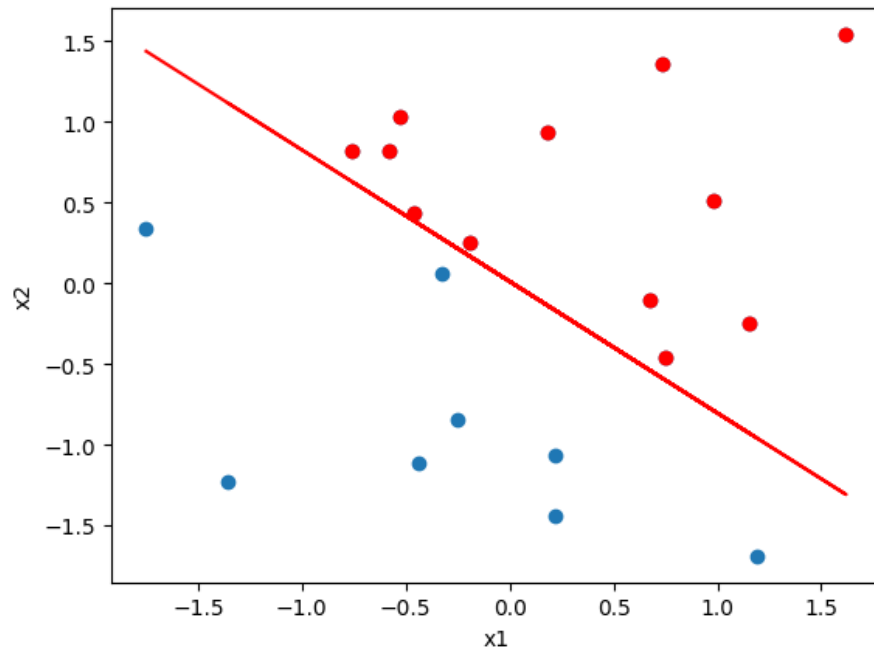
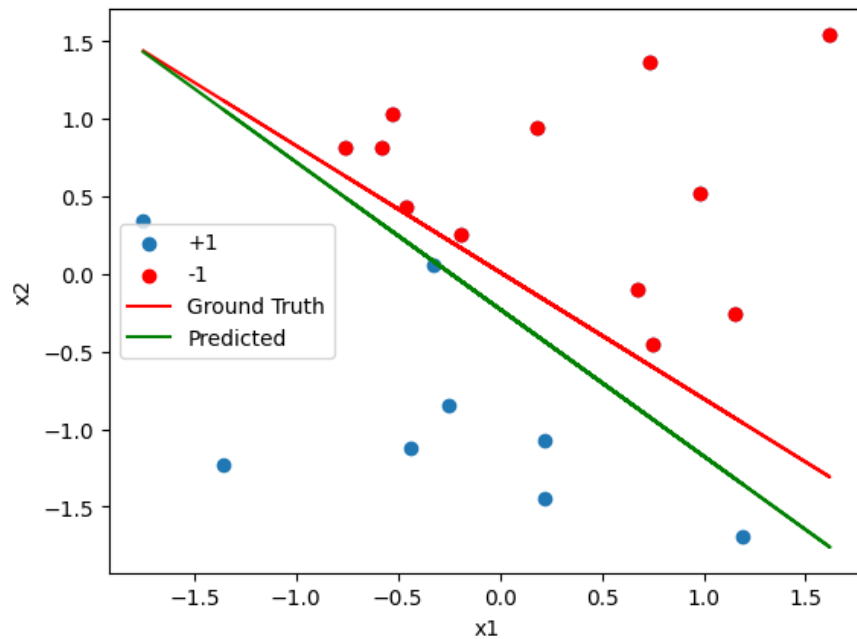


Q2 Part A:



Q2 Part B:

