On-app model Description:

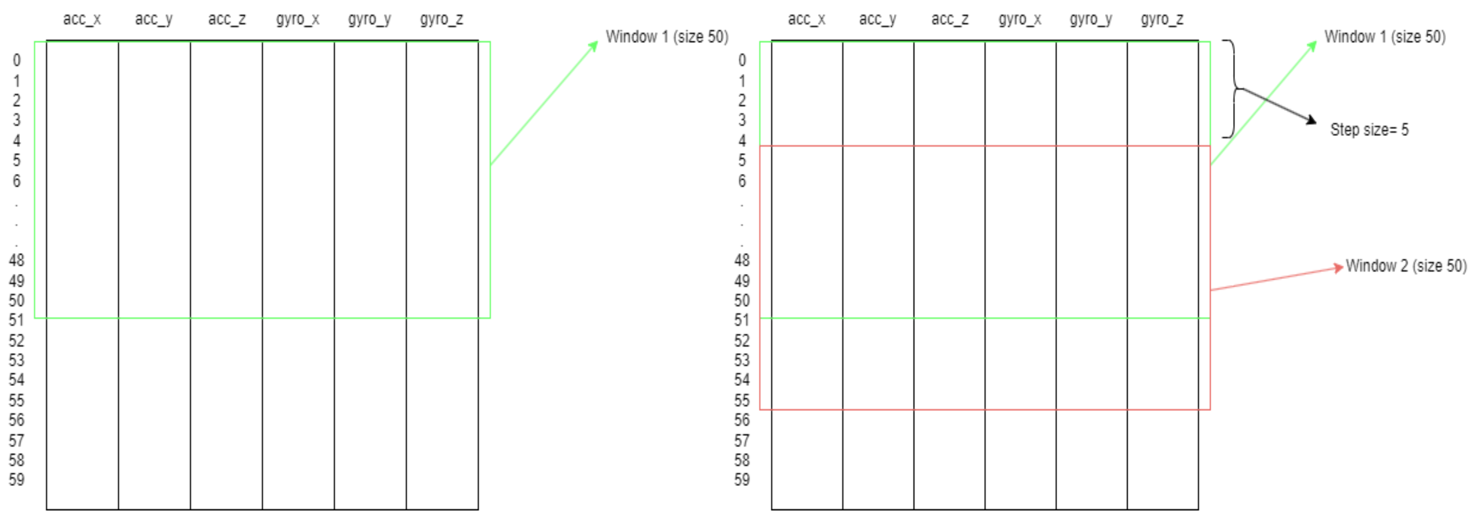
Data Preprocessing:

Data Labelling:

- The provided data for all the subjects were concatenated to create a singular dataframe
- Two new columns are added to the data
  - A column for the physical activity to differentiate between the 11 physical activities (Normal walking, Shuffle walking, Running, Ascending stairs, Descending stairs, Miscellaneous movement, Sitting/standing, Lying down on back, Lying down on stomach, Lying down on right, Lying down on left).
  - And another column to differentiate between the 4 breathing activities for the stationary activities (Normal, Coughing, Hyperventilating, Other)
  - In the Breathing activities column, the dynamic activities (Normal walking, Shuffle walking, Running, Ascending stairs, Descending stairs, Miscellaneous movement) are assigned a value of -1 to differentiate them from the stationary ones.

Data Processing (Before passing to the model):

- **Sliding Window:** The data is segmented using a sliding window with size of 50 and an overlap of 90%. This means that the data has a step size of 5. The label for each segment are extracted by taking the mode of the labels in a segment. This can be visualised by the diagram:



- **One-hot encoding:** One-hot encoding is then performed on the labels, this is done by applying the function `to_categorical()`

Model Architecture:

We use two models, one for the classification of the physical activity, and one for the breathing activity that is used only for the stationary activities, otherwise the breathing is set to normal. However, it is important to note that on the app, you don't have to specify the model you would like to use and the use of the models is done automatically.

The train-test split is set to 80-20. Where the test data is treated as the validation data.

First level model (Physical activities):

The first level model features a sequence of three Conv1D layers, each equipped with 64 filters, a kernel size of 3, and 'relu' activation, enhanced with L2 regularization to mitigate overfitting. Following each Conv1D layer, batch normalization is employed for stabilizing learning, coupled with a dropout rate of 0.2 to further prevent overfitting.

After the three Conv1D layers, the data is then flattened and passed through two dense layers, the first one having 128 nodes with the final layer's size matching the number of output classes (11) and employing a 'softmax' activation.

The model is compiled with an Adam optimizer and 'categorical_crossentropy' loss function, focusing on accuracy as the evaluation metric. Training is conducted with a batch size of 64, utilizing both training and validation datasets while ensuring the training data is shuffled.

Second level model (Breathing activities):

The second model, similarly, begins with a sequence of three Conv1D layers, each having 64 filters and a kernel size of 3, with 'relu' activation. These layers are complemented with batch normalization to maintain stable learning and a dropout rate of 0.2 after each normalization step, which helps in reducing overfitting. The architecture then transitions to a Flatten layer, preparing the data for dense processing.

The flattened output feeds into two sequential dense layers, however this time, the first has 128 nodes and the second has 32, both using 'relu' activation. This is additionally followed by a final dense layer, tailored to match the number of output classes (4) that employs a 'softmax' activation.

The model is compiled using the Adam optimizer and 'categorical_crossentropy' as the loss function, which is ideal for multi-class classification scenarios. The primary metric for performance evaluation is accuracy. Training is conducted over a batch size of 128, incorporating both training data and validation data while ensuring the training data is shuffled.

Model Performance:

The model is capable of successfully classifying all the activities, however it cannot distinguish between the different types of breathing activities that are grouped together as ‘other’, i.e., singing, talking, eating, laughing, classifying them as ‘other’.

The accuracy of the first level model came out to: 98.6%

The accuracy of the second level model came out to: 78.61%

This gives us the combined accuracy of the model to be: 89.52%