Machine Learning Set 3

Total points 20/40

The respondent's email (**vaibhav.mum.dbda@gmail.com**) was recorded on submission of this form.

✗ Student ID * ···/1

230340325073 ✗

✓ What is the primary goal of linear regression? 1/1

- ○ a) Classification
- ◉ b) Regression ✓
- ○ c) Clustering
- ○ d) Dimensionality reduction

✓ Which of the following is a key assumption of linear regression? 1/1

- ○ a) Non-linearity
- ○ b) Heteroscedasticity
- ◉ c) Homoscedasticity ✓
- ○ d) Multicollinearity

✓ Which evaluation metric is commonly used to assess the performance of a linear regression model? 1/1

- a) Accuracy
- b) F1 score
- c) R-squared ✓
- d) Precision

✗ In linear regression, what is the term used to represent the predicted values? 0/1

- a) Labels
- b) Targets ✗
- c) Outputs
- d) Predictors

Correct answer

- c) Outputs

✗ Which of the following is true about the residuals in linear regression? 0/1

- (●) a) Residuals should be normally distributed. ✗
- ( ) b) Residuals should be evenly spaced.
- ( ) c) Residuals should have a constant variance.
- ( ) d) Residuals should be correlated with the predictors.

Correct answer

- (●) c) Residuals should have a constant variance.

✗ Which technique can be used to handle multicollinearity in linear regression? 0/1

- ( ) a) Principal Component Analysis (PCA)
- (●) b) Support Vector Machines (SVM) ✗
- ( ) c) k-nearest neighbors (k-NN)
- ( ) d) Ridge regression

Correct answer

- (●) d) Ridge regression

✗ What is the purpose of the intercept term in linear regression? 0/1

- (●) a) To control overfitting ✗
- ( ) b) To center the data
- ( ) c) To account for the bias in the data
- ( ) d) To increase the model complexity

Correct answer

- (●) c) To account for the bias in the data

✓ What is the equation for a simple linear regression line? 1/1

- (●) a) y = mx + b ✓
- ( ) b) y = ax^2 + bx + c
- ( ) c) y = e^(mx + b)
- ( ) d) y = a / (bx + c)

✓ Which technique is used to find the best-fitting line in linear regression? 1/1

- (●) a) Gradient Descent ✓
- ( ) b) K-means clustering
- ( ) c) Decision Trees
- ( ) d) Naive Bayes

✗ What does the coefficient of determination (R-squared) represent in linear regression? 0/1

- ○ a) The percentage of variance in the target variable explained by the predictors.
- ◉ b) The accuracy of the model in predicting the target variable. ✗
- ○ c) The sum of squared errors between the predicted and actual values.
- ○ d) The number of predictors in the model.

Correct answer

- ◉ a) The percentage of variance in the target variable explained by the predictors.

✗ What is the purpose of feature scaling in linear regression? 0/1

- ○ a) To improve the interpretability of the coefficients.
- ◉ b) To handle missing values in the dataset. ✗
- ○ c) To convert categorical variables into numerical ones.
- ○ d) To ensure that variables are on a similar scale.

Correct answer

- ◉ d) To ensure that variables are on a similar scale.

✓ Which of the following is not a type of linear regression? 1/1

- ○ a) Simple linear regression
- ○ b) Multiple linear regression
- ○ c) Polynomial regression
- ◉ d) Logistic regression ✓

✗ Which of the following is used to measure the statistical significance of the coefficients in linear regression? 0/1

- ○ a) T-statistic
- ◉ b) Z-score ✗
- ○ c) F-statistic
- ○ d) P-value

Correct answer

- ◉ a) T-statistic

✗ What is the purpose of residual analysis in linear regression? 0/1

- ○ a) To estimate the model parameters.
- ○ b) To assess the linearity assumption.
- ○ c) To visualize the relationship between variables.
- ○ d) To calculate the mean squared error (MSE).

