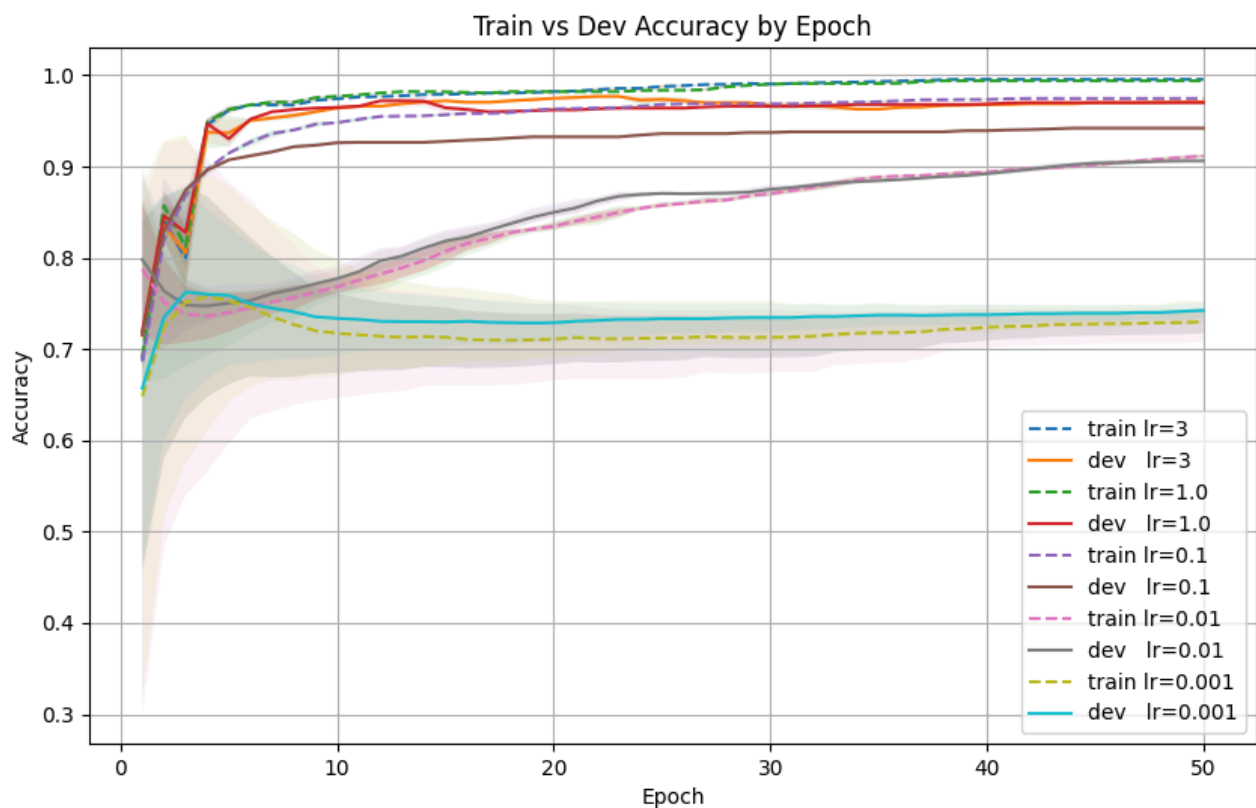


Summary

For this project, we implemented logistic regression and trained it using batch gradient descent. We tried five learning rates, ran each one five times with different random initializations, and tracked accuracy on both the training and development sets over 50 epochs. The plot produced below shows not only the mean accuracy per epoch, but also the full range (min/max) with the lighter fill and ± 1 std variability across runs with the darker fill.

1.



2. To summarize our findings, it is seen clearly in the graph that as the learning rate increases, so does accuracy and the speed at which it converges. For example, you can see that the accuracy for a learning rate of 0.001 has very little increase in accuracy as the number of epochs increases, and it plateaus around 75% for both data sets. On the contrary, the learning rate of 3 displays very rapid convergence for both data sets,

reaching over 95% accuracy before 10 epochs, and by the time we hit 50 epochs for this rate, both data sets were near 100% accuracy. Based on these findings we can conclude that increased learning rate leads to faster convergence and final accuracy, but perhaps more instability, as we can observe some up and down trends in the first few epochs in the higher learning rates, whereas in the smaller learning rates we see pretty gradual changes with smoother curves at the start, suggesting more stability. Additionally, I do not believe there to be any significant overfitting, because both the training and development curves for all learning rates are pretty close together and differ by only a couple percentage points at max. As for early stopping, we could potentially advise it for the higher learning rates 1 and 3, since a very high accuracy is reached around epochs 10-15, and the growth after that is minimal, so early stopping might be beneficial to save time and resources.

3. Krish and I worked together to write the required code needed to pass the autograder, and I (Christopher) am responsible for the non extra credit portion of this report. Krish did the initial experimental work but I was responsible for fine tuning it and making sure all required aspects of the experimentation were met. Krish will be responsible for the extra credit work and that portion of the report as well.
4. All extra code written to run experiments is in the main.py file uploaded and commented out.