

CSI5155 Assignment 1 - Supervised Learning - Report

In the next pages, you will find the Confusion Matrix, Recall, Precision and ROC Curve for each model. This document is organized model by model. For example, Chocolate, Chocolate Undersampled, Chocolate Oversampled, etc.

One note, the Confusion Matrix follows this format:

[TP , FP]

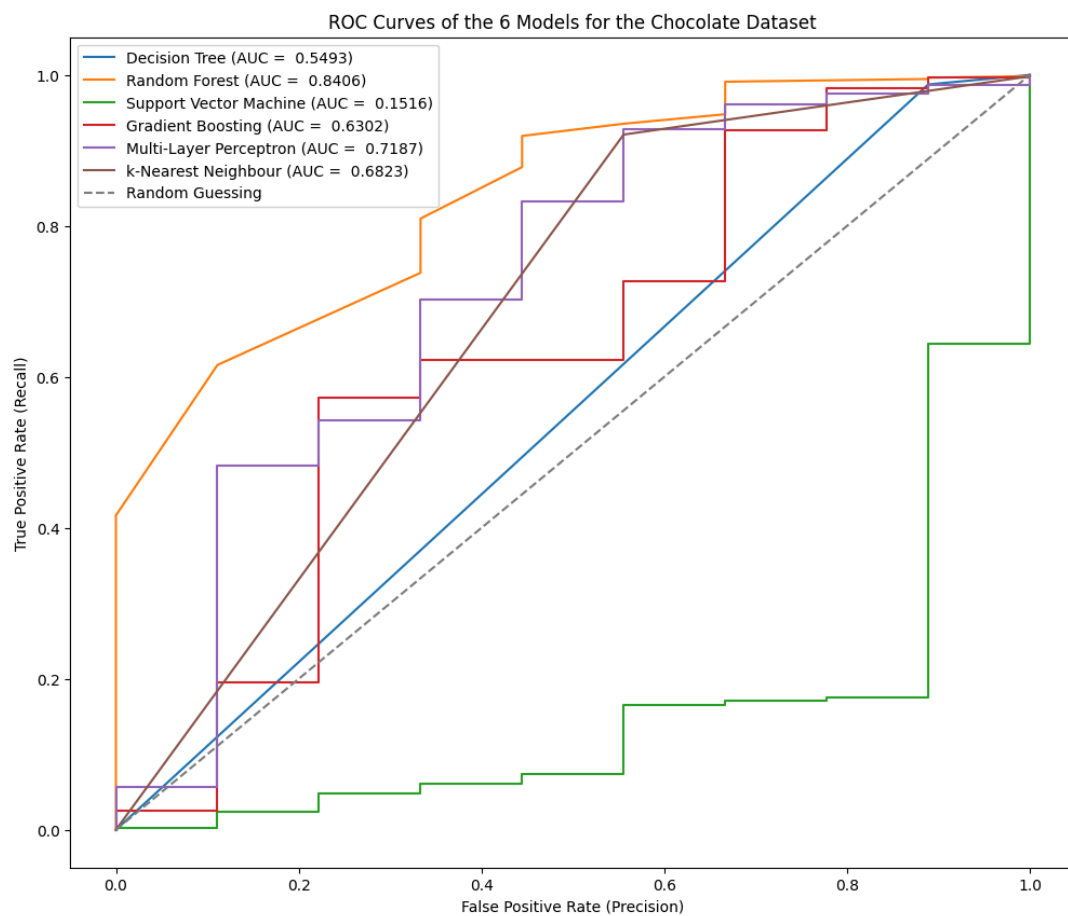
[FN, TN]

Table of Contents:

Chocolate Dataset.....	1
Chocolate Dataset with Under Sampling.....	2
Chocolate Dataset with Over Sampling.....	3
Chocolate Dataset with Combined Sampling.....	4
Magic Mushroom Dataset.....	5
Magic Mushroom Dataset with Under Sampling.....	6
Magic Mushroom Dataset with Over Sampling.....	7
Magic Mushroom Dataset with Combined Sampling.....	8

