



BREAST CANCER DIAGNOSIS



Guilherme Casal 102587
João Rodrigues 102487

INDEX

01

Model Evaluation

02

Data explanation

03

Previous work

04

Logistic Regression

05

Neural Network

06

Results Comparison

07

Conclusion

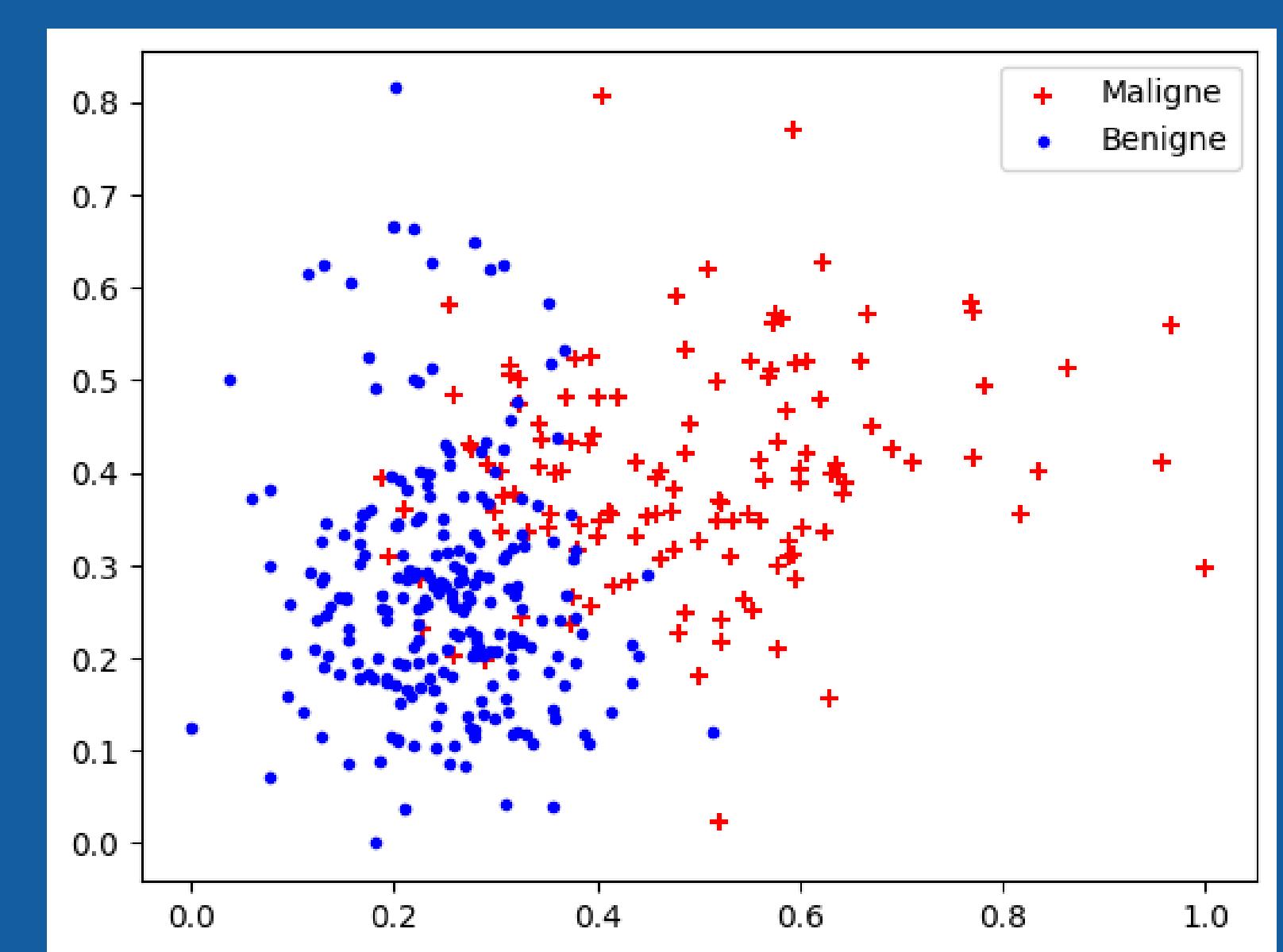
2

MODEL EVALUATION

- K-Fold Cross Validation
- Confusion Matrix
- Roc Curve and AUC

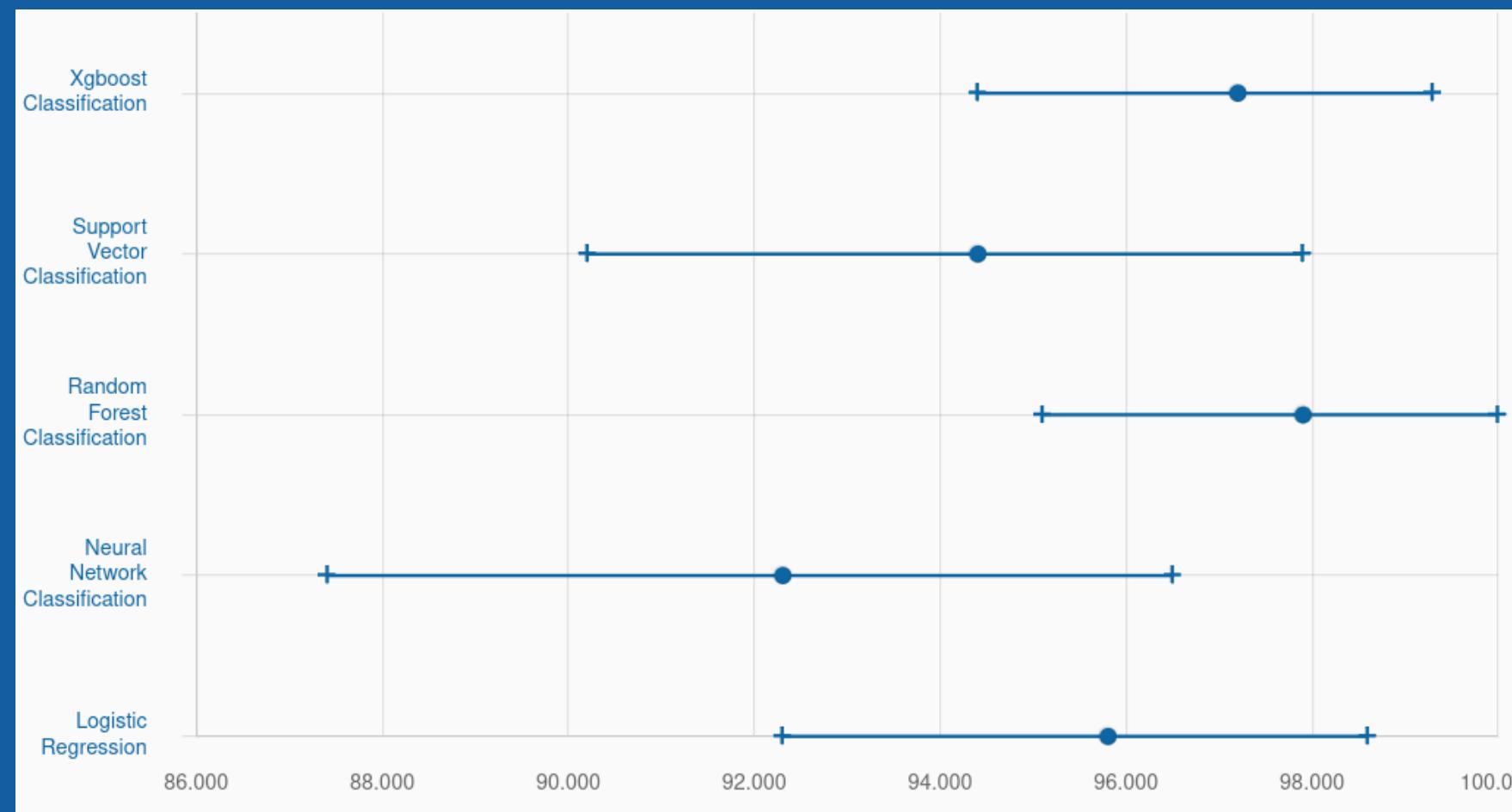
DATA EXPLANATION

- 32 features
 - Tumor size, shape and texture
 - Compactness, smoothness and symmetry
 - Benigne or Maligne Diagnosis



