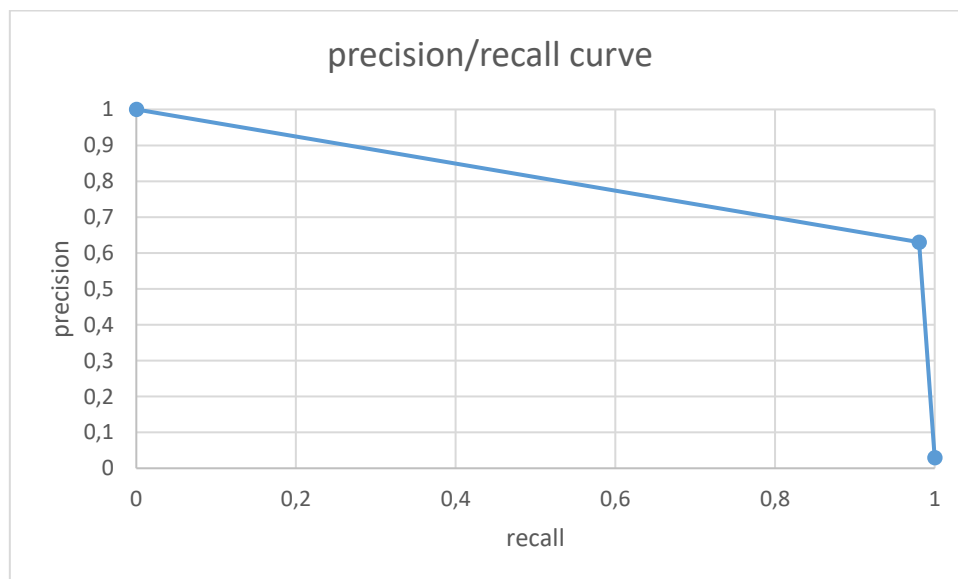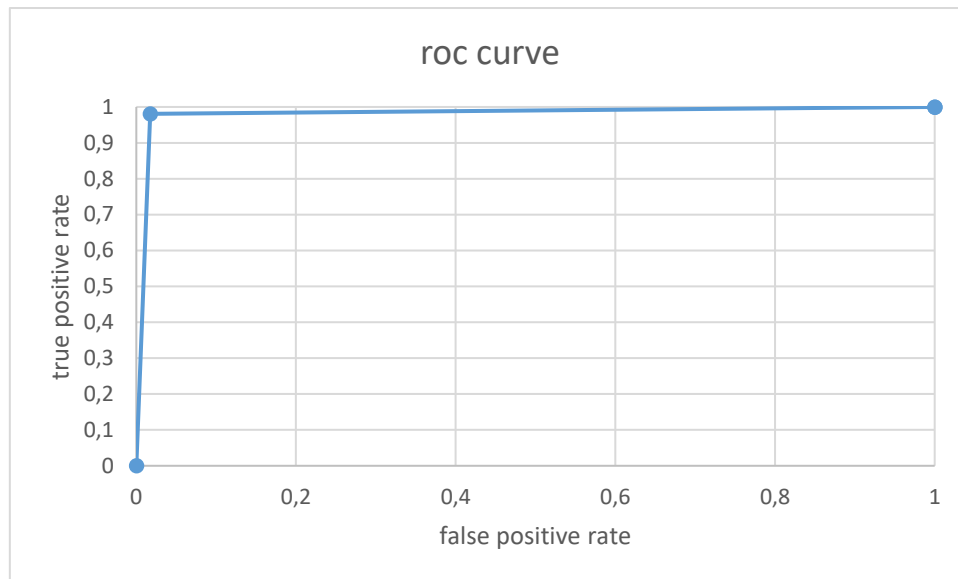**LogisticRegressionTrainAndTest1_2Sample100_500_1000_iteration100**

```
anzahl links true (training): 50
anzahl links true (test): 1903
anzahl links false (training): 50
anzahl links false (test): 62681
--> gesamtanzahl der eingelesen trainingsdaten: 100
--> gesamtanzahl der eingelesen testdaten: 64584
model coefficients: [25.11829476579835,-
6.234802699944511,11.073323323731753]
model threshold: 0.5
coefficient titel: 25.11829476579835
coefficient artist: -6.234802699944511
coefficient album: 11.073323323731753
confusion matrix:
61585.0  1096.0
37.0     1866.0
true positive: 1866.0
true negative: 61585.0
false positive: 37.0
false negative: 1096.0
accuracy: 0.9824569552830422
precision: 0.9824569552830422
f measure: 0.9824569552830422
weighted f measure: 0.9842915874441603
weighted precision: 0.9885144244883024
recall: 0.9824569552830422
weighted recall: 0.9824569552830422
weighted true positive rate: 0.9824569552830422
weighted false positive rate: 0.01938530243446314
precision-recall-curve: [(0.0,1.0),
(0.9805570152390962,0.6299797434166104), (1.0,0.029465502291589247)]
roc-curve: [(0.0,0.0), (0.017485362390517064,0.9805570152390962),
(1.0,1.0), (1.0,1.0)]
```
