INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 5, May 2026 | ISSN: 2456-4184 | IJNRD.ORG

# AI-Powered Smart Interview System Using NLP, Speech Recognition, and Emotion Detection

**1st Roshan. D**
Department of advanced Computing Science
Vels Institute of Science, Technology & Advanced Studies

**2nd Dr. P. Arivazhagan M.C.A., Ph.D.,**
Department of advanced Computing Science
Vels Institute of Science, Technology & Advanced Studies

**ABSTRACT -The hiring process is an essential part of hiring the right candidate for the right job; however, the traditional hiring method through an interview is a biased approach and an inefficient one. This paper introduces a smart interview system using AI for the evaluation of the candidate by utilizing the capabilities of Natural Language Processing, speech recognition, and emotion detection. The system starts by evaluating the candidate's resumes through the use of skills and job role prediction. Based on this approach, the questions for the candidate are generated dynamically**

**During the course of the interview, the candidate's answers are recorded through the use of speech recognition. In addition to this, the candidate's facial expressions are also detected through the use of a Convolutional Neural Network (CNN) for emotion detection. In addition to this, the candidate's answer is also evaluated through the use of TF-IDF vectorization and cosine similarity with reference answers for a performance score. Moreover, the candidate's emotional state is also detected through the use of emotion detection.**

**All the results are stored and visualized through the use of Microsoft Power BI. The proposed system is an efficient and unbiased approach for the evaluation of the candidate.**

## 1. INTRODUCTION

Recruitment is an important factor in any organization, and the conventional method of recruitment through interviews is a laborious process, which may also be influenced by human bias, thereby leading to a lack of uniformity in the results. In order to overcome the problems that are being faced, the proposed project aims at introducing a new method of recruitment through the implementation of an AI-based smart interview process, which will help in making the process efficient and unbiased. In this process, first, the resume of the candidate will be analyzed, and based on the skill sets determined, a suitable role will be assigned to the candidate. Further, appropriate questions will be generated for the candidate. During the interview process, the spoken answers given by the candidate will be converted into text format through a speech recognition system, and at the same time, the facial expressions of the candidate will also be captured through a camera, which will help in understanding the emotions of the candidate, whether he is confident, stressed, or nervous, during the interview process. Further, the answers given by the candidate will be compared.

## 2. LITERATURE SURVEY

The recruitment process has been improved significantly with the help of Artificial Intelligence, which led to the development of automated systems for the evaluation of the candidates. In the past, the automated systems were used for the evaluation of the resume of the candidates by making use of Natural Language Processing (NLP) techniques, where the skills, keywords, and experience of the candidate were identified and matched with the requirement for the position. However, the automated systems were also restricted in the sense that the performance of the candidate was not taken into consideration for the evaluation process.

Recently, automated interview systems were designed for the evaluation of the candidates by making use of chatbots for the interview process. Although the use of the automated system saves a lot of time for the hiring manager, the automated system mainly focuses on the text-based responses of the candidate and does not consider the behavior and emotions of the candidate, where the facial expressions and speech of the candidate are not considered for the evaluation process.

Recently, emotion detection has also been improved with the help of computer vision and the use of deep learning algorithms, specifically Convolutional Neural Networks (CNN). These algorithms are used for emotion detection, i.e., happiness, sadness, anger, and even neutrality. However, this is not widely used in recruitment systems.

Old techniques used for evaluating the answer given by the candidate were based on keyword matching, which is not reliable. However, using techniques such as TF-IDF and cosine similarity, results are better, where the meaning of the answer given by the candidate is matched with a reference answer. However, this is not used in combination with other techniques.

Generally, most of the existing systems are designed for a particular purpose, i.e., resume analysis, emotion detection, answer evaluation, etc. There are very few systems that combine these features in a single system, which shows the need for a complete solution.

## 3. PROPOSED METHODALOGY

In this project, a smart interview system using AI technology will be developed for automating the recruitment process and evaluating the candidates efficiently. The proposed system will analyze resumes, generate questions, track the behavior of the candidate, and finally produce results based on multiple criteria. The main idea behind this project is to make the interview process more efficient, accurate, and unbiased.

Resume Analysis

The first step is resume data collection. This is done by simply uploading the resume. The resume is then processed using NLP techniques. The important information is retrieved from the resume. The irrelevant information is removed to make it more accurate.

Formula for Text Processing

Processed Text = Raw Text - Stop Words

The resume is analyzed, and according to the features, a suitable job role is predicted for the candidate.