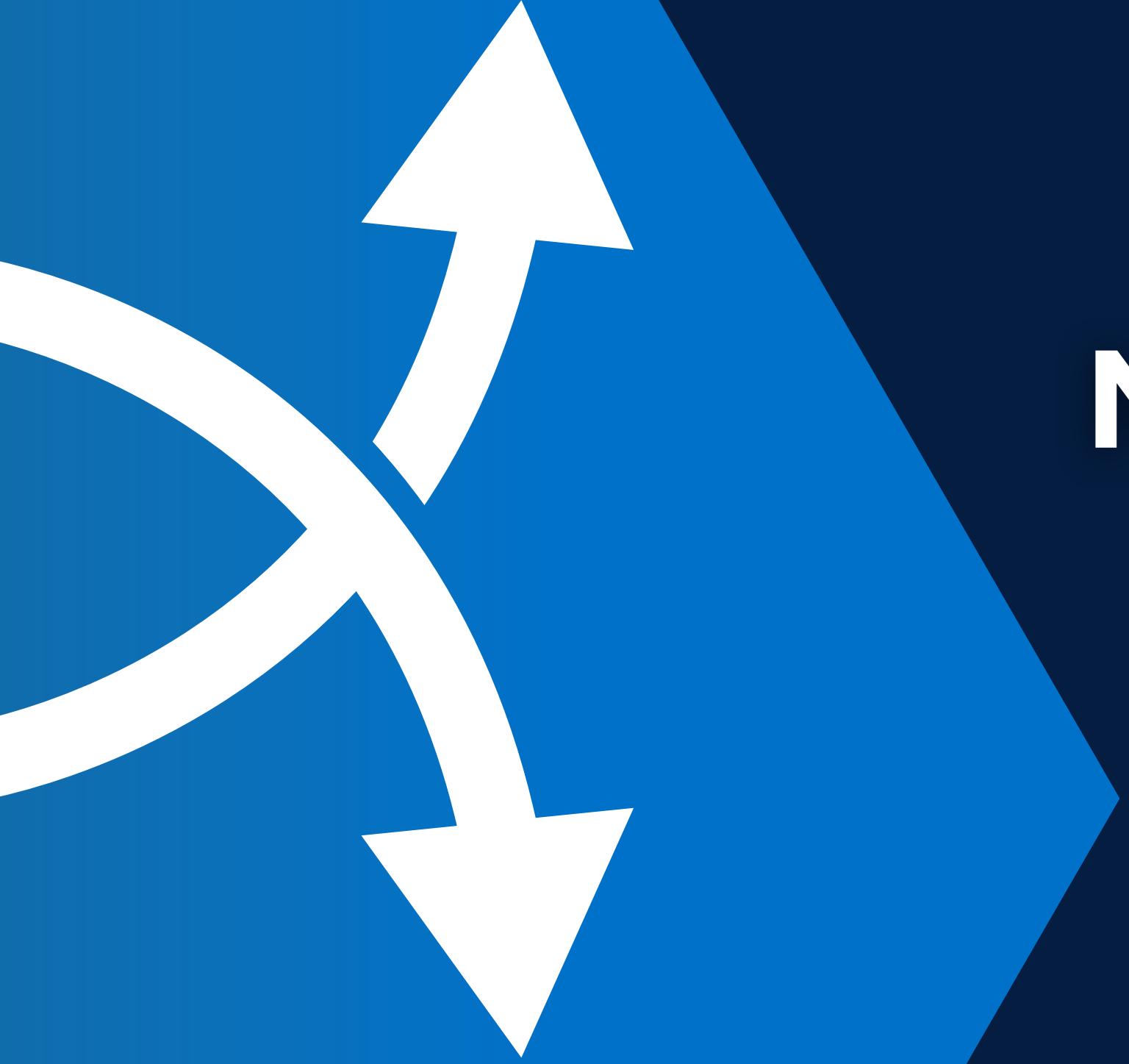
PREVIOUS WORK

UCI WEBSITE



ARTICLES

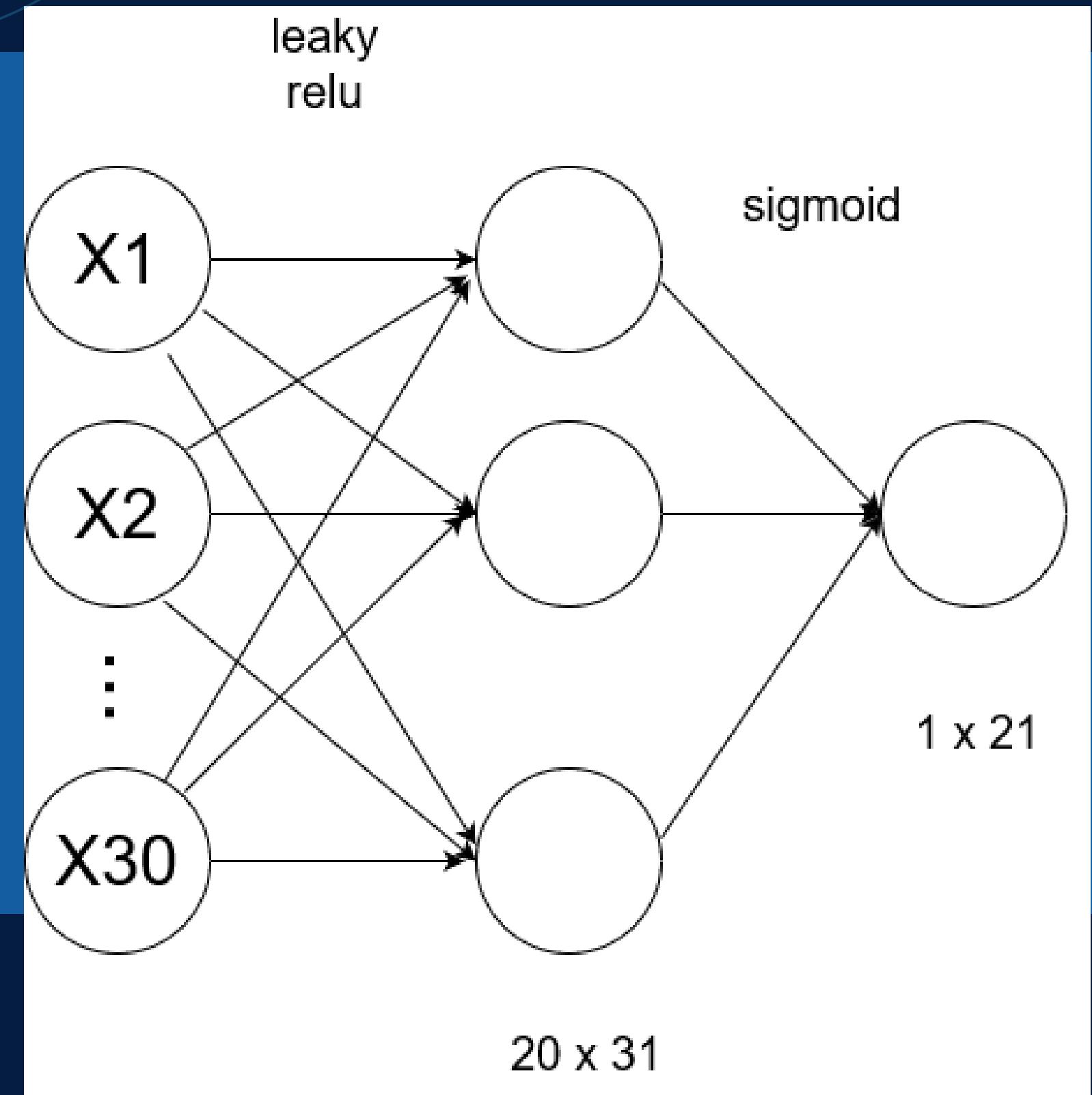
Classifier Type	Performance	
	F-measure	Accuracy
VGG16	98.2	98.2
Random forest	99%	98.9%
AdaBoost	75%	83%
Logistic regression	97%	97.7%
Decision trees	75%	83%
Naïve Bayes	76%	74%



