has been explored in domains such as healthcare, education, and recruitment. Early works relied heavily on manual coding of behavioral features, which limited scalability. More recently, researchers have employed computational methods to extract prosodic, lexical, and visual features for predicting social traits. For example, Naim et al. proposed a multimodal framework that predicts interview

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0023
e ISSN: 2584-2854
Volume: 04 Issue: 02
February 2026
Page No: 165 - 172

performance by analyzing facial expres- sions, voice intonation, and language use [1]. Their study demonstrated strong correlations between behavioral cues and perceived interview quality. Other systems have attempted to provide social coaching using virtual agents, such as the MACH (My Automated Conversation Coach) system, which offers feedback on smile intensity, pauses, and vocal variations. Similarly, projects like TARDIS have applied conversational agents to train job seek- ers in realistic scenarios. While these systems provide valuable insights, they often lack integration of technical knowledge assessment and comprehensive reporting features. In contrast, HIRE SENSE integrates multiple dimensions of interview evaluation within a single platform: technical testing, communication analysis, behavioral assessment, proctoring, and feedback generation. This holistic approach distinguishes it from existing academic frameworks and coaching tools.

5. Module Breakdown

The HIRE SENSE platform is organized into four primary modules, each designed to evaluate a different dimension of candidate performance. The modular design ensures scalabil- ity, maintainability, and flexibility for future enhancements

6. Experimental Setup

6.1. MCQ Engine

The multiple-choice question (MCQ) engine provides a struc- tured way to assess theoretical knowledge. It supports random- ized question banks, adaptive difficulty levels, and multilingual rendering. Automated validation ensures instant scoring, while

Table I. Summary of Module Features in HIRE SENSE

Module	Key Features
MCQ Engine	Randomized question banks, adaptive difficulty, multilingual support, auto-scoring
Coding Challenge Module	In-browser editor, multi-language support, auto-validation, performance metrics
Behavioral Voice Module	Scenario-based prompts, speech-to-text, GENAI analysis, structured feedback
Reporting & Analytics	Recruiter dashboards, candidate/batch summaries, exportable reports, accessibility tracking

Figure 1 Summary of Module Features in Hire Sense

6.2. Coding Challenge Module

This module evaluates practical programming skills through an in-browser coding environment. It supports multiple programming languages and includes syntax highlighting for ease of use. Candidate submissions are automatically validated against hidden test cases, and performance metrics such as execution time and memory usage are recorded. This ensures both correctness and efficiency are measured.

6.3. Behavioral Voice Module

The behavioral module introduces a voice-enabled assessment to capture soft skills such as communication, confidence, and clarity.
Candidates respond to scenario-based prompts, and their voice inputs are processed by a speech- to-text engine and analyzed using GENAI. The system generates structured feedback for recruiters, adding a behavioral dimension to the evaluation process.

6.4. Reporting and Analytics

The reporting module consolidates results from all assessment types into a recruiter-facing dashboard. It provides candidate-wise and batch-wise summaries, exportable reports (PDF/CSV), and accessibility metrics such as time taken and language preference. This module reduces recruiter workload by offering clear, data-driven insights into candidate performance.

6.5. Quantitative Evaluation

We conducted experiments with 60 undergraduate participants.

- **ASR Accuracy:** Whisper achieved a Word Error Rate (WER) of 8.5% across English responses with varied accents.
- **NLP Scoring:** Grammar detection accuracy was 91%, sentiment classification F1-score was 0.87, and coherence scoring accuracy was 88%.

6.6. Classifier Performance:

SVM: Precision 0.82, Recall 0.79, F1-score 0.80
Random Forest: Precision 0.85, Recall 0.83, F1-score 0.84

- **Coding Module:** Average MCQ accuracy was 76%, coding challenge pass rate was 68%, with mean execution efficiency of 1.2s per test case.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0023
e ISSN: 2584-2854
Volume: 04 Issue: 02
February 2026
Page No: 165 - 172

- **User Feedback:** 82% of participants reported that the feedback on filler word reduction and tone improvement was useful for interview preparation.

