Problem 1

You are hired by a very exclusive fashion brand to create an automated system to assess the quality of their jeans. For the reputation of the brand, it is very important that defective jeans are identified before leaving the factory and do not end up on the shop shelves. After some work, you create a model that produces the following confusion matrix:

	Predicted negative	Predicted positive
Target negative	11500	30
Target positive	10	50

1. Compute accuracy, precision and recall of the following model

2. Given the nature of the problem, what metric is most important? Why?

3. The cost of a false positive is $80 (products marked positive will need to be manually assessed by a human). The expected cost associated with a false negative is $1000 of lost profits for bad publicity. Does this model protect the client from potential losses? If not, how would you fix it?

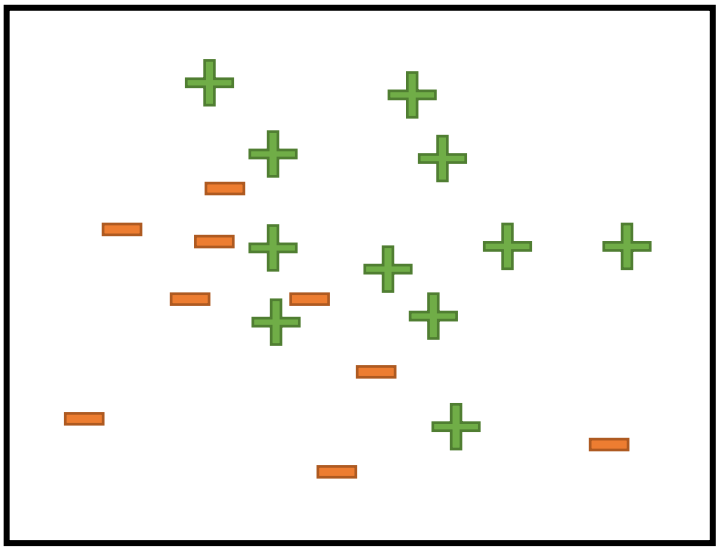
You are also asked to produce a model to predict expected sales from different stores. You know that different locations can have a large difference in profits. You decide to use a linear model, and end up with the following candidates:

	Training			Validation		
	MSE	RMSE	MAPE	MSE	RMSE	MAPE
Model 1	64224196	8014	14.3%	127531849	11293	19.1%
Model 2	192515625	13875	5.6%	211208089	14533	8.4%

What model would you choose, and why? Using the metrics, convince the customer that yours is the best choice.

Problem 2

You are working on a model to predict which students will successfully complete a computer science program. The sample distribution across 2 features looks like this:



1. On the picture, draw to the best of your capabilities:
   - The decision boundary produced by a well calibrated SVM RBF
   - The decision boundary produced by SVM RBF with too high values of C and gamma
   - The decision boundary produced by a linear SVM

2. SVM RBF is the most promising model, and you want to do a good job tuning the hyperparameters C and gamma. However, the model needs to be ready in 1 week, and every training/validation session takes 3 hours. How many different values of C and gamma could you test using grid search? Could you do better with randomized search?

3. After you are done with the tuning of the hyperparameters, you find a combination that seems to yield much better accuracy than any other. Are you confident that picking this combination will yield the best results on unseen data? Why or why not?