1. **Random Forest Algorithm:**

Random forest, also known as random decision forest algorithm, is an ensemble learning method used for classification and regression. It works by creating a multitude of decision trees at training time and finally outputting the mean prediction of each tree for regression, or mode of all the individual trees for classification problems.

Random forest, being a supervised learning algorithm, upon getting the input data builds a forest, which is ensemble of decision trees which are trained using bagging method. Bagging works in the principle that a combination of models improves the chance of getting overall better results. The random forest algorithm builds multiple decision trees and aggregates the results to get more stable and accurate results. This idea of combining the predictions of several base estimators is called ensemble method. The goal of ensemble method is to improve the robustness and generalizability of the model when compared to a single estimator.

We’ve used the scikit-learn library’s RandomForestClassifier for the prediction of the quality of wine. The figure below contains the performance metrics observed on classification:

Table

Description automatically generated

1. **Stochastic Gradient Descent Classifier:**

Stochastic Gradient Descent is a type of linear classifier which is optimized by using stochastic gradient descent. In this, we use either Support Vector Machines or Logistic Regression as machine learning model for generating loss function and optimization method to minimize or maximize it.

Stochastic Gradient Descent, unlike models, is an optimization technique. It’s just a way of training a model. Stochastic Gradient Descent is a very simple and efficient approach to fitting linear classifiers and regressors under convex loss functions (like SVM and LR). It uses gradient descent, which is an iterative algorithm that starts at a random point on a function and traverses down it’s slope till it reaches the lowest point of that function.

We’ve used SGDclassifier from the scikit-learn library for classification. The figure below contains the performance metrics observed on classification:

Table

Description automatically generated

1. **Support Vector Machine**

Support Vector Machine is an algorithm that can be used for both regression and classification problems, but it’s mostly used for classification. It is a supervised machine learning algorithm. It works by plotting each datapoint in n-dimensional space and then finding the hyperparameter plane that separates the plotted data points into classes. The number of features determine the dimensions of the hyperplane.

It uses SVM kernels, which are functions used to form high dimensional space taking low dimensional input space. It comes into use when we have data set which isn’t linearly separable. Kernels, in essence converts non-separable problem to a separable problem.

Support Vector Machines are very effective in higher dimensional spaces. They are very memory efficient, and we can use different kernels for the decision function.

We have used SVC from the scikit-learn library for our classification problem. The figure below contains the performance metrics observed on classification:

Table

Description automatically generated

After optimizing, the results were improved:

Table

Description automatically generated

**Performance Comparison: | <<Doesn’t look relevant>>**

Runtime Comparison:

1st Random Forest | 2nd SGD | 3rd Support Vector Machines

Chart, line chart

Description automatically generated

**Reference:**

1. <https://en.wikipedia.org/wiki/Random_forest>
2. <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>
3. https://en.wikipedia.org/wiki/Stochastic\_gradient\_descent
4. <https://scikit-learn.org/stable/modules/sgd.html>
5. https://towardsdatascience.com/stochastic-gradient-descent-clearly-explained-53d239905d31
6. https://en.wikipedia.org/wiki/Support\_vector\_machine
7. <https://www.geeksforgeeks.org/support-vector-machine-algorithm/>
8. <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>