Acc = early stopping acc ; no valid set

Models are trained on MPS so cpu or cude acc are likely to be better

More channels = more overfeating hence lower test acc

Aléatoire bayes : 3.6

Linéaire : ≃ 4.6

CNN 2 Layer dropout final de 0.5 puis linear ≃5

CNN 2 Layer layer norm dropout final de 0.5 puis linear sur 270 channels ≃ 6.9

Après le max overfeating, accuracy = 5.4 +-1 grande entropy dans les outputs sur le test set.

Trainacc = 99.97 loss = 0.0022 (crossentropy)

self, input\_channels=185, conv\_out\_channels=64, kernel\_size=3,

n\_conv\_layers=1, transformer\_layers=3, transformer\_heads=4,

transformer\_emb\_size=64, dropout=0.1, batch\_size=64

lr=10\*\*(-4), weight\_decay=0.01

400 epochs meilleur à la 10 eme acc de 5 overfeating trainacc encore entrain de monter

New data en dropant les channel un peu bizard jutilise les meg pas les plan

Linear vers 4.6

CNN 2 Layer dropout final de 0.5 puis linear ≃ 6.25

En droppant les channles j’ai

Linear vers 4.7+ batch 128

CNN 2 Layer dropout final de 0.5 puis linear ≃ 6.25 ; monte à 6.6 (d’un entrainement à un autre perf complétement différentes)

Avec layer norm baisse les perfs ; quoi que en fait l’apprentissage sur le test est bc plus long ??? atteint des 6.25+

Avec un seul norm layer je monte à 6.92 sur une epoch qui généralise bien sur les autres dataset mais est-ce du testset overfit ?

Confiance avec hyperparamtuning 7% acc