Breast Cancer Prediction

Faculty of Computers & Artificial Intelligence

Selected 1

Repository Link: https://github.com/mahmoudhaney/BreastCancerPrediction

Team 15

	ID	NAME	LEVEL	Department
1	202000186	آية عماد طلعت حرب	3	CS
2	202000376	ساره وليد سيد	3	CS
3	202000649	فيروز جمعه منصور محمد	3	CS
4	202000677	كريمه ابو حسيبه محارب	3	CS
5	202000861	محمود هاني سعيد	3	CS
6	202000895	مصطفى جمال عبد الحكيم	3	CS
7	202001010	نورهان طارق عبدالهادي	3	CS

Dataset Information

Name

Breast Cancer Prediction

Classes

id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean

compactness_mean	concavity_mean	concave_points_mean	symmetry_mean

fractal_dimension_mean	radius_se	texture_se	perimeter_se	area_se	smoothness_se

compactness_se	concavity_se	concave_points_se	symmetry_se	fractal_dimension_se

radius_worst	texture_worst	perimeter_worst	area_worst	smoothness_worst

compactness_worst	concavity_worst	concave_points_worst

symmetry_worst	fractal_dimension_worst

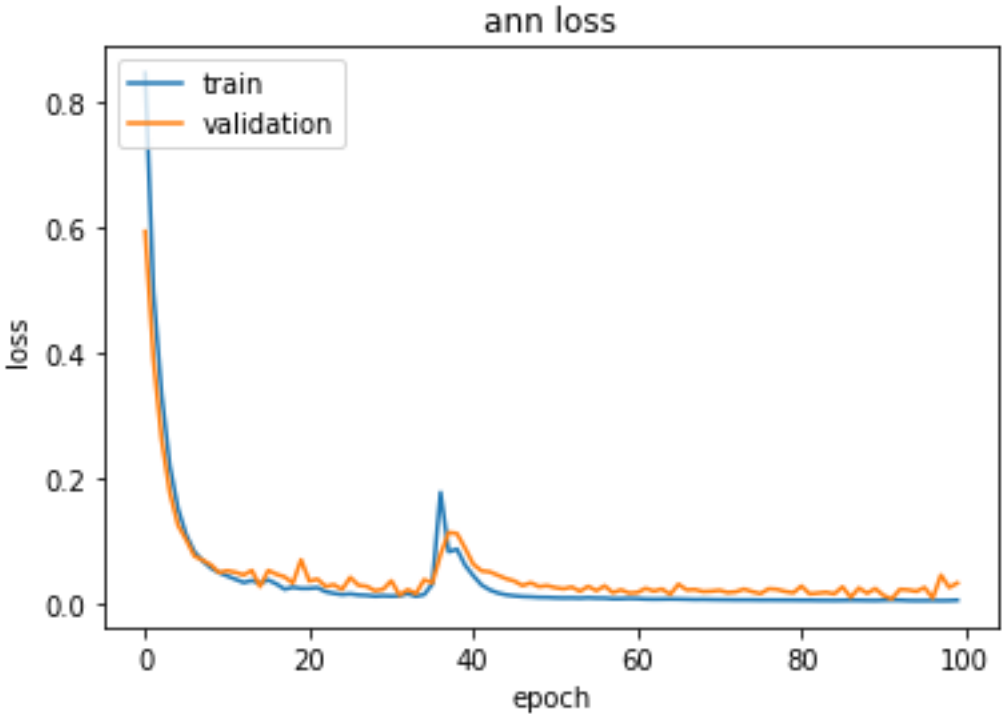
Total Dataset Columns → **32 Columns**

Total Dataset Rows → **570 Rows**

