**Interview Question**

1. What is Normalization & Standardization and how is it helpful?

Ans:

**Normalization**: Normalization rescales the values of numeric features to a common scale without distorting differences in the ranges of values. It typically involves scaling the values of a feature to a range between 0 and 1.

**Standardization**: Standardization transforms the features of a dataset to have a mean of 0 and a standard deviation of 1. It centers the data around 0 and scales it by the standard deviation.

**How they are helpful**:

1. **Improved performance of machine learning algorithms**: Many machine learning algorithms perform better when the features are on a similar scale or have a standard normal distribution. Normalization and standardization help achieve this, improving the performance and convergence of these algorithms.
2. **Reduced computational complexity**: Scaling the features to a common scale can reduce the computational complexity of some algorithms, especially those sensitive to the scale of features. For instance, distance-based algorithms like k-nearest neighbors (KNN) can give biased results if features are not scaled.
3. **Easier interpretation**: Normalized and standardized features are often easier to interpret because they are on a common scale, making it easier to compare the effects of different features on the model's output.
4. **Regularization**: In regularization techniques like ridge regression and LASSO, normalization and standardization can help ensure that all features are penalized equally, preventing the model from being biased towards features with larger magnitudes.
5. What techniques can be used to address multicollinearity in multiple linear regression?

Ans

1. **Feature Selection**:
   * Remove one or more of the highly correlated independent variables from the model. This approach reduces the complexity of the model and eliminates the multicollinearity issue.
   * Use domain knowledge or statistical techniques (e.g., stepwise regression, forward selection, backward elimination) to select the most relevant features and exclude redundant ones.
2. **Principal Component Analysis (PCA)**:
   * PCA is a dimensionality reduction technique that transforms the original variables into a new set of uncorrelated variables called principal components.
   * By retaining only a subset of principal components that explain most of the variance in the data, multicollinearity can be reduced while preserving most of the information.
3. **Ridge Regression**:
   * Ridge regression adds a penalty term to the ordinary least squares (OLS) estimation to shrink the coefficients towards zero.
   * This penalty term helps to stabilize the coefficients and mitigate the effects of multicollinearity, especially when there are many correlated predictors.
4. **LASSO (Least Absolute Shrinkage and Selection Operator) Regression**:
   * LASSO regression also adds a penalty term to the OLS estimation but uses the 𝐿1*L*1​ norm of the coefficients as the penalty term.
   * LASSO tends to set some coefficients to exactly zero, effectively performing variable selection and reducing the impact of multicollinearity.
5. **Elastic Net Regression**:
   * Elastic Net combines the penalties of ridge regression and LASSO, allowing for a more flexible regularization approach.
   * It addresses multicollinearity by shrinking the coefficients and performing variable selection simultaneously.
6. **Variance Inflation Factor (VIF)**:
   * Calculate the VIF for each independent variable, which measures how much the variance of an estimated regression coefficient is increased due to multicollinearity.
   * Variables with high VIF values (typically above 5 or 10) are considered to be highly collinear and may need to be addressed using one of the aforementioned techniques.
7. **Data Collection and Preprocessing**:
   * Collect more data if possible to reduce the impact of multicollinearity.
   * Standardize or normalize the variables to a common scale, which can sometimes alleviate multicollinearity.