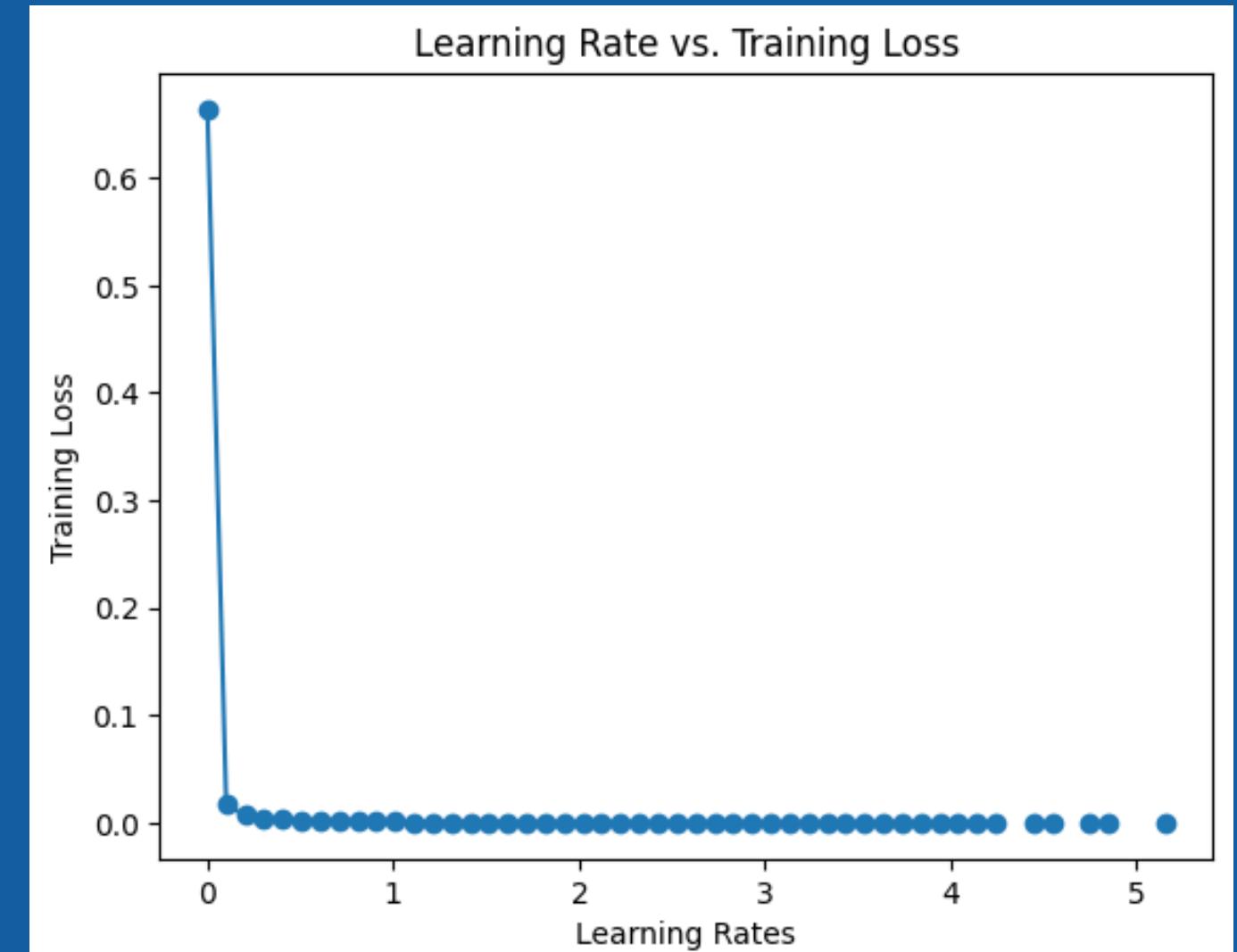
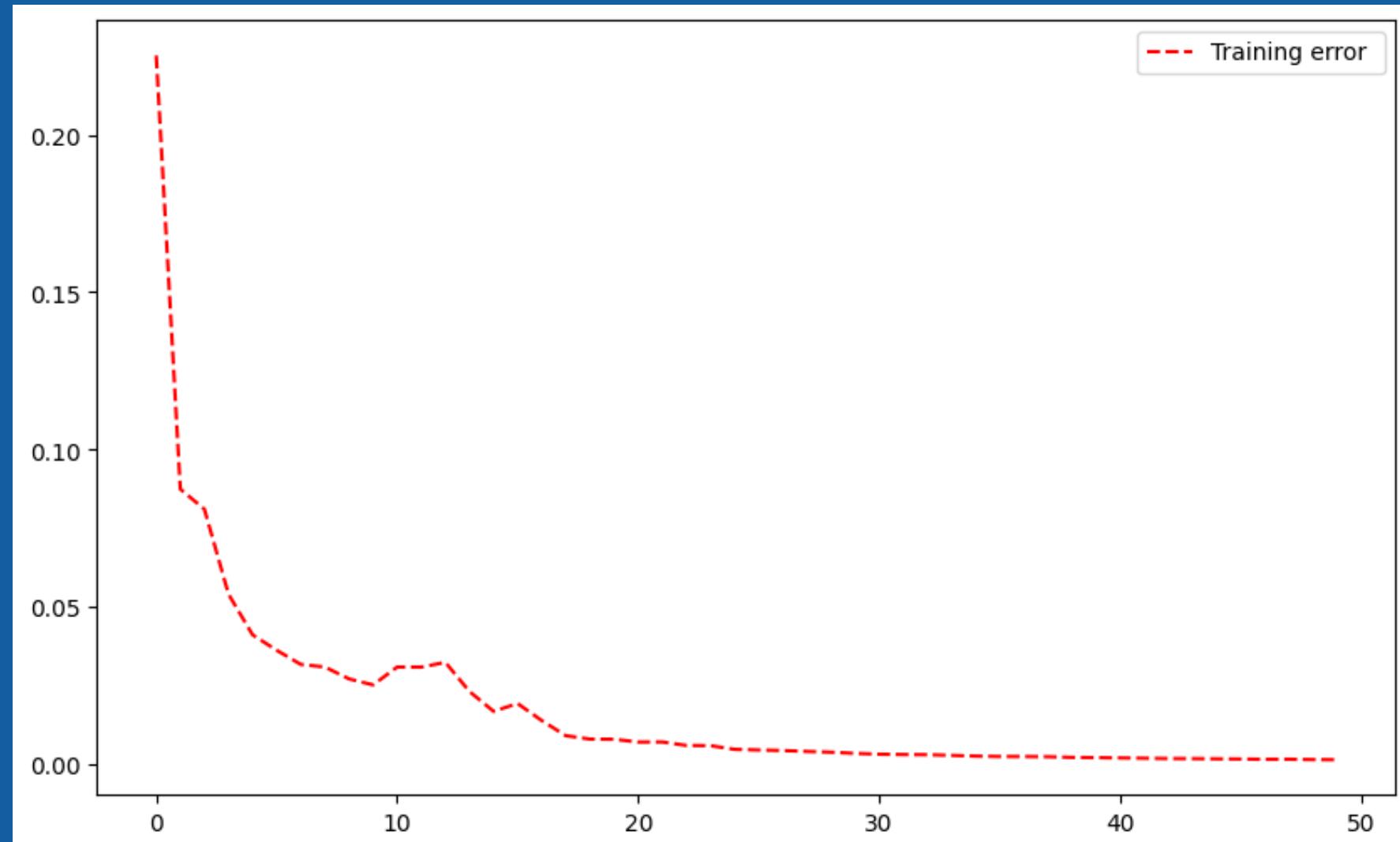
NEURAL NETWORK

