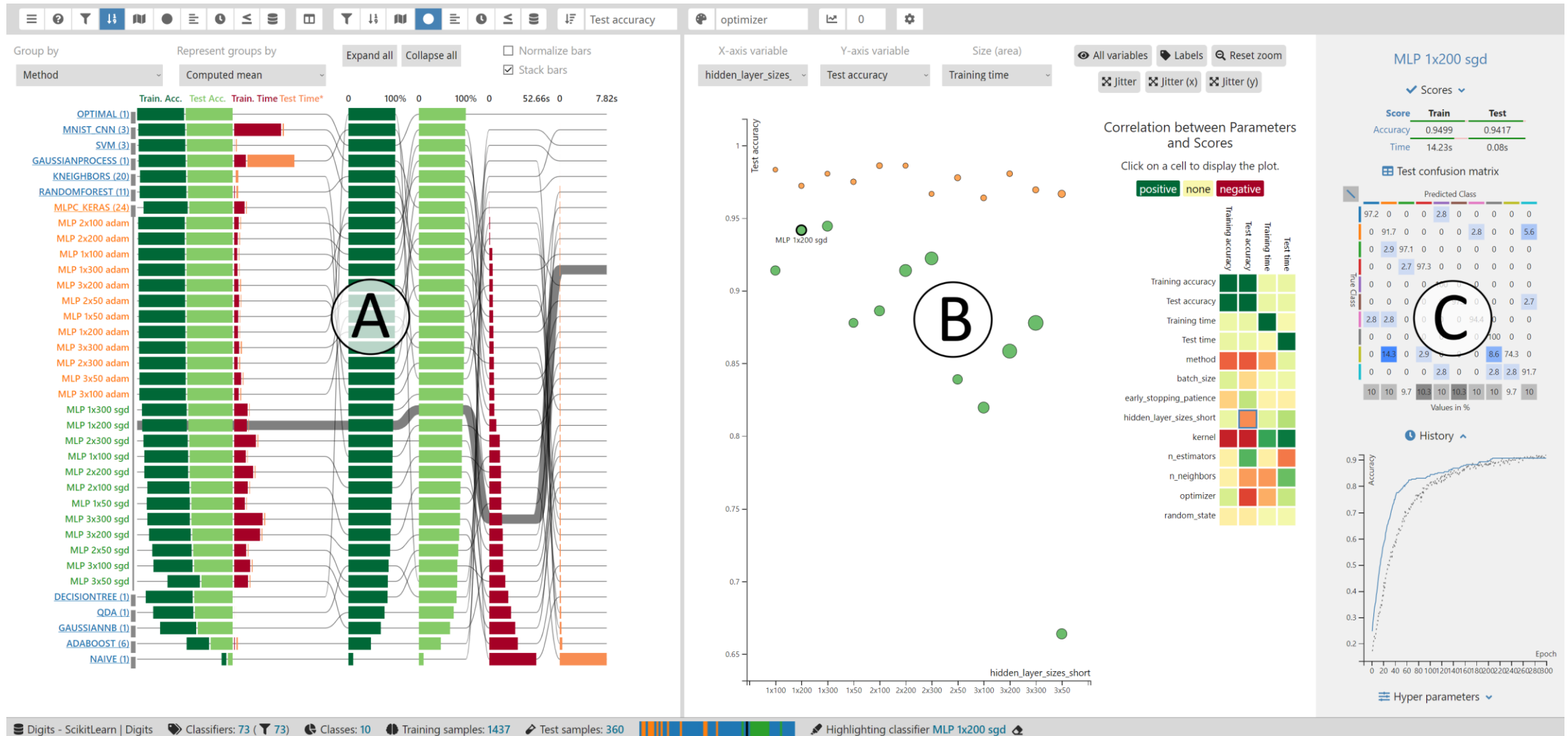


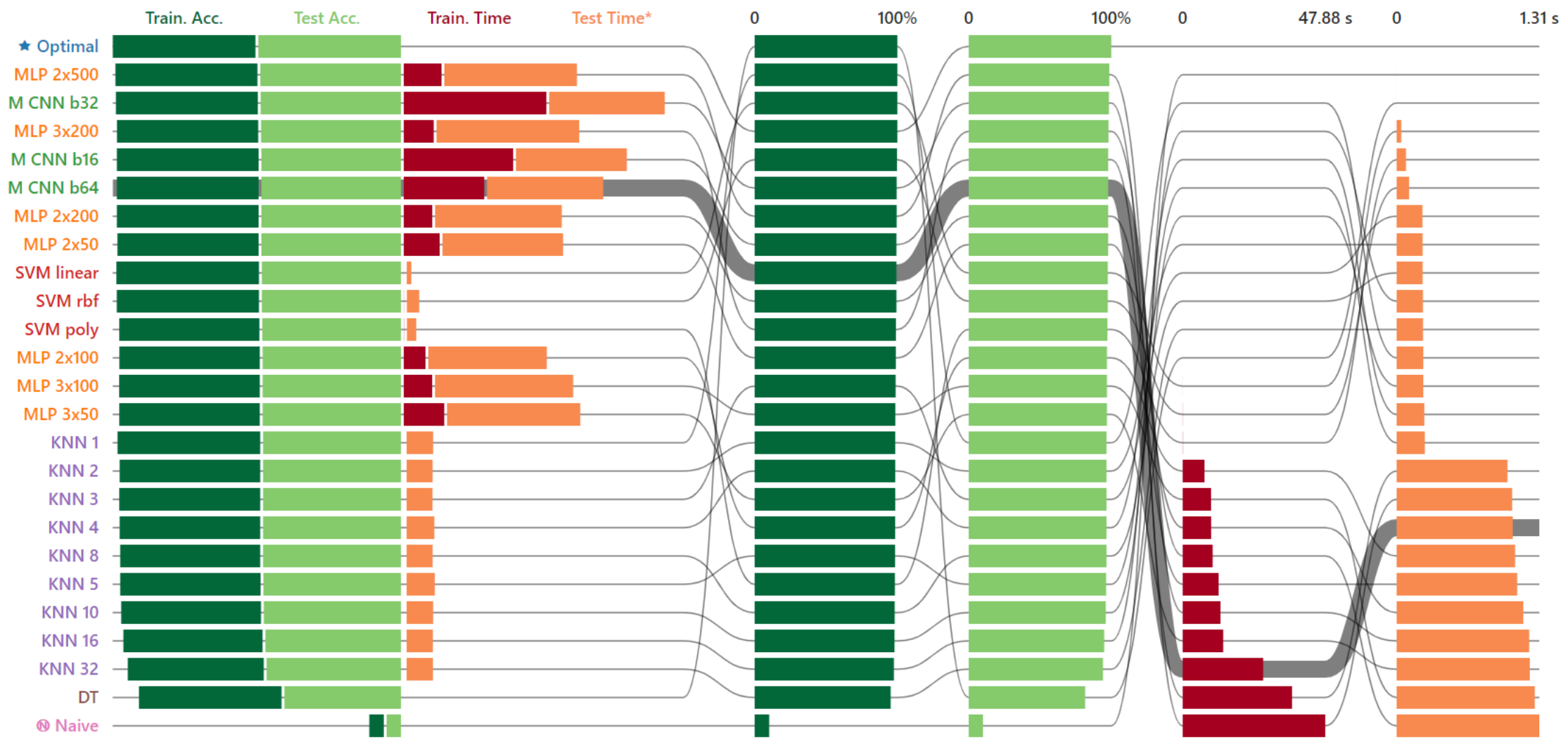
ClaVis: An Interactive Visual Comparison System for Classifiers

Overview



ClaVis can show two views (A and B) at the same time, information on the current highlight is shown in a sidebar (C)