B. Question Generation

Once the job role is identified, the system retrieves the relevant interview questions from the dataset. The questions are retrieved either by using a predefined rule or by random selection, as in a real interview scenario

.C. Emotion Detection

During the course of the interview, the system observes the candidate through the webcam. Face detection is done, and the emotion is detected using deep learning to recognize a happy face, a neutral face, a nervous face, and a sad face.

D. Speech Recognition

The candidate answers the question posed to him/her by the system. The system captures the audio input and converts it to text using various speech recognition techniques.

E. Answer Evaluation

The system evaluates the answer given by the candidate according to the similarity between the answer and the predefined answer using various text similarity techniques.

Formula (TF-IDF):

TF-IDF = Term Frequency * Inverse Document FrequencyFormula (Cosine Similarity):

Similarity = (A * B) / (|A| * |B|)

According to the similarity, the final score for the answer will be generated.

F. Result Storage

All the information, i.e., the name of the candidate, the job role, answers, score, and emotion, is stored for future use.

G. Dashboard Visualization

The stored results are shown in a graphical format by using a dashboard, which gives a clear idea about the performance, scores, and emotions of the candidates.

H. Model Evaluation

The performance of the system can be evaluated on the basis of the accuracy and effectiveness of the evaluation.
Formula (Accuracy): Accuracy = Correct Evaluations / Total Evaluations
On the whole, the system incorporates various technologies like NLP, computer vision, and speech recognition to provide a comprehensive and intelligent system for the interview.


## 4. ARCHITECTURE DIAGRAM

Fig 1. Architecture diagram

## 5. INPUT

The system requires input data in various forms to operate effectively with a complete evaluation of the candidate. The input data for the system consists of the candidate's resume in PDF or TXT format. The resume consists of information such as skills possessed by the candidate, education, and experience, which is processed using NLP to predict a suitable job role for the candidate. The second input data for the system consists of a question-answer dataset in CSV format containing questions related to a particular role with reference answers provided.

For evaluating the candidate's behavior, the emotion dataset is used for training the model with the FER dataset for emotion detection such as happy, sad, angry, etc., based on the facial expression of the candidate. The system also has the capability to take audio input using a microphone for recording the answer provided by the candidate based on the input provided in the resume using speech recognition. At the same time, video input is also taken using a webcam for detecting the emotion based on the facial expression of the candidate.

## 6. PSEUDOCODE

START

Upload candidate resume Extract text from resume emove stop words and clean data Identify skills, education, and experience Predict job role using the extracted features Load question answer data Select questions to be asked to the candidate Initialize webcam for emotion detection Initialize microphone for speech FOR each question Display question to candidate Capture candidate's audio Convert speech to text Capture facial expressions using webcam Detect emotion (happy, sad, nervous, etc.) Compare answer with the correct answer Calculate similarity score Generate score for the answer END FOR Store candidate data (name, role, answer, scores, emotion) Save data to a CSV file Generate visualization for the dashboard Display results END

## 7. OUTPUT

```
[5]: dfspd.read_csv('ai_interview_dataset_largu.csv')

[6]: df.head(10)
```

[6]:

| | Role | Question | Answer |
|---|---|---|---|
| 0 | Data Scientist | What is Machine Learning? | This is a concise sample answer explaining: wh... |
| 1 | Data Scientist | Explain supervised learning. | This is a concise sample answer explaining: ex... |
| 2 | Data Scientist | Explain unsupervised learning. | This is a concise sample answer explaining: ex... |
| 3 | Data Scientist | What is overfitting? | This is a concise sample answer explaining: wh... |
| 4 | Data Scientist | What is underfitting? | This is a concise sample answer explaining: wh... |
| 5 | Data Scientist | What is feature engineering? | This is a concise sample answer explaining: wh... |
| 6 | Data Scientist | What is cross validation? | This is a concise sample answer explaining: wh... |
| 7 | Data Scientist | Explain bias vs variance. | This is a concise sample answer explaining: ex... |
| 8 | Data Scientist | What is a confusion matrix? | This is a concise sample answer explaining: wh... |
| 9 | Data Scientist | What is precision and recall? | This is a concise sample answer explaining: wh... |

Fig
2. Load The Dataset

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 360 entries, 0 to 359
Data columns (total 4 columns):
 #   Column      Non-Null Count  Dtype 
---  ------      --------------  ----- 
 0   Role        360 non-null    object
 1   Question    360 non-null    object
 2   Answer      360 non-null    object
 3   Unnamed: 3  1 non-null      object
dtypes: object(4)
memory usage: 11.4+ KB
```