✓ Which of the following statements is true about outliers in linear regression? 1/1

- ○ a) Outliers have no impact on the regression model.
- ○ b) Outliers always increase the accuracy of the model.
- ◉ c) Outliers can significantly influence the regression line. ✓
- ○ d) Outliers can only occur in the target variable, not the predictors.

✓ Which algorithm is commonly used for solving linear regression problems when the number of predictors is large? 1/1

- ○ a) Linear Discriminant Analysis (LDA)
- ○ b) Support Vector Regression (SVR)
- ◉ c) Lasso regression ✓
- ○ d) K-means clustering

✓ What is the difference between simple linear regression and multiple linear regression? 1/1

- a) Simple linear regression can handle multiple predictors, while multiple linear regression can only handle one predictor.
- b) Simple linear regression is used for classification, while multiple linear regression is used for regression.
- c) Simple linear regression has a higher model complexity than multiple linear regression.
- d) Simple linear regression has one predictor, while multiple linear regression has multiple predictors. ✓

✓ Which of the following statements is true about the residuals in linear regression? 1/1

- a) The residuals should be perfectly correlated with the predictors.
- b) The residuals should follow a normal distribution. ✓
- c) The residuals should be larger than the target variable.
- d) The residuals should be equal to the target variable.

✓ Which technique can be used to handle missing values in linear regression? 1/1

- ○ a) Removing the entire row with missing values
- ○ b) Replacing missing values with the mean of the variable
- ○ c) Ignoring the missing values and using the available data
- ◉ d) All of the above ✓

✓ What is the purpose of cross-validation in linear regression? 1/1

- ◉ a) To assess the performance of the model on unseen data. ✓
- ○ b) To increase the number of predictors in the model.
- ○ c) To reduce the bias in the model.
- ○ d) To decrease the model complexity.

✗ What is the purpose of the cost function in linear regression? 0/1

- ○ a) To measure the accuracy of the model's predictions.
- ○ b) To estimate the model parameters.
- ○ c) To calculate the R-squared value.
- ○ d) To assess the linearity assumption.

✓ Which of the following algorithms can be used to solve linear regression problems analytically (without iterative optimization)? 1/1

- a) Gradient Descent
- b) Support Vector Regression (SVR)
- c) Ordinary Least Squares (OLS) ✓
- d) K-means clustering

✓ What is the difference between the mean squared error (MSE) and the root mean squared error (RMSE) in linear regression? 1/1

- a) MSE is the squared difference between predicted and actual values, while RMSE is the square root of MSE. ✓
- b) MSE measures the error in the model, while RMSE measures the variability of the target variable.
- c) MSE is the average error, while RMSE is the maximum error.
- d) MSE is used for binary classification, while RMSE is used for multi-class classification.

✓ In linear regression, what is the assumption regarding the relationship between the predictors and the target variable? 1/1

- a) Linearity ✓
- b) Normality
- c) Independence
- d) Equal variance

✗ Which technique is used to handle overfitting in linear regression? 0/1

- (●) a) Feature scaling ✗
- ( ) b) Regularization
- ( ) c) Gradient Descent
- ( ) d) Principal Component Analysis (PCA)

Correct answer

- (●) b) Regularization

✗ Which of the following statements is true about the p-value in linear regression? 0/1

- ( ) a) A lower p-value indicates a stronger correlation between the predictors.
- ( ) b) The p-value represents the percentage of variance explained by the predictors.
- (●) c) A higher p-value indicates a significant relationship between the predictors and the target. ✗
- ( ) d) The p-value measures the accuracy of the model's predictions.

Correct answer

- (●) d) The p-value measures the accuracy of the model's predictions.

✓ Which of the following is a disadvantage of linear regression? 1/1

- a) It is computationally expensive.
- b) It assumes a linear relationship between predictors and the target. ✓
- c) It cannot handle categorical variables.
- d) It is not suitable for large datasets.