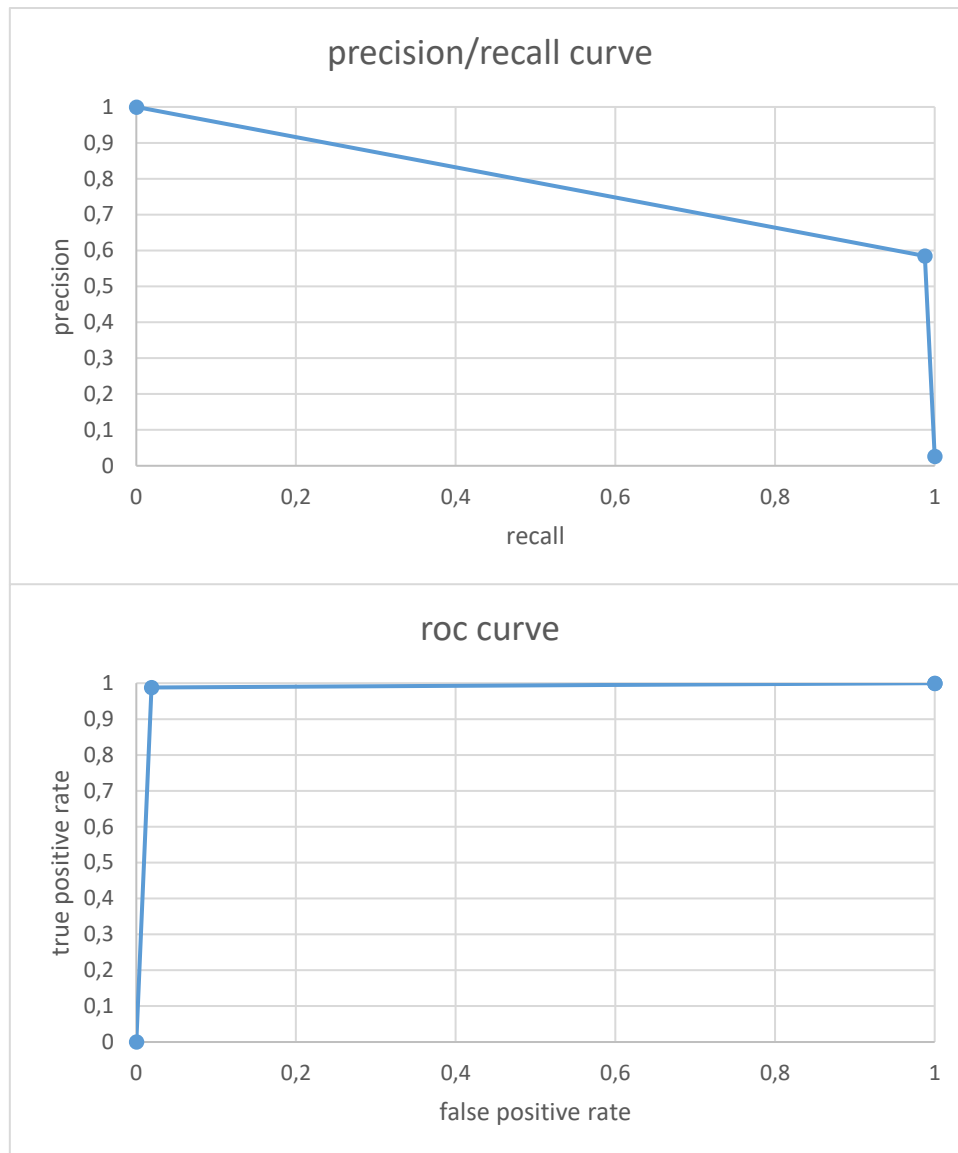
LogisticRegressionTrainAndTest1_2Sample100_500_1000_iteration500

```
anzahl links true (training): 250
anzahl links true (test): 1703
anzahl links false (training): 250
anzahl links false (test): 62481
--> gesamtanzahl der eingelesen trainingsdaten: 500
--> gesamtanzahl der eingelesen testdaten: 64184
model coefficients: [28.528647263152724,-
5.283015286470511,13.560836462675132]
model threshold: 0.5
coefficient titel: 28.528647263152724
coefficient artist: -5.283015286470511
coefficient album: 13.560836462675132
confusion matrix:
61286.0  1195.0
21.0      1682.0
true positive: 1682.0
true negative: 61286.0
false positive: 21.0
false negative: 1195.0
accuracy: 0.9810544684033404
precision: 0.9810544684033404
f measure: 0.9810544684033404
weighted f measure: 0.9833928009495024
weighted precision: 0.9886456794027211
recall: 0.9810544684033404
weighted recall: 0.9810544684033404
weighted true positive rate: 0.9810544684033404
weighted false positive rate: 0.012511462920983214
precision-recall-curve: [(0.0,1.0),
(0.9876688197298884,0.5846367744177963), (1.0,0.026533092359466534)]
roc-curve: [(0.0,0.0), (0.01912581424753125,0.9876688197298884), (1.0,1.0),
(1.0,1.0)]
```
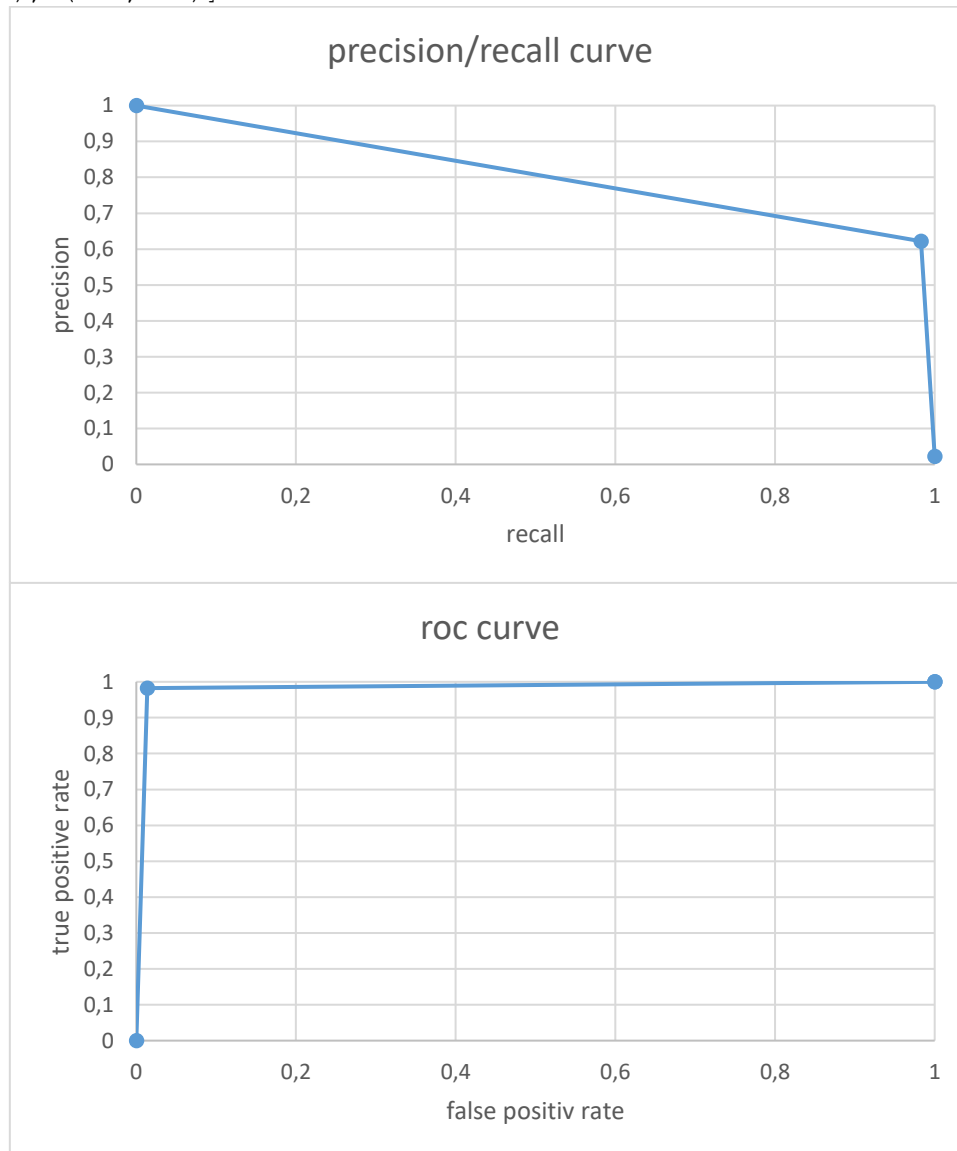
precision/recall curve

