

1 VC Dimension

1.1 Linear

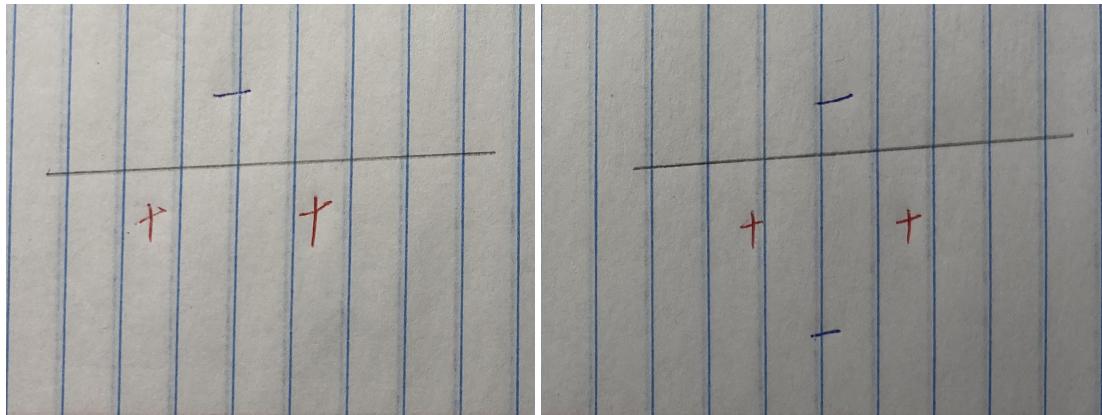


Abbildung 1: Not shattered on the left, shattered on the right.

As two plots show, adding a negative mark below two positives will shatter the linear classifier. Thus, 3 is the maximum dimension for the linear line.

1.2 Triangle

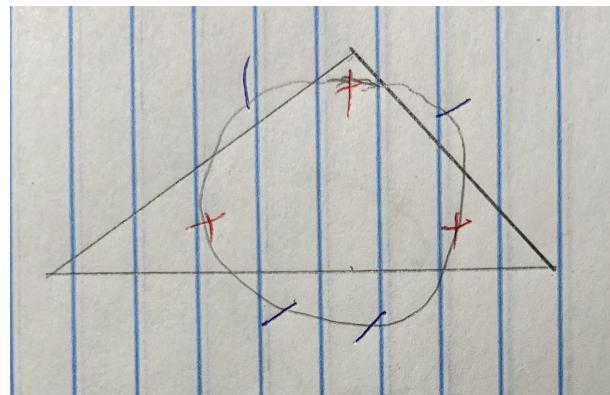


Abbildung 2: Not shattered with 7 points.

As the figure is showing, having a positive between two negatives will shatter the triangular classifier. Therefore, it will be shattered by 8 points as triangle only has three edges, it cannot contain the positive point within the shape when it is added between negatives. Thus, 7 is the maximum VC dimension of the triangular classifier.

