CS 4375

ASSIGNMENT 2

Names of students in your group:

Shruti Bindingnavile

Number of free late days used: _1_____

Note: You are allowed a **total** of 4 free late days for the **entire semester**. You can use at most 2 for each assignment. After that, there will be a penalty of 10% for each late day.

Please list clearly all the sources/references that you have used in this assignment.

https://archive.ics.uci.edu/dataset/94/spambase

```
activation="tanh", solver="sgd", learning_rate_init = 0.001, max_iter=1000, hidden_layer_sizes=(15,2)
```

```
Training:
              precision    recall  f1-score   support

           0       0.89      0.59      0.71      2222
           1       0.59      0.88      0.71      1458

    accuracy                           0.71      3680
   macro avg       0.74      0.74      0.71      3680
weighted avg       0.77      0.71      0.71      3680

Testing:
              precision    recall  f1-score   support

           0       0.89      0.61      0.72       566
           1       0.58      0.88      0.70       355

    accuracy                           0.71       921
   macro avg       0.73      0.74      0.71       921
weighted avg       0.77      0.71      0.71       921
```

```
activation="logistic", solver="sgd", learning_rate_init = 0.01, max_iter=1000,
hidden_layer_sizes=(20,2)
```

```
Training:
              precision    recall  f1-score   support

           0       0.67      0.93      0.78      2222
           1       0.74      0.31      0.43      1458

    accuracy                           0.68      3680
   macro avg       0.70      0.62      0.61      3680
weighted avg       0.70      0.68      0.64      3680

Testing:
              precision    recall  f1-score   support

           0       0.68      0.94      0.79       566
           1       0.74      0.28      0.41       355

    accuracy                           0.69       921
   macro avg       0.71      0.61      0.60       921
weighted avg       0.70      0.69      0.64       921
```

```
activation="logistic", solver="sgd", learning_rate_init = 0.0001, max_iter=1000, hidden_layer_sizes=
(15,2))
```

```
Training:
              precision    recall  f1-score   support

           0       0.60      1.00      0.75      2222
           1       0.00      0.00      0.00      1458

    accuracy                           0.60      3680
   macro avg       0.30      0.50      0.38      3680
weighted avg       0.36      0.60      0.45      3680

Testing:
              precision    recall  f1-score   support

           0       0.61      1.00      0.76       566
           1       0.00      0.00      0.00       355

    accuracy                           0.61       921
   macro avg       0.31      0.50      0.38       921
weighted avg       0.38      0.61      0.47       921
```

```
activation="logistic", solver="sgd", learning_rate_init = 0.01, max_iter=1000, hidden_layer_sizes=
(100,2)
```

```
Training:
              precision    recall  f1-score   support

           0       0.80      0.81      0.81      2222
           1       0.71      0.70      0.70      1458

    accuracy                           0.77      3680
   macro avg       0.76      0.76      0.76      3680
weighted avg       0.77      0.77      0.77      3680

Testing:
              precision    recall  f1-score   support

           0       0.79      0.82      0.80       566
           1       0.69      0.66      0.68       355

    accuracy                           0.76       921
   macro avg       0.74      0.74      0.74       921
weighted avg       0.76      0.76      0.76       921
```

```
activation="tanh", solver="sgd", learning_rate_init = 0.01, max_iter=1000, hidden_layer_sizes= (100,2)
```

Training:

	precision	recall	f1-score	support
0	0.60	1.00	0.75	2222
1	0.60	0.00	0.01	1458
accuracy			0.60	3680
macro avg	0.60	0.50	0.38	3680
weighted avg	0.60	0.60	0.46	3680

Testing:

	precision	recall	f1-score	support
0	0.62	1.00	0.76	566
1	1.00	0.00	0.01	355
accuracy			0.62	921
macro avg	0.81	0.50	0.38	921
weighted avg	0.76	0.62	0.47	921

```
activation="relu", solver="sgd", learning_rate_init = 0.01, max_iter=1000, hidden_layer_sizes= (100,2)
```

Training:

	precision	recall	f1-score	support
0	0.60	1.00	0.75	2222
1	0.00	0.00	0.00	1458
accuracy			0.60	3680
macro avg	0.30	0.50	0.38	3680
weighted avg	0.36	0.60	0.45	3680