✓ What is the purpose of feature selection in linear regression? 1/1

- a) To remove outliers from the dataset.
- b) To preprocess the data before fitting the model.
- c) To identify the most relevant predictors for the target variable. ✓
- d) To estimate the model parameters.

✓ Which of the following techniques can be used to handle outliers in linear regression? 1/1

- a) Removing the outliers from the dataset.
- b) Transforming the target variable.
- c) Using robust regression techniques.
- d) All of the above. ✓

✗ Which statistical test can be used to determine whether the overall linear regression model is significant? 0/1

- a) T-test
- b) Z-test ✗
- c) F-test
- d) Chi-square test

Correct answer

- c) F-test

✗ What is the purpose of residual plots in linear regression? 0/1

- a) To visualize the relationship between predictors.
- b) To assess the linearity assumption.
- c) To estimate the model parameters. ✗
- d) To calculate the R-squared value.

Correct answer

- b) To assess the linearity assumption.

✓ Which of the following statements is true about the multicollinearity problem in linear regression? 1/1

- ○ a) Multicollinearity increases the stability of the model.
- ◉ b) Multicollinearity decreases the interpretability of the coefficients. ✓
- ○ c) Multicollinearity has no impact on the model's performance.
- ○ d) Multicollinearity can be solved by removing outliers from the dataset.

✗ Which technique can be used to handle heteroscedasticity in linear regression? 0/1

- ○ a) Principal Component Analysis (PCA)
- ◉ b) Support Vector Machines (SVM) ✗
- ○ c) Weighted Least Squares (WLS)
- ○ d) Naive Bayes

Correct answer

- ◉ c) Weighted Least Squares (WLS)

✗ What is the purpose of the Durbin-Watson statistic in linear regression? 0/1

- ○ a) To measure the autocorrelation of the residuals.
- ◉ b) To assess the normality assumption of the residuals. ✗
- ○ c) To calculate the mean squared error (MSE).
- ○ d) To estimate the model parameters.

Correct answer

- ◉ a) To measure the autocorrelation of the residuals.

✓ Which of the following statements is true about the adjusted R-squared in linear regression? 1/1

- ○ a) The adjusted R-squared always increases when more predictors are added to the model.
- ◉ b) The adjusted R-squared penalizes the model for including unnecessary predictors. ✓
- ○ c) The adjusted R-squared is always lower than the regular R-squared.
- ○ d) The adjusted R-squared is equal to the mean squared error (MSE).

✗ Which algorithm can be used to solve linear regression problems when the predictors are highly correlated? 0/1

- ○ a) Principal Component Analysis (PCA)
- ◉ b) Lasso regression ✗
- ○ c) Decision Trees
- ○ d) K-means clustering

Correct answer

- ◉ a) Principal Component Analysis (PCA)

✗ Which of the following statements is true about the residuals in linear regression? 0/1

- ○ a) The residuals should be perfectly correlated with the target variable.
- ◉ b) The residuals should follow a uniform distribution. ✗
- ○ c) The residuals should have a linear relationship with the predictors.
- ○ d) The residuals should have a mean of zero.

Correct answer

- ◉ d) The residuals should have a mean of zero.

✗ Which technique can be used to handle non-linearity in linear regression? 0/1

- a) Polynomial regression
- b) Support Vector Machines (SVM)
- c) K-means clustering
- d) Ridge regression ✗

Correct answer

- a) Polynomial regression

✗ What is the purpose of stepwise regression in linear regression? 0/1

- a) To select the best subset of predictors for the model.
- b) To handle missing values in the dataset.
- c) To estimate the model parameters. ✗
- d) To calculate the mean squared error (MSE).

Correct answer

- a) To select the best subset of predictors for the model.

Submission Tracking ID - DO NOT CHANGE *

This is important for tracking purposes. Do not change. Any change in this will make your submission void

230340325073

Google Forms