ECE 498 HW5

Name: Shuyue Lai
NetID: shuyuel2

Problem 1:

1. For each of the above perform 5-fold cross-validation of the data with 80%-20% train test split. Report the mean accuracy, precision and recall on the test set. Report precision and recall for the positive class (pulsar). Also report the respective standard deviations. Provide the results in a table.

Linear SVM:

	Accuracy		Precision		Recall	
C	mean	std	mean	std	mean	std
0.1	0.97838	0.00254	0.94569	0.01702	0.81128	0.01487
1	0.97938	0.00221	0.94246	0.01414	0.82493	0.02889
10	0.97927	0.00134	0.93970	0.01263	0.82593	0.01589

Decision Trees:

	Accuracy		Precision		Recall	
Max Depth	mean	std	mean	std	mean	std
3	0.97793	0.00352	0.90676	0.02184	0.84585	0.03192
4	0.97867	0.00209	0.91653	0.00799	0.84371	0.03016
6	0.97737	0.00251	0.90543	0.02379	0.84069	0.02626

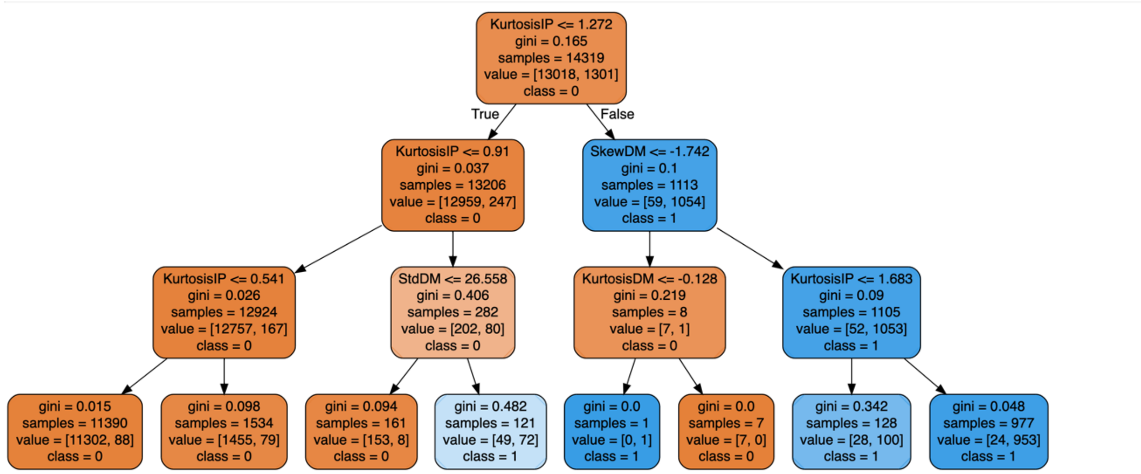
Random Forests:

	Accuracy		Precision		Recall	
Number of Estimator	mean	std	mean	std	mean	std
5	0.97837	0.00294	0.93668	0.01284	0.82049	0.01313
11	0.97820	0.00111	0.93730	0.01640	0.81724	0.01684
13	0.97882	0.00231	0.93934	0.00879	0.82192	0.02496

2. Explain the high value of accuracy compared to precision and recall.
   Accuracy = (TP + TN) / (TP + TN + FP + FN)
   Precision = TP / (TP + FP)
   Recall = TP / (TP + FN)
   According to the result, it showed that the dataset is unbalanced. Negative class is much more than positive class.

3. Which classifier performs the best? Explain the reason why the picked classifier performed the best.
   Since the result of accuracy is similar, Linear SVM(C==1) is best according to all three values: accuracy, precision and recall.

4. Visualize a decision tree (depth=3) from one iteration of the cross-validation. You can visualize using the following code for ease. As a part of your submission, turn in an image of the generated tree.

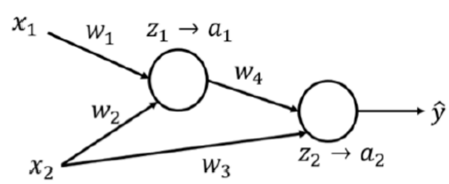


Problem 2:

1. $z_1 = w_1^T x_1 + w_2^T x_2 + b_1 \quad a_1 = \sigma(z_1)$

$z_2 = w_4^T a_1 + w_3^T x_2 + b_2 \quad a_2 = \hat{y} = \sigma(z_2)$

2. $L = \|\hat{y}_i - y_i\|^2$



$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a_2}\frac{\partial a_2}{\partial w_1}$$

$$= \frac{\partial L}{\partial a_2}\frac{\partial a_2}{\partial z_2}\frac{\partial z_2}{\partial w_1}$$

$$= \frac{\partial L}{\partial a_2}\frac{\partial a_2}{\partial z_2}\frac{\partial z_2}{\partial a_1}\frac{\partial a_1}{\partial w_1}$$

$$= \frac{\partial L}{\partial a_2}\frac{\partial a_2}{\partial z_2}\frac{\partial z_2}{\partial a_1}\frac{\partial a_1}{\partial z_1}\frac{\partial z_1}{\partial w_1}$$

$$= 2(\hat{y}-y)\,\sigma(z_2)\sigma(-z_2)\,w_4\,\sigma(z_1)\,\sigma(-z_1)\,x_1$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial a_2}\frac{\partial a_2}{\partial w_2}$$

$$= \frac{\partial L}{\partial a_2}\frac{\partial a_2}{\partial z_2}\frac{\partial z_2}{\partial w_2}$$

$$= \frac{\partial L}{\partial a_2}\frac{\partial a_2}{\partial z_2}\frac{\partial z_2}{\partial a_1}\frac{\partial a_1}{\partial z_1}\frac{\partial z_1}{w_2}$$

$$= 2(\hat{y}-y)\,\sigma(z_2)\,\sigma(-z_2)\,w_4\,\sigma(z_1)\,\sigma(-z_1)\,x_2$$

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial a_2}\frac{\partial a_2}{\partial z_2}\frac{\partial z_2}{\partial w_3}$$

$$= 2(\hat{y}-y)\,\sigma(z_2)\,\sigma(-z_2)\,x_2$$

$$\frac{\partial L}{\partial w_4} = \frac{\partial L}{\partial a_2}\frac{\partial a_2}{\partial z_2}\frac{\partial z_2}{\partial w_4}$$

$$= 2(\hat{y}-y)\,\sigma(z_2)\,\sigma(-z_2)\,a_1$$

3. $$w_1^{(t+1)} = w_1^{(t)} - \eta\frac{\partial L}{\partial w_1^{(t)}}$$