Initial Configuration

$$\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)]$$

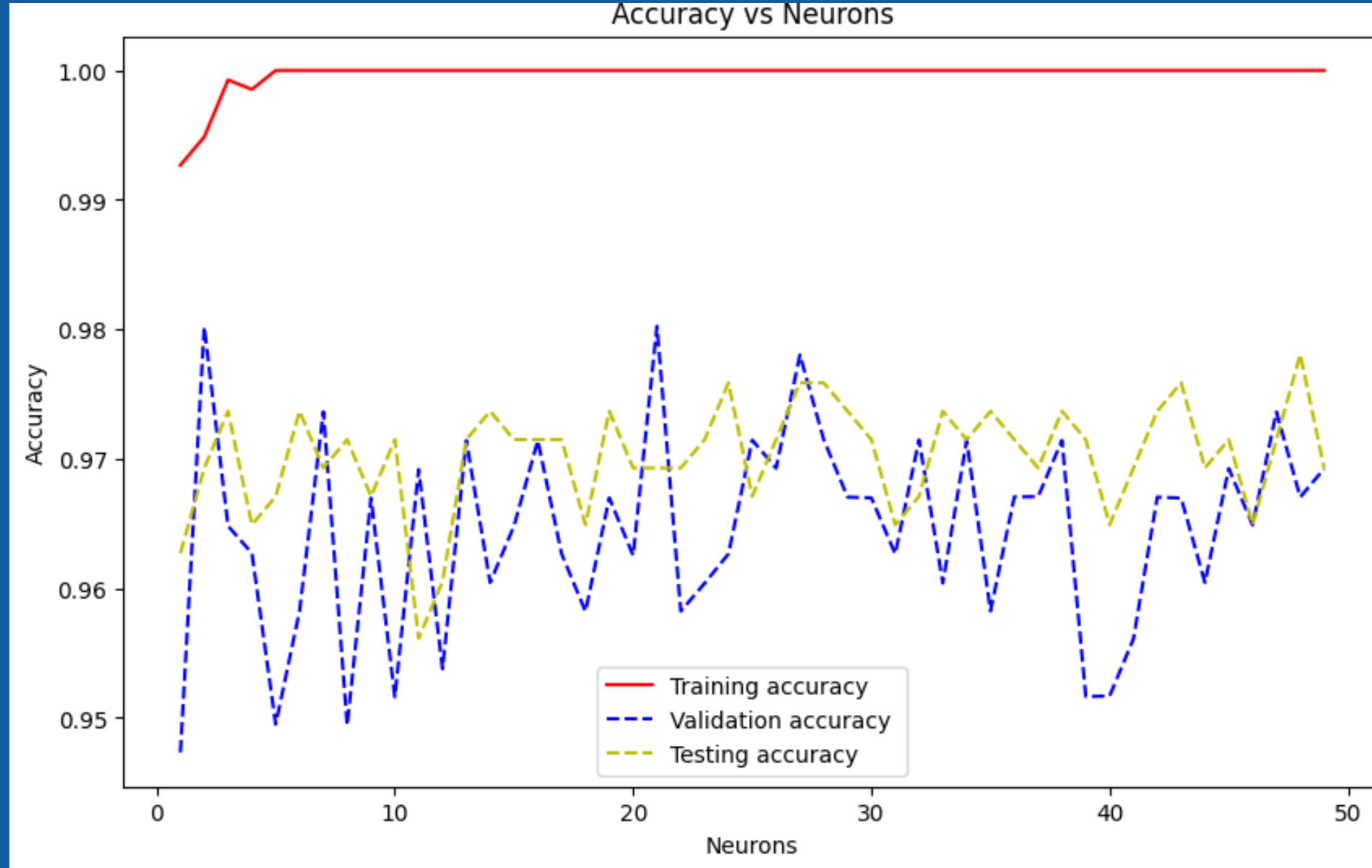


Learning Rate



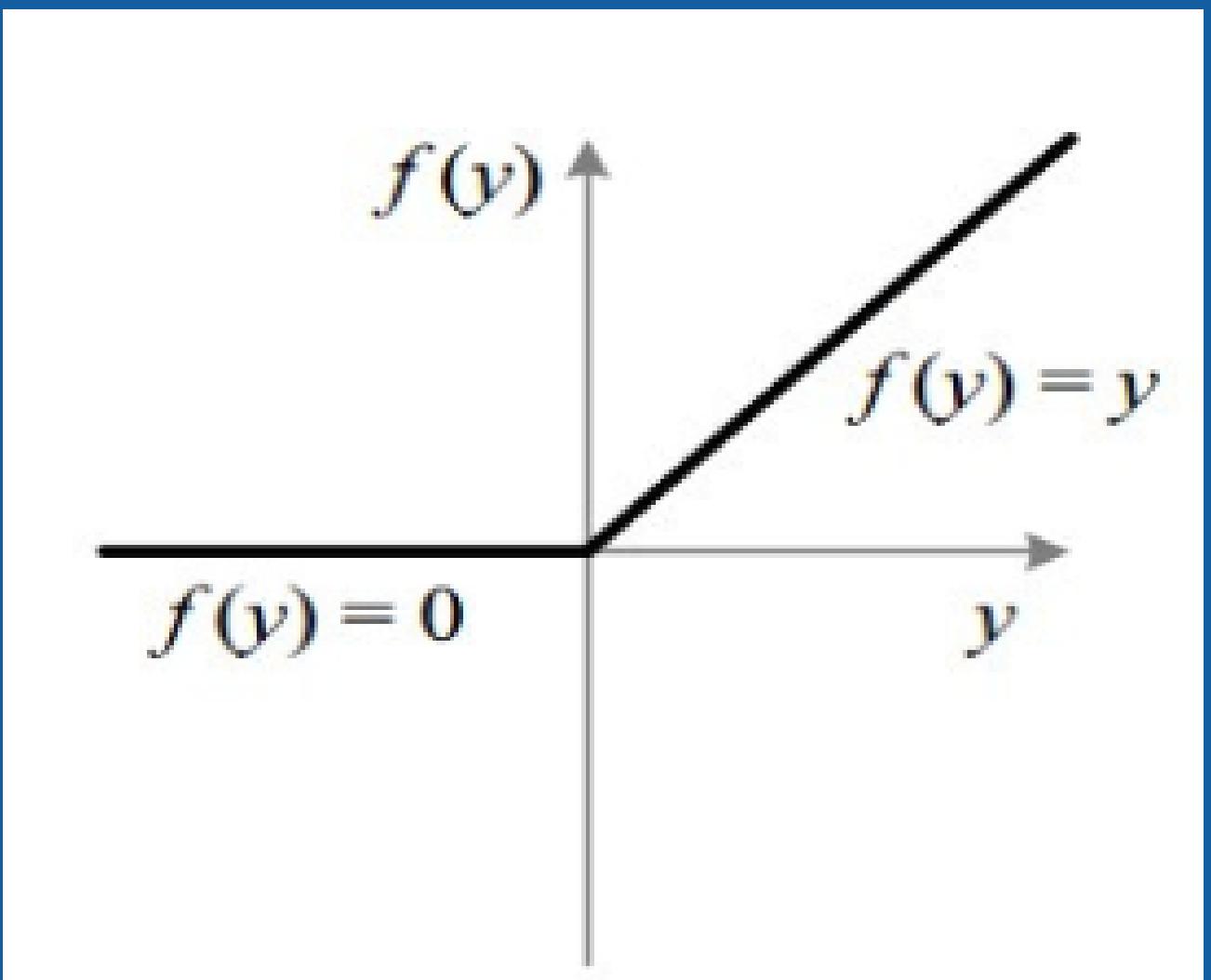
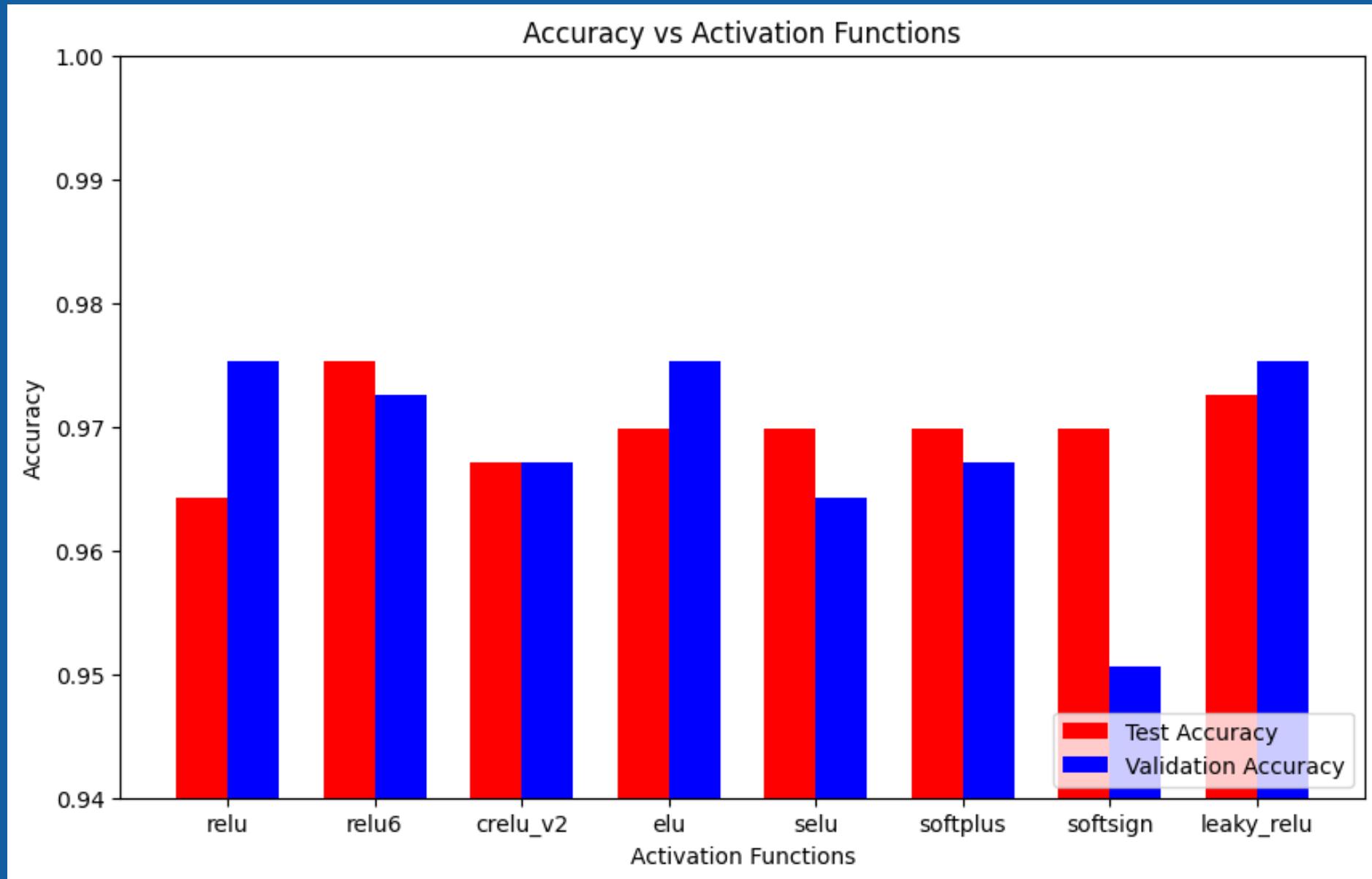
acc	sen	spe
0.973684	0.941860	0.992958
1.000000	1.000000	1.000000
0.958236	0.956359	0.963281