The Ranking View

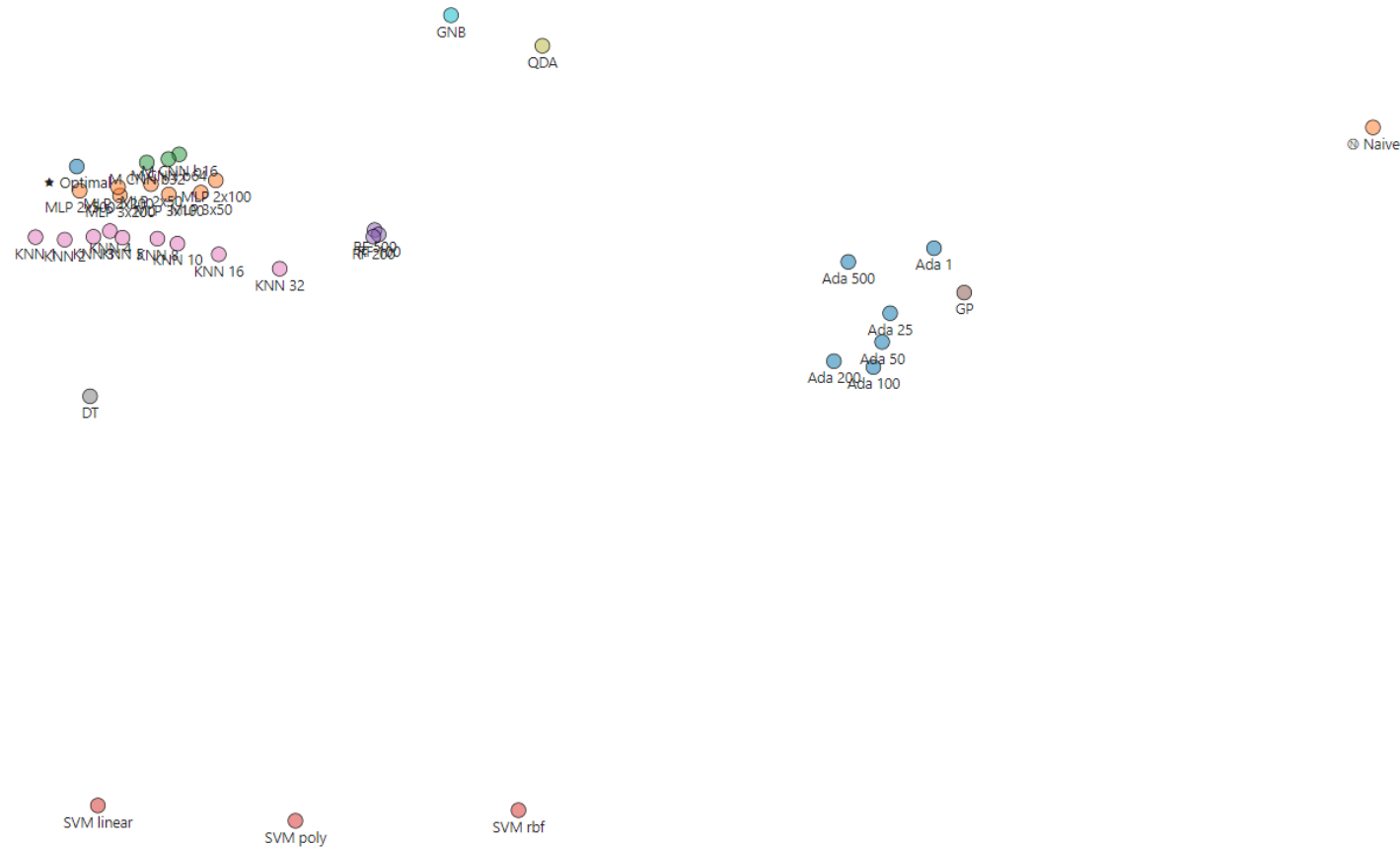


The Scatterplot View