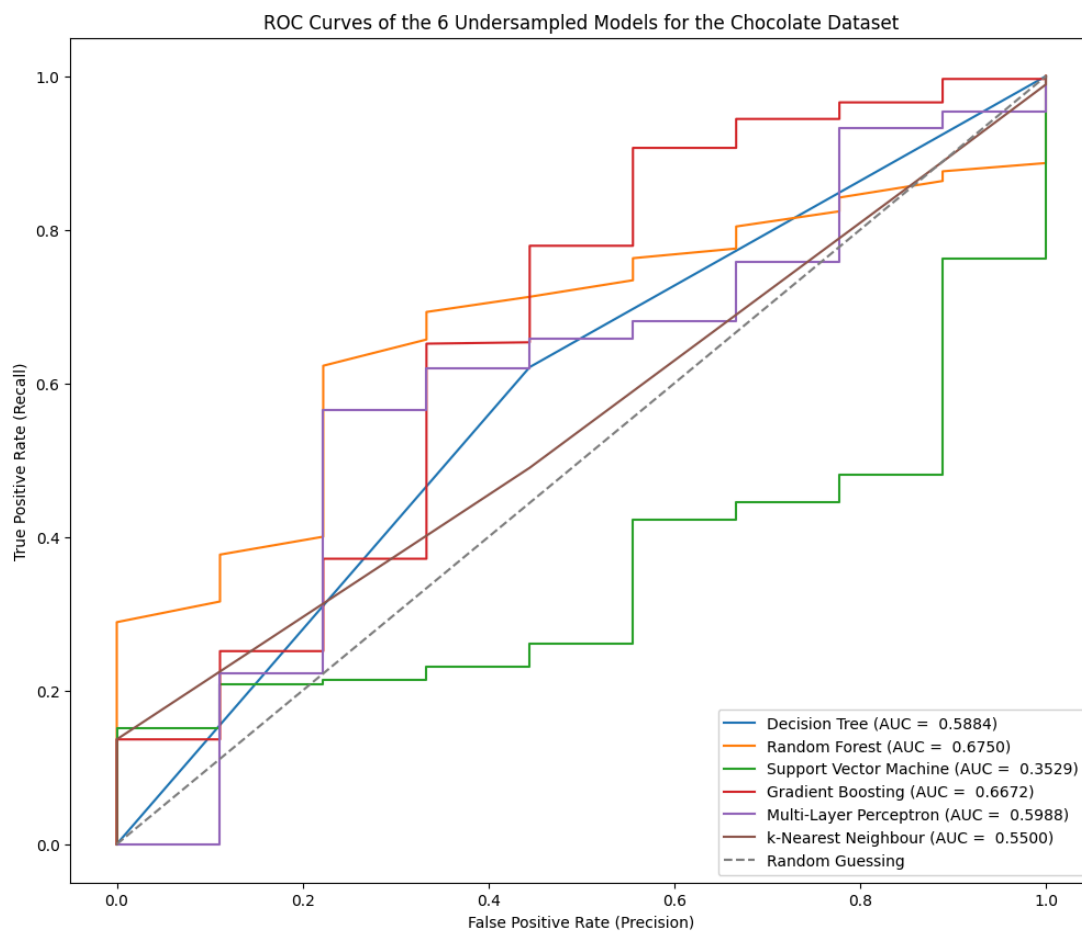
Chocolate Dataset

Model	Confusion Matrix	Recall	Precision
Decision Tree	[550, 8] [7, 1]	0.9874	0.9857
Random Forest	[557, 9] [0, 0]	1.0000	0.9841
Support Vector Machine	[557, 9] [0, 0]	1.0000	0.9841
Gradient Boosting	[555, 8] [2, 1]	0.9964	0.9858
Multi-Layer Perceptron	[554, 9] [3, 0]	0.9946	0.9840
k-Nearest Neighbor	[557, 9] [0, 0]	1.0000	0.9841



