1) Please show all your work in the Python Jupyter notebook.

Ans: I have showcased the maximum features, techniques, visualization etc. in the Python program.

2) Using data visualization tools, please explain how we can understand the data structure.

Ans: To apply different graphs and parsing different parameters to know the overall structure of the data. To understand this, firstly study the business objective, then moved to data understanding part , where to find the different relationships pattern , to know the correlation and significance of the observations, related attributes , records etc.

3) Please explain if dimensionality reduction is required/possible or not. How did you check?

Ans: I have applied dimensionality reduction while using PCA method. To purpose behind is to use transformation done in PCA linear, though the relationship of X and Y remains linear. Reducing the number of input variables in training data. To check redundant and highly corelated applied to reduce the random variables under consideration, by obtaining principal variables. To validate the same while applying feature selection and extraction.

4) Please explain eigen-vectors and eigen-values and their importance.

Ans: Eigenvectors make understanding linear transformations easy. They are the "axes" (directions) along which a linear transformation acts simply by "stretching/compressing" and/or "flipping"; eigenvalues give you the factors by which this compression occurs. The more directions you have along which you understand the behavior of a linear transformation, the easier it is to understand the linear transformation; so you want to have as many linearly independent eigenvectors as possible associated to a single linear transformation.

5) Which classification methods are you using? How do you decide among different methods?

Ans: I have used Random Forest and Logistic Regression. Looking at the dataset logistic regression performs better when the number of noise variables is less than or equal to the number of explanatory variables and random forest has a higher true and false positive rate as the number of explanatory variables increases in a dataset. So, these points are present in data to handle through these classification methods.

6) Please provide a confusion matrix and explain how it can help us to check the reliability of the result.

Ans: Confusion matrix is used as metrics for model evaluation purpose and this is one of the parameters of metrics. In this assignment it is based on the dependent variable and predicted variable and then checked the accuracy score of these parameters where actual and predicated represented. Also, Precision, Recall, F1 -score and Support are the major key attributes and every attribute have the specific outcome of the result regarding Actual and Predicted. To check the reliability of confusion matrix result, the dependent variable and predicated variable and other attributes of classifier considered and then calculate overall score.

7) Please provide the learning curve and explain how it can help us in determining whether the model is being over-fit or under-fit.

Ans: Learning curve is plot of learning y-axis over experience y-axis. In this assignment different learning curves explained the impacts on different hyperparameters. To validate the same for Max features, min simple leaf, min samples split etc. After that optimal hyperparameters on grid search explaining all these hyperparameters together, the plot explained the training and testing bias regards to fitting position for determining the values to these hyperparameters.

8) When do you consider adding the "regularization parameter" to the model? and how it will help to improve the model performance?

Ans: When overfitting observed in model, the regularization will be useful. Regularization penalizes the coefficients. To improve the model performance we can apply different methods of regularization like L1 , L2. As increase the regularization parameter, optimization function will have to choose a smaller theta to minimize the total cost. The overall loss function as in your example above consists of an error term and a regularization term that is weighted by lambda function, the regularization parameter.

9) Please briefly explain how reinforcement-learning can be utilized in fraud detection models.

Ans: Algorithm must spot the element that does not fit in with the group. It may be a flawed product, potentially fraudulent transaction or any other event associated with breaking the norm. Apply the supervised learning, simulation, acting, learning and experience can apply.

10) Please describe when to use logistic sigmoid, tanh, and Fourier as a basis function.

Ans: we use sigmoid function is because it exists between **(0 to 1).** Therefore, it is especially used for models where we must **predict the probability** as an output. Since probability of anything exists only between the range of **0 and 1,** sigmoid is the right choice. Tanh is also like logistic sigmoid but better. The range of the tanh function is from (-1 to 1). tanh is also sigmoidal (s - shaped). The advantage is that the negative inputs will be mapped strongly negative and the zero inputs will be mapped near zero in the tanh graph The Fourier basis is a simple, principled basis function scheme for linear value function approximation. The Fourier basis is a simple, principled basis function scheme for linear value function approximation.