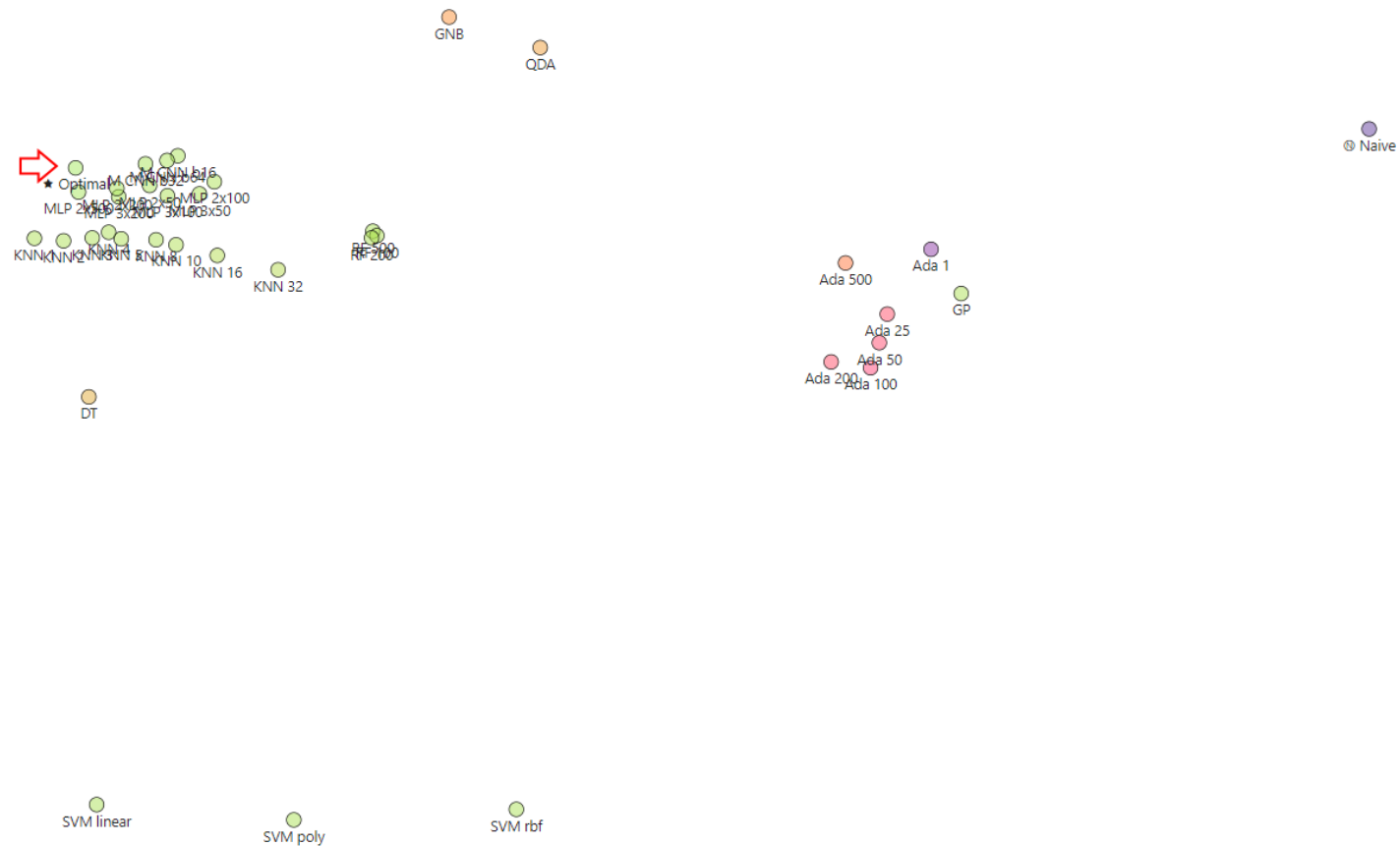
The Similarity Map

- Colored by algorithm: We see classifiers are clustered