*(b)* ***(2 points****) Do you intuitively think that one boundary is better than another? It may be possible to use such an intuition to invent method that uses multiple learning algorithms and combine the results, using your intuition as a prior probability. Explore this line of thought. This is a difficult question, but try to be creative!*

We see that the 1-nearest neighbour classifier and the decision tree classifiers do not misclassify instances, whereas plain and logistic regression do misclassify some instances.

We see however that quadratic logistic regression misclassifies only one instance whereas plain logistic regression misclassifies two instances. We see that both of the points that were misclassified in the plain logistic regression lie on the sides, therefore quadratic logistic regression would not overfit the data.

Although that 1-nearest neighbour classifier and decision trees do not misclassify instances, they are high varience classifiers, meaning that slight changes in the training set change the decision boundary a lot, thus that the decision boundaries most likely will generalize less well on the test set.

Therefore it would be a good idea to combine these classifiers to obtain a classifier that and misclassifies no or little instances and generalizes well on the test set.

An option what we could do is take the “average” of the Knn classifier and the quadratic logistic regression classifier. However, there might be a better option.

What we could also do, is first make “preprocess” our data, using logistic regression. We then get x classes, say x=2, so we get in this example two classes. We know that there might be instances that are being misclassified, so we create a confusing matrix.
We leave the instances that are correctly classified, so the true positives and true negatives. On the false positives and false negatives we run our decision tree classifier. As the decision tree classifier classifies every instance correctly (but has high varience), all these initially wrongly classified instances are now correctly classified.

The combination of these two classifiers makes sure that at least a big part of our classification is not sensitive to high variance, since the largest part is being classified using logistic regression. Only for the parts that logistic regression cannot classify, we use our decision tree classifier. So we have a classifier who is not very much overfitting the data and does not misclassify instances