N° of Neurons

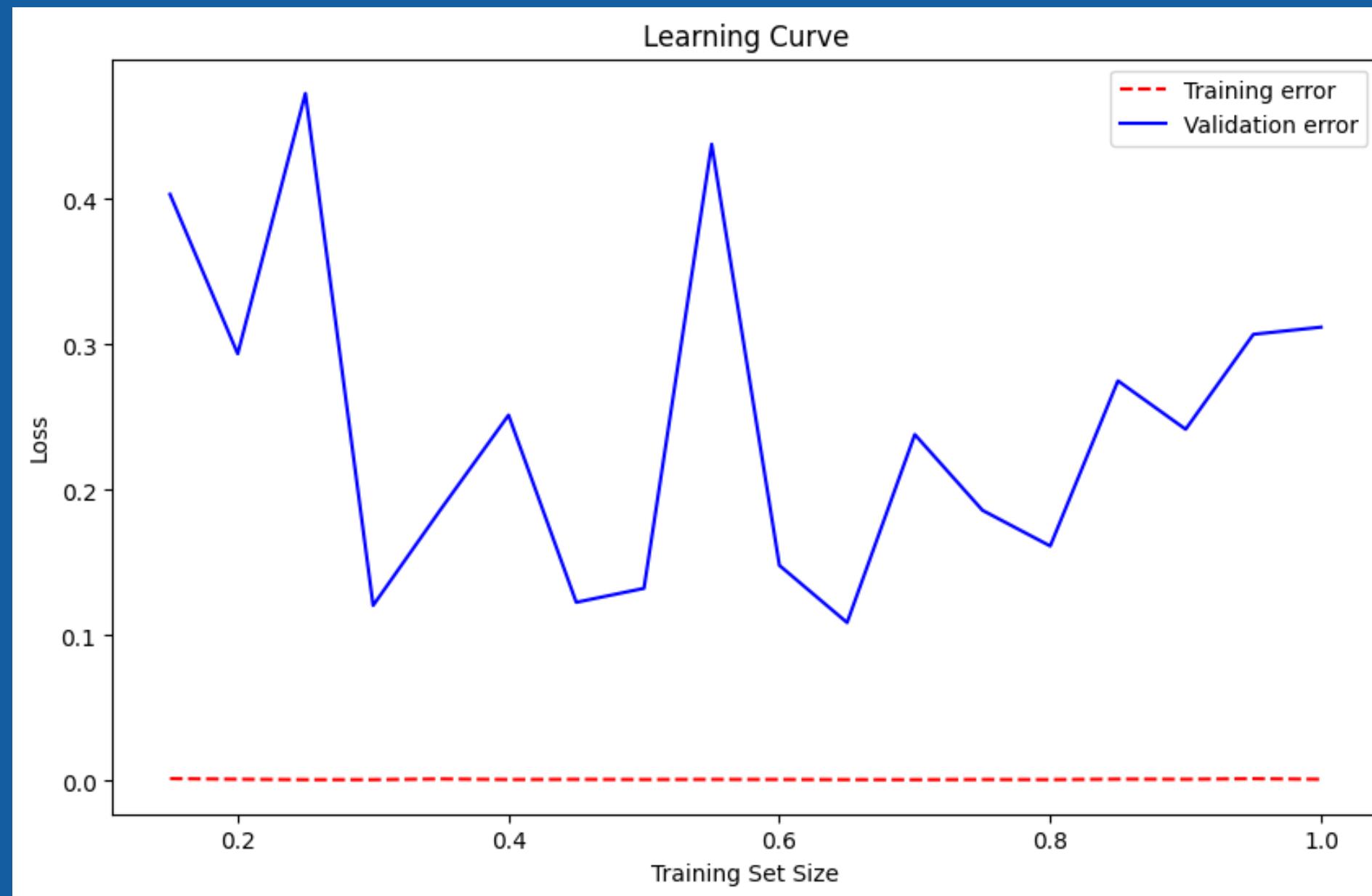


	acc	sen	spe
test	0.975877	0.959302	0.985915
train	1.000000	1.000000	1.000000
val	0.964835	0.954000	0.973499