by algorithm

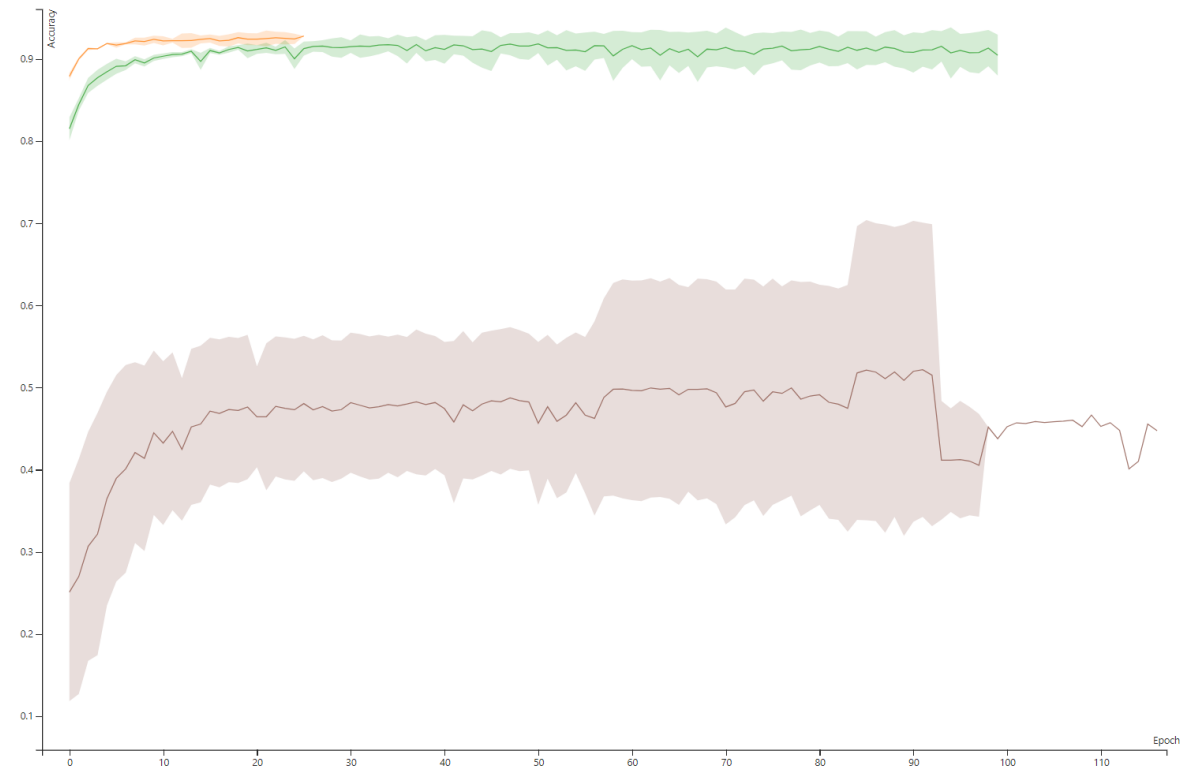
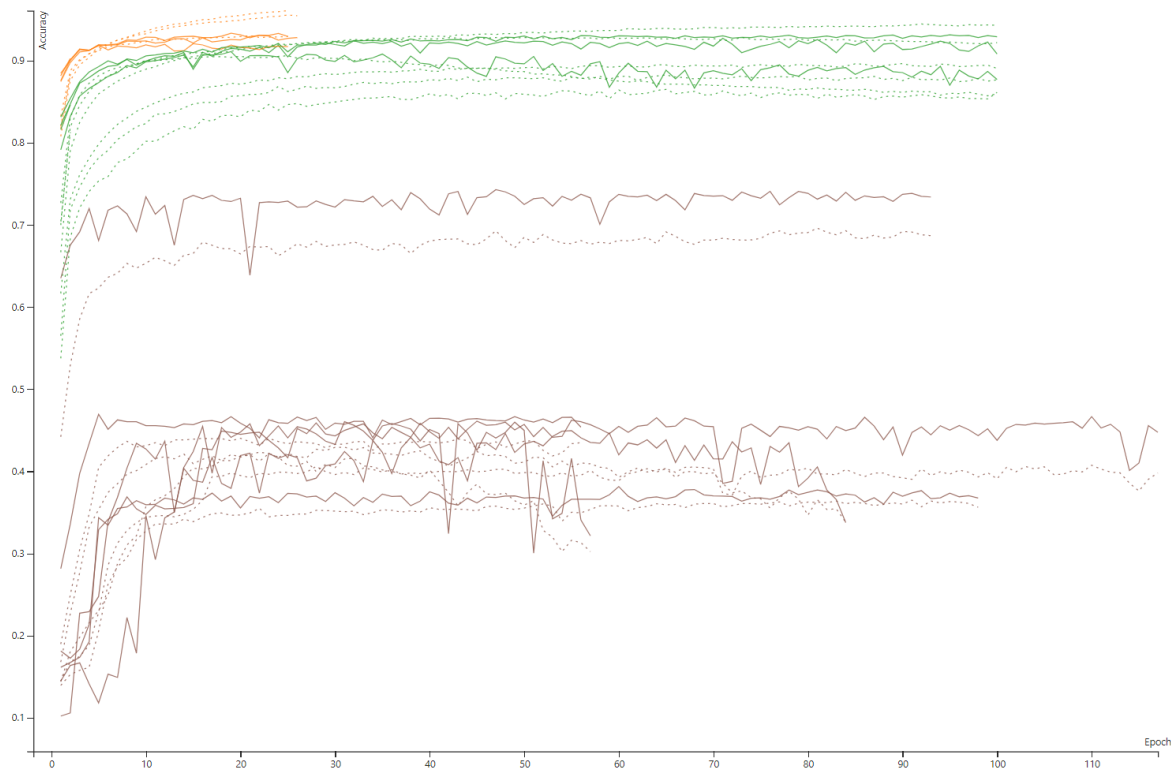