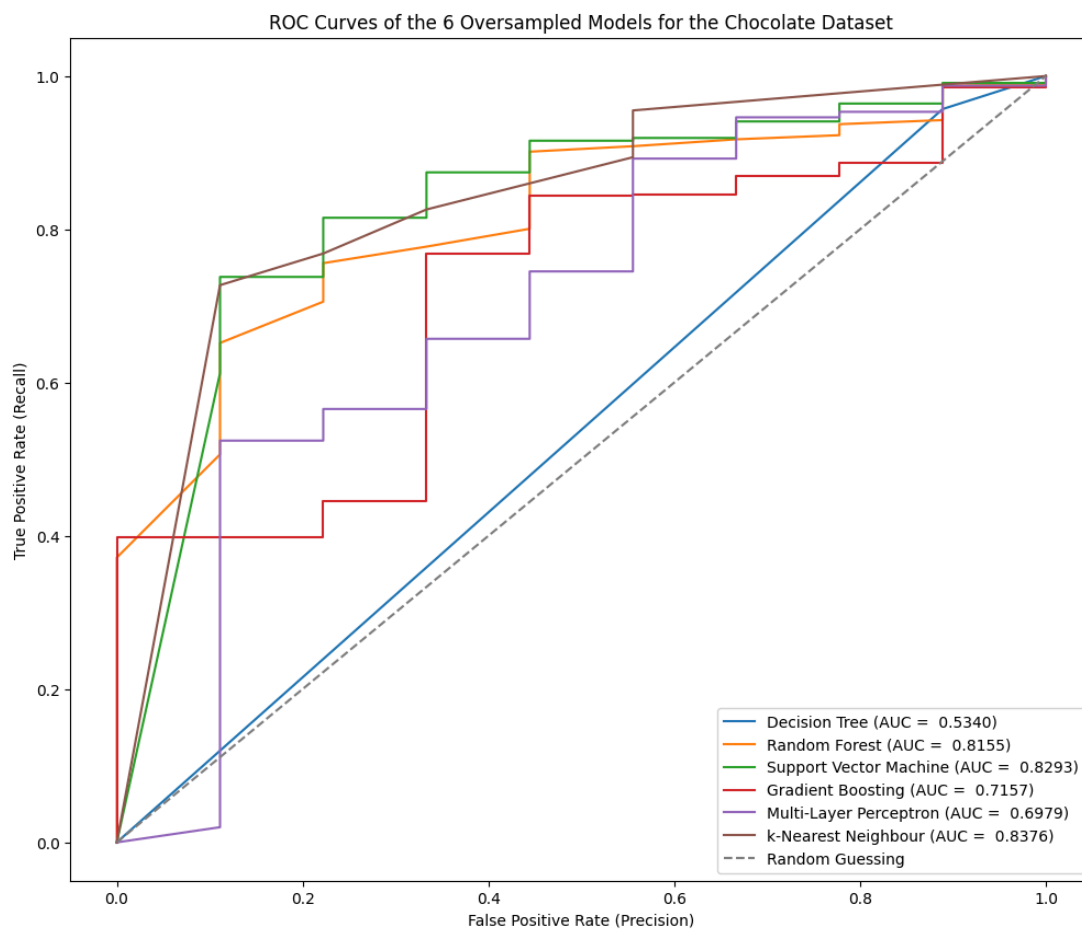
Chocolate Dataset with Under Sampling

Model	Confusion Matrix	Recall	Precision
Decision Tree	[346, 4] [211, 5]	0.6212	0.9886
Random Forest	[325, 2] [232, 7]	0.5835	0.9939
Support Vector Machine	[363, 4] [194, 5]	0.6517	0.9891
Gradient Boosting	[322, 3] [235, 6]	0.5781	0.9908
Multi-Layer Perceptron	[330, 3] [227, 6]	0.5925	0.9910
k-Nearest Neighbor	[273, 4] [284, 5]	0.4901	0.9856



