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of another human being, but my work provided there is certainty regarding it unless known to refer to published and unpublished sources.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

List of Figures

List of Tables