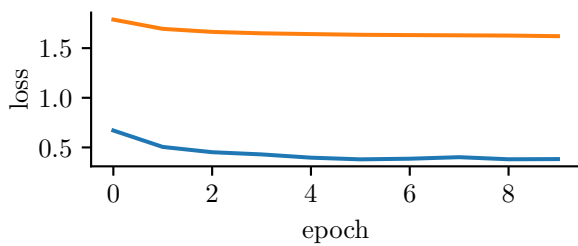


COMP6248: Lab Exercise 4

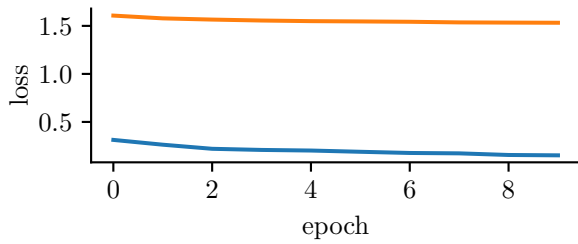
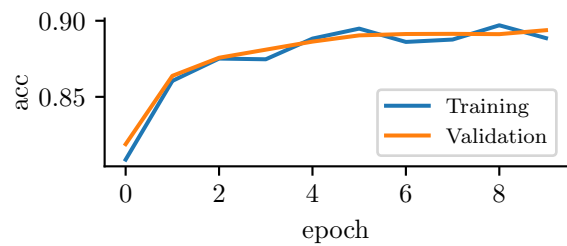
David Jones (dsj1n15@soton.ac.uk)

Task: Wide MLPs on MNIST

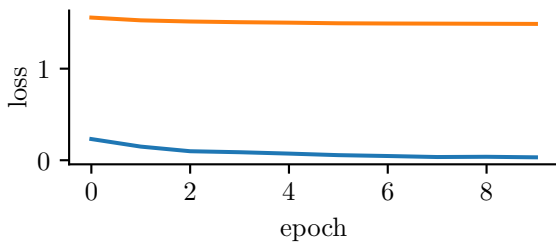
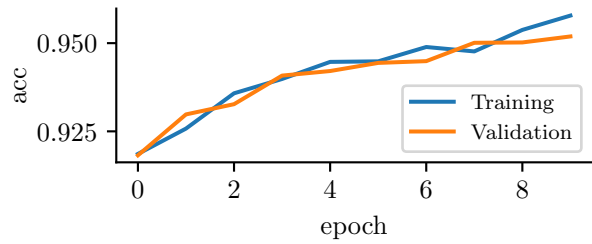
Overfitting is typically visible either by: a validation accuracy that is dramatically lower than training accuracy, or a validation accuracy that decreases as epoch count is increased. This example does not heavily exhibit either of these phenomena, likely due to the MNIST dataset being largely linearly separable and there being a small number of extractable features present. That being said, the experimental results show that around 25 hidden units the training accuracy does start to slightly diverge from validation accuracy indicating that features are being learnt from the training set that are not present in the validation set. However, the model never fully learns the training dataset even with the 250,000 hidden unit, 200 epoch model.



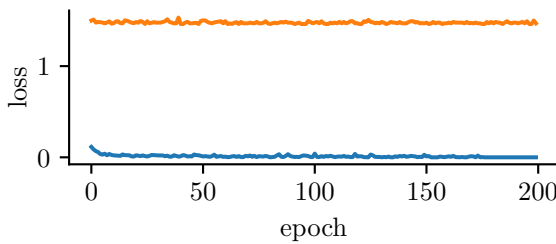
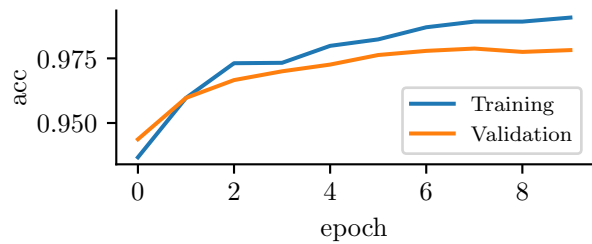
(a) Hidden Units: 5 : $\text{Acc}_{\text{test}} = 89\%$, $\text{Acc}_{\text{val}} = 89\%$



(b) Hidden Units: 25 : $\text{Acc}_{\text{test}} = 96\%$, $\text{Acc}_{\text{val}} = 95\%$



(c) Hidden Units: 200 : $\text{Acc}_{\text{test}} = 99\%$, $\text{Acc}_{\text{val}} = 98\%$



(d) Hidden Units: 250,000 : $\text{Acc}_{\text{test}} = 99\%$, $\text{Acc}_{\text{val}} = 98\%$

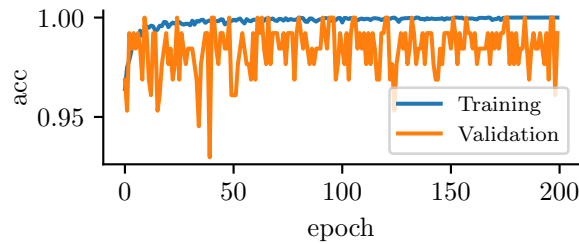


Figure 1: Accuracy and loss plots for varying hidden unit count.