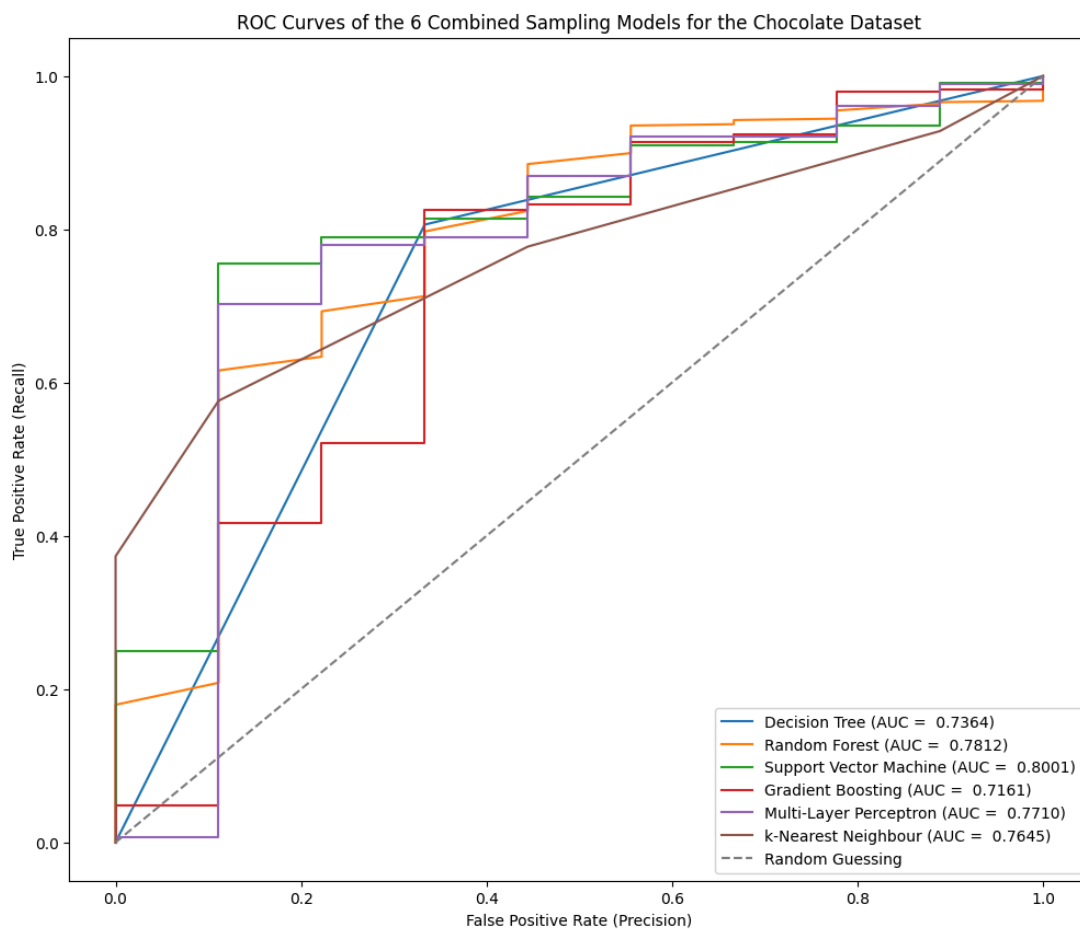
Chocolate Dataset with Over Sampling

Model	Confusion Matrix	Recall	Precision
Decision Tree	[533, 8] [24, 1]	0.9569	0.9852
Random Forest	[557, 9] [0, 0]	1.0000	0.9841
Support Vector Machine	[525, 7] [32, 2]	0.9425	0.9868
Gradient Boosting	[552, 9] [5, 0]	0.9910	0.9840
Multi-Layer Perceptron	[546, 8] [11, 1]	0.9803	0.9856
k-Nearest Neighbor	[460, 3] [97, 6]	0.8259	0.9935