roc curve

LogisticRegressionTrainAndTest1_2Sample100_500_1000_iteration1000

```
anzahl links true (training): 500
anzahl links true (test): 1453
anzahl links false (training): 500
anzahl links false (test): 62231
--> gesamtanzahl der eingelesen trainingsdaten: 1000
--> gesamtanzahl der eingelesen testdaten: 63684
model coefficients: [18.86427657243907,-4.17363628120943,11.76508990036071]
model threshold: 0.5
coefficient titel: 18.86427657243907
coefficient artist: -4.17363628120943
coefficient album: 11.76508990036071
confusion matrix:
61362.0   869.0
25.0      1428.0
true positive: 1428.0
true negative: 61362.0
false positive: 25.0
false negative: 869.0
accuracy: 0.9859619370642547
```

```
precision: 0.9859619370642547
f measure: 0.9859619370642547
weighted f measure: 0.9874937647502086
weighted precision: 0.9909703846630789
recall: 0.9859619370642547
weighted recall: 0.9859619370642547
weighted true positive rate: 0.9859619370642547
weighted false positive rate: 0.017131819699543745
precision-recall-curve: [(0.0,1.0),
(0.9827942188575362,0.6216804527644754), (1.0,0.02281577790339803)]
roc-curve: [(0.0,0.0), (0.013964101492825119,0.9827942188575362),
(1.0,1.0), (1.0,1.0)]
```
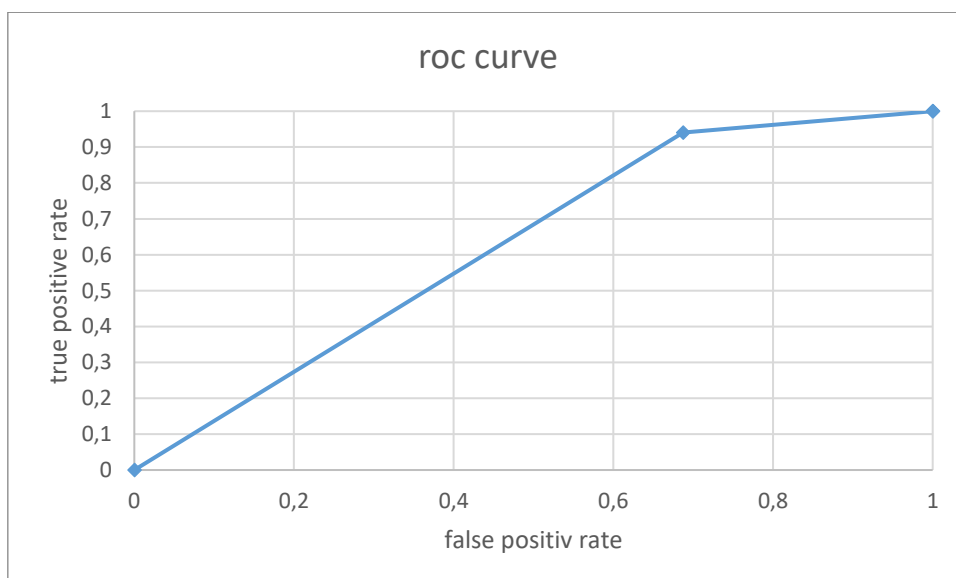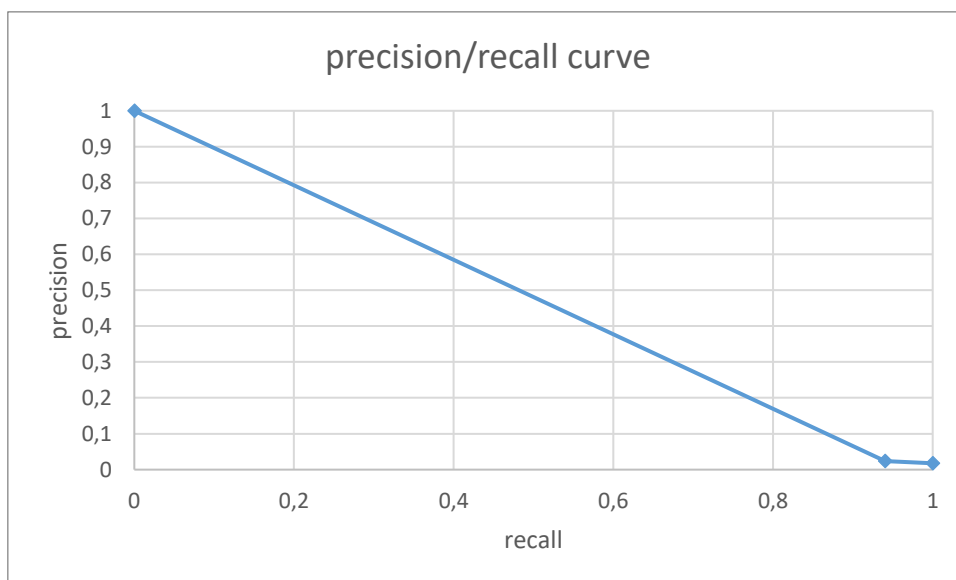
precision/recall curve