```
df.shape
```

```
(360, 4)
```

```
df.duplicated().sum()
```

```
0
```

```
df.isnull().sum()
```

```
Role           0
Question       0
Answer         0
Unnamed: 3   359
dtype: int64
```

Fig 3. Preprocessing

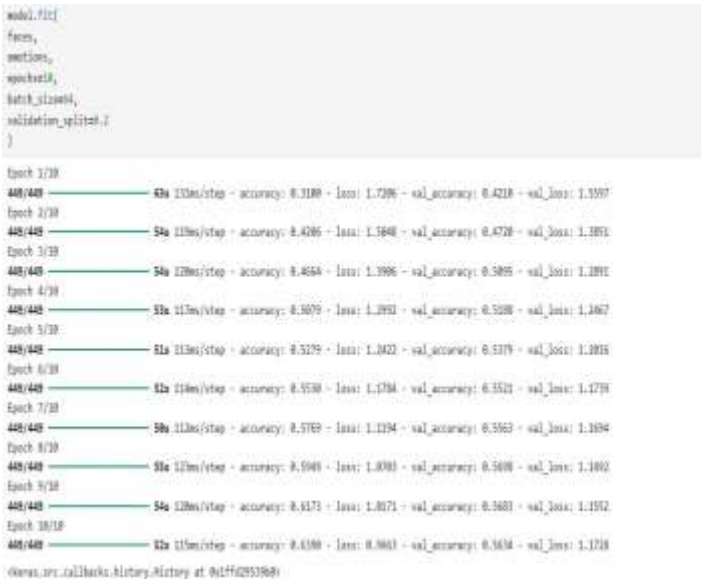


Fig 4 CNN Model Training For Emotion Detection


Fig 5. Captured Audio


Fig 6. AI Smart Interview Dashboard

## 8. RESULT AND DISCUSSION

The proposed AI-based interview system was implemented and tested using real-time inputs from the users.
In the proposed system, during execution, the speech of the candidate was captured by the speech recognition module, and the speech was converted into text format. For example, the speech of the candidate was correctly recognized by the proposed system as follows:

Then, the proposed system would fetch the corresponding answer for the given question, which was “Explain batch prediction?” The answer had explained the batch prediction as follows:

“Batch prediction is a way to generate prediction for large volume of data at once.”In order to assess the candidate’s response, the system had utilized a TF-IDF vectorization method coupled with a cosine similarity approach. This method assesses the semantic similarity between the candidate’s response and the reference response. Based on this approach, the system had computed the similarity score as follows: Score: 12.0%

Fig 7. Analyzing The Answer

The overall results show that the proposed system is a reliable and automated method of evaluating interview responses. The proposed system saves human resources and ensures consistency in evaluation. Therefore, it can be a useful tool in today’s recruitment environment.

## 9. CONCLUSION

This project intends to develop a smart interview system using AI, which makes the recruitment process easier and faster for a firm or organization. It can analyze a resume and recognize the skills possessed by the candidate, thereby offering a suitable position for the candidate. Additionally, it can ask suitable questions, record the answers given by the candidate in the form of voice, and evaluate them.

It also includes a camera that can recognize the facial expressions of the candidate, thereby understanding the level of confidence possessed by the candidate. It also evaluates the answers given by the candidate through simple text comparisons, which can be used for providing unbiased results. All the results are stored and displayed in a suitable form for easy understanding.

This project can be considered a simple, efficient, and reliable way of conducting an interview with the help of AI.

## 10. REFERENCE

[1] L. Liu, Z. Yang, and X. Wang, “Artificial Intelligence in Recruitment Systems: A Review,” *IEEE Access*, vol. 9, pp. 123456–123470, 2021.

[2] H. Han, A. Jain, and F. Wang, “Automated Resume Screening Using Natural Language Processing,” *International Journal of Computer Applications*, vol. 182, no. 10, pp. 25– 30, 2020.

[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *IEEE CVPR*, 2016.

[5] P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” *IEEE CVPR*, 2001.

[6] Y. LeCun, Y. Bengio, and G. Hinton, “Deep Learning,” *Nature*, vol. 521, pp. 436–444, 2015.

[7] OpenCV Library, "Open Source Computer Vision Library," 2023.

[8] TensorFlow, "An End-to-End Open Source MachineLearning Platform," Google, 2023.

[9] D. Jurafsky and J. H. Martin, *Speech and Language.