2 Multi-armed Bandits

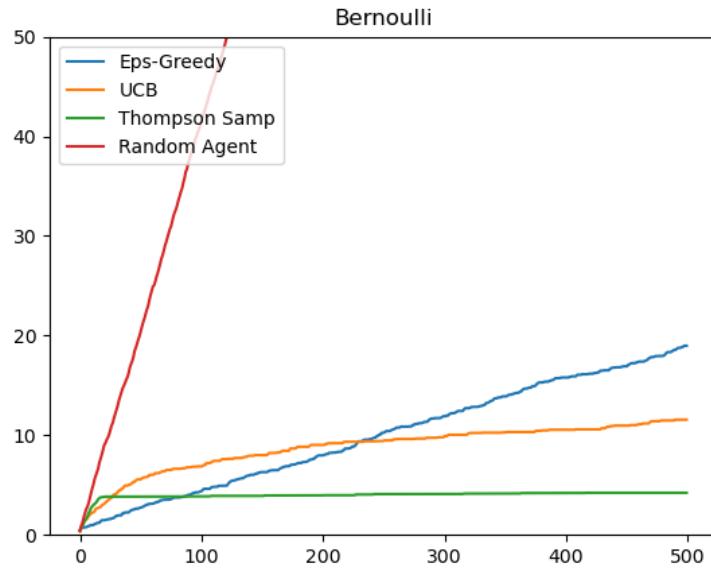


Abbildung 3: Four models under Bernoulli Env. Avg.Cumulative Regret vs. Rounds

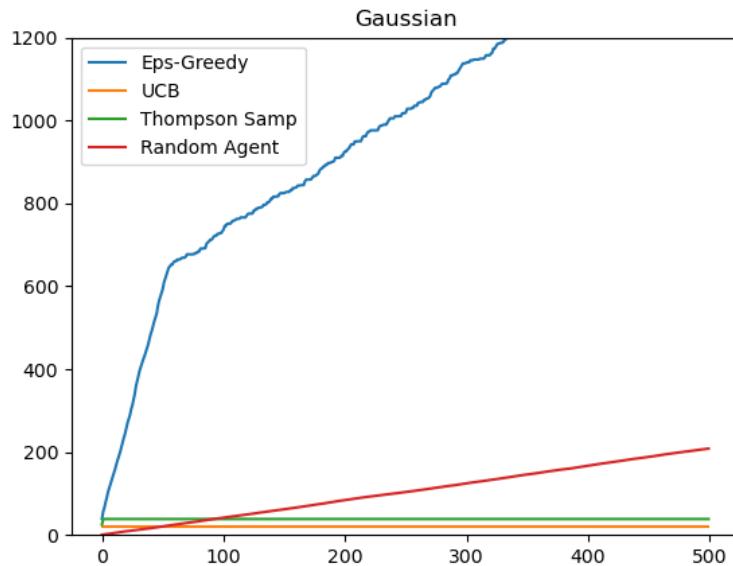


Abbildung 4: Four models under Gaussian Env. Avg.Cumulative Regret vs. Rounds

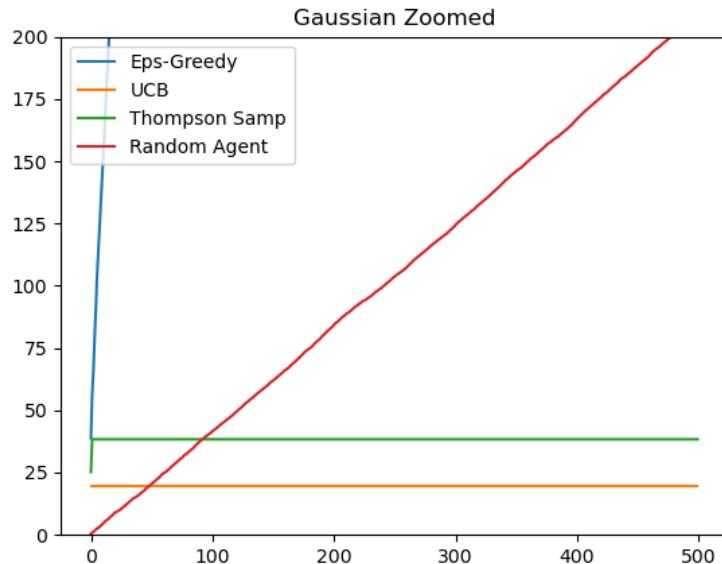


Abbildung 5: Four models under Gaussian Env. Avg.Cumulative Regret vs. Rounds

Under Bernoulli environment, all of the models showed classic curve of cumulative regret. While epsilon greedy had the linear and largest regret, ucb had second largest regret, and Thompson sampling showed least largest regret among four models.

Under Gaussian environment, models acted little bit differently. Strangely, epsilon greedy showed way higher regret. It is even higher than the random agent. While ucb acted normally, due to consistency on reward by arms than Bernoulli environment, it was able to find the optimal arm really quickly, and we can see almost a straight line of a regret on the plot. Similarly, Thompson sampling showed almost flat line of regret after few steps of lower cumulative regret on early rounds of training.