Calvin AI/ML Handout | *Feb 27, 2026*   Name: _____
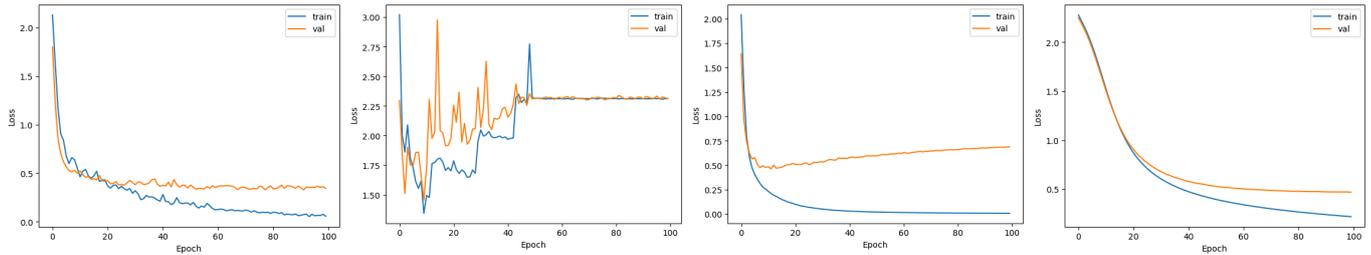
1. Fill in the blanks below with "an input feature", "a parameter value", "layer outputs", "loss value":

   A **gradient** is the slope of the tangent line on a plot where the x axis is _____ and the y axis is _____. A **weight** in a linear layer is the slope of the tangent line on a plot where the x axis is _____ and the y axis is _____.

   Sketch an example of both of these plots.

2. Here are four plots of loss by epoch for an MNIST classifier. Sort them from best to worst and explain why.
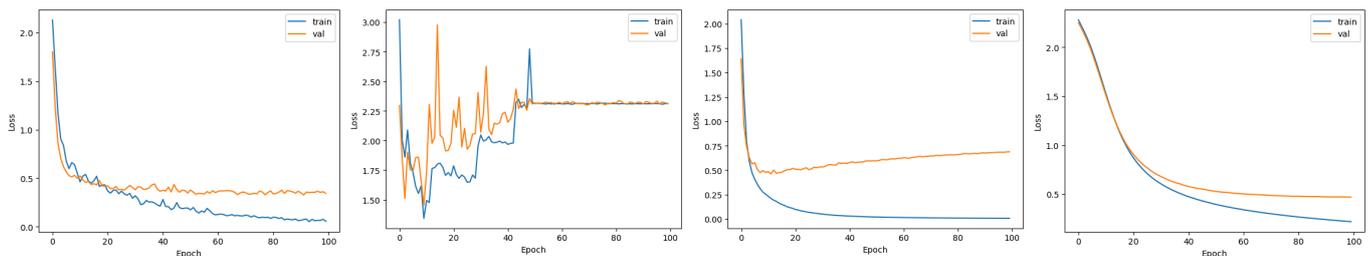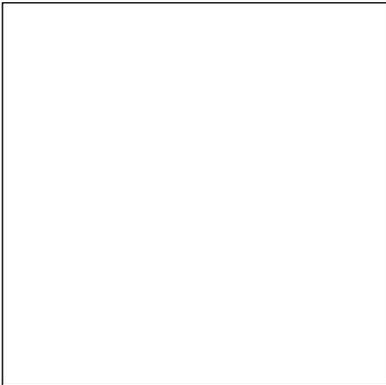

3. Consider one of the letter images you collected. List a few ways that you should / could change that image (ideally in a fully automated way / by writing a program) that would not change the letter it represents.

To think about for next time

1. Can we use accuracy as loss in SGD? Why or why not?

2. Why is "surprise" a good intuition for log-loss (categorical cross-entropy) but not MSE? What are the units of each?

3. (a) Place the following words in the box below so that similar words are closer together: *father*, *mother*, *son*, *daughter*. (b) Once you've done that, now try to add *brother*. Explain why the flat paper makes this hard; what would you do if you had another dimension to work with?