

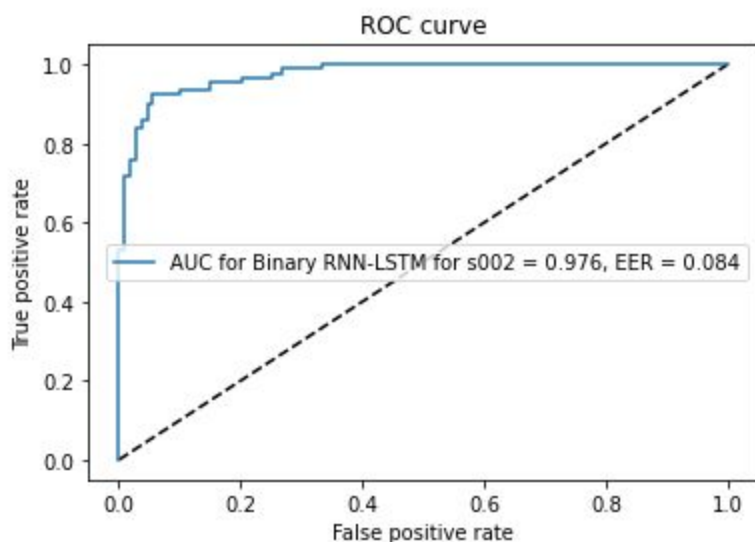
1. RNN

RNNs are used in applications where data is sequential i.e. the order in which data appears is significant in addition to other data parameters. In our case, the order of key-press and key-release is also considered along with Up-down, down-down and hold latencies of each pressed key.

Here we explore RNNs for both binary as well as multiclass classification of keystroke data. Both the models are trained using the CMU keystroke dynamics benchmark dataset which is given in tabular (2D) format. We reshape this dataset into 3D format to represent sequential view.

1.1. Binary Classification :

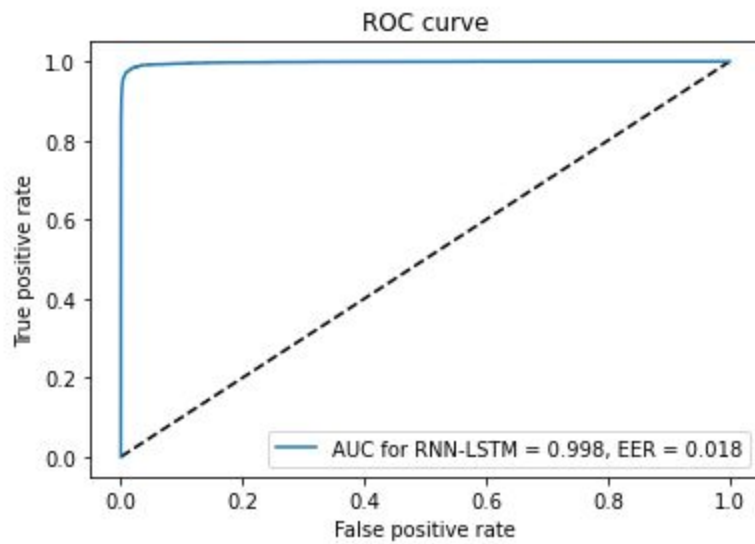
Here, given a keystroke sequence and a subject, our goal is to determine if the keystroke sequence belongs to the given subject or not. Here we choose a subject and make a dataset for that subject. This dataset consists of 400 positive samples of that subject and randomly chosen 400 samples typed by other subjects (negative samples). We then split this dataset into train and test subparts. We train the model using the training dataset and evaluate its performance on the testing dataset. The model's ROC curve for subject s002, trained over 75 epochs on 600 samples (which include positive and negative samples) is as follows.



This plot is plotted for test data.

1.2. Multi class Classification :

Here, given a keystroke sequence, our goal is to determine the subject to which this keystroke belongs. This entire dataset is taken and split for training and testing. The model's ROC curve when trained over 60 epochs on 16320 samples and tested for 4080 samples. Its ROC curve plotted over test data is as follows



As seen from the above curves, RNNs work better for multiclass classification. The reason for this can be intuitively attributed to availability of limited data in case of binary model training. The advantage of using multiclass classification is that a single model is capable of authenticating/verifying all subjects. Also, RNNs are advantageous in case of authenticating sequences of variable length.