The Similarity Map

- Colored by accuracy: Classifiers closer to Optimal are generally better



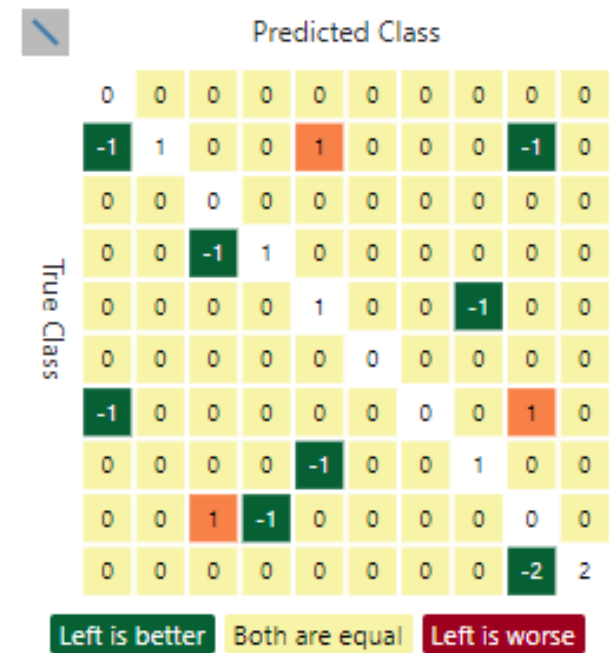
The History View



To avoid clutter when displaying many histories at once (left) we allow the user to switch to a mean and confidence interval visualization (right).