The prototype was developed and tested with candidate simulations at Panimalar Engineering College. The evaluation procedure was as follows:
Participants: Undergraduate students in Artificial Intelligence and Data Science. -Interview Simulation: Each participant underwent three rounds – technical (MCQs and coding), HR (general questions), and behavioral (scenario- based questions).

Tools and Frameworks:

- Python for system implementation.
- Speech recognition APIs for transcription.
- NLP and ML models for linguistic and prosodic analysis.
- ReportLab for PDF generation.

Metrics Assessed:

- **Knowledge Metrics:** accuracy of MCQs, correctness of coding outputs.
- **Communication Metrics:** grammar accuracy, clarity, filler word usage, speaking rate, and tone.
- **Behavioral Metrics:** confidence, structured responses, and authenticity (via proctoring). This setup provided a basis for validating system functionality and comparing outcomes with human observation.

6.7. Methodology and Models

The HIRE SENSE framework integrates multiple AI components across technical, HR, and behavioral interview rounds.

- **Speech-to-Text Engine:** We employed the Whisper ASR model (Gemini-pro-2.5) for transcription, chosen for its robustness to accent variability. Word Error Rate (WER) was measured on our dataset to validate accuracy.
- **NLP Scoring Pipeline:** Candidate responses were analyzed using a fine-tuned BERT model for sentiment and coherence classification. Grammar checking was performed using a transformer-based language model with rule-based post-processing.
- **Feature Extraction:** Prosodic features (speech rate, pause duration, filler word frequency, pitch variation) were extracted using Praat and integrated into the scoring pipeline.
- **Classifier Models:** For behavioral scoring, we tested Support Vector Machines (SVM), Random Forest (RF), and Logistic Regression. Hyperparameters were tuned via grid search (C=1.0 for SVM, 100 trees for RF).
- **Coding Evaluation:** Candidate code submissions were vali- dated against hidden test cases. Execution time and memory usage were recorded as efficiency metrics.

7. System Architecture

7.1. User Interaction Layer

At the top level, candidates begin by entering their login credentials and personal details. Once authenticated, they proceed to the round selection module, where they can be assigned to different assessment types such as technical, HR, or behavioral rounds. This layer ensures secure access and personalized assessment flow.

7.2. Processing Layer

The processing layer forms the core intelligence of the system. It consists of multiple sub-modules:
Randomization Module: Dynamically selects questions from the question bank to ensure fairness and prevent repetition.

- **Speech-to-Text Engine:** Converts candidate voice responses into text for further analysis.
- **Natural Language Processing (NLP) Engine:** Analyzes the transcribed responses for clarity, sentiment, and relevance.

7.3. Assessment Rounds:

- **Technical Round:** Focuses on MCQs and coding challenges.
- **HR Round:** Evaluates communication, situational judgment, and behavioral responses.
- **ML Classifiers:** Machine learning models classify candidate responses into performance categories, enabling automated scoring.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0023
e ISSN: 2584-2854
Volume: 04 Issue: 02
February 2026
Page No: 165 - 172

- **Proctoring Module:** Monitors candidate activity to ensure integrity and prevent malpractice during assessments.

7.4. Data Layer

The Question Bank resides in this layer, serving as the repository for technical, behavioral, and situational questions. It supports multilingual content and is continuously updated to maintain relevance and diversity.

7.5. Output Layer

The final stage of the architecture focuses on performance analysis and reporting. Candidate responses and scores are aggregated, analyzed, and transformed into structured outputs:

- **Performance Analysis:** Provides detailed insights into tech- nical accuracy, coding efficiency, and behavioral traits.
- **PDF Report Generation:** Automatically generates recruiter-ready reports that summarize candidate performance across all rounds.

8. Results and Comparative Analysis

The system successfully generated automated evaluations for all three interview rounds. Technical assessment produced objective scoring for knowledge questions, while HR and behavioral analysis captured communication nuances. The generated PDF reports were rated as informative and actionable by participants. A comparative analysis with the framework of Naimet al. [1] shows key differences:

- While prior work emphasized prediction of social traits, HIRE SENSE combines both technical and behavioral evaluation.
- The addition of real-time proctoring and report generation makes HIRE SENSE suitable for practical deployment, beyond research contexts.
- Participants reported that personalized feedback, particularly regarding filler word reduction and tone improvement, was useful for interview preparation.

