Question 1

a. What does linear regression try to optimize?

Linear Regression tries to optimize the fit between predicted values and actual values of a continuous variable, by minimizing the Mean Squared Error (MSE) cost function.

b. Is it possible to use linear regression to represent quadratic equations? Explain with an example.

Yes, it is possible to use linear regression to represent quadrating equations.

Example below:

$y = w0 + w1.x + w2.t$ where $t = x^2$

Even if the above equation is not linear in the variable x, it is linear in the coefficients w0, w1, w2 and hence possible to use linear regression technique.

c. Why is it crucial to detect and remove outliers?

It is crucial to detect and remove outliers as they can otherwise impose undesired influence on the parameters of the distribution, thereby negatively effecting the ability to draw accurate and reliable statistical analysis and prediction models from the same.

d. **What is feature scaling? When is it required?**

Feature scaling is the process of normalizing the range of independent variables/features. It is needed when the features in a dataset have different units or vastly different ranges, to ensure each feature contributes fairly to the model's learning process.

e. **State two differences between linear regression and logistic regression.**

Linear regression helps predict a continuous variable, with an outcome in the form of a linear equation to predict the dependent variable and is suitable for supervised learning problem in ML. On the other hand, Logistic regression helps predict a discrete variable (rather probability of possible outcome, often binary) and is suitable for classification problem in ML. Linear regression minimizes Mean Square Error cost function while logistic regression minimizes Cross Entropy Loss function

f. Why is the Mean Square Error cost function unsuitable for logistic regression?

Mean Square Error cost function is unsuitable for logistic regression, as it does not align with the probabilistic nature of the logistic regression model. That is, the desired outcome for logistic regression is probability ranging from 0 to 1 while MSE may produce values outside this range. Also, MSE poses convergence issues in the context of logistic regression.

g. What can be inferred if the cost function initially decreases but then increases or gets stuck at a high value?

The value chosen for learning rate alpha is possibly too high or too low and needs to be adjusted.

h. Describe two ways to perform multi-class classification using logistic regression.

One-Vs-All / One-Vs-Rest and One-Vs-One are two possible ways to perform multi-class classification using logistic regression

Question 2

Answer:

	w0	w1	w2
After Iteration 1	0.0182	2.3056	1.3394
After Iteration 2	0.0167	2.2223	1.3004

Initial Mean Squared Error	8834.6
Final Mean Squared Error	303.1813