roc curve

LogisticRegressionTrain1_2AndTest1_3_Sample100_500_1000_iteration100

```
anzahl links true (training): 50
anzahl links false (training): 50
--> gesamtanzahl der eingelesen trainingsdaten: 100
--> gesamtanzahl der eingelesen testdaten: 68551
model coefficients: [25.118294765798346,-
6.234802699944525,11.073323323731751]
model threshold: 0.5
coefficient titel: 25.118294765798346
```
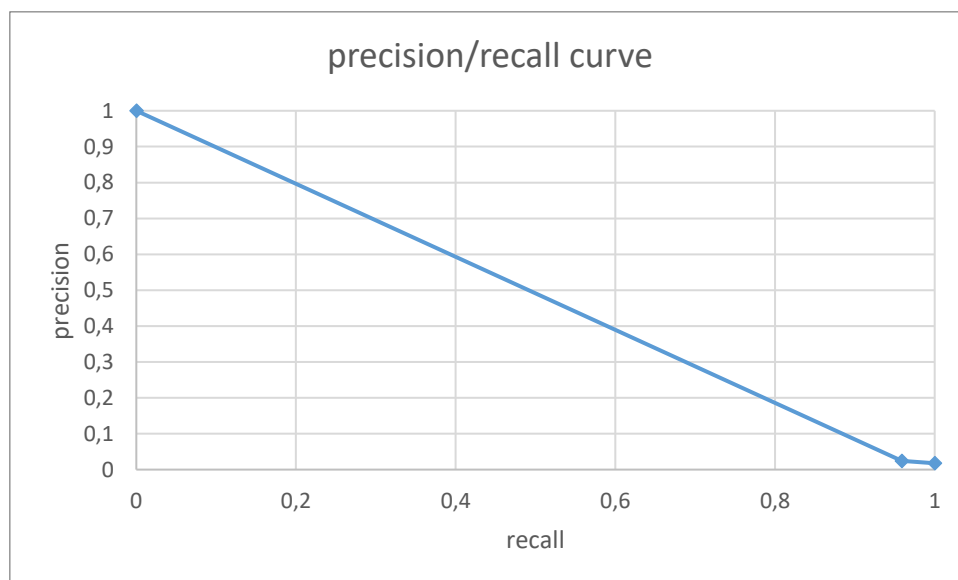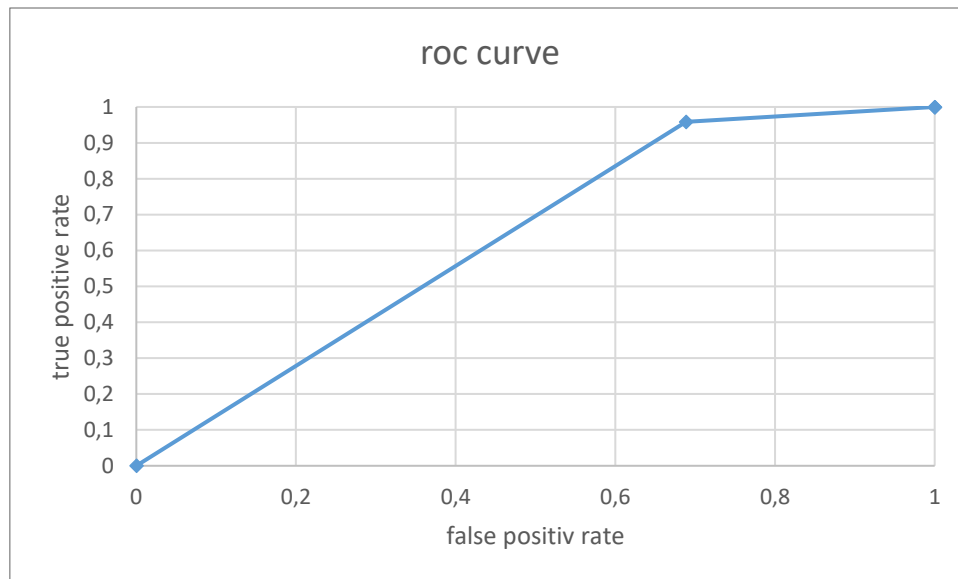
```
coefficient artist: -6.234802699944525
coefficient album: 11.073323323731751
confusion matrix:
21064.0   46276.0
72.0      1139.0
true positive: 1139.0
true negative: 21064.0
false positive: 72.0
false negative: 46276.0
accuracy: 0.32389024230135227
precision: 0.32389024230135227
f measure: 0.32389024230135227
weighted f measure: 0.46856777151595885
weighted precision: 0.9794123523198089
recall: 0.32389024230135227
weighted recall: 0.32389024230135227
weighted true positive rate: 0.32389024230135227
weighted false positive rate: 0.07054452537182032
precision-recall-curve: [(0.0,1.0),
(0.9405450041288191,0.024021933987134873), (1.0,0.01766567956703768)]
roc-curve: [(0.0,0.0), (0.6871992871992872,0.9405450041288191), (1.0,1.0),
(1.0,1.0)]
```


precision/recall curve

roc curve

LogisticRegressionTrain1_2AndTest1_3_Sample100_500_1000_iteration500

```
anzahl links true (training): 250
anzahl links false (training): 250
--> gesamtanzahl der eingelesen trainingsdaten: 500
--> gesamtanzahl der eingelesen testdaten: 68551
model coefficients: [28.528647263152706,-
5.283015286470504,13.56083646267513]
model threshold: 0.5
coefficient titel: 28.528647263152706
coefficient artist: -5.283015286470504
coefficient album: 13.56083646267513
confusion matrix:
20983.0  46357.0
50.0      1161.0
true positive: 1161.0
true negative: 20983.0
false positive: 50.0
false negative: 46357.0
accuracy: 0.32302956922583187
precision: 0.32302956922583187
f measure: 0.32302956922583187
weighted f measure: 0.4673263759531606
weighted precision: 0.9804307216649718
recall: 0.32302956922583187
weighted recall: 0.3230295692258319
weighted true positive rate: 0.3230295692258319
weighted false positive rate: 0.05271989920517919
precision-recall-curve: [(0.0,1.0),
(0.9587118084227911,0.02443284650027358), (1.0,0.01766567956703768)]
roc-curve: [(0.0,0.0), (0.6884021384021384,0.9587118084227911), (1.0,1.0),
(1.0,1.0)]
```
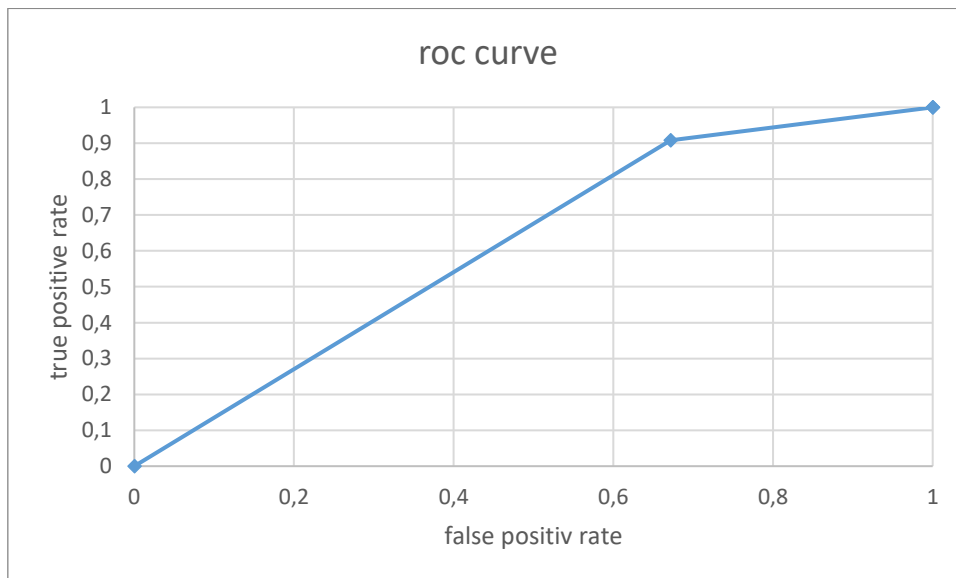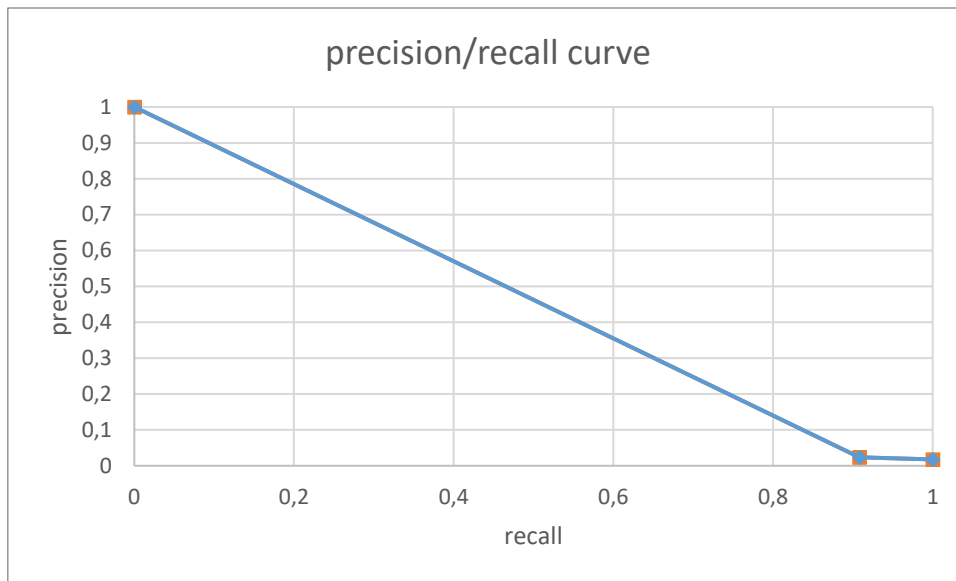

