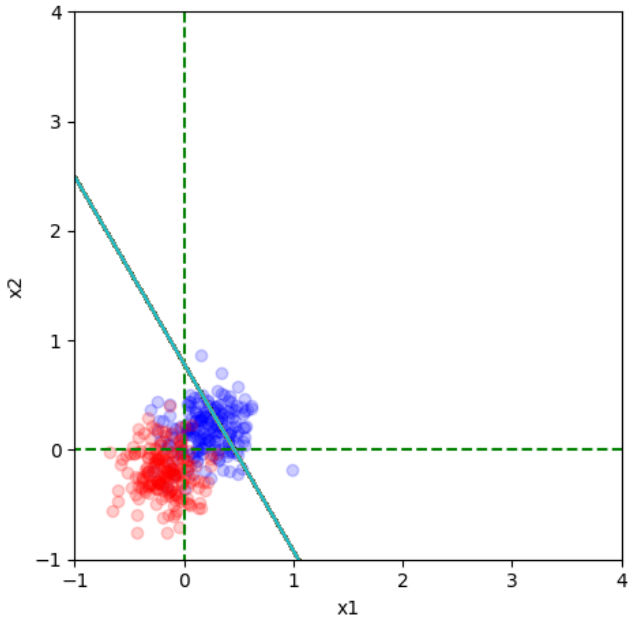
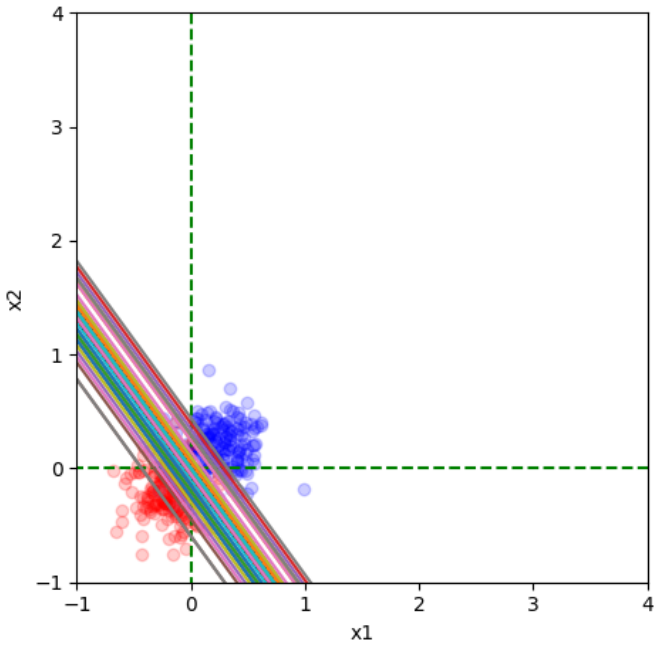
Problem 1.

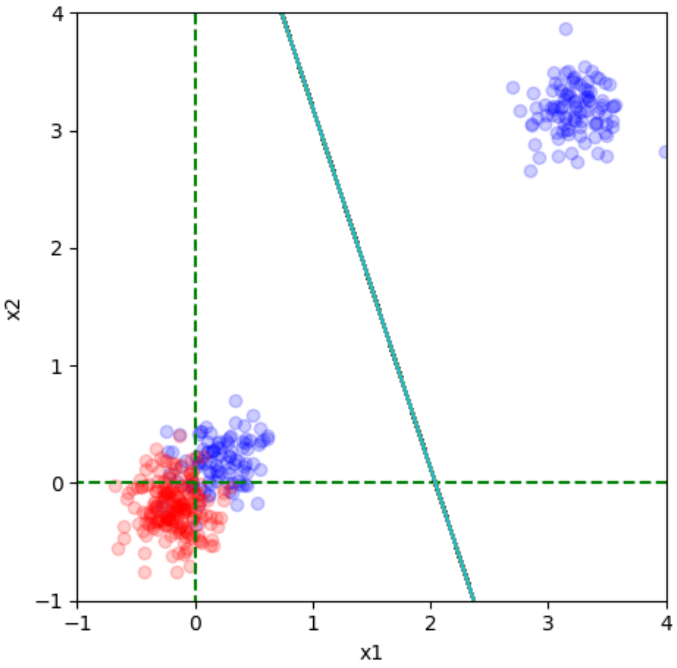
1. Training accuracy of linear regression on Dataset A :0.555