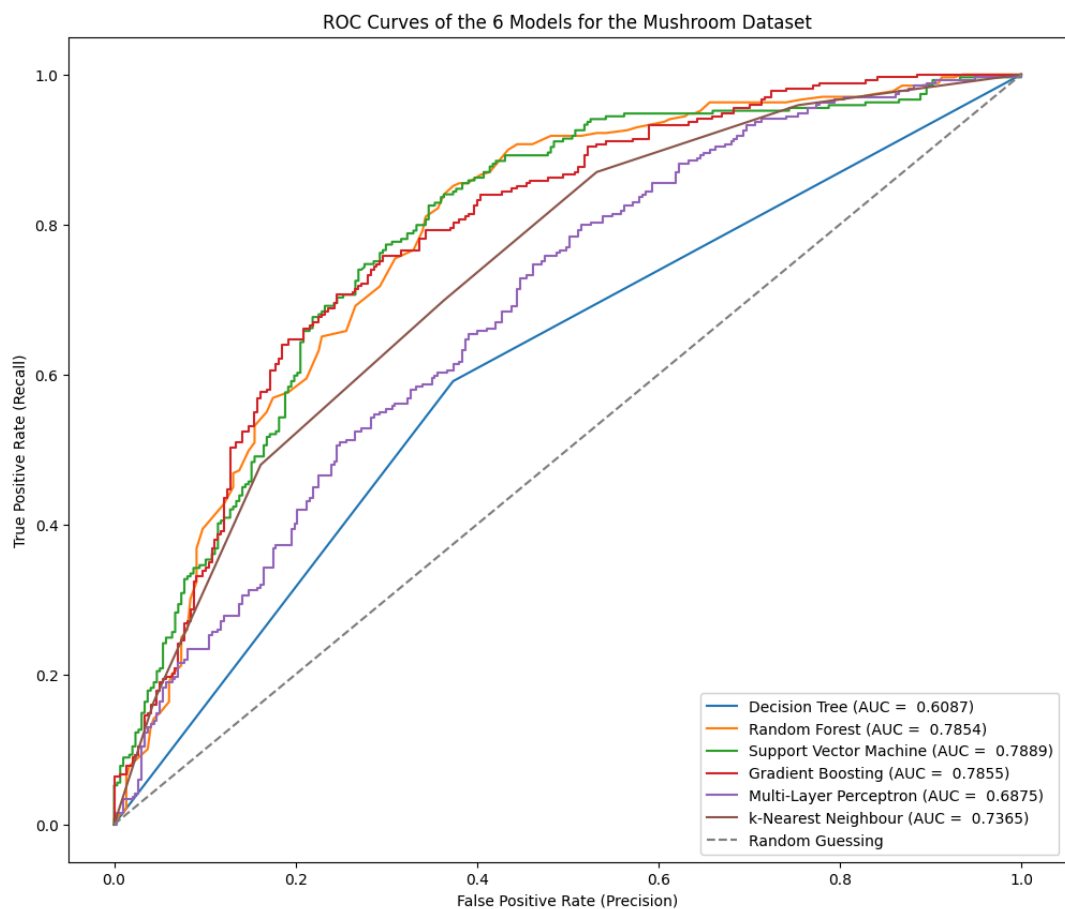
Chocolate Dataset with Combined Sampling

Model	Confusion Matrix	Recall	Precision
Decision Tree	[449, 3] [108, 6]	0.8061	0.9934
Random Forest	[521, 5] [36, 4]	0.9354	0.9905
Support Vector Machine	[482, 5] [75, 4]	0.8654	0.9897
Gradient Boosting	[501, 5] [56, 4]	0.8995	0.9901
Multi-Layer Perceptron	[485, 5] [72, 4]	0.8707	0.9898
k-Nearest Neighbor	[321, 1] [236, 8]	0.5763	0.9969