precision/recall curve

LogisticRegressionTrain1_2AndTest1_3_Sample100_500_1000_iteration1000

```
anzahl links true (training): 500
anzahl links false (training): 500
--> gesamtanzahl der eingelesen trainingsdaten: 1000
--> gesamtanzahl der eingelesen testdaten: 68551
model coefficients: [18.864276572439138,-
4.173636281209467,11.765089900360733]
model threshold: 0.5
coefficient titel: 18.864276572439138
coefficient artist: -4.173636281209467
coefficient album: 11.765089900360733
confusion matrix:
22105.0   45235.0
111.0     1100.0
true positive: 1100.0
true negative: 22105.0
false positive: 111.0
false negative: 45235.0
accuracy: 0.3385070969059532
precision: 0.3385070969059532
f measure: 0.3385070969059532
weighted f measure: 0.48575421104209204
weighted precision: 0.9778455721646886
recall: 0.3385070969059532
weighted recall: 0.33850709690595326
weighted true positive rate: 0.33850709690595326
weighted false positive rate: 0.10190730394777878
precision-recall-curve: [(0.0,1.0),
(0.9083402146985962,0.023740153231898133), (1.0,0.01766567956703768)]
roc-curve: [(0.0,0.0), (0.6717404217404217,0.9083402146985962), (1.0,1.0),
(1.0,1.0)]
```
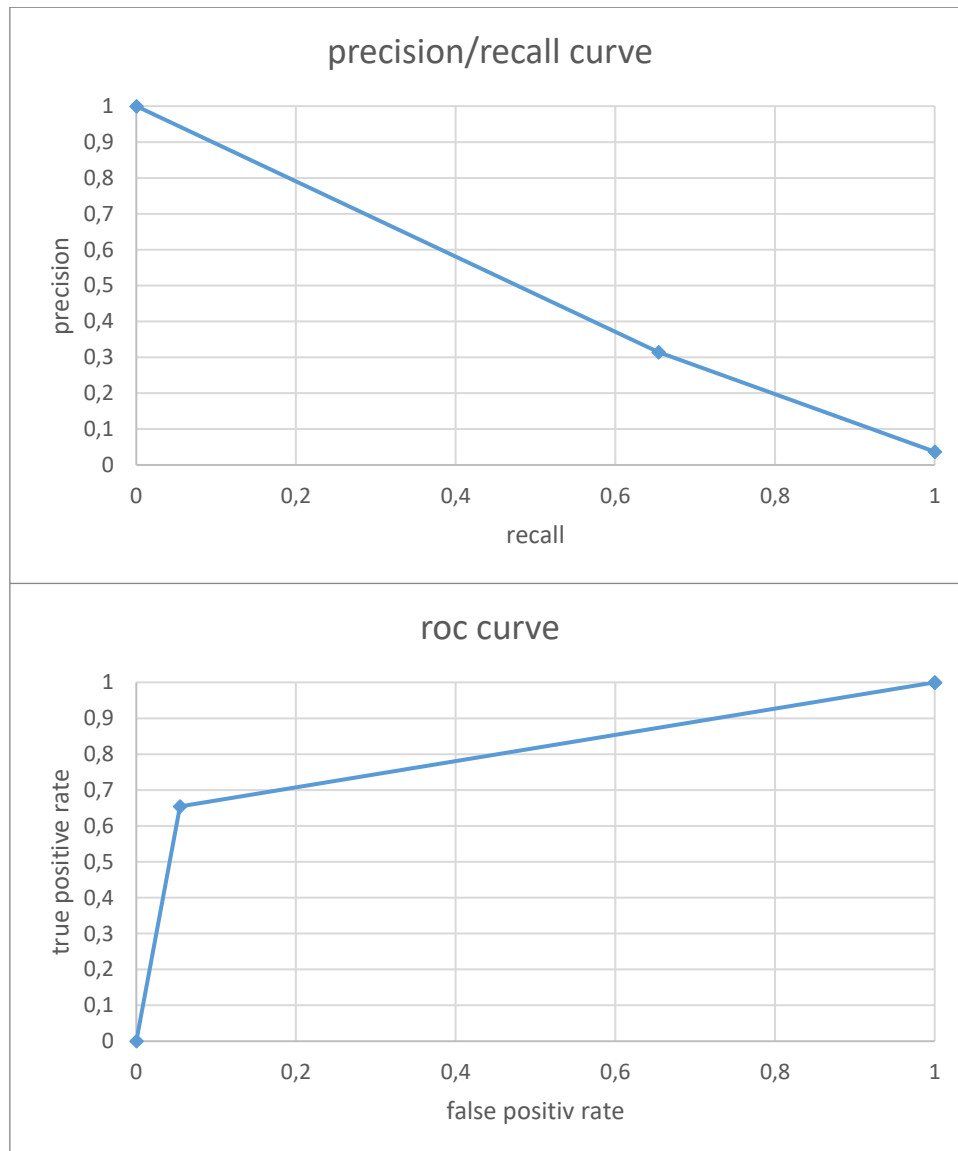
precision/recall curve

