1. Why did you choose the particular algorithm?

I have Choose SVM algorithm for classification of data. There are many advantages that this algorithm provides. Which are as following,

Assume that some data points are assigned to one of two classes, and the purpose is to determine which class a new data point will be assigned to. In classification problem we try to predict Hyperplane that best separates the classes. The best hyperplane is the one that reflects the greatest distance between the two classes, or margin. As a result, we chose the hyperplane to maximise the distance between it and the nearest data point on each side. If such a hyperplane exists, it is referred to as a maximum-margin hyperplane, and the linear classifier it creates is referred to as a maximum-margin classifier.

- In scenarios with a lot of dimensions, this method works well.
- It saves memory by using a subset of training points termed support vectors in the decision function.
- For the decision functions, several kernel functions can be given, as well as bespoke kernels.

2. What are the different tuning methods used for the algorithm?

Grid Search:

Grid search is a common approach for tweaking hyperparameters. It conducts an exhaustive search on the collection of hyperparameters that users have specified. This method is the simplest and produces the most accurate forecasts. Users can determine the best combination using this tweaking process. Grid search is useful for a variety of hyperparameters, although it has a limited search space.

Random search:

Random search is a simple enhancement over grid search. A randomised search over hyper-parameters from particular distributions over conceivable parameter values is used in this method. The search continues until the necessary level of precision is achieved. Random search is similar to grid search, however it has been shown to produce superior results. The method is frequently used as an HPO baseline to assess the efficacy of newly developed algorithms.

Gradient-based optimisation:

Gradient-based optimisation is a method for optimising multiple hyperparameters based on the gradient of a machine learning model selection criterion with respect to the hyperparameters. When the training criterion's differentiability and continuity constraints are met, this hyperparameter tuning mechanism can be used. Despite being more effective than grid search, random search is still a computationally intensive strategy.

3. Did you consider any other choice of algorithm? Why or why not?

I consider two algorithms Stochastic Gradient Descent (SGD) and Support Vector Machine (SVM).

SVM is a special linear-model. From a theoretical view it's a convex-optimization problem and we can get the global-optimum in polynomial-time. There are many different optimization-approaches.

- **SVM:**

1. Precision score = 0.8700786628461699
2. Recall score = 0.8379508625196027
3. F1 score = 0.8453629004016364
4. Accuracy score = 0.8379508625196027

SGD is based optimizer which can optimize many different convex-optimization problems. SGD is a general optimization method.

- **SGD:**

1. Precision score = 0.7408039146781437
2. Recall score = 0.7569262937794041
3. F1 score = 0.680354658157873
4. Accuracy score = 0.7569262937794041

I tried to do Hyper-tunning for both the algorithms and in both cases I got better results in SVM, that’s why I choose SVM classification algorithm.

4. What is the accuracy?

- Accuracy = 0.8379508625196027

5. What are the different types of metrics that can be used to evaluate the model?

- **Precision score:**

  The precision is calculated as “tp / (tp + fp)”, where tp represents the number of true positives and ‘fp’ represents the number of false positives. Precision is the capacity of the classifier to not label a sample that is negative as positive.

- **Recall score:**

  The ratio “tp / (tp + fn)”, where tp is the number of true positives and ‘fn’ is the number of false negatives, is the recall. The recall refers to the classifier's capacity to locate all positive samples.

- **F1 score:**

  The F1 score is a harmonic mean of precision and recall, with a maximum of 1 and a minimum of 0. Precision and recall make the same amount of difference in the F1 score.

- **Accuracy score:**

  This function computes subset accuracy in multilabel classification: the set of labels predicted for a sample must exactly match the corresponding set of labels in y true.