

# 1. PENDAHULUAN

## 1.1. Latar Belakang

Identifikasi gender adalah salah satu masalah krusial dalam bidang analisis wicara. Hal tersebut dapat dilacak dari data akustik seperti pitch, median, dan frekuensi yang dapat diklasifikasikan menggunakan metode machine learning [1], [2], [3]. Klasifikasi ini penting untuk pengenalan suara yang dapat diterapkan pada beberapa aplikasi [4], [5]. Menentukan jenis kelamin pembicara mengurangi beban komputasi sistem tersebut untuk pemrosesan lebih lanjut. Metode khas untuk pengenalan gender dari ucapan sangat bergantung pada ekstraksi fitur dan proses klasifikasi [6], [7].

Pengklasifikasi Deep Long Short Term Memory (Deep LSTM) adalah jenis penting dari jaringan saraf berulang (RNN) yang menangani dependensi jarak jauh lebih efisien daripada RNN [8]. Baru-baru ini, RNN dengan model akustik Memori Jangka Pendek Panjang (RNN-BLSTM) dua arah telah terbukti memberikan kinerja yang sangat baik [9].

Penelitian ini berfokus pada pembangunan sistem klasifikasi gender berdasarkan sinyal suara dengan menggunakan lebih sedikit data suara. Ekstraksi fitur yang digunakan adalah MFCC, yang biasa digunakan dalam pengenalan suara karena memiliki akurasi yang tinggi [10], [11], [12]. Klasifikasi dikembangkan menggunakan BLSTM, salah satu arsitektur RNN yang memberikan akurasi tinggi [13], [14].

## 1.2. Perumusan Masalah

*Speech* klasifikasi menggunakan Bahasa Indonesia memiliki akurasi yang kurang besar dengan hanya menggunakan dataset yang sedikit.

### **1.3. Tujuan**

Tujuan Penelitian ini adalah membangun model segmentasi ucapan Bahasa Indonesia berakurasi tinggi menggunakan dataset kecil.

### **1.4. Batasan Masalah**

Data yang digunakan untuk percobaan adalah 1000 ujaran atau rekaman suara: 500 laki-laki dan 500 perempuan. Ucapan memiliki sampling frekuensi 16 kHz, tingkat kuantisasi 16 bit, dan saluran mono. Mereka disimpan di ekstensi WAV.