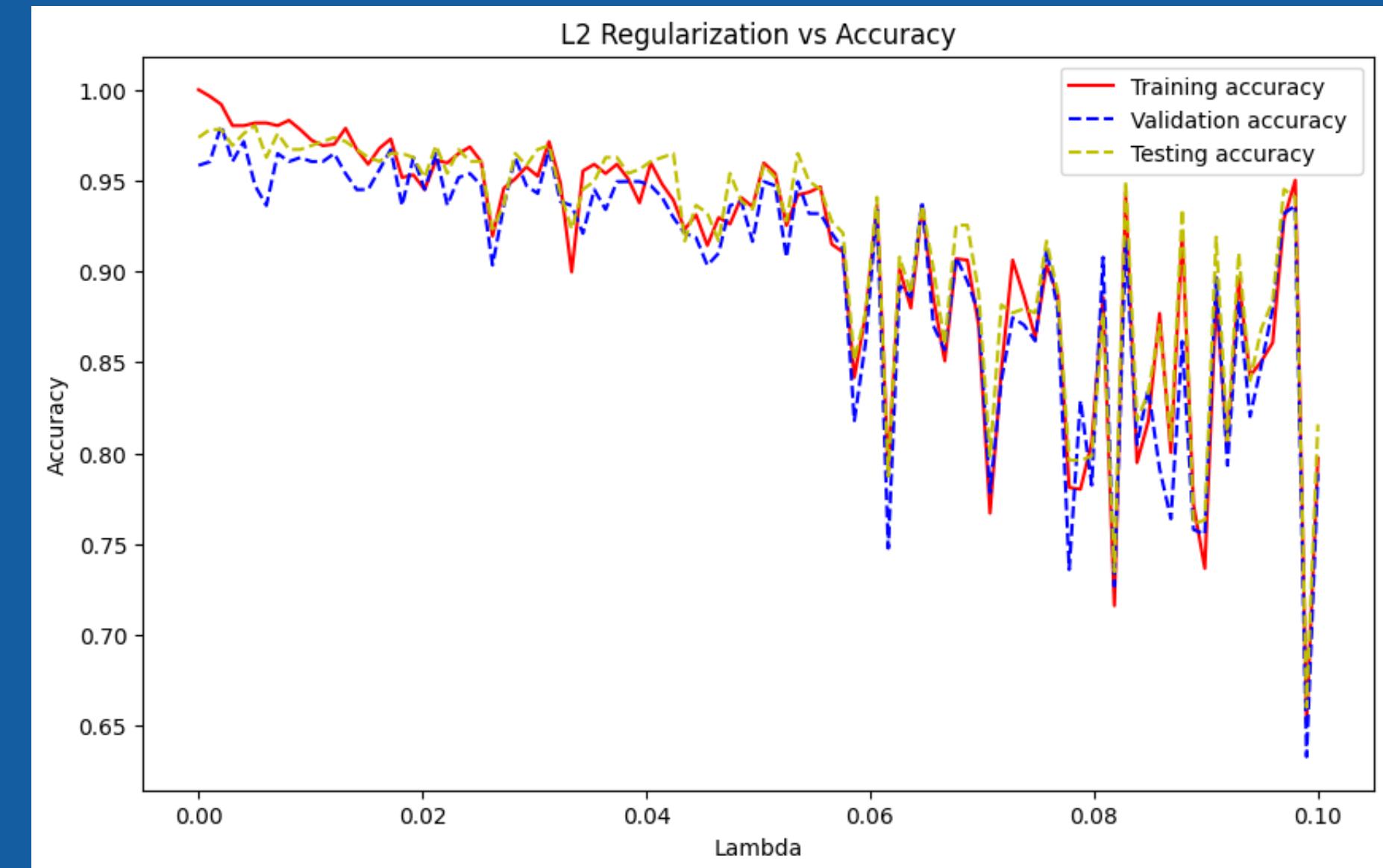
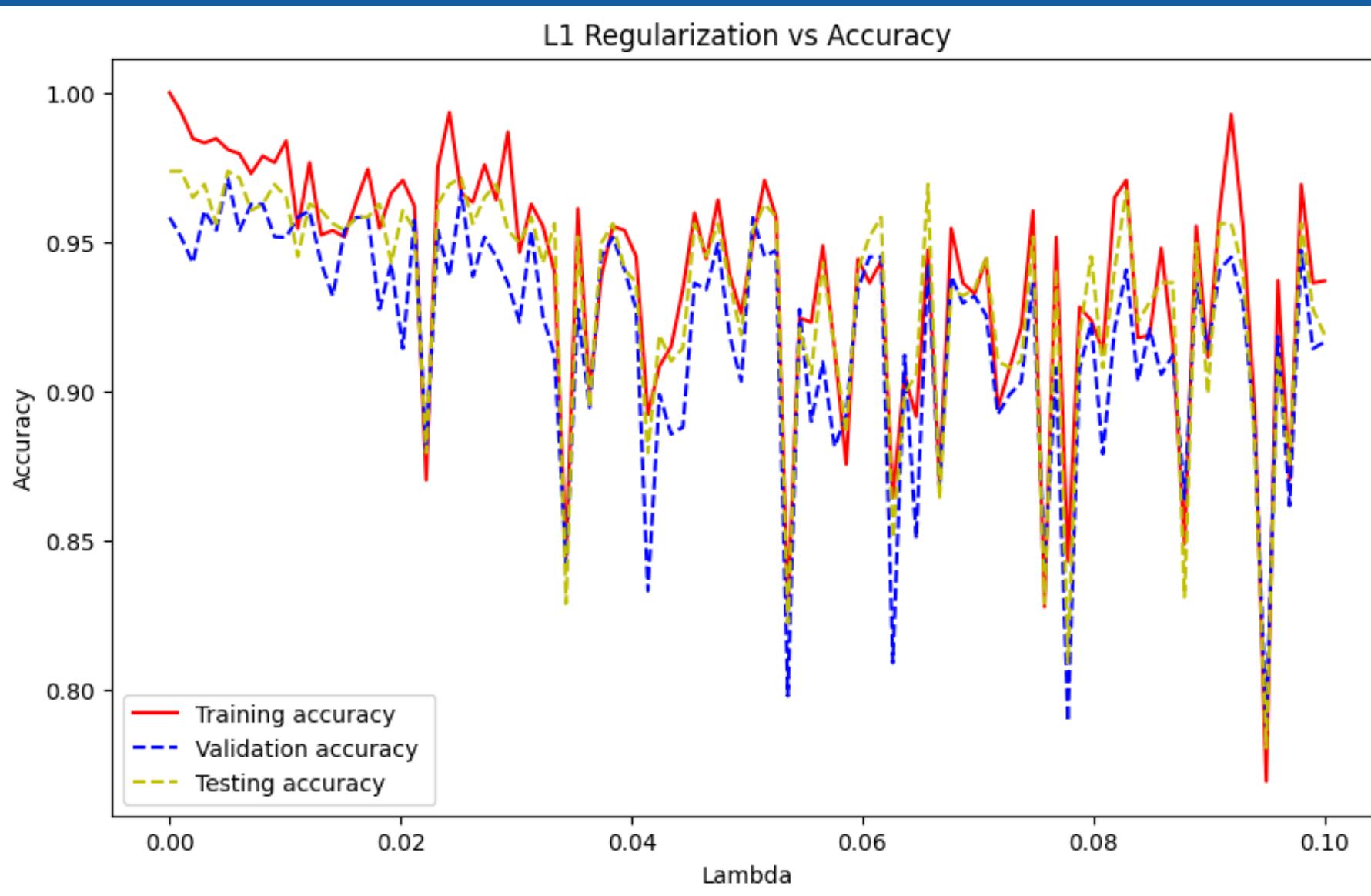
Activation function



Regularization



Regularization



Regularization

	acc	sen	spe
test	0.962719	0.930233	0.982394
train	0.978017	0.950581	0.994213
val	0.960429	0.937104	0.975846

Table 4. L1 Results

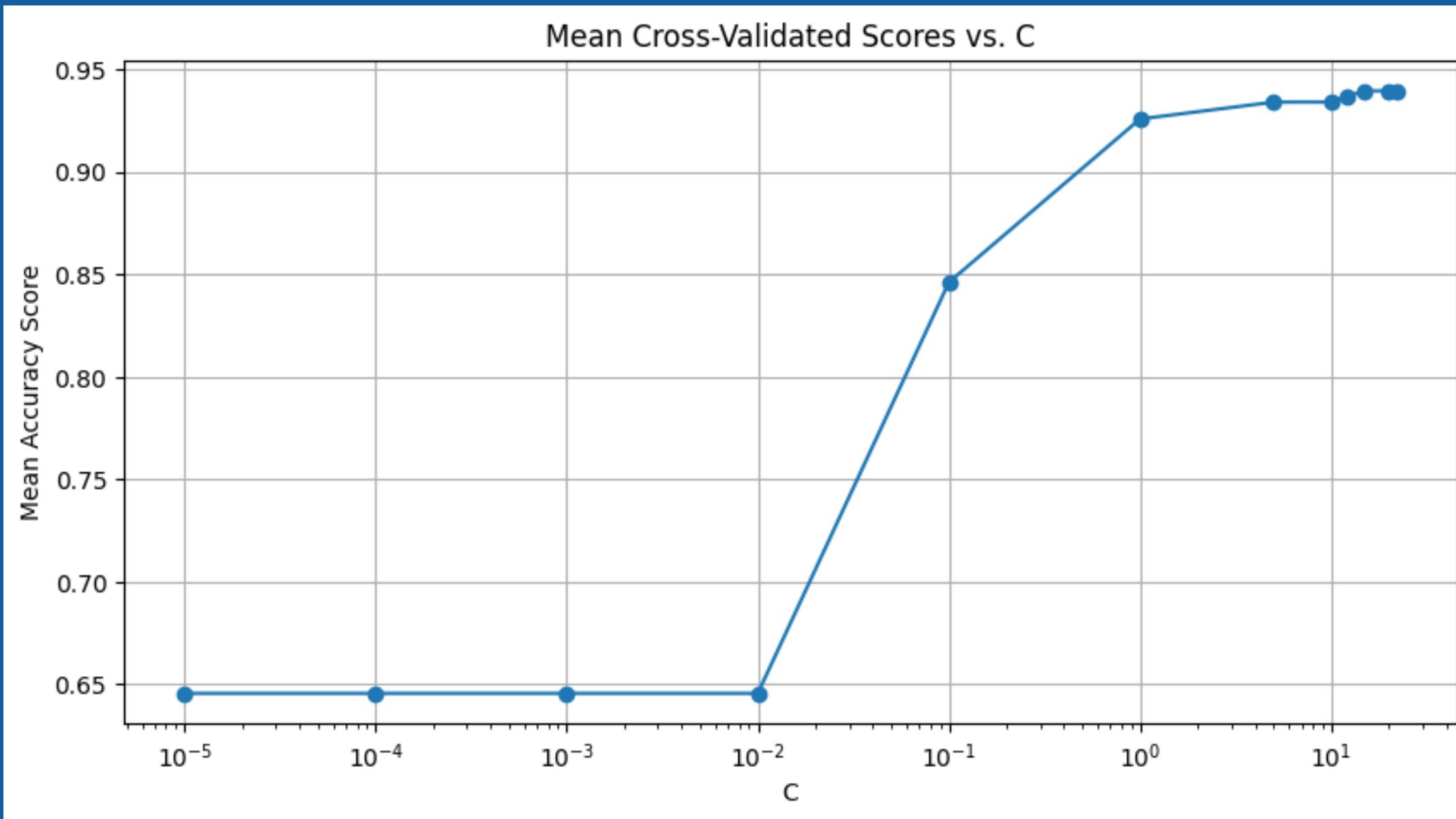
	acc	sen	spe
test	0.980263	0.953488	0.996479
train	0.992673	0.980307	1.000000
val	0.973645	0.942593	0.992802