roc curve

LogisticRegressionTrain1_2AndTest1_4_Sample100_500_1000_iteration100

```
anzahl links true (training): 50
anzahl links false (training): 50
--> gesamtanzahl der eingelesen trainingsdaten: 100
--> gesamtanzahl der eingelesen testdaten: 32274
model coefficients: [25.118294765798304,-
6.234802699944521,11.073323323731735]
model threshold: 0.5
coefficient titel: 25.118294765798304
coefficient artist: -6.234802699944521
coefficient album: 11.073323323731735
confusion matrix:
29378.0  1702.0
413.0    781.0
true positive: 781.0
true negative: 29378.0
false positive: 413.0
false negative: 1702.0
accuracy: 0.9344673731176799
precision: 0.9344673731176799
f measure: 0.9344673731176799
weighted f measure: 0.9452599935469562
```

```
weighted precision: 0.96129050251567
recall: 0.9344673731176799
weighted recall: 0.9344673731176798
weighted true positive rate: 0.9344673731176798
weighted false positive rate: 0.3351254252832697
precision-recall-curve: [(0.0,1.0),
(0.6541038525963149,0.31453886427708416), (1.0,0.036995724112288526)]
roc-curve: [(0.0,0.0), (0.05476190476190476,0.6541038525963149), (1.0,1.0),
(1.0,1.0)]
```
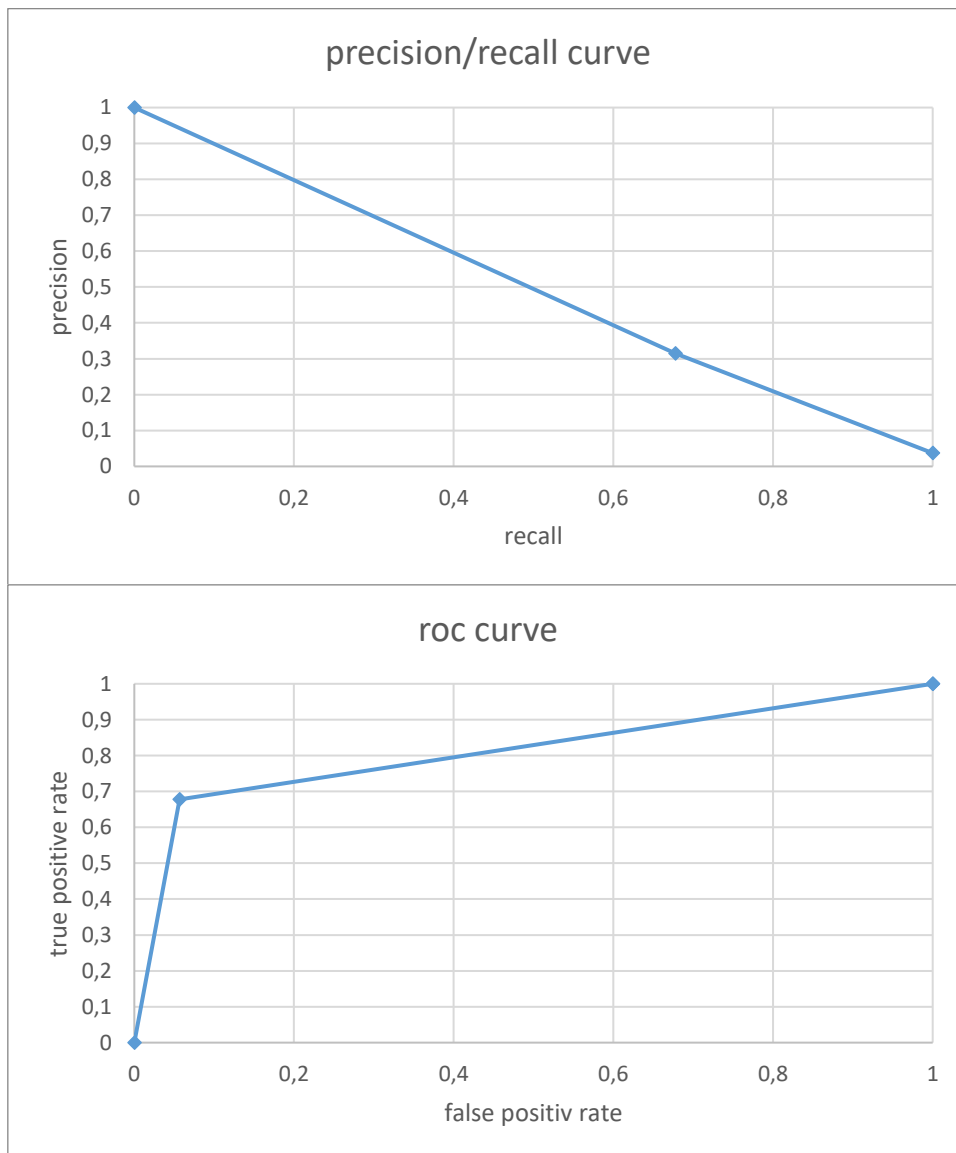

LogisticRegressionTrain1_2AndTest1_4_Sample100_500_1000_iteration500

```
anzahl links true (training): 250
anzahl links false (training): 250
--> gesamtanzahl der eingelesen trainingsdaten: 500
--> gesamtanzahl der eingelesen testdaten: 32274
model coefficients: [28.528647263152706,-
5.283015286470501,13.560836462675116]
model threshold: 0.5
coefficient titel: 28.528647263152706
coefficient artist: -5.283015286470501
coefficient album: 13.560836462675116
```
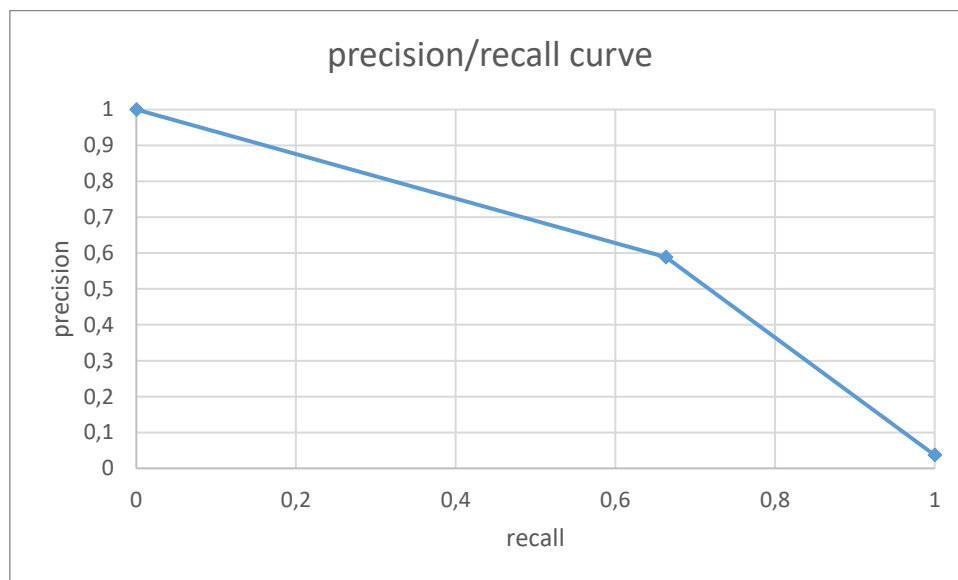
```
confusion matrix:
29317.0  1763.0
385.0    809.0
true positive: 809.0
true negative: 29317.0
false positive: 385.0
false negative: 1763.0
accuracy: 0.9334448782301543
precision: 0.9334448782301543
f measure: 0.9334448782301543
weighted f measure: 0.9448668780286101
weighted precision: 0.9621584076422911
recall: 0.9334448782301543
weighted recall: 0.9334448782301543
weighted true positive rate: 0.9334448782301543
weighted false positive rate: 0.31261502109376454
precision-recall-curve: [(0.0,1.0),
(0.6775544388609716,0.3145412130637636), (1.0,0.036995724112288526)]
roc-curve: [(0.0,0.0), (0.056724581724581724,0.6775544388609716),
(1.0,1.0), (1.0,1.0)]
```
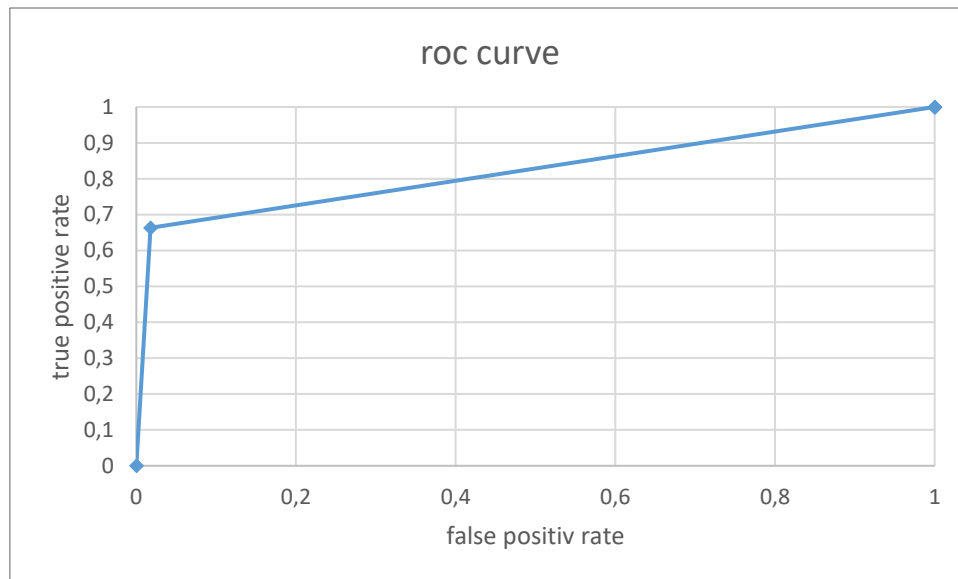