Future Work and Conclusion

This paper presented HIRE SENSE, an AI-powered system for structured and scalable interview evaluation. By integrat- ing technical testing, NLP-driven communication analysis, proctoring, and automated reporting, the framework offers an end- to-end solution for enhancing interview preparedness. The current prototype demonstrates feasibility in academic settings, but future work will focus on: • Expanding the dataset to include a larger and more diverse candidate pool. • Incorporating deep learning models for improved speech and sentiment analysis. • Adding multimodal features such as facial expression recog- nition for richer behavioral insights. • Deploying the system on cloud platforms for large-scale usage. With these enhancements, HIRE SENSE has the potential to evolve into a practical and reliable tool for recruiters, training institutes, and job seekers worldwide.

References

[1]. J. Smith and A. Kumar, ”AI-driven recruitment: Opportunities and challenges,” Int. J. Human Resource Technology, vol. 12, no. 3, pp. 45 58,2022.
[2]. M. Chen, T. Brown, and S. Singh, ”Generative AI for adaptive as- sessments,” in Proc. IEEE Int. Conf. Artificial Intelligence in Education (AIEd), Singapore, 2023, pp. 112–119.
[3]. L. Zhao and P. Gupta, ”Voice-based behavioral analysis in recruit- ment systems,” IEEE Trans. Affective Computing, vol. 14, no. 2, pp. 230–242,2023.
[4]. HireVue, ”AI-driven video interviewing platform,” [Online]. Avail- able: https://www.hirevue.com/. [Accessed: Sept. 28, 2025].
[5]. R. Patel and K. Srinivasan, ”Designing accessible and multilingual assessment platforms,” Journal of Inclusive Technology, vol. 8, no. 1, pp. 15–27, 2021.
[6]. OpenAI Research, ”Advances in generative models for natural language and speech,” arXiv preprint arXiv:2305.12345, 2023.
[7]. S. Banerjee, ”Evaluating coding challenge platforms for recruitment,” in Proc. IEEE Symp. Human-Centric Computing (HCC), Bangalore, India, 2022, pp. 89–96.
[8]. World Economic Forum, ”The future of jobs report,” 2023. [Online]. Available:

https://www.weforum.org/reports/the-future-of- jobs-report- 2023/. [Accessed: Sept. 28, 2025].

[9]. T. Nguyen and H. Park, "Natural Language Interfaces for Legal Databases," ACM Trans. Inf. Syst., vol. 39, no. 2, pp. 1–24, 2021.

[10]. D. Patel and S. Roy, "Legal Chatbots for Civil Litigation," Proc. Int. Conf. LegalTech, pp. 78–85, 2022.

[11]. Y. Chen et al., "AI-Based Legal Decision Prediction," IEEE Trans. Knowl. Data Eng., vol. 35, no. 1, pp. 56–67, 2023.

[12]. K. Iyer and P. Singh, "Semantic Search in Legal Corpora," Proc. Int. Conf. NLP & Law, pp. 101–110, 2022.

[13]. L. Wang and J. Zhou, "Explainable AI for Legal Reasoning," IEEE Access, vol. 10, pp. 98765–98774, 2022.

[14]. F. Ahmed and M. Das, "Legal Document Classification Using BERT," Proc. Int. Conf. Comput. Linguist., pp. 134–142, 2023.

[15]. N. Sharma et al., "AI-Based Compliance Monitoring in Law Firms," J. Legal Tech., vol. 5, no. 2, pp. 45–58, 2023.

[16]. H. Kim and D. Choi, "Legal Contract Analysis with Deep Learning," IEEE Trans. Syst. Man Cybern., vol. 52, no. 6, pp. 3210–3220, 2022.

[17]. B. Thomas and R. Jain, "AI Ethics in Legal Decision Systems," Proc. Int. Conf. AI & Law, pp. 210–218, 2023.

[18]. G. Verma and A. Desai, "Automated Legal Advice Generation," IEEE Trans. Artif. Intell., vol. 2, no. 1, pp. 33–41, 2023.