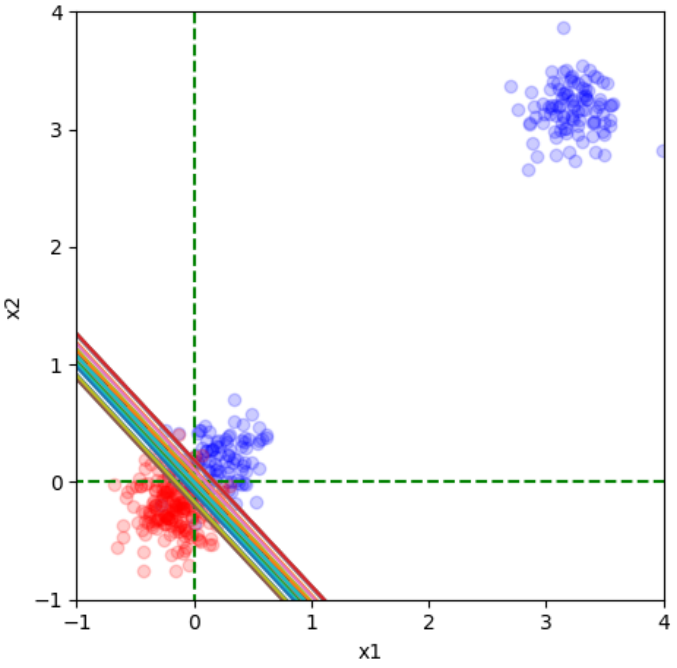
2. Training accuracy of logistic regression on Dataset A: 1.0



3. Training accuracy of linear regression on Dataset B:0.75

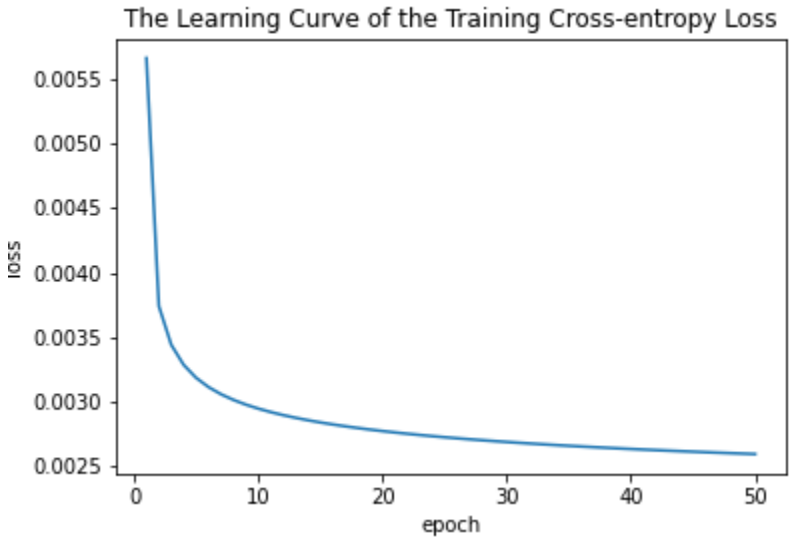


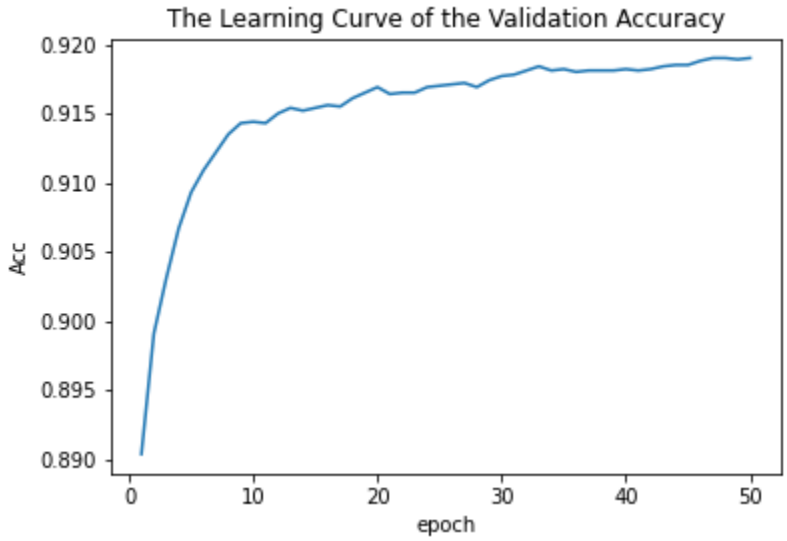
4. Training accuracy of logistic regression on Dataset B:0.9825



Problem 2:

1. The number of epoch that yields the best validation performance: 46
2. The validation performance (accuracy) in that epoch: 0.919
3. The test performance (accuracy) in that epoch: 0.9232





Question: Would adding the L2 penalty (weight decay) to Cross-entropy loss perform better in this classification task? If so, what is the number of weight decay.

Experiment: Keep all the hyperparameters unchanged, and add following code:

```
def l2_penalty(w):
    return (w**2).sum() / 2
```

And loss would be:

```
loss = cross_entropy(y, t_hat) + decay*l2_penalty(W)
```

Then, set the weight decay to {1.0, 0,1, 0,001},
See how the test accuracy is for each decay.

Conclusion: add L2 regularization has no effect on this task.
.