Table 3. L2 Results



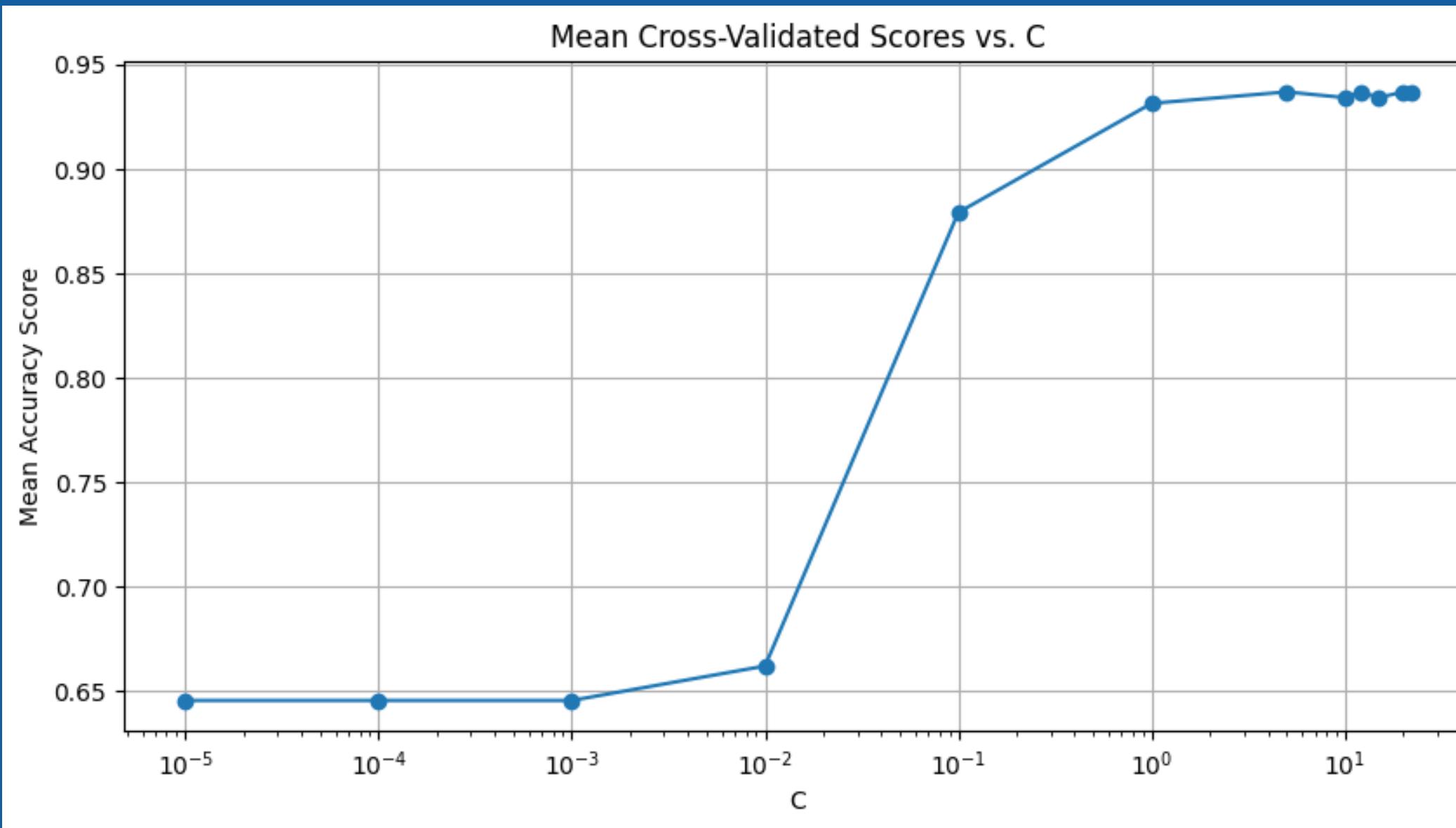
LOGISTIC REGRESSION

First Approach



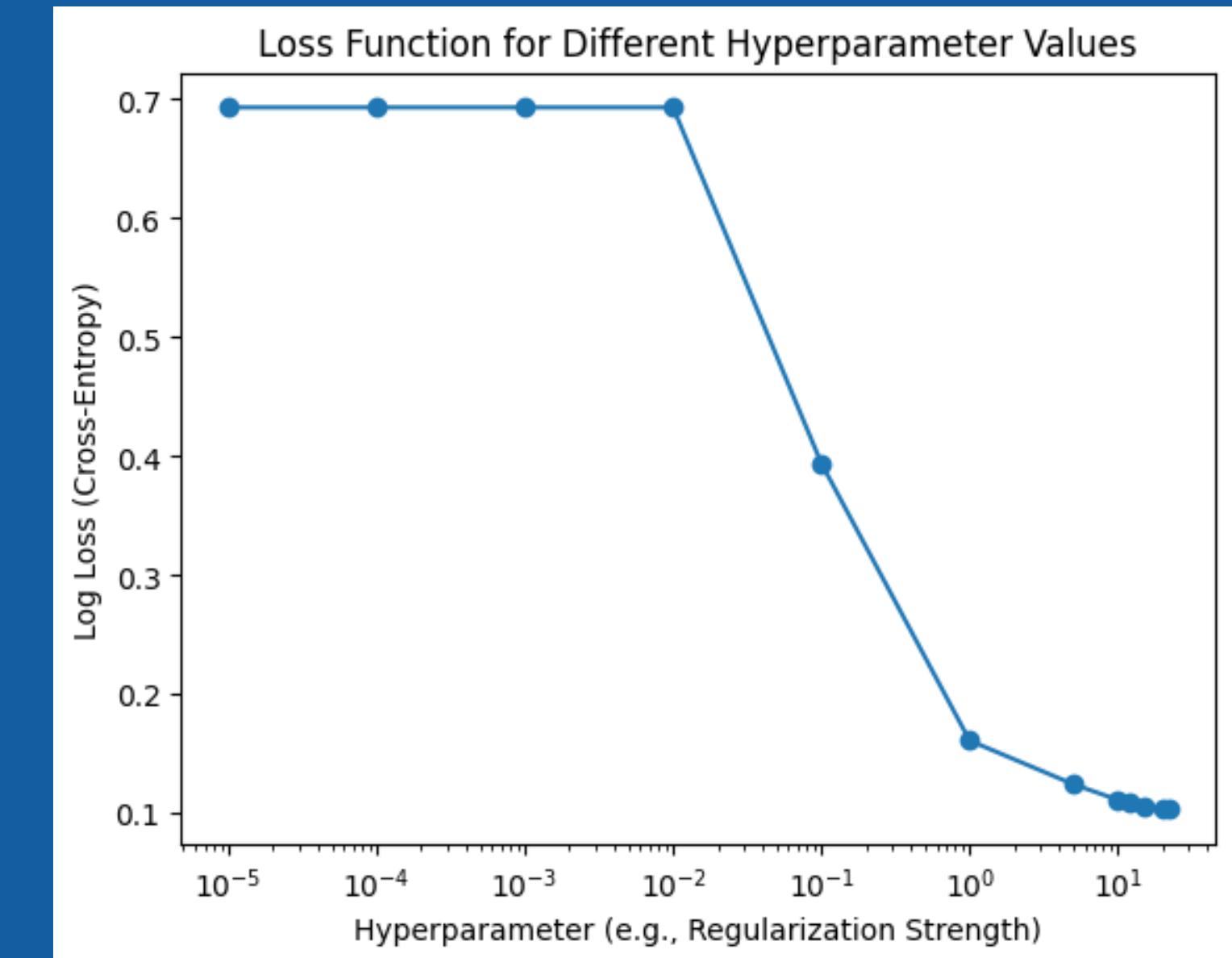
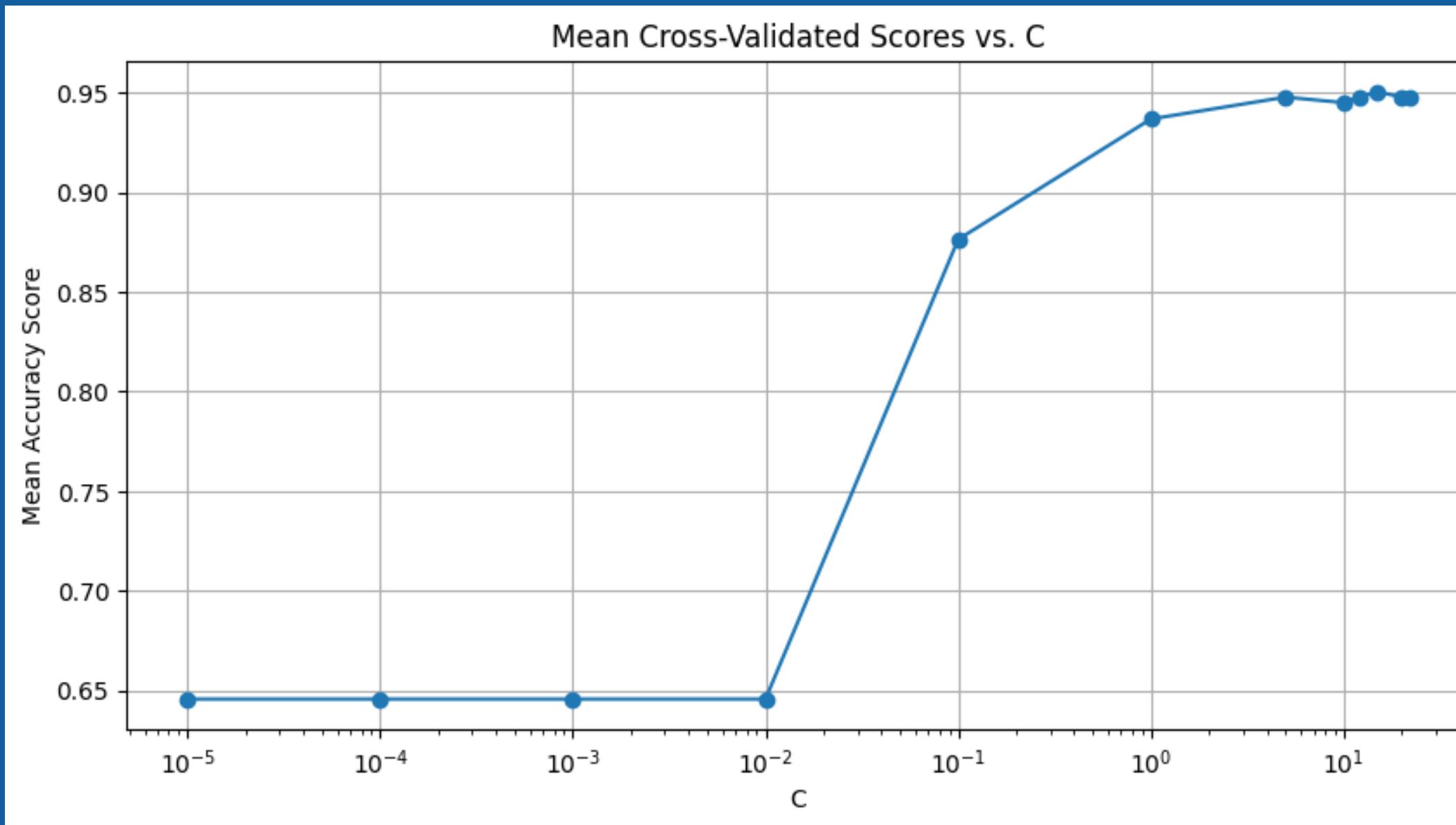
	acc	sen	spe
Train	0.94780220	0.88372093	0.98297872
Test	0.94736842	0.91489362	0.97014925
Val	0.91208791	0.88888889	0.92727273

Selected Features



	acc	sen	spe
