

Paper id	BMEiCON2022-003
Title	Classification of Depression Audio Data by Deep Learning
Author	Phanomkorn Homsiang
Paper topics	
Abstract	
<p>Due to many factors such as anxiety from contracting the disease and concern about the socio economic impacts, Thai people have accumulated stress and are at risk of depression. The diagnosis of depression can be primarily assessed by testing the assessments such as PHQ 8, PHQ-9, and CES-D. The applied deep learning technology in medicine has received research interest and has been developing. In this research, we tried the classification of depression and non-depression audio datasets with the implementation of 4 model architectures: 1D CNN, 2D CNN, LSTM, and GRU. By converting wave audio format (WAV) of Daic-woz database to the Mel frequency cepstrum (MFC). We have done the training and evaluated the 4 model architectures and compared the results between non-augmented and augmented datasets. The highest accuracy was obtained from 1D CNN with a non-data augmentation of 95%, and a 2D CNN with a data augmentation of 75%. These results confirm that human voices can differentiate between depression and non-depression.</p>	