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

A feature is an individual measurable property or characteristic of a phenomenon being observed. Example: In a house price prediction model, features might include the size, number of bedrooms, and location of the house.

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

Feature construction is required when raw data needs transformation for better model performance, dealing with missing values, enhancing interpretability, or incorporating domain knowledge to create meaningful attributes.

3. Describe how nominal variables are encoded.

Nominal variables are encoded using techniques like one-hot encoding, which creates binary columns for each category, or label encoding, which assigns a unique integer to each category.

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

Numeric features are converted to categorical features through binning or discretization, which involves dividing the range of the numeric feature into intervals and assigning each interval a category label.

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

The wrapper approach evaluates feature subsets using a specific model and performance metric. Advantages: high accuracy, tailored to the model. Disadvantages: computationally expensive, risk of overfitting.

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

A feature is irrelevant if it doesn't contribute to the target variable's prediction. Irrelevance can be quantified by low correlation with the target or minimal contribution to model performance metrics.

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

A feature is redundant if it provides no additional information beyond what's already captured by other features. Criteria include high correlation with other features and similar contribution to model performance.

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

Distance measurements include Euclidean distance, Manhattan distance, Minkowski distance, cosine similarity, and Hamming distance, each assessing similarity based on different mathematical formulations.

9. State the difference between Euclidean and Manhattan distances.

Euclidean distance measures the straight-line distance between two points in a multi-dimensional space, while Manhattan distance measures the sum of absolute differences along each dimension.

10. Distinguish between feature transformation and feature selection.

Feature transformation alters the feature space, like scaling or PCA. Feature selection chooses a subset of existing features. Transformation changes feature representation, selection reduces dimensionality.

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

1. SVD (Singular Value Decomposition): A matrix factorization technique reducing dimensionality by decomposing data into singular vectors and values, used in noise reduction and data compression.

2. Receiver Operating Characteristic (ROC) Curve: A graphical plot illustrating the diagnostic ability of a binary classifier, displaying the trade-off between true positive rate and false positive rate.