Train	0.93681319	0.87596899	0.97021277
Test	0.94736842	0.91489362	0.97014925
Val	0.94505495	0.91666667	0.96363636

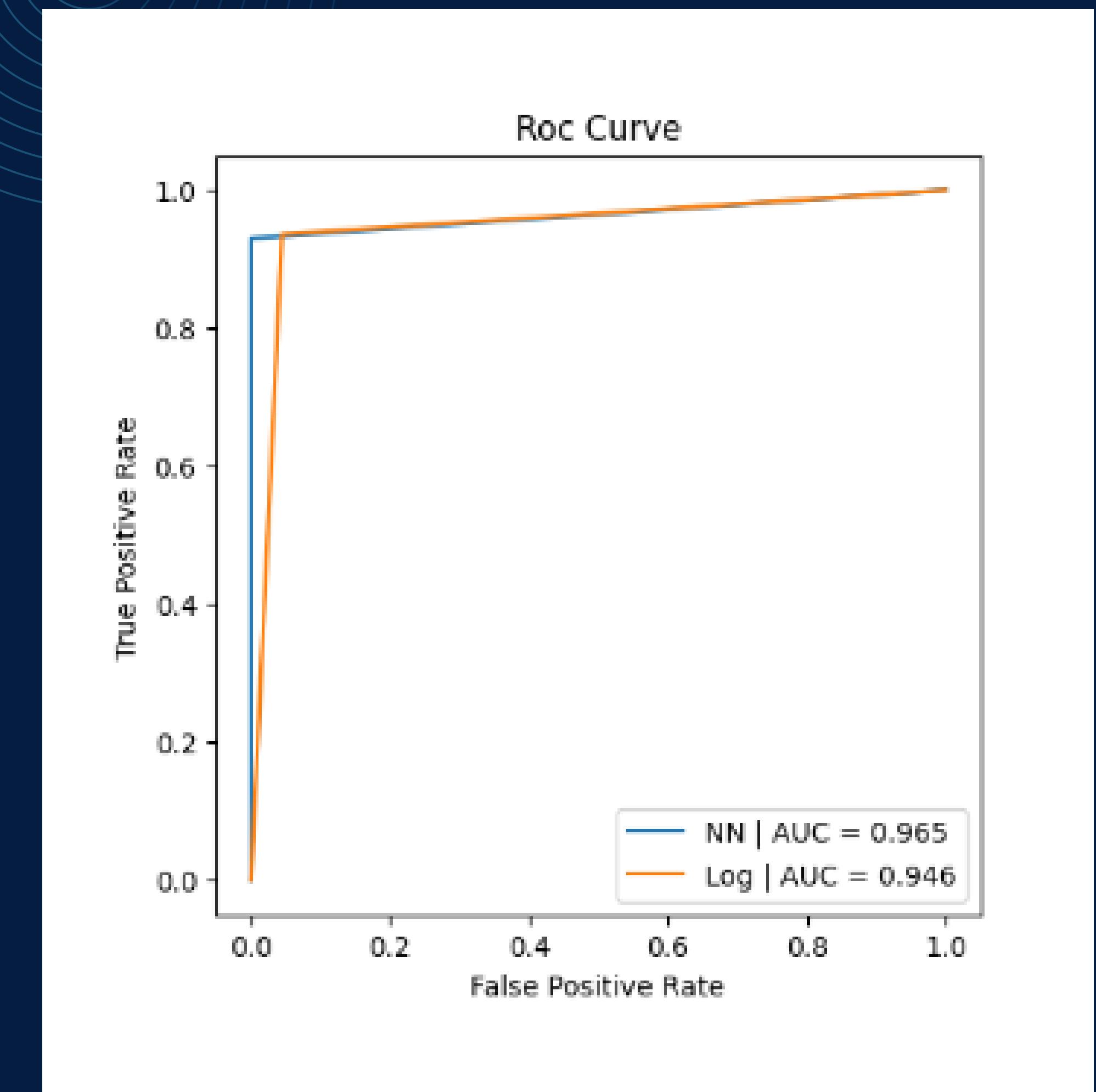
Regularization



Final Results of Logistic Regression

	acc	sen	spe
Train	0.95879121	0.93023256	0.97446809
Test	0.97802198	1.000000	0.96363636
Val	0.94736842	0.93617021	0.95522388

Results Comparison



Conclusion

