Speech Emotion Recognition

- Rawabi Alharbi
- Dimah Albunayyih

Problem:

Speech Emotion Recognition (SER) is the act of recognizing human emotion from speech. This is the advantage of the fact that voice often reflects underlying emotion through tone and pitch. SER can be beneficial for call centers that are giving health care support for elderly people and emergency call centers. The proposed project is based on speech emotion recognition is giving greater priority to calls with emotions such as fear, anger, and sadness, and less priority to calls with neutral speech and happiness.

Design:

The four datasets provided on Kaggle were stored as CSV files. The first process was clean the datasets, made them as one format and concatenate them. Then we did EDA to understand the data. After that we started the CNN to predict the speech emotion.

Algorithm:

We applied different Augmentation techniques (Noise injection, stretching and pitch), and applied feature extraction techniques (zero crossing rate, chroma_stft, MFCC, RMS and melSpectogram) to extract the most important features from the audio files and then we applied neural network algorithms using convolution and dense layers with (relu and softmax) activation functions.

TOOLS:

- Technologies: Python, Jupyter notebook.
- Libraries: NumPy, Pandas, Sklearn, librosa, Matplotlib, Seaborn

Communication:

We shared our work in our GitHub accounts

- **https://github.com/RawabiKhalaf/speech-emotion-recognition.git**
- **https://github.com/DimahAlbunayyih/T5_bootcamp_projects.git**