

## ABSTRACT

*This essay discusses about speaker recognition, a system that recognize subject identity by their voice. First, subject's voice features are extracted using MFCC (Mel-Frequency Cepstrum Coefficients) method. Steps in the MFCC are pre-emphasis, framing, windowing, FFT (Fast Fourier Transform), mel scaling and DCT (Discrete Cosine Transform), which produce feature vector called cepstrums. These cepstrums are then modelled using GMM (Gaussian Mixture Model). Steps in th GMM are Expectation-step and Maximization-step, which produce gaussian distribution along with its parameters, mean ( $\mu$ ) and variance ( $\sigma^2$ ) which are different for every subjects. Classification step is done by comparing between training data parameters and testing data parameters. If the comparison gets high score, it means two datas are match, vice versa.*

*Previous research done by student group at Preston University and Jinnah Women University, Pakistan, with title "Speaker Identification Using GMM with MFCC" gets accuracy score 87,5% using feature extraction method MFCC, clustering method K-Means, modelling method GMM and classification by log probability. In this essay, we will pass the clustering step and classify by doing comparation between gaussian distribution using parameter mean ( $\mu$ ) and variance ( $\sigma^2$ ), which is the fastest and easiest way. In this essay, we do the best thing to get as close as possible with previous research accuracy score knowing that the classification step can be called the 'rough' one in GMM usage so we are not expecting high, even though there are so many factors that can influence the simulation accuracy.*