1. What are the key tasks that machine learning entails? What does data pre-processing imply?

2. Describe quantitative and qualitative data in depth. Make a distinction between the two.

3. Create a basic data collection that includes some sample records. Have at least one attribute from each of the machine learning data types.

4. What are the various causes of machine learning data issues? What are the ramifications?

5. Demonstrate various approaches to categorical data exploration with appropriate examples.

6. How would the learning activity be affected if certain variables have missing values? Having said that, what can be done about it?

7. Describe the various methods for dealing with missing data values in depth.

8. What are the various data pre-processing techniques? Explain dimensionality reduction and function selection in a few words.

9.

i. What is the IQR? What criteria are used to assess it?

ii. Describe the various components of a box plot in detail? When will the lower whisker surpass the upper whisker in length? How can box plots be used to identify outliers?

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

1. Data collected at regular intervals

2. The gap between the quartiles

3. Use a cross-tab

1. Make a comparison between:

1. Data with nominal and ordinal values

2. Histogram and box plot

3. The average and median