Paper ID:	1571072883
Paper Title:	The significance of time duration and feature extraction of voice signal dataset for depression classification
Authors:	Patteera Tongnopparat (King Mongkut's Institute of Technology, Ladkrabang, Thailand); Komsan Kiatrungrit (Mahidol University, Thailand); Treesukon Treebupachatsakul and Suvit Poomrittigul (King Mongkut's Institute of Technology Ladkrabang, Thailand)
Email:	64011533@KMITL.AC.TH

Patteera Tongnopparat (King Mongkut's Institute of Technology, Ladkrabang, Thailand); Komsan Kiatrungrit (Mahidol University, Thailand); Treesukon Treebupachatsakul and Suvit Poomrittigul (King Mongkut's Institute of Technology Ladkrabang, Thailand)

64011533@KMITL.AC.TH

Abstract

This project aims to study the effect of time duration of audio datasets and feature extractions for depression classification of machine learning models. The three different time durations of audio signals including dataset (1): 1-6 min, dataset (2): 1 min, and dataset (3): 30 seconds were investigated. There were 5 machine learning models applied with 12 feature extractions including 4 features of time domains, 4 features of frequency domains, and 4 features of time-frequency domains. The statistic significant of time duration of dataset and feature extraction were examined by one-way ANOVA compared to a p-critical of 0.05. The results of performance metrics indicate the highest accuracy, and f1-score was achieved by logistic regression with MFCC on dataset (1), which was 88.24% and 85.71%, respectively. Moreover, the precision and recall could reach 1 by several models and feature extraction, especially in frequency domains and time-frequency domains. Also, the time durations and feature extraction on each dataset didn't have statistical significance on model performance.