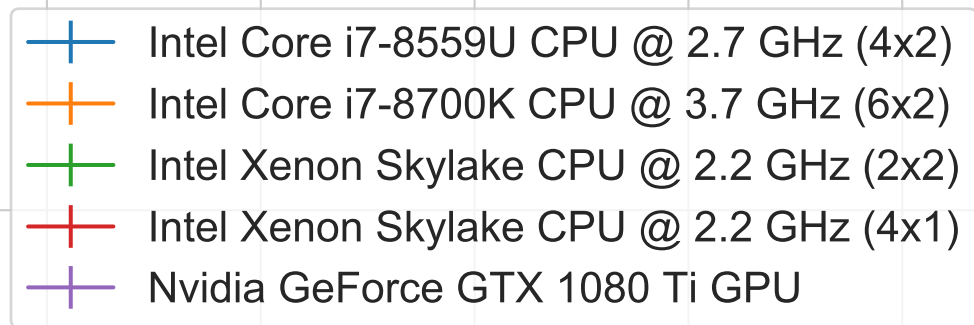


Train time

Time increase / Initial time

40
30
20
10
0



ReLU

ReLU
Ensemble

ReLU
Ensemble
Embed

ReLU
Ensemble
Embed
Fix

ReLU
Ensemble
Embed
Aug

ReLU
Ensemble
Embed
Aug
Cosine

Swish
Ensemble
Embed
Aug
Cosine

Swish
Ensemble
Embed
Aug
SWA

Swish
Ensemble
Embed
Aug
Onecycle

Swish
Ensemble
Embed
Aug
Onecycle
Dense

Solution