**Question 1**

Rahul built a logistic regression model with a training accuracy of 97% and a test accuracy of 48%. What could be the reason for the gap between the test and train accuracies, and how can this problem be solved?

This may be due to complexity of model and overfitting. This can be resolved by making model simple and generic. Care should also be taken that model not very simpler.There should always be a trade off between variance and Bias.

**Question 2**

List at least four differences in detail between L1 and L2 regularisation in regression.

Lasso regression uses L1 regularization. Ridge uses L2 regularization.

In L1 regularization sum magnitude of coefficients is added to error tem to prevent the model becoming more complex. In L2 sum of Squares of magnitudes are added.

The coefficients that are significant in Lasso becomes zero and in ridge they are near to zero but not zero.

Lasso helps in feature selection as it shrinks the insignificant coefficients while ridge does not.

**Question 3**

Consider two linear models:

L1: y = 39.76x + 32.648628

And

L2: y = 43.2x + 19.8

Given the fact that both the models perform equally well on the test data set, which one would you prefer and why?

Both have similar degree and complexity wise also looks similar.

I prefer L1 as the Y value results in small value for same X that we substitute in given equations.

Even though the constant value is high in L1 the coefficient of X is less which decreases whole Y value when multiplied with any positive X value.

**Question 4**

How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?

The model should be simple and should not be over fit. The model should never fail on test data after performing well on training data. The model should never have chance of access to test data while learning. Always use regularization methods and hyper parameters to prevent model becoming highly complex. To make model robust it should not be very simple or very complicate to get overfit on training data. There should be a trade off between Bias and variance.

**Question 5**

You have determined the optimal value of lambda for ridge and lasso regression during the assignment. Now, which one will you choose to apply and why?

I choose lasso regression. In ridge regression the values are alpha is near zero. When it is 0 it is like normal linear regression and error term is more significant which can make model more complex.

In Lasso the alpha is around 100 and it also acts as feature selection method by shrinking the coefficients to zero which makes our work easy in identifying coefficients that can be used in building model.