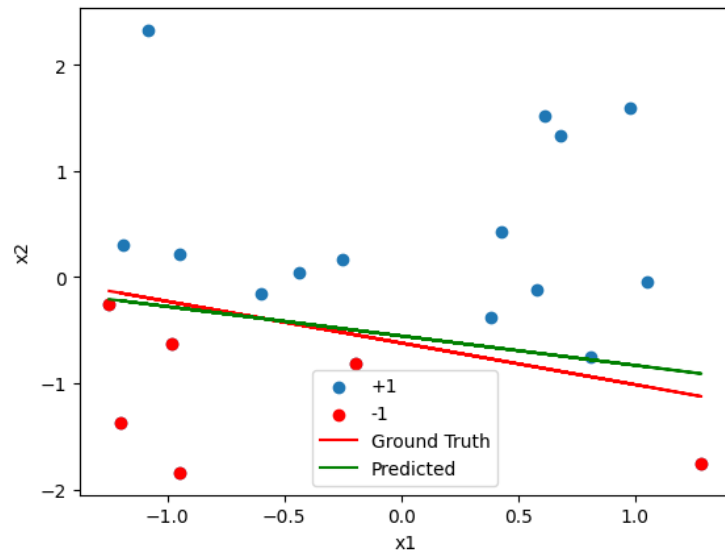


Model: [-7.36959582 -7.78309663 -1.76645938]

Steps: 57

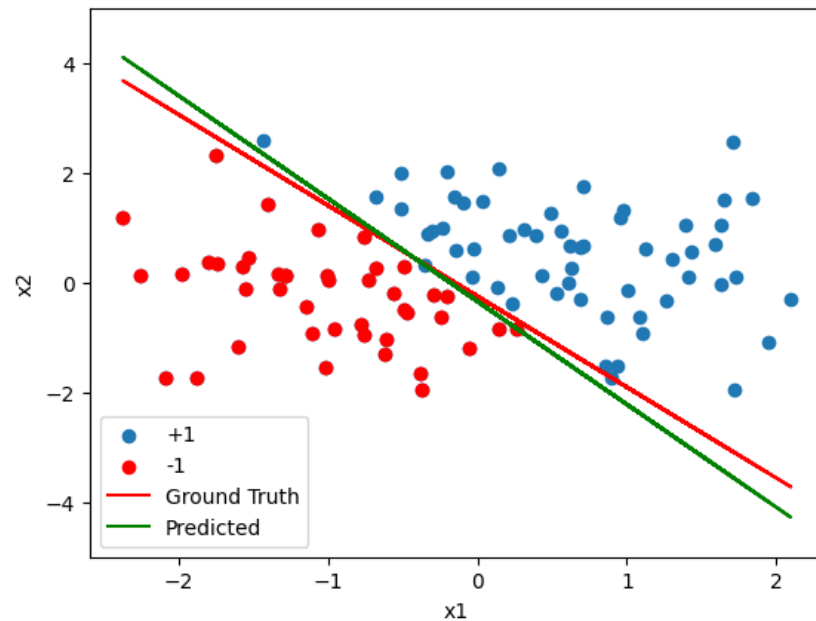
The models fit are slightly different yet both perfectly separate the data. They are about at the extremes of this.

Q2 Part C



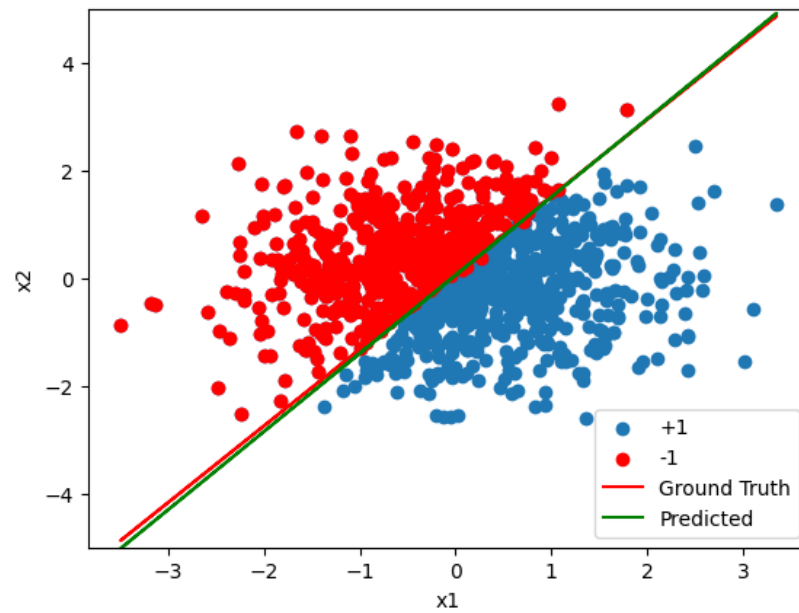
In this one the ground truth separates with decent amount of space while predicts just about touches the closest +1 and -1 points. This is different from part b but shows a similar agreement.

