Lab Exercise 1
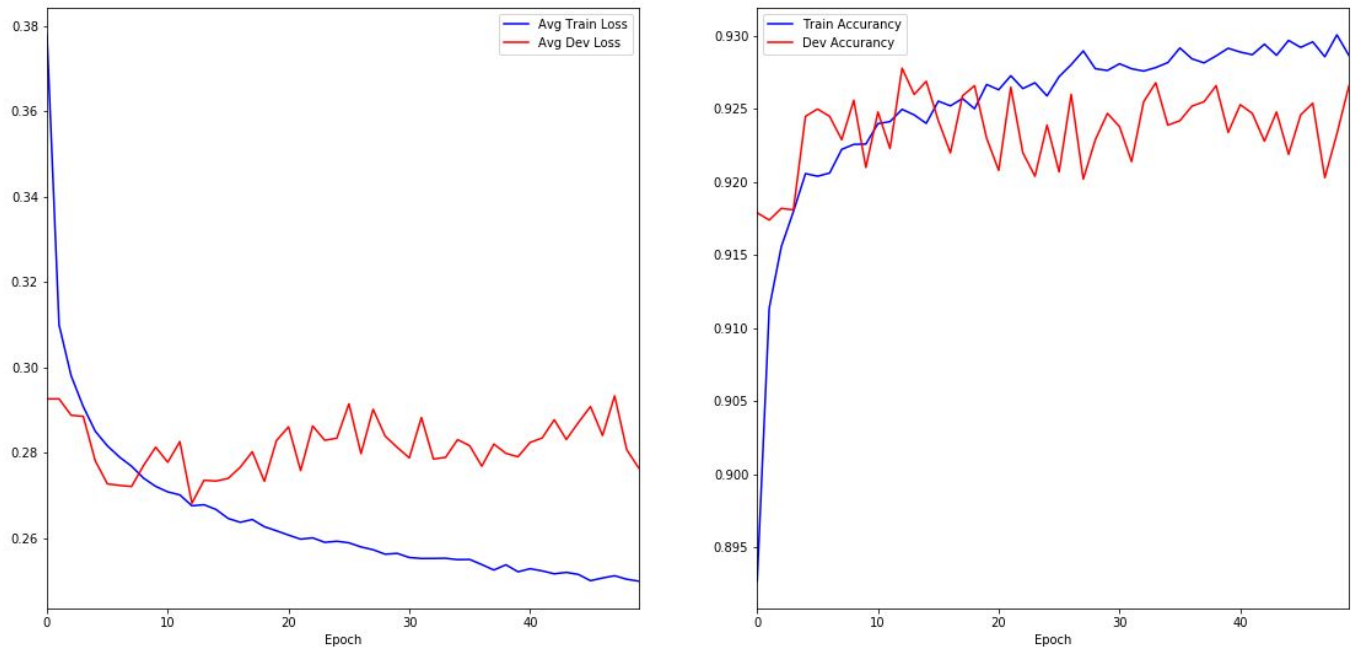
Edoardo CONTE <edo.conte96@gmail.com>
Pratham SOLANKI <pratham.c.solanki@gmail.com>

Model 1

- **Static** Learning Rate (Step Size): 0.01
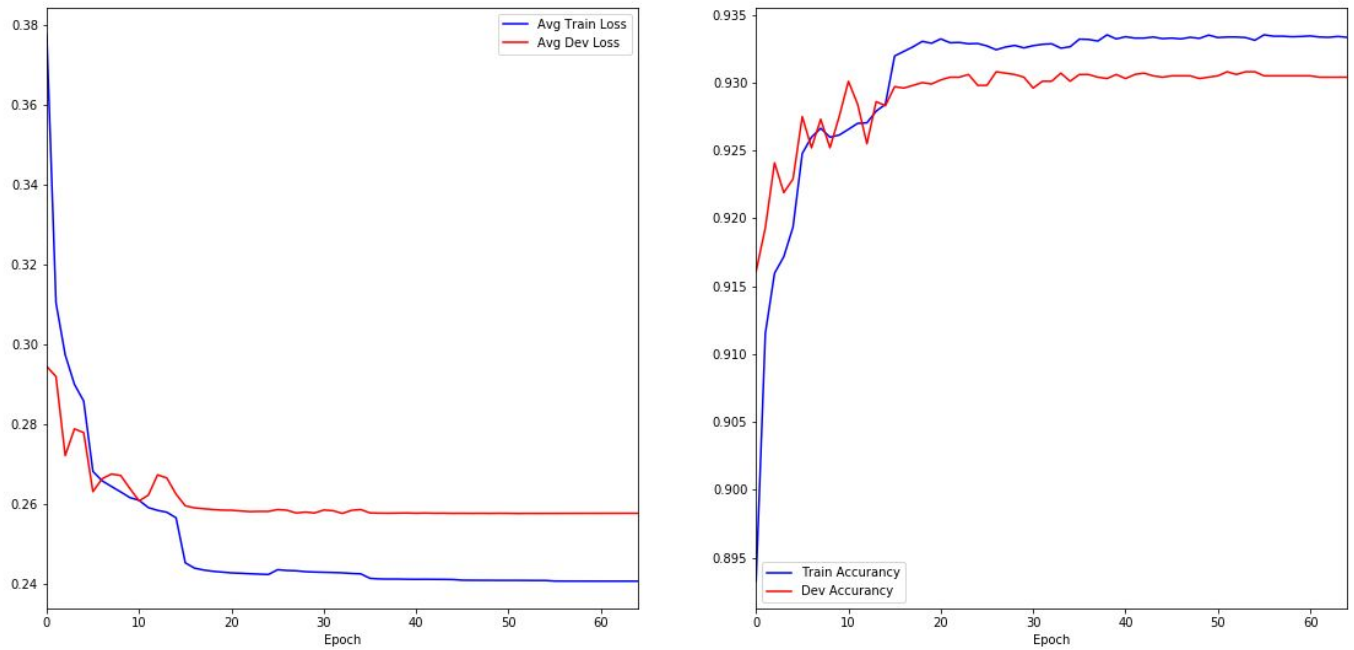- Number of Epochs: 50

Static Step Size 0.01

- Observations:
  - Consistently decreasing training loss (+ increasing training accuracy), but not similar results on the dev data. Hence overfitting.

Model 2

- **Dynamic** Learning Rate (Step Size): {0.01, 0.005, 0.001, 0.0005, 0.0001, 0.00005, 0.00001}
- Number of Epochs: 65

Dynamic Step Size between [0.01,0.00001]

- Observations:
  - Decaying learning rate works perfectly and we actually converge (negligible change in loss and accuracy over several epochs).

Comparison on Test set

Model	Accuracy	Avg. Loss
1	0.925	0.290
2	0.926	0.267

As we can see from the above results, the model with decaying learning rate performs better on the test set.