Reinforcement Learning?-homework

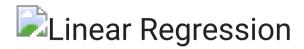
DAY2:LINEAR ** REGRESSION:

15/6/2024

1. What is Linear Regression?

- Answer: Linear regression is a supervised ml algortihm used to model the relationship between a dependent variable (target) and one or more independent variables (predictors).
- The goal is to find the best-fitting line through the data points that predicts the target variable.

Linear Regression Example



2. What are the Assumptions Made by Linear Regression?

- Answer:
- Linearity: The relationship between the independent variable(s) (X) and the dependent variable (Y) is linear.
- Homoscedasticity: The variance of the residuals (errors) is constant for any value of the independent variable(s).
- Independence: The observations are independent of each other.
- Normality: For any fixed value of the independent variable(s), the dependent variable is normally distributed.

3. Explain the mathematical intuition of linear regression

- Answer:
- Imagine we have a bunch of dots on a graph. Each dot represents something we want to predict (like house prices=y) and something that helps us predict it (like house size=x).
- Linear regression is like drawing a straight line through these dots.
 This line helps us make predictions for new data points. The goal is to find the line that best fits all the dots.

- Mathematically, this line is described by an equation:
- [Y = beta_0 + beta_1X]
- Here,
- (Y) is what we want to predict (like house price).
- (X) is what helps us predict (like house size).
- (\beta_0) is where the line crosses the Y-axis (the intercept).
- (\beta_1) is how steep the line is (the slope).

By finding the best (\beta_0) and (\beta_1), we can draw a line that predicts (Y) as accurately as possible based on (X). So, linear regression is about finding this optimal /best fit line to understand and predict relationship

4. Advantages of linear regression?

- Linear regression performs
 exceptionally well for linearly
 separable data
- Easy to implement and train the model

5. Disadvantages of linear regression?

- Sometimes Lot of Feature
 Engineering Is required
- If the independent features are correlated it may affect performance
- It is often quite prone to noise and overfitting

6.Gradient Descent:

- Gradient Descents an optimization algorithm commonly used in machine learning, including in the context of linear regression, to find the optimal parameters (coefficients) of a model by minimizing a cost function.
- In linear regression, gradient descent
 adjusts the slope and intercept of the
 regression line to minimize the
 difference between predicted and actual
 values. It iteratively improves the model's
 accuracy by moving towards the
 parameters that reduce prediction errors.
- Steps Involved in Gradient Descent:
- Initialization: Start with initial values for the coefficients (often set randomly or to zero).

- Calculate Error: Compute the difference between the predicted values (using current coefficients) and the actual values (from the dataset).
- Cacluate the Gradient: Calculate the gradient (derivative) of the cost function with respect to each coefficient. The gradient indicates the direction of steepest ascent of the cost function.
- Update Coefficients: Adjust the coefficients in the opposite direction of the gradient to minimize the cost function. This involves multiplying the gradient by a learning rate (α), which determines the size of the steps taken towards the minimum.

 Iterate: Repeat steps 2-4 until the algorithm converges to a minima.
 Convergence is achieved when the changes to the coefficients are minimal or when a predetermined number of iterations is reached.

Cost Function vs Residual Error: Explanation with Example

Residual Error:

- Definition: Residual error, also known simply as residuals, difference between actual data points and predicted data points
- Example: Suppose we have a linear regression model that predicts house prices based on their size. For a particular house:

- Observed house price (actual value):
 300 Predictedhouseprice(free 280)
- Residual error = Actual value Predicted value = 300-280 = \$20
- Ideally, we want residuals to be as small as possible, indicating accurate predictions.

Cost Function:

- Definition: A cost function (or loss function) measures the difference between predicted values by the model and actual values across all data points.
- Example: Consider we have multiple
 houses in our dataset. The cost function
 aggregates the residuals for all houses to
 compute the overall model performance.
 - Suppose we have three houses:

- House 1: Actual price
 300,000, Predicted price 280,000 (Residual \$20,000)
- House 2: Actual price
 400, 000, Predictedprice
 390,000 (Residual \$10,000)
- House 3: Actual price
 250, 000, Predictedprice
 260,000 (Residual -\$10,000)
- Cost function (Mean Squared Error, for example) = Average of squared residuals:
 - MSE = (\frac{1}{3}[(20,000)^2 + (10,000)^2 + (-10,000)^2])
- It provides a single number that represents the model's performance across all predictions.

DAY2 END