

Custom model explore in below regions

Adding more linear layers into model test the model with 3 linear layers

Adding new layer type called batch normalization layers which help the linear layer outputs to be normalized vectors , since most deep learning model advantages from batchnorm layers

Adding new layer called dropout layers which randomly drop certain connection in the linear layer , this type of layers helps to reduce the model overfitting

Make the linear layers much wider with hidden size 256 which helps to capture more robust feathers from the input dataset .

Tried with different optimizers such as SGD , Adamax , Adam , which Adam performance is the best result giving optimizer .

Use the relu activation function , in deep learning paradigm relu is the best performing activation function than others , which helps to improve the results .