Task 5: Detecting class-specific Errors

- The confusion matrix shows which classes get confused how often
- We provide
 - Confusion matrices for single classifiers
 - A difference matrix for two selected classifiers (see figure)
 - A mean matrix for multiple selected classifiers



Task 5: Detecting class-specific Errors

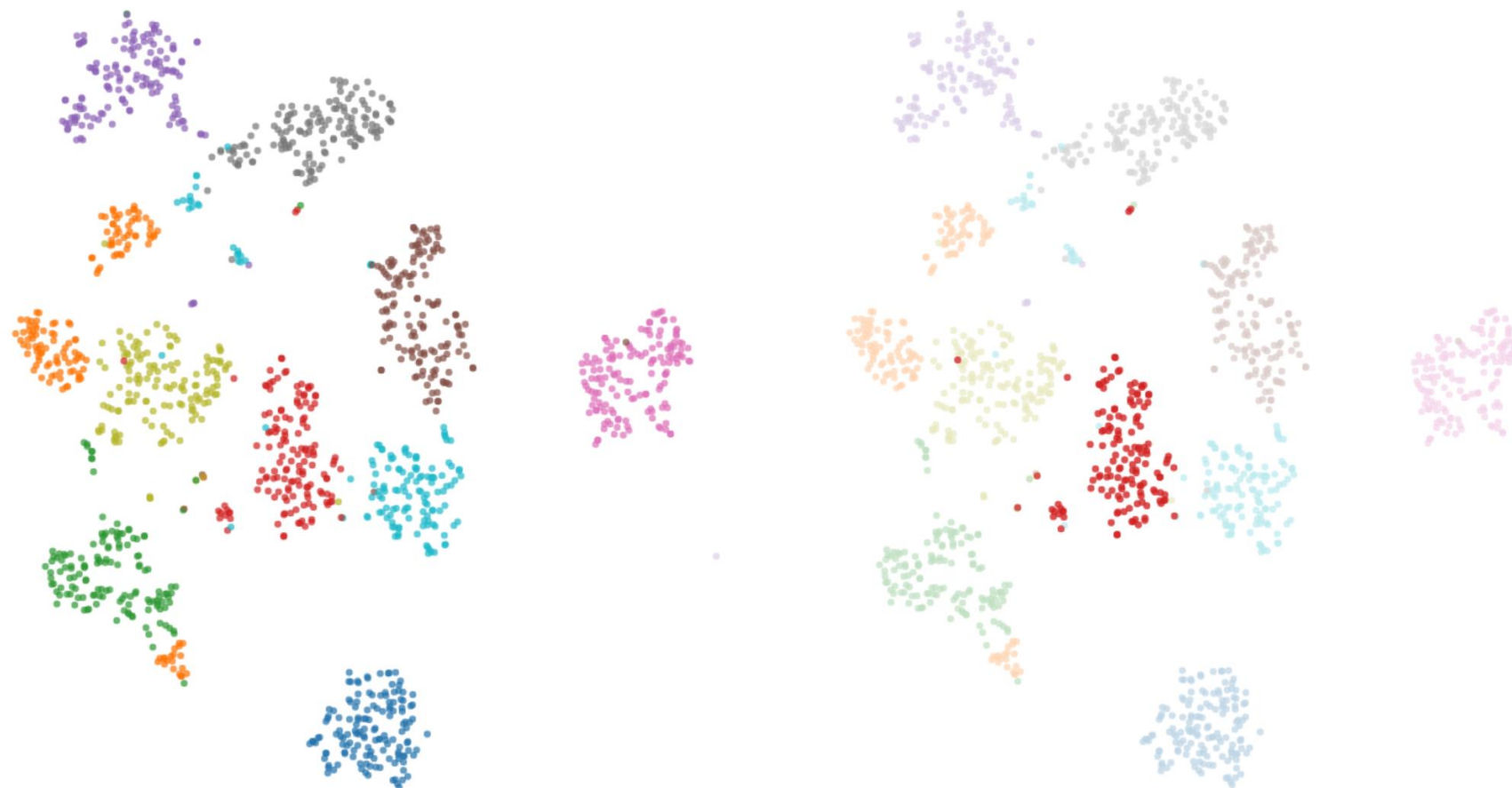
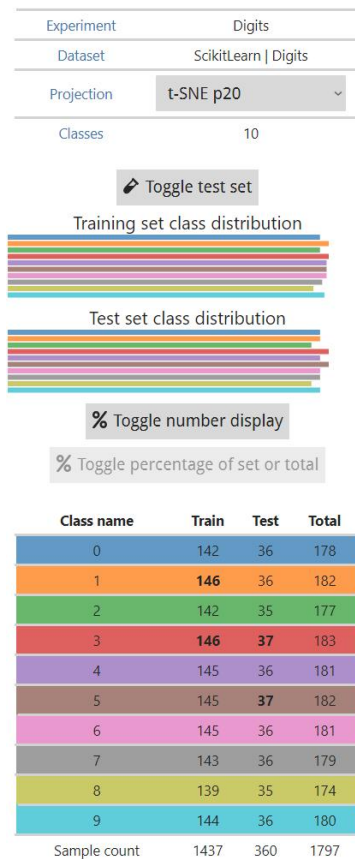
		Predicted Class				
True Class		728	59	115	137	1
		150	1636	466	98	69
		23	108	8152	559	98
		28	14	301	1696	9
		1	16	129	11	886
1040		2419	8940	2048	1043	
		Sample counts				