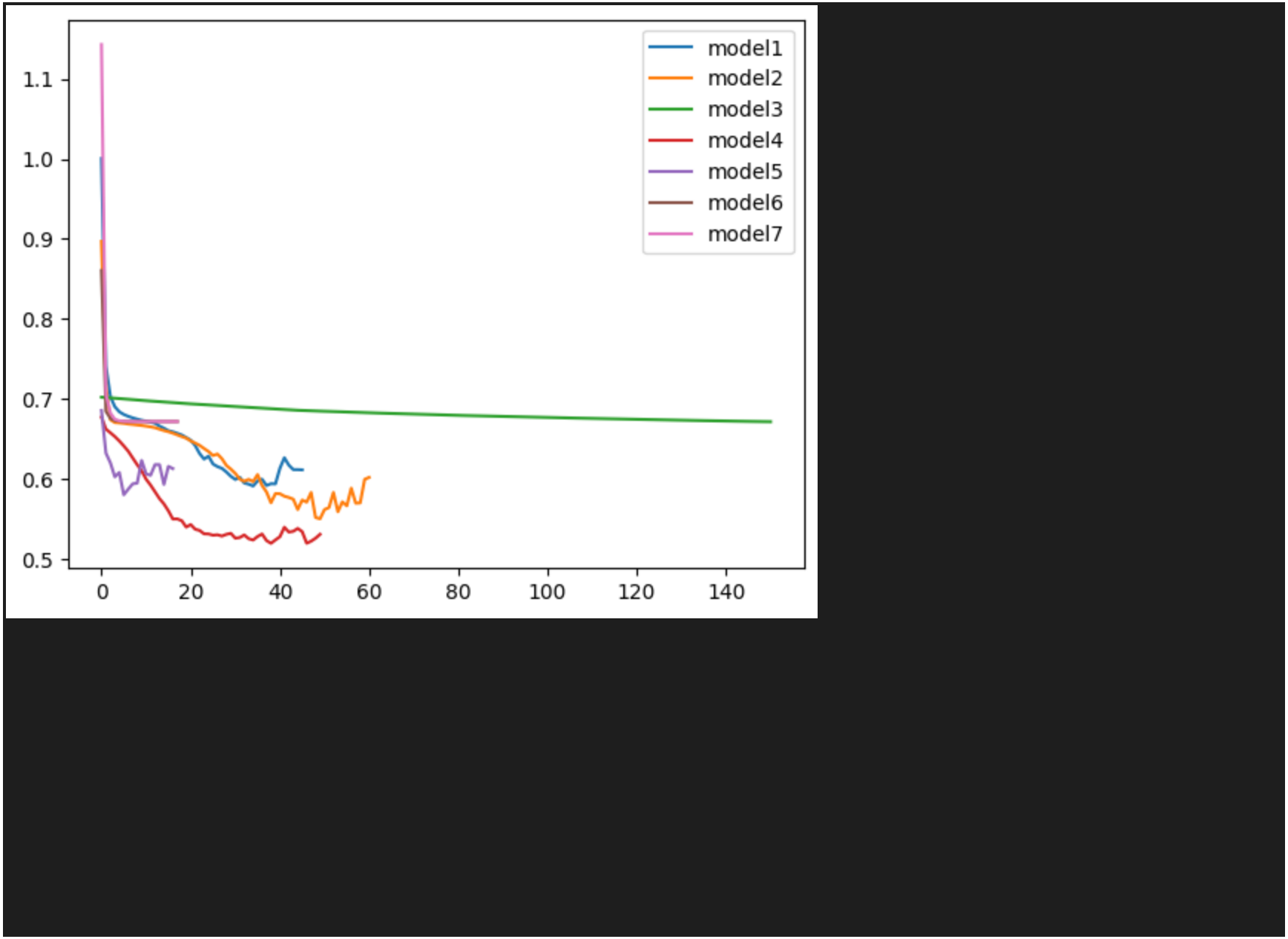
Testing:

	precision	recall	f1-score	support
0	0.61	1.00	0.76	566
1	0.00	0.00	0.00	355
accuracy			0.61	921
macro avg	0.31	0.50	0.38	921
weighted avg	0.38	0.61	0.47	921

```
activation="logistic", solver="sgd", learning_rate_init = 0.000000001, max_iter=1000,
hidden_layer_sizes= (100,2))
```

```
Training:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00      2222
           1       0.40      1.00      0.57      1458

    accuracy                           0.40      3680
   macro avg       0.20      0.50      0.28      3680
weighted avg       0.16      0.40      0.22      3680

Testing:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       566
           1       0.39      1.00      0.56       355

    accuracy                           0.39       921
   macro avg       0.19      0.50      0.28       921
weighted avg       0.15      0.39      0.21       921
```

Activation	Learning rate	Epochs	Hidden layer	Training accuracy	Testing accuracy
logistic	0.01	1000	(20,2)	0.72	0.72
tanh	0.001	10000	(15,2)	0.68	0.67
logistic	.0001	1000	(15,2)	0.6	0.61
logistic	.01	1000	(100,2)	0.77	0.76
tanh	.01	1000	(100,2)	0.71	0.71
relu	.01	1000	(100,2)	0.60	0.61
logistic	0.000000001	1000	(100,2)	0.40	0.39


model1
model2
model3
model4
model5
model6
model7
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0
20
40
60
80
100
120
140

The activation function that did the best was the sigmoid (or logistic function). This is what I expected because the range of the sigmoid function goes from 0 to 1, making it a good candidate for a binary classification model such as this one.