1) We first convert the multi-label datasets to binary. In order to do that, we pick one of the labels and consider it as a positive label (+1) and we consider all of the other labels as -1. Then we do the classification using the perceptron algorithm. We do this for every possible label value. Below is the obtained results for the tumor and mushroom datasets.

The perceptron algorithm performed well on both of the datasets. When compared to the previous algorithms, such as Decision Tree and KNN, the perceptron performed well especially on the tumor dataset. The reason is that we classified a very simplified version of the tumor dataset which is a binary dataset. Compared to a multi-label dataset with 22 labels, a binary dataset is much easier to classify.

Furthermore, the number of epochs slightly effected the results: 100 epochs gave better results compared to 1 epoch and 1000 epochs was just slightly better than 100 epochs.

The Primary Tumor Dataset:

Label	Epochs = 1			Epochs = 100			Epochs = 1000		
	Train Acc	Val Acc	Test Acc	Train Acc	Val Acc	Test Acc	Train Acc	Val Acc	Test Acc
1	86.98	78.82	81.18	93.49	80.00	84.71	92.90	78.82	78.82
2	98.82	98.82	100.00	98.82	96.47	98.82	98.82	92.94	98.82
3	65.09	69.41	55.29	96.45	91.76	94.12	98.22	95.29	95.29
4	97.63	95.29	92.94	100.00	94.12	90.59	100.00	94.12	91.76
5	84.62	82.35	74.12	83.43	80.00	85.88	85.21	95.29	87.06
6	100.00	100.00	98.82	100.00	100.00	98.82	100.00	100.00	98.82
7	96.45	92.94	97.65	94.08	80.00	91.76	96.45	91.76	97.65
8	98.22	98.82	97.65	99.41	97.65	98.82	99.41	97.65	98.82
9	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
10	100.00	98.82	98.82	100.00	98.82	98.82	100.00	98.82	98.82
11	90.53	90.59	90.59	89.35	89.41	92.94	91.12	90.59	89.41
12	93.49	95.29	94.12	97.04	95.29	95.29	95.86	94.12	90.59
13	97.63	97.65	98.82	99.41	91.76	91.76	99.41	92.94	91.76
14	91.72	90.59	97.65	90.53	83.53	96.47	84.62	82.35	84.71
15	100.00	98.82	98.82	100.00	98.82	98.82	100.00	98.82	98.82
16	99.41	100.00	100.00	100.00	98.82	98.82	100.00	100.00	100.00
17	97.63	97.65	95.29	98.22	96.47	95.29	97.63	97.65	94.12
18	91.72	90.59	91.76	95.86	92.94	89.41	95.27	91.76	91.76
19	97.63	97.65	98.82	100.00	92.94	89.41	100.00	96.47	97.65
20	99.41	98.82	100.00	100.00	98.82	98.82	100.00	98.82	100.00
21	99.41	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
22	95.27	94.12	92.94	97.63	94.12	91.76	100.00	94.12	90.59

The Mushroom Data Set:

Label	Epochs = 1			Epochs = 100			Epochs = 1000		
	Train Acc	Val Acc	Test Acc	Train Acc	Val Acc	Test Acc	Train Acc	Val Acc	Test Acc
e	98.70	98.08	98.28	100.00	99.85	99.90	100.00	100.00	100.00
p	99.88	99.75	99.85	100.00	99.95	100.00	100.00	99.90	99.90

2) We then run the perceptron algorithm on the Mysterious dataset. Below is the results. The results are not good. The best test accuracy is 58.8 percent. When we run the previous algorithms such as KNN, we will get a much better results around 97 percent accuracy for the test dataset.

The Mysterious Data Set:

Label	Epochs = 1			Epochs = 100			Epochs = 1000		
	Train Acc	Val Acc	Test Acc	Train Acc	Val Acc	Test Acc	Train Acc	Val Acc	Test Acc
y	58.40	58.80	58.80	48.00	46.00	51.20	41.80	41.60	40.80
n	52.20	54.00	48.80	50.20	50.40	44.80	50.40	48.80	52.00

We tried to improve the accuracy by generating new features from the existing ones. The new features were created from tuples of length 2 of the existing features. In other words, we get the binary product of the existing features and add them as new features to the existing ones. This improved the results dramatically as demonstrated below. We also tried the ternary products of the features, but it did not get a better accuracy. The obtained accuracy is similar to the accuracy of the previous algorithms such as KNN.

Epochs = 1			Epochs = 100			Epochs = 1000		
Train Acc	Val Acc	Test Acc	Train Acc	Val Acc	Test Acc	Train Acc	Val Acc	Test Acc
97.20	98.00	98.00	98.00	96.00	97.60	98.00	96.80	97.60