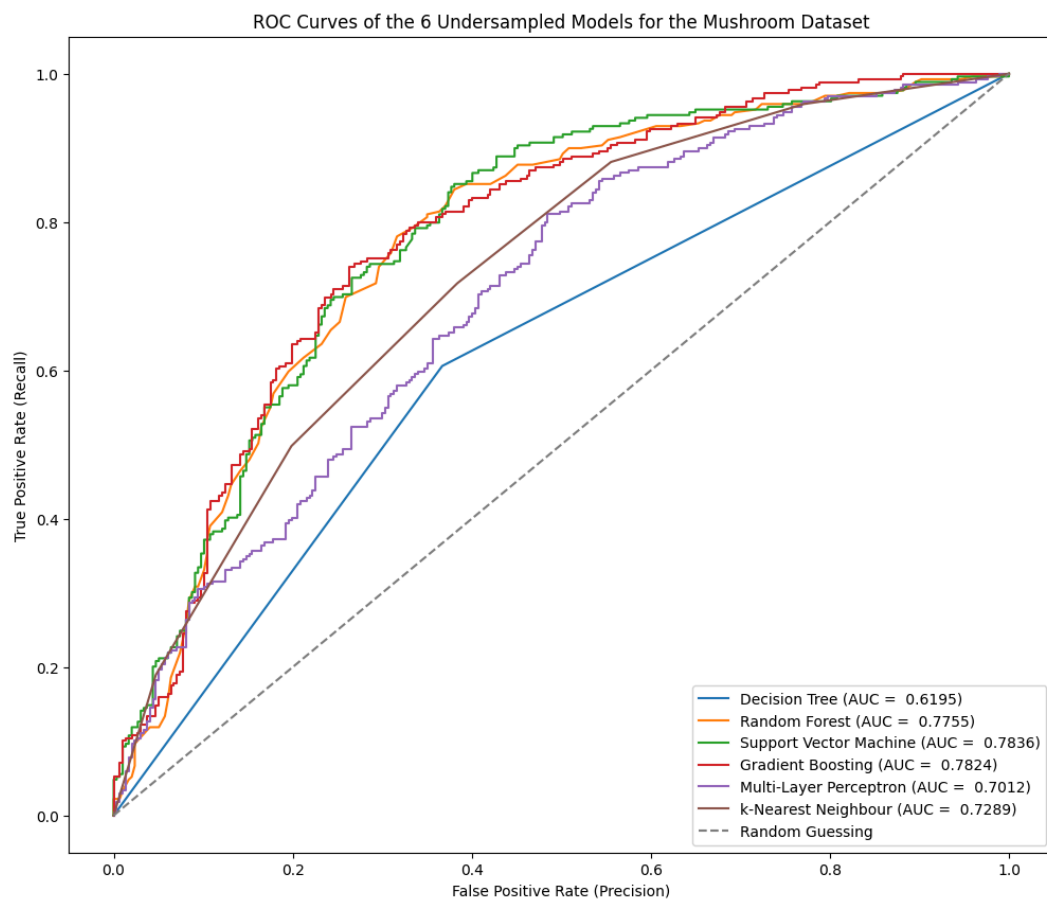
Magic Mushroom Dataset

Model	Confusion Matrix	Recall	Precision
Decision Tree	[159, 111] [110, 186]	0.5911	0.5889
Random Forest	[206, 98] [63, 199]	0.7658	0.6776
Support Vector Machine	[202, 87] [67, 210]	0.7509	0.6990
Gradient Boosting	[206, 95] [63, 202]	0.7658	0.6844
Multi-Layer Perceptron	[162, 107] [107, 190]	0.6022	0.6022
k-Nearest Neighbor	[188, 108] [81, 189]	0.6989	0.6351



