



A Flask-Based System for Measuring and Analyzing Confidence in Interviewee Speech Using Speech Recognition Technology

H.P. Dangalla
MS23009240
MSc in Information System

A THESIS
SUBMITTED TO
SRI LANKA INSTITUTE OF INFORMATION TECHNOLOGY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE IN INFORMATION SYSTEMS

December 2025

I hereby testify that I have read this thesis, and that I think it is quite satisfactory in extent and quality, as a thesis in the degree of Master of Science.

Prof. Anuradha Jayakody

Approved for MSc. Research Project:

MSc. Programme Co-ordinator, SLIIT

Approved for MSc:

Head of Graduate Studies, FoC, SLIIT

DECLARATION

This is to ensure that it is not the work of another human being, but my work provided there is certainty regarding it unless known to refer to published and unpublished sources. No version of the work has been submitted to the Sri Lanka Institute of Information Technology or to any other institution, to ascertain the evaluation of the work to any other end.



Sign:

H.P Dangalla MS23009240

Date: 05/01/2026

ABSTRACT

A Flask-Based System for Measuring and Analyzing Confidence in Interviewee Speech Using Speech Recognition Technology

H.P Dangalla

MSc. In Information Systems

Supervisor :Prof Anuradha Jayakody

December 2025

Confidence is vital in the interview process, as it is a key determinant of credibility and competence. Nevertheless, conventional approaches to evaluating confidence are highly dependent on human judgment, which brings in bias and variability. This article suggests an effective machine learning platform which uses convolutional neural networks (CNNs) and long short-term memory networks (LSTMs) to overcome these shortcomings, offering an objective and scalable method to determine confidence levels in speech during interviews. In this architecture, CNNs extract the spatial characteristics of audio spectrograms, paying attention to the key prosodic variations in pitch and tone that act as confidence indicators. Meanwhile, LSTMs learn the time-varying behavior of these features, enabling the system to identify change in speech rate and time-varying pauses. These models can jointly identify speech as confident or non-confident with 92.5 percent accuracy on labeled data. This system is more precise, recalls higher, and has a better F1 score than current methods. Although the model demonstrates potential in confidence detection, it struggles with extrapolating across accents and languages due to overfitting. But it has a lot of potential in the future as a tool. To overcome future challenges, more diverse datasets and sophisticated methods such as data augmentation and transfer learning can be implemented to enhance the adaptability of the system. Such a framework might be of immense use in practical situations when conducting job interviews, educational evaluations, and coaching in speech delivery, giving consistent, objective measures of confidence. The resultant system might help enhance fairer judgments, offer constructive criticism to applicants, and contribute to making informed choices, benefiting the science of affective computing. It also paves the way to scalable, real-time solutions that could improve human-AI interaction and enhance communication dynamics in various areas.

ACKNOWLEDGEMENT

I have been fortunate to have outstanding advisors at the Sri Lanka Institute of Information Technology, despite their differing views. One such mentor, Prof. Anuradha Jayakody, was truly exceptional. He not only listened attentively to my ideas but also provided constructive criticism that helped improve my work. Moreover, He granted me the freedom to pursue my own projects, allowing me to use his research account time, which was a tremendous opportunity for growth.

Table of Contents

DECLARATION	3
ABSTRACT.....	4
ACKNOWLEDGEMENT	5
1. Introduction.....	9
1.1. Research Background	9
1.2. Problem Statement.....	10
1.3. Research Goals/Objectives	11
1.4. Importance of Research	12
1.5. Research Scope and Constraints	14
Chapter 2. Literature Review	17
2.1. Current Approaches in Confidence Detection	17
2.2. Application of Neural Networks in Speech Analysis	21
2.3. Integrating Multimodal Approaches for Confidence Detection	28
2.4. Identified Gaps in Literature.....	30
3. Conceptual Framework.....	35
3.1. Utilizing Machine Learning for Confidence Detection	35
3.2. Structure of the Proposed Framework	37
Chapter 4. Methodology	41
4.1. Design of the Study.....	41
4.2. Data Gathering Techniques.....	43
4.3. Techniques for Speech Feature Extraction	45
4.4. Development of the Confidence Detection Model	49
4.5. Testing and Validating the Model.....	52
5. Results and Findings	56
5.1. Evaluation of Model Performance	56
5.2. Examination of Key Speech Features	60
5.3. Issues and Misclassification Analysis.....	64
5.4. Comparative Analysis with Existing Methods	68
6. Discussion.....	72
6.1. Importance of Findings	72
6.2. Implications for Theory	75
6.3. Real-World Applications	76
6.4. Limitations and Prospects for Future Research	79
7. Conclusion	82

7.1. Key Findings Recap.....	82
7.2. Contribution to Knowledge.....	83
7.3. Practical Recommendations.....	85
8. References.....	88
10. Appendices.....	92
Appendix A: Description of the Dataset.....	92
Appendix B: Architecture and Hyperparameter Settings	92
Appendix C: Evaluation Metrics and Confusion Matrix	92
Appendix D: Process of Feature Extraction.....	93
Appendix E: Methods for Data Preprocessing and Augmentation.....	93

List of Figures

Figure 1 - Overview of the Machine Learning Model Framework	13
Figure 2 - Annual Publications on Machine Learning and Speech Recognition.....	19
Figure 3 - Flowchart for Speech Data Analysis Process	22
Figure 4 - Visual Representation of Speech Signal Features.....	25
Figure 5 - Distribution of Flowchart.....	40
Figure 6 - User Interface of the Confidence Measurement Application.....	41
Figure 7 - Formula for Model Accuracy Calculation	56
Figure 8 - Mathematical Expression for Precision Calculation	57
Figure 9 - Recall Calculation Formula for Model Evaluation	57
Figure 10 - F1 Score Formula for Performance Evaluation	58
Figure 11 - Dataset Used for Voice-Based Confidence Detection	92
Figure 12 - Architecture and Hyperparameters of the Confidence Detection Model.....	92
Figure 13 - Confusion Matrix and Key Evaluation Metrics	92
Figure 14 - Process of Feature Extraction from Audio Data	93
Figure 15 - Steps in Data Preprocessing and Augmentation	93

List of Tables

Table 1 - Challenges in Existing Confidence Detection Models.....	32
---	----