LogisticRegressionTrain1_2AndTest1_4_Sample100_500_1000_iteration1000

```
anzahl links true (training): 500
anzahl links false (training): 500
--> gesamtanzahl der eingelesen trainingsdaten: 1000
--> gesamtanzahl der eingelesen testdaten: 32274
model coefficients: [18.864276572439174,-
4.1736362812094345,11.76508990036077]
model threshold: 0.5
coefficient titel: 18.864276572439174
coefficient artist: -4.1736362812094345
coefficient album: 11.76508990036077
confusion matrix:
30526.0  554.0
402.0    792.0
true positive: 792.0
true negative: 30526.0
false positive: 402.0
false negative: 554.0
accuracy: 0.9703786329553201
precision: 0.9703786329553201
f measure: 0.9703786329553201
weighted f measure: 0.9712286361972906
weighted precision: 0.9722558709402982
recall: 0.9703786329553201
weighted recall: 0.97037863295532
weighted true positive rate: 0.97037863295532
weighted false positive rate: 0.324887017865715
precision-recall-curve: [(0.0,1.0),
(0.6633165829145728,0.5884101040118871), (1.0,0.036995724112288526)]
roc-curve: [(0.0,0.0), (0.017824967824967824,0.6633165829145728),
(1.0,1.0), (1.0,1.0)]
```
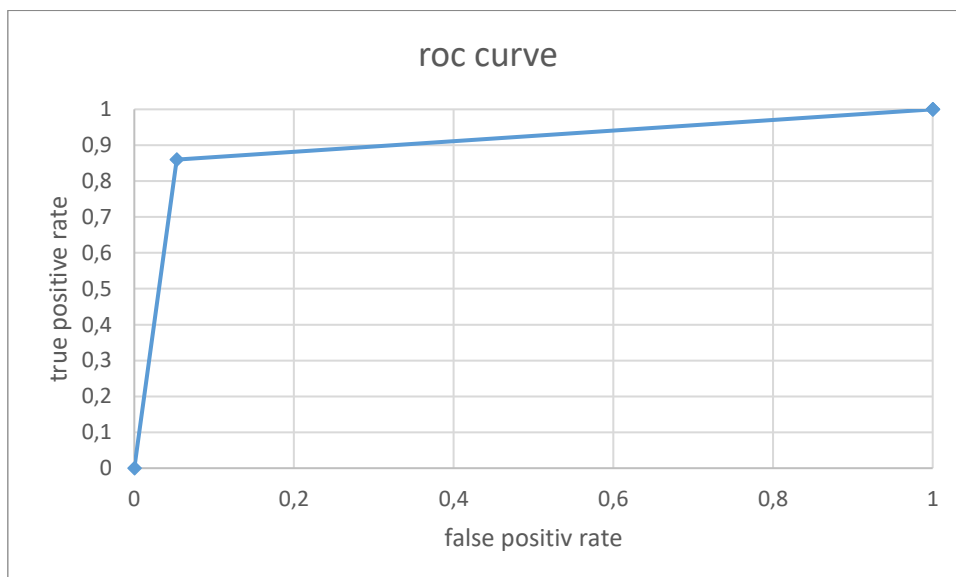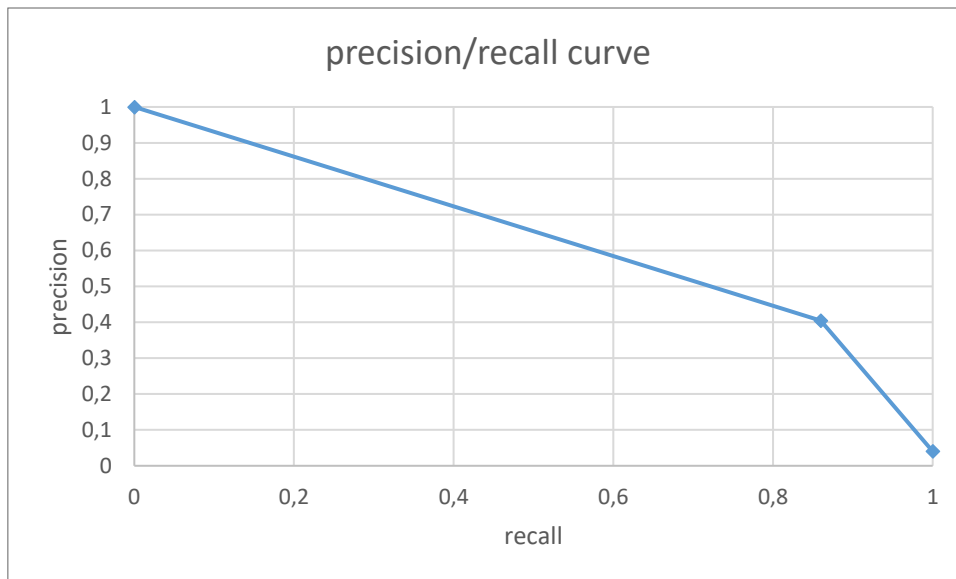

