UNIT 4 ASSIGNMENT

Introduction to Linear Models

## Instructions

The questions below will prepare you for future interviews as they relate to concepts discussed throughout the unit. You’ve practiced these concepts in the coding activities, exercises and coding portion of the assignment. Now, let’s formulate your programming into well-thought responses.

Except as indicated, use this document to record all your assignment work and responses to any questions. At a minimum, you will need to turn in a digital copy of this document to your facilitator as part of your assignment completion. You may also have additional supporting documents that you will need to submit. Your facilitator will provide feedback to help you work through your findings.

**Note:** Though your work will only be seen by those grading the course and will not be used or shared outside the course, you should take care to obscure any information you feel might be of a sensitive or confidential nature.

*Begin your assignment by completing the questions below. Directions to submit your work can be found on the assignment page. Information about the grading rubric is available on any of the course assignment pages online. Do not hesitate to contact your facilitator if you have any questions about the assignment.*

Unit 4 Written Portion

# Logistic Regression

Answer the questions below about linear models.

## Questions:

1. What is a linear model? What are the advantages and disadvantages of linear models?

|  |
| --- |
| A linear model assumes a linear relationship between features and label. The model is good with small data sets, simple, fast to train/implement/predict, and high interpretability. However, it would not fit for non-linear or complex relationships, and could be sensitive to outliers. |

1. What type of supervised learning problem is logistic regression best suited for? Give an example of a problem you would use a logistic regression model for. Explain what you are trying to predict.

|  |
| --- |
| Predict whether an online customer will purchase the items in their cart. 1 indicates the customer will buy the products, and 0 means they will not. The features would be age, stay time, search history, coupon/promotion existence, etc. |

1. Describe the training phase of a logistic regression model: explain the intuition behind using gradient descent algorithm and the use of loss functions.

|  |
| --- |
| The training phase of a logistic regression model involves finding the optimal weights and intercept that minimize prediction error by using an optimization algorithm like gradient descent. Gradient descent iteratively updates the model parameters by calculating the gradient of the loss function and adjusting the parameters in the direction that reduces this loss, ensuring improved prediction accuracy. |

1. Explain the purpose of using regularization when training a logistic regression model.

|  |
| --- |
| We use regularization to prevent overfitting by adding a penalty term to the loss function. Regularization discourages the model from relying too heavily on any single feature or from having excessively large weights. |

1. Explain which linear model and accompanying loss function you would use for a classification problem and for a regression problem.

|  |
| --- |
| I would use logistic regression and binary cross-entropy loss function for a classification problem because the model prediction would be between 0 and 1 and loss function punishes deviations from the true class labels. For a regression problem, I would use linear regression and mean squared error because it predicts continuous values, and MSE minimizes the squared differences between these predictions and the actual target values. |

*To submit this assignment, please refer to the instructions in the course*.