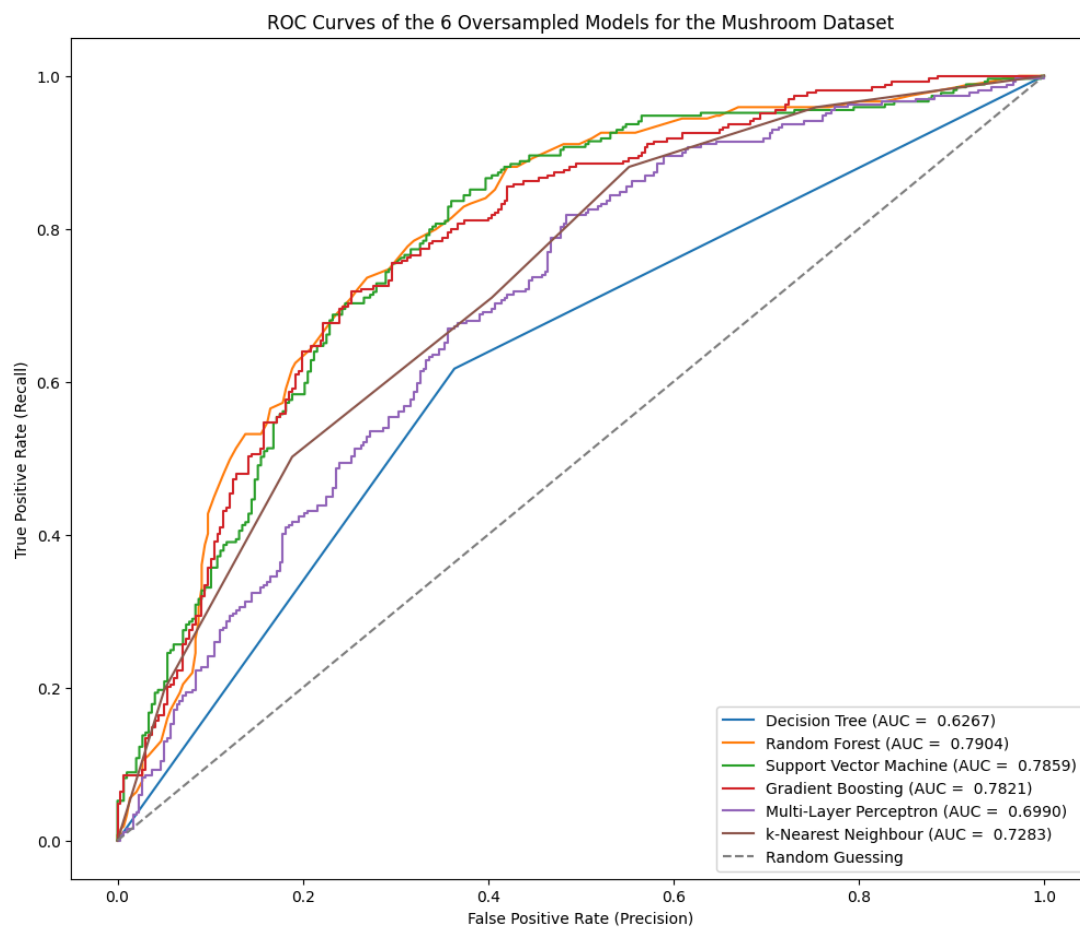
Magic Mushroom Dataset with Under Sampling

Model	Confusion Matrix	Recall	Precision
Decision Tree	[163, 109] [106, 188]	0.6059	0.5993
Random Forest	[213, 99] [56, 198]	0.7918	0.6827
Support Vector Machine	[213, 100] [56, 197]	0.7918	0.6805
Gradient Boosting	[214, 100] [55, 197]	0.7955	0.6815
Multi-Layer Perceptron	[168, 106] [101, 191]	0.6245	0.6131
k-Nearest Neighbor	[193, 114] [76, 183]	0.7175	0.6287



Magic Mushroom Dataset with Over Sampling

Model	Confusion Matrix	Recall	Precision
Decision Tree	[166, 108] [103, 189]	0.6171	0.6058
Random Forest	[205, 90] [64, 207]	0.7621	0.6949
Support Vector Machine	[208, 96] [61, 201]	0.7732	0.6842
Gradient Boosting	[208, 99] [61, 198]	0.7732	0.6775
Multi-Layer Perceptron	[175, 106] [94, 191]	0.6506	0.6228
k-Nearest Neighbor	[191, 120] [78, 177]	0.7100	0.6141



Magic Mushroom Dataset with Combined Sampling

Model	Confusion Matrix	Recall	Precision
Decision Tree	[163, 109] [106, 188]	0.6059	0.5993
Random Forest	[213, 99] [56, 198]	0.7918	0.6827
Support Vector Machine	[213, 100] [56, 197]	0.7918	0.6805
Gradient Boosting	[214, 100] [55, 197]	0.7955	0.6815
Multi-Layer Perceptron	[168, 106] [101, 191]	0.6245	0.6131
k-Nearest Neighbor	[193, 114] [76, 183]	0.7175	0.6287

