3. SVM and LOOCV ✱ PS 3 ✱

a) LOOCV: Train SVM on all samples but 1 and test on last one. Then average the errors.

SVM: $y^{(i)}(w^T x + b)$

α_i will be 0 except for 3 support vectors

$$w^T x + b = \sum_{i=1}^{m} \alpha_i y^{(i)} \langle x^{(i)}, x \rangle + b$$

Data is linearly separable => No error necessary if trained on entire set.

If support vector is left out in one CV iteration it's possible that it is misclassified

If no SV is left out but other point, then there won't be a misclassification (because optimum of convex optimization problem isn't affected by leaving out non-active constraints ie $\alpha_i = 0$)

Max errors: $|SV|$

Total runs: m

$$\Rightarrow \hat{\varepsilon}_{LOOCV} \leq \frac{|SV|}{m}$$

b) Generalizes, yes. Because data is still separable. So same argument applies!