Assignment - 03

Q1

a. Linear regression tries to optimize the coefficients of the linear equation to minimize the difference between predicted and actual values, typically measured by Mean Squared Error (MSE). It finds the best-fitting line to predict the target variable.

b. Yes, Linear regression can represent quadratic equations by including polynomial terms (e.g., x2x^2x2) as features.
For example, the equation y=ax2+bx+cy = ax^2 + bx + cy=ax2+bx+c can be modeled by linear regression with features xxx and x2x^2x2.

c. To detect and remove outliers is crucial because they can significantly skew the model's estimates and lead to inaccurate predictions. Outliers can dominate the model's optimization process, resulting in poor generalization.

d. Feature scaling (or normalization) transforms features to a common range, usually between 0 and 1, to prevent features with large ranges from dominating the model. It's required when using algorithms sensitive to feature scales, like gradient descent.

e.

Point	**Linear Regression**	**Logistic Regression**
Cost function	Mean Squared Error (MSE)	Logistic/Log loss function
Target variable	Predicts continuous outcomes/variables	Predicts categorical outcomes (binary classification)

f. The Mean Square Error cost function is unsuitable for logistic regression because it's designed for continuous outputs, whereas logistic regression predicts probabilities (0 to 1).
Log loss is a better fit, as it measures the difference between predicted probabilities and actual classes.

g. If the cost function initially decreases but then increases or gets stuck at a high value, it may indicate overfitting, underfitting, or poor model initialization. It can also suggest issues with the optimization algorithm or learning rate.

h. Two ways to perform multi-class classification using logistic regression are as follows:

- One-vs-Rest (OvR): Train a separate binary classifier for each class, treating it as a positive class and all others as negative.

- Softmax Regression (Multinomial Logistic Regression): Extend logistic regression to handle multiple classes directly by modeling the probabilities of each class with a softmax function.

Q2

	w0	w1	w2
Initial	0	1	1
After Iteration 1	0.0182	2.3056	1.3394
After Iteration 2	0.0167	2.2223	1.3004

When using the formulae with 1/m :-

Initial Mean squared error	8834.6000
Final Mean squared error	303.1812

When using the formulae with 1/2m :-

Initial Mean squared error	4417.3000
Final Mean squared error	151.5906