

Machine Learning

1. C
2. D
3. D
4. A
5. B
6. A,D
7. A,D
8. A
9. B,C
10. When a new predictor improves the model beyond what would be predicted by probability, the adjusted R-squared increases to account for the new variables. Otherwise, it stays the same. In contrast, it will fall off when a predictor enhances the model by less than what would be expected by chance.
11. The simplest type of regression is linear regression, in which the model's selection of weights is completely unaffected. This means that, during the training stage, the model may give a lot of weight to a feature if it believes that it is particularly essential. In tiny datasets, this can occasionally result in overfitting.
Lasso is a variant of linear regression in which the model is penalised for the total absolute value of the weights. As a result, the absolute values of weight will typically decrease and many will tend to be zeros.
12. A set of multiple regression variables' variance inflation factor (VIF) is a measure of how multicollinear they are.
A VIF of three or less is not cause for alarm. As VIF increases, your regression results become less credible.
13. We scale the data before feeding it to the model to guarantee that the gradient descent goes smoothly towards the minima and that the gradient descent steps are updated at the same rate for all features.
14. There are three error measures that are typically used for measuring and reporting a regression model's performance: Error Squared Mean (MSE). Error Mean Squared (RMSE) (RMSE). Absolute Mean Error (MAE)
15. Accuracy : 88%, Precision : 95%, Recall/Sensitivity : 8%, Specificity : 96%

SQL WorkSheet

1. A, C, D
2. A, C, D
3. B
4. C
5. B
6. B
7. A
8. C
9. D
10. A
11. Denormalization is the process of taking a normalised schema and rendering it non-normalized.
12. A database cursor is a pointer to a specific row inside the results of a query. The pointer can be shifted from row to row. You may even be able to move the cursor to the previous row, depending on the type of cursor.
13. Different types of queries:
 - Select Query
 - Action query (Append, update, delete, make table)
 - Parameter Query
 - Aggregate Query (sum, avg, min, max, first, last, group by)
14. Constraints are the rules that are imposed on a table's data columns. These are used to restrict the type of data that may be entered into a table. This guarantees that the data in the database is accurate and reliable.

Constraints could be applied at the column or table level. Column level constraints apply to a single column, whereas table level constraints apply to the entire table.
15. When a new record is placed into a table, auto-increment automatically generates a unique number.

This is frequently the primary key field that we want to be produced automatically whenever a new record is inserted.

Statistics

1. C
2. A
3. A
4. C
5. A
6. B
7. C
8. B
9. B
10. Histograms are a type of bar graph that displays a bar for a range of data values rather than a single value. A box plot is a type of data visualisation in which a box is drawn over a number line to indicate the data's interquartile range.
11. In that sequence, prioritise objectives, investigate which metrics reliably predict their achievement, and identify which activities influence predictors.
12. Statistical significance is frequently calculated in conjunction with statistical hypothesis testing.
13. There are no log-normal or Gaussian distributions for exponential distributions.
14. Income is a great example of when the median should be used instead of the mean since its distribution is skewed.
15. The likelihood function is a fundamental notion in statistical inference. It reveals how probable a given population is to yield an observed sample.