Regression Assignment - 1

February 22, 2024

[ ]:
```
"""Q1. Explain the difference between simple linear regression and multiple␣
 ↪linear regression. Provide an example of each.

    Ans: Simple linear regression is a statistical method that examines the␣
 ↪relationship between two continuous variables.
        It involves one independent variable and one dependent variable, and␣
 ↪aims to find a linear relationship between them.
        An example of simple linear regression is analyzing the relationship␣
 ↪between a person's height and weight.
        Multiple linear regression, involves examining the relationship␣
 ↪between multiple independent variables and one dependent variable.
        It aims to find a linear relationship between the independent␣
 ↪variables and the dependent variable.
        An example of multiple linear regression is analyzing the relationship␣
 ↪between a person's salary and their level of education, years of experience,␣
 ↪and age.
"""
```

[ ]:
```
"""Q2. Discuss the assumptions of linear regression. How can you check whether␣
 ↪these assumptions hold in a given dataset?

    Ans: Linear regression assumes that there is a linear relationship between␣
 ↪the independent and dependent variables, the residuals are normally␣
 ↪distributed,
        the variance of the residuals is constant across the range of the␣
 ↪independent variable, and there is no multicollinearity among independent␣
 ↪variables.
        These assumptions can be checked by analyzing residual plots, Q-Q␣
 ↪plots, and correlation matrices.
"""
```

[ ]:
```
"""Q3. How do you interpret the slope and intercept in a linear regression␣
 ↪model? Provide an example using a real-world scenario.

    Ans: In a linear regression model, the slope represents the rate of change␣
 ↪of the dependent variable with respect to the independent variable,
```

```
        while the intercept represents the value of the dependent variable␣
↪when the independent variable is zero.
        For example, in a model predicting house prices based on square␣
↪footage, the slope represents the increase in price per square foot,
        and the intercept represents the base price of a house with zero␣
↪square footage.
"""
```

[ ]:
```
"""Q4. Explain the concept of gradient descent. How is it used in machine␣
↪learning?

   Ans: Gradient descent is a method used in machine learning to find the best␣
↪values for the parameters of a model by iteratively adjusting them to␣
↪minimize
        the error between predicted and actual outcomes. It does this by␣
↪calculating the gradient of the error function with respect to the␣
↪parameters and
        taking small steps in the direction of steepest descent until it␣
↪reaches a minimum.
"""
```

[ ]:
```
"""Q5. Describe the multiple linear regression model. How does it differ from␣
↪simple linear regression?

   Ans: Multiple linear regression is a statistical method that helps us␣
↪understand the relationship between multiple independent variables and a␣
↪dependent variable.
        Unlike simple linear regression, which only uses one independent␣
↪variable, multiple linear regression considers several independent variables␣
↪that can impact
        the dependent variable.
"""
```

[ ]:
```
"""Q6. Explain the concept of multicollinearity in multiple linear regression.␣
↪How can you detect and address this issue?

   Ans: Multicollinearity is a problem in multiple linear regression where the␣
↪independent variables are highly correlated with each other.
        This can cause issues in accurately estimating the effect of each␣
↪independent variable on the dependent variable.
        To detect and address multicollinearity, one can use correlation␣
↪matrices and variance inflation factors (VIF) to identify highly correlated␣
↪variables and
        remove them or use regularization techniques to reduce their impact.
"""
```

[ ]:
```
"""Q7. Describe the polynomial regression model. How is it different from␣
↪linear regression?

   Ans: Polynomial regression is a type of regression analysis where the␣
↪relationship between the independent variable (X) and the dependent variable␣
↪(Y)
        is modeled as an nth degree polynomial. It differs from linear␣
↪regression as it allows for a curved relationship between X and Y, rather␣
↪than assuming a straight line.
"""
```

[ ]:
```
"""Q8. What are the advantages and disadvantages of polynomial regression␣
↪compared to linear regression? In what situations would you prefer to use␣
↪polynomial regression?

   Ans: Advantages:
           1. Polynomial regression can capture non-linear relationships␣
↪between variables that linear regression cannot.
           2. Polynomial regression can provide a better fit to the data than␣
↪linear regression.
        Disadvantages:
           1. Polynomial regression can be more complex than linear␣
↪regression, which can lead to overfitting if the model is not carefully␣
↪tuned.
           2. Polynomial regression can be sensitive to outliers.
        Situations where polynomial regression is preferred:
           1. When there is a non-linear relationship between the dependent␣
↪variable and independent variable(s).
           2. When a higher degree of accuracy is required in the predictions.
           3. When there is enough data to fit a complex model without␣
↪overfitting.
"""
```