Q1 to Q15 are subjective answer type questions, Answer them briefly.
1. R-squared or Residual Sum of Squares (RSS) which one of these two is a better measure of
goodness of fit model in regression and why?

R-squared and Residual Sum of Squares (RSS) boh can be used for measuring goodness of fit model in regresion.

R-squared is a statistical measure that represents the proportion of the variance for a
dependent variable that's explained by an independent variable or variables in a regression model.
In other words, R-squared measures the extent to which changes in the dependent variable can be predicted
by changes in the independent variable(s). r-squared value lies between 0-1 ,and higher R-squared values indicate a better fit of the regression model to the data. Therefore, R-squared is often used to compare different models and select the best one.

On the other hand, Residual Sum of Squares (RSS) measures the difference between the observed values of the dependent variable
and the predicted values by the model. It represents the sum of the squared differences between the actual and predicted values of the dependent
variable. The goal is to minimize the residual sum of squares to obtain a better model fit.

In terms of determining the goodness of fit of a model,
R-squared is generally considered a better measure than RSS.
This is because R-squared provides an overall measure of the proportion of variance in the dependent variable
that is explained by the model, whereas RSS only measures the magnitude of the residuals.
Additionally, R-squared is a standardized measure and ranges from 0 to 1, making it easy to compare the fit of different models.
In contrast, the magnitude of the RSS value depends on the scale of the dependent variable and can't be easily compared across models.

However, it's worth noting that neither R-squared nor RSS is a perfect measure of model fit.
R-squared can be influenced by outliers or data points that don't fit the model well,
while RSS doesn't take into account the number of variables or degrees of freedom in the model.
Therefore, it's important to consider multiple metrics when evaluating a regression model's goodness of fit.

2. What are TSS (Total Sum of Squares), ESS (Explained Sum of Squares) and RSS (Residual Sum

of Squares) in regression. Also mention the equation relating these three metrics with each other.
ans.
TSS = ESS + RSS, where TSS is Total Sum of Squares, ESS is Explained Sum of Squares and RSS is Residual Sum of Suqares.
The aim of Regression Analysis is explain the variation of dependent variable Y

3. What is the need of regularization in machine learning?
we use regularization in machine learning to properly fit a model onto our test set.
Regularization techniques help reduce the chance of overfitting and help us get an optimal model.

4. What is Gini–impurity index?
ans.
Gini Impurity is a measurement used to build Decision Trees to determine how the features of a dataset should split nodes to form the tree.
it varies between 0 to 1,and where 0 represents purity of the classification and 1 denotes random distribution of elements
among various classes. A Gini Index of 0.5 shows that there is equal distribution of elements across some classes.

5. Are unregularized decision-trees prone to overfitting? If yes, why?
ans.
when it comes to decision trees the thing is, it makes very few assumptions about training data
(linear model assumes that the data you will be feeding will be linear).
If we don't constraint it, the tree will adapt itself to the training data, which will lead to overfitting.

6. What is an ensemble technique in machine learning?
ans.
it is machine learning technique which enhances accuracy and resilience in forecasting by merging predictions from multiple models.
which aims to mitigate errors or biases that may exist in individual models by leveraging the collective intelligence of the ensemble.

7. What is the difference between Bagging and Boosting techniques?
ans.
bagging and boosting both model recieves equal weight but bagging reduces varience and boosting reduces bias, in boosting models are weighted based on their performance.

8. What is out-of-bag error in random forests?
ans.
out of bag is a method of measuring the prediction error of random forests, boosted decision trees,
and other machine learning models utilizing bootstrap aggregating (bagging).

9. What is K-fold cross-validation?
k- fold cross validation method is used to validate predictive models. inthis method given dataset is divided into k subsets or folds.
This model is trained and evaluated k times, using a different fold as the validation set each time.

10. What is hyper parameter tuning in machine learning and why it is done?
hyperparameter control data structure,performance, and functions.it is mainly used to control model performance for optimum result.

11. What issues can occur if we have a large learning rate in Gradient Descent?
ans.
if learning rate in gradient descent is large then algorithm may overshoot the minimum.

12. Can we use Logistic Regression for classification of Non-Linear Data? If not, why?
ans.
Logistic regression has traditionally been used to come up with a hyperplane that separates the feature space into classes.
 But if we suspect that the decision boundary is nonlinear we may get better results by attempting some nonlinear functional
 forms for the logit function.

13. Differentiate between Adaboost and Gradient Boosting.
ans.

Adaboost is computed with a specific loss function and becomes more rigid when comes to few iterations.
But in gradient boosting, it assists in finding the proper solution to additional iteration modeling problem as
it is built with some generic features.

14. What is bias-variance trade off in machine learning?
ans.
bias varience trade off in machine learning is used relationship between a model's complexity, the accuracy of its predictions,
and how well it can make predictions on previously unseen data that were not used to train the model.

15. Give short description each of Linear, RBF, Polynomial kernels used in SVM.
RBF:
RBF is a popular kernel function used in various kernelized learning algorithms.
In particular, it is commonly used in support vector machine classification.

Linear:
Linear Kernel is used when the data is Linearly separable, that is, it can be separated using a single Line.
it mostly used when there is large number of features in dataset.

polynomial:
commonly used with support vector machines (SVMs) and other kernelized models, that represents the similarity of vectors (training samples)
in a feature space over polynomials of the original variables, allowing learning of non-linear models.