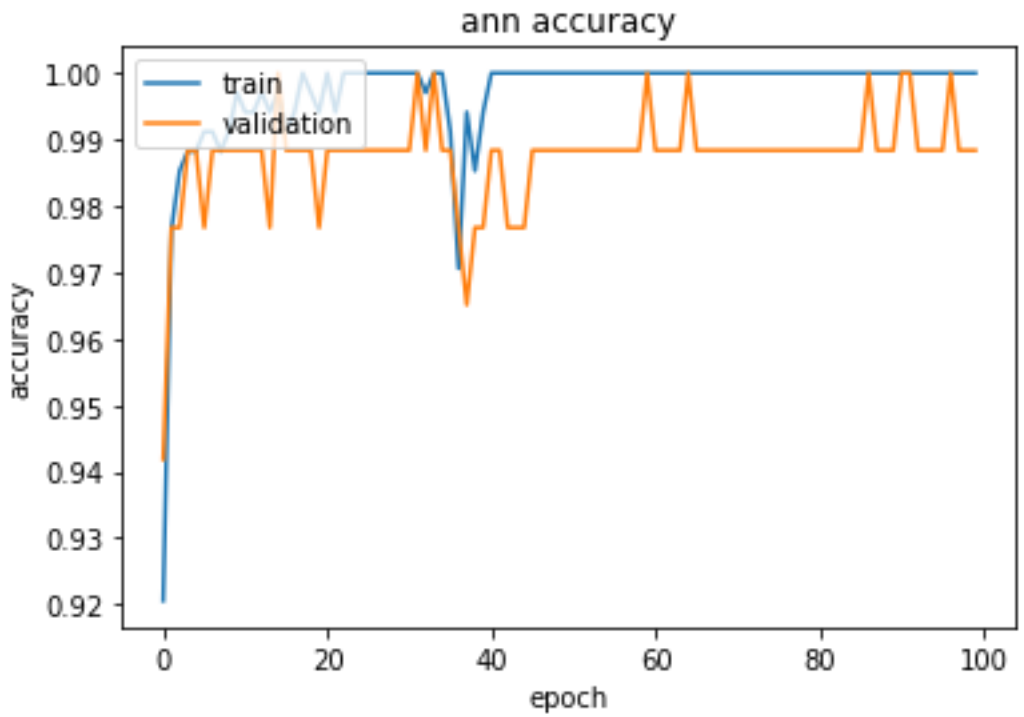
Training Data → **426**

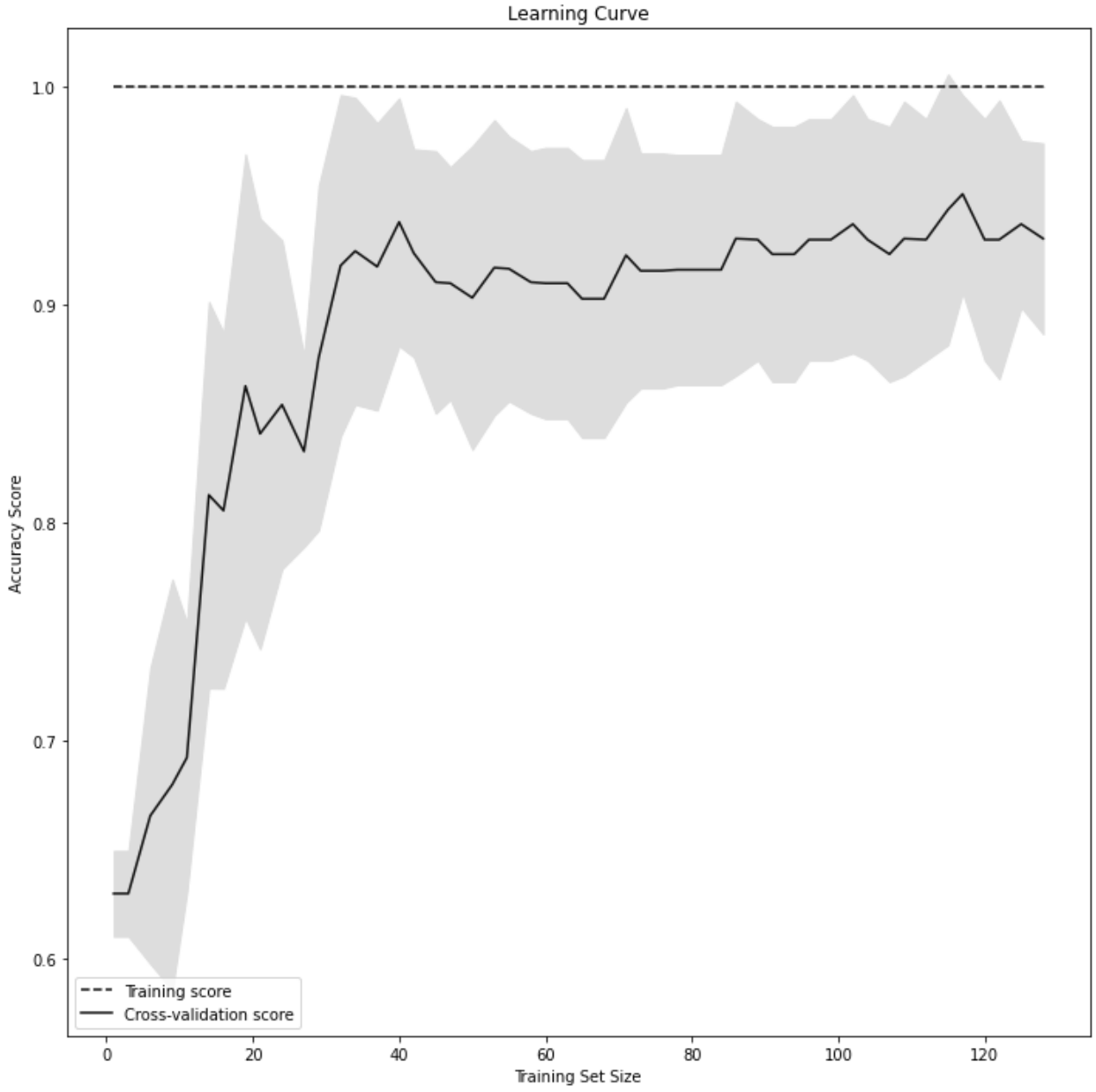
Validation Data → **143**

Loss Curve

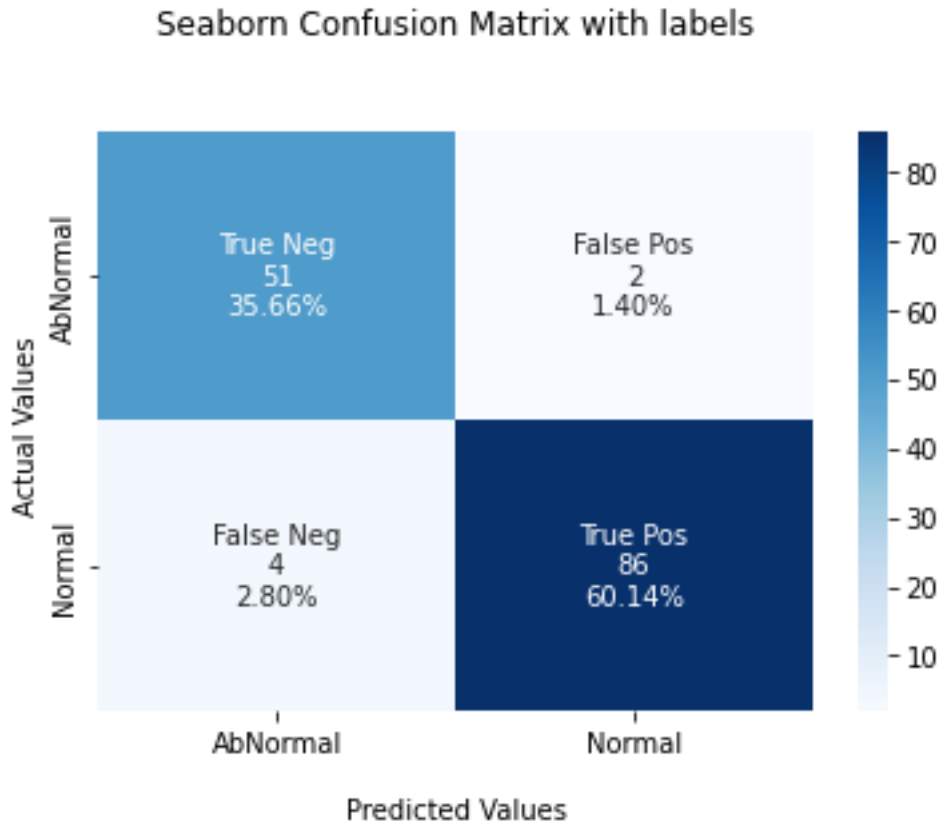

ann loss
train
validation
loss
epoch
0.0
0.2
0.4
0.6
0.8
0
20
40
60
80
100

Accuracy Curve

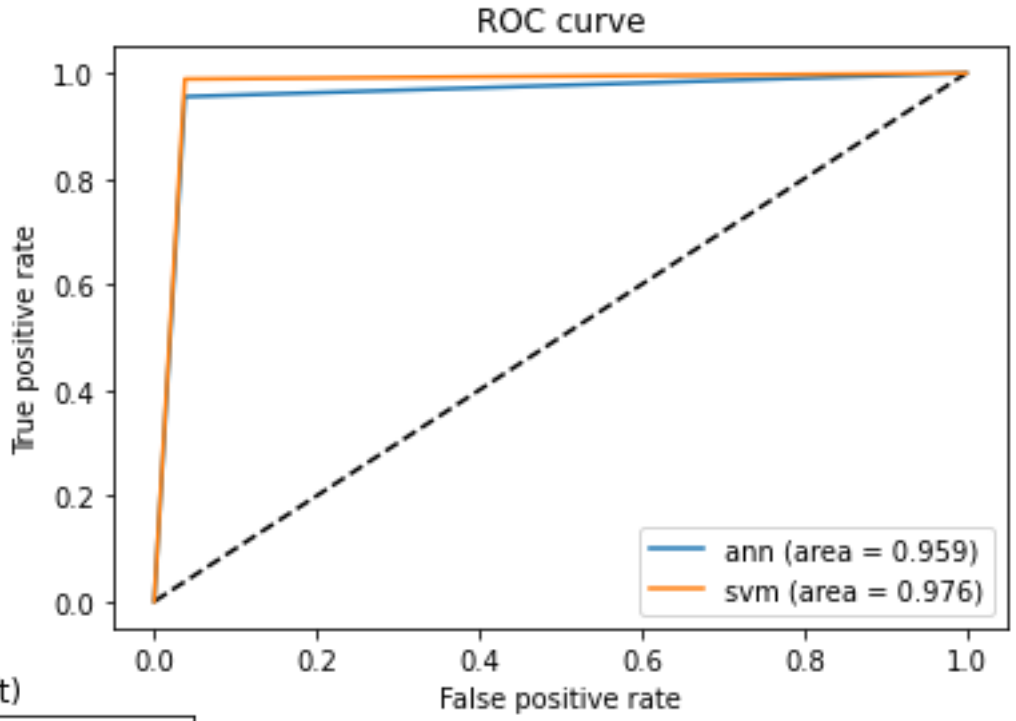


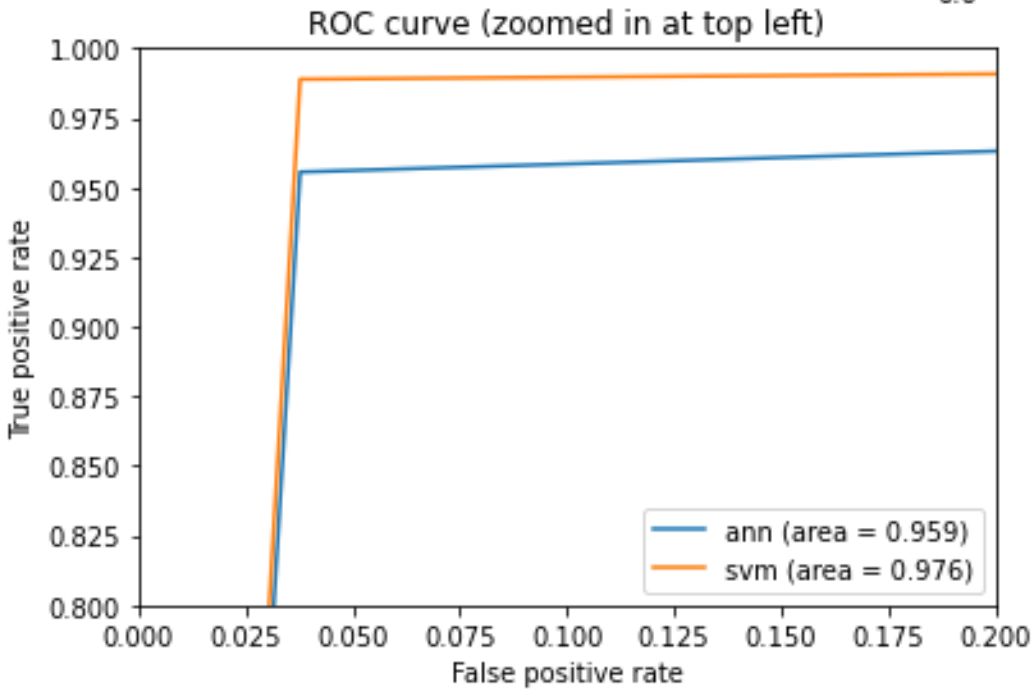

Learning Curve
Accuracy Score
Training Set Size
1.0
0.9
0.8
0.7
0.6
0
20
40
60
80
100
120
Training score
Cross-validation score

Confusion Matrix



ROC Curve


ROC curve
True positive rate
False positive rate
0.0
0.2
0.4
0.6
0.8
1.0
ann (area = 0.959)
svm (area = 0.976)


ROC curve (zoomed in at top left)
True positive rate
False positive rate
1.000
0.975
0.950
0.925
0.900
0.875
0.850
0.825
0.800
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
0.200
ann (area = 0.959)
svm (area = 0.976)

Hyperparameters

SVM & Logistic Regression Models

Before

Model Accuracy: 94%

Without Standardization

Training Data: 60%

After

Model Accuracy: 97%

With Standardization (-1, 1)

Training Data: 75%