Q2 Part D



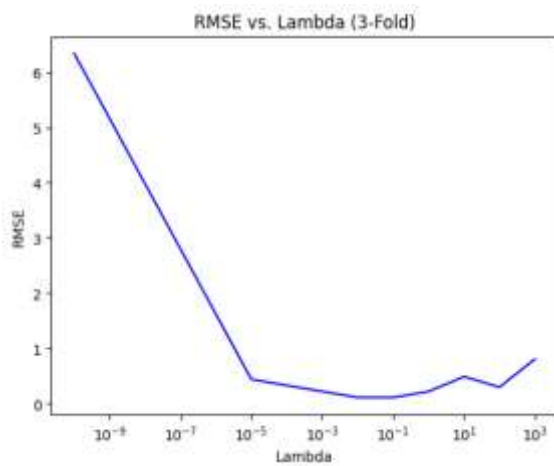
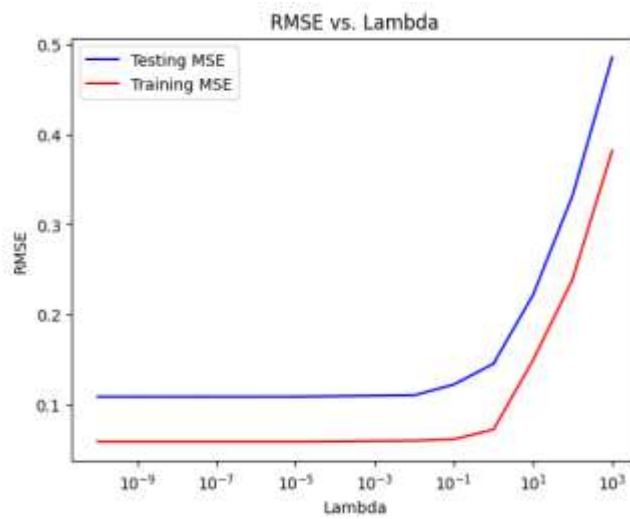
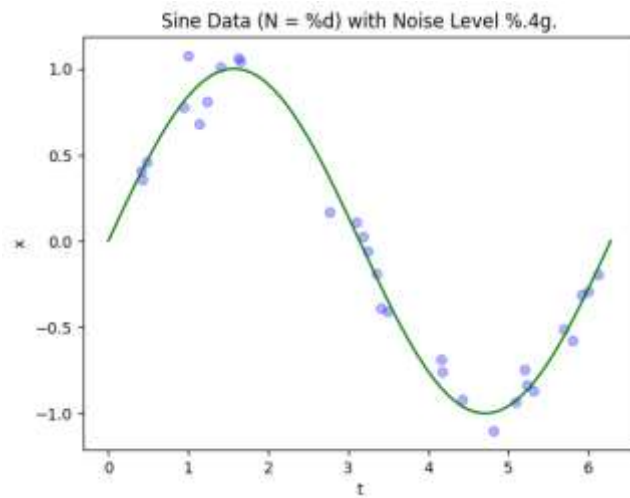
Since there are so many data points it makes the region in between the two categories smaller so the lines have to be in better agreement

Q2 Part e



Following the same logic as used in part d, the two models are almost exactly in agreement.

Q5



From the graph it appears that a lambda value of 0.1 looks best and 10^{-6} is the worst.