Curvature Loss - 400k data, CNN([LayerNorm(), CELU x2(), Linear()]) training validation 10^{-1} testing model minimum validation value epoch 10^{-2} 10^{-3} 500 1000 1500 2000 2500 3000 **Epoch**