1. What exactly is a feature? Give an example to illustrate your point.

Ans: Features are the basic building blocks of datasets. The quality of the features in your dataset has a major impact on the quality of the insights you will gain when you use that dataset for machine learning.

Additionally, different business problems within the same industry do not necessarily require the same features, which is why it is important to have a strong understanding of the business goals of your data science project.

2. What are the various circumstances in which feature construction is required?

Ans: The features in your data will directly influence the predictive models you use and the results you can achieve. Your results are dependent on many inter-dependent properties. You need great features that describe the structures inherent in your data. Better features means flexibility.

3. Describe how nominal variables are encoded.

Ans: Nominal data is made of discrete values with no numerical relationship between the different categories — mean and median are meaningless. Animal species is one example. For example, pig is not higher than bird and lower than fish.

4. Describe how numeric features are converted to categorical features.

Ans: Converting categorical features into numeric features using domain knowledge. For example, we are given a list of countries and say we know the distance to these countries from India then we can replace it with distance from India. So, every country can be represented as its distance from India.

5. Describe the feature selection wrapper approach. State the advantages and disadvantages of this approach?

Ans: Wrapper methods measure the "usefulness" of features based on the classifier performance. In contrast, the filter methods pick up the intrinsic properties of the features (i.e., the "relevance" of the features) measured via univariate statistics instead of cross-validation performance.

The wrapper classification algorithms with joint dimensionality reduction and classification can also be used but these methods have high computation cost, lower discriminative power. Moreover, these methods depend on the efficient selection of classifiers for obtaining high accuracy

6. When is a feature considered irrelevant? What can be said to quantify it?

Ans: Features are considered relevant if they are either strongly or weakly relevant, and are considered irrelevant otherwise.

Irrelevant features can never contribute to prediction accuracy, by definition. Also to quantify it we need to first check the list of features, There are three types of feature selection:

- **Wrapper methods** (forward, backward, and stepwise selection)
- **Filter methods** (ANOVA, Pearson correlation, variance thresholding)
- **Embedded methods** (Lasso, Ridge, Decision Tree).

7. When is a function considered redundant? What criteria are used to identify features that could be redundant?

Ans: If two features {X1, X2} are highly correlated, then the two features become redundant features since they have same information in terms of correlation measure. In other words, the correlation measure provides statistical association between any given a pair of features.

Minimum redundancy feature selection is an algorithm frequently used in a method to accurately identify characteristics of genes and phenotypes

8. What are the various distance measurements used to determine feature similarity?

Ans: Four of the most commonly used distance measures in machine learning are as follows:

- Hamming Distance.
- Euclidean Distance
- Manhattan Distance.

9. State difference between Euclidean and Manhattan distances?

Ans: Euclidean & Hamming distances are used to measure similarity or dissimilarity between two sequences. Euclidean distance is extensively applied in analysis of convolutional codes and Trellis codes.

Hamming distance is frequently encountered in the analysis of block codes

10. Distinguish between feature transformation and feature selection.

Ans: Feature selection is for filtering irrelevant or redundant features from your dataset. The key difference between feature selection and extraction is that feature selection keeps a subset of the original features while feature extraction creates brand new ones.

11. Make brief notes on any two of the following:

1.SVD (Standard Variable Diameter Diameter)

2. Collection of features using a hybrid approach

3. The width of the silhouette

4. Receiver operating characteristic curve

1 . SVD (Standard Variable Diameter Diameter)

The SVD is used widely both in calculation of other matrix operations, such as matrix inverse but also as a data reduction method in machine learning.SVD can also be used in least Squares Linear Regression, image compression and denoising data.

2. Collection of features using a hybrid approach

Multiple simple algorithm work together to complement and augment each other.

Together they can solve problems that alone they were not designed to solve.

3. The width of the silhouette

The silhouette algorithm is one of the many algorithms to determine the optimal number of clusters for an unsupervised learning technique. In the sinhouette algorithm, we assume that the data has already been clustered into k clusters by a clustering technique.

4. Receiver operating characteristic curve

4. Receiver operating characteristic curve

A Receiver Operating Characteristic Curve (ROC) is a Standard Technique for Summarizing Classifier Performance over a range of trade-offs between true positive (Tp) and False Positive (FP) error rates.