precision/recall curve

LogisticRegressionTrain1_2AndTest1_5_Sample100_500_1000_iteration100

```
anzahl links true (training): 50
anzahl links false (training): 50
--> gesamtanzahl der eingelesen trainingsdaten: 100
--> gesamtanzahl der eingelesen testdaten: 48656
model coefficients: [25.118294765798318,-
6.234802699944524,11.073323323731735]
model threshold: 0.5
coefficient titel: 25.118294765798318
coefficient artist: -6.234802699944524
coefficient album: 11.073323323731735
confusion matrix:
44219.0  2477.0
275.0    1685.0
true positive: 1685.0
true negative: 44219.0
false positive: 275.0
false negative: 2477.0
accuracy: 0.9434396580072345
precision: 0.9434396580072345
f measure: 0.9434396580072345
weighted f measure: 0.9529287580888075
weighted precision: 0.9700941925780598
recall: 0.9434396580072345
weighted recall: 0.9434396580072345
weighted true positive rate: 0.9434396580072345
weighted false positive rate: 0.13679100916959422
precision-recall-curve: [(0.0,1.0),
(0.8596938775510204,0.4048534358481499), (1.0,0.04028280170996383)]
roc-curve: [(0.0,0.0), (0.05304522871338016,0.8596938775510204), (1.0,1.0),
(1.0,1.0)]
```
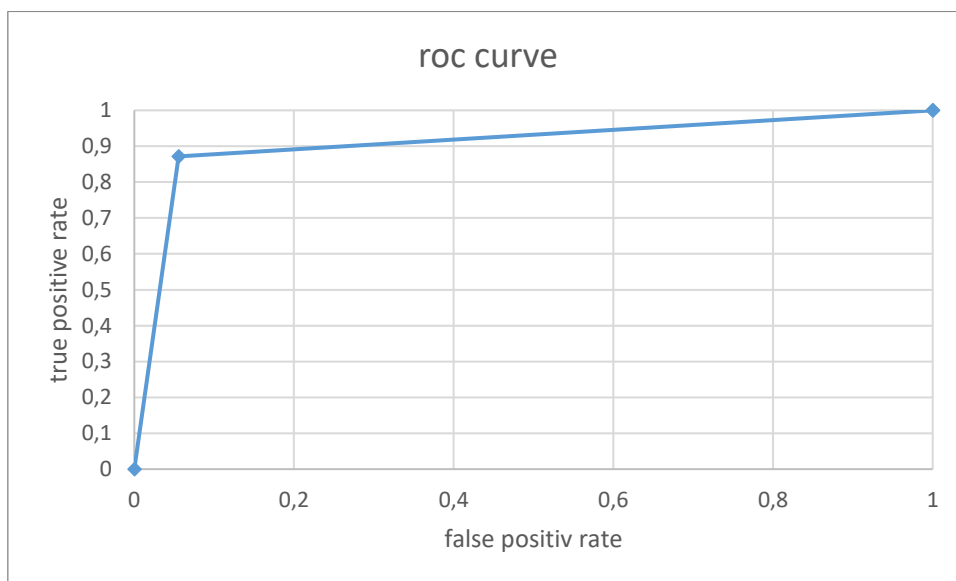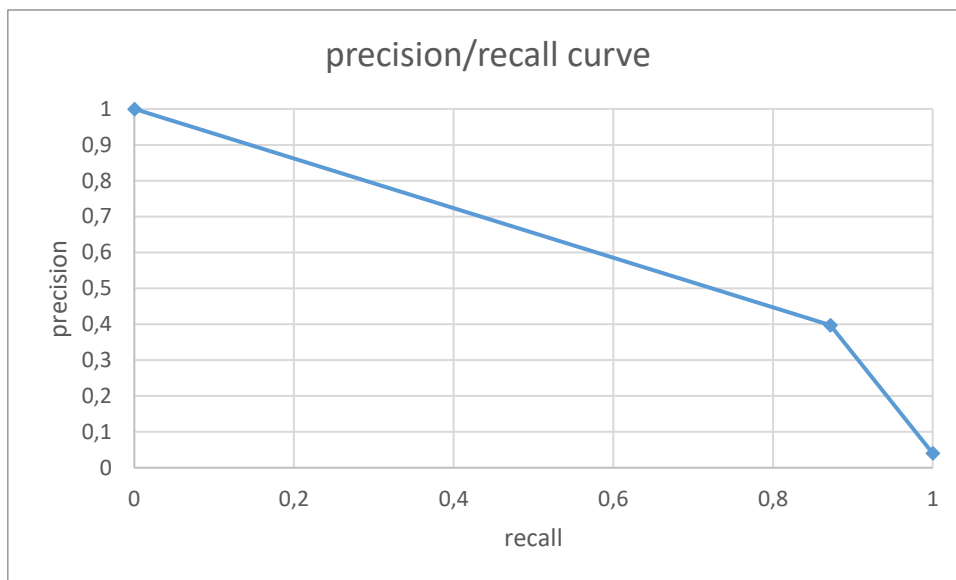
precision/recall curve

roc curve

LogisticRegressionTrain1_2AndTest1_5_Sample100_500_1000_iteration500

```
anzahl links true (training): 250
anzahl links false (training): 250
--> gesamtanzahl der eingelesen trainingsdaten: 500
--> gesamtanzahl der eingelesen testdaten: 48656
model coefficients: [28.528647263152703,-
5.28301528647052,13.560836462675129]
model threshold: 0.5
coefficient titel: 28.528647263152703
coefficient artist: -5.28301528647052
coefficient album: 13.560836462675129
confusion matrix:
44103.0  2593.0
251.0    1709.0
true positive: 1709.0
true negative: 44103.0
false positive: 251.0
false negative: 2593.0
accuracy: 0.9415488326208484
precision: 0.9415488326208484
f measure: 0.9415488326208484
weighted f measure: 0.951727512642968
```

```
weighted precision: 0.9702887717889148
recall: 0.941548326208484
weighted recall: 0.941548326208484
weighted true positive rate: 0.941548326208484
weighted false positive rate: 0.12513943864225302
precision-recall-curve: [(0.0,1.0),
(0.871938775510204,0.39725708972570895), (1.0,0.04028280170996383)]
roc-curve: [(0.0,0.0), (0.055529381531608704,0.871938775510204), (1.0,1.0),
(1.0,1.0)]
```

precision/recall curve

roc curve

LogisticRegressionTrain1_2AndTest1_5_Sample100_500_1000_iteration1000

```
anzahl links true (training): 500
anzahl links false (training): 500
--> gesamtanzahl der eingelesen trainingsdaten: 1000
--> gesamtanzahl der eingelesen testdaten: 48656
model coefficients: [18.86427657243914,-
4.173636281209466,11.765089900360735]
model threshold: 0.5
coefficient titel: 18.86427657243914
coefficient artist: -4.173636281209466
coefficient album: 11.765089900360735
```

```
confusion matrix:
45468.0  1228.0
260.0    1700.0
true positive: 1700.0
true negative: 45468.0
false positive: 260.0
false negative: 1228.0
accuracy: 0.9694179546201908
precision: 0.9694179546201908
f measure: 0.9694179546201908
weighted f measure: 0.9722859776208589
weighted precision: 0.9776486824488482
recall: 0.9694179546201908
weighted recall: 0.9694179546201908
weighted true positive rate: 0.9694179546201908
weighted false positive rate: 0.12836877154109994
precision-recall-curve: [(0.0,1.0),
(0.8673469387755102,0.5806010928961749), (1.0,0.04028280170996383)]
roc-curve: [(0.0,0.0), (0.026297755696419393,0.8673469387755102),
(1.0,1.0), (1.0,1.0)]
```