		Predicted Class				
True Class		728	59	115	137	1
		150	1636	466	98	69
		23	108	8152	559	98
		28	14	301	1696	9
		1	16	129	11	886
1040		2419	8940	2048	1043	
		Sample counts				

		Predicted Class				
True Class		70	5.7	11.1	13.2	0.1
		6.2	67.6	19.3	4.1	2.9
		0.3	1.2	91.2	6.3	1.1
		1.4	0.7	14.7	82.8	0.4
		0.1	1.5	12.4	1.1	84.9
		6.7	15.6	57.7	13.2	6.7
		Values in %				

Left to right: Default view, without diagonal, percentage instead of sample count.
Class sizes are shown in gray boxes below for reference.

The Dataset View



Left to right: Projection options and class statistics, a projection of the dataset, highlighting class '3' reveals outliers

Comparison View

Select two classification results to compare them:

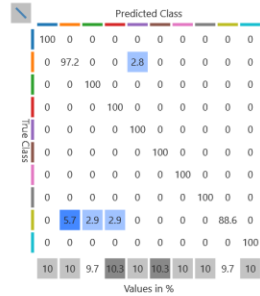
- ★ Optimal
- ✓ MLP 2x100 adam
- ✓ MLP 2x200 adam
- M CNN b32
- MLP 1x100 adam
- MLP 1x300 adam
- MLP 3x200 adam
- M CNN b64
- MLP 2x50 adam
- M CNN b16
- MLP 1x50 adam
- SVM linear
- SVM rbf
- MLP 1x200 adam
- RF 400
- RF 450
- RF 500
- SVM poly
- KNN 13
- MLP 3x300 adam
- RF 300
- RF 350
- KNN 1
- KNN 12
- KNN 2
- KNN 3
- KNN 4
- KNN 7
- KNN 8

MLP 2x100 adam

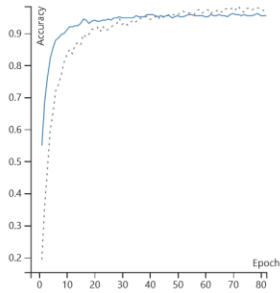
✓ Scores

Score	Train	Test
Accuracy	0.9903	0.9861
Time	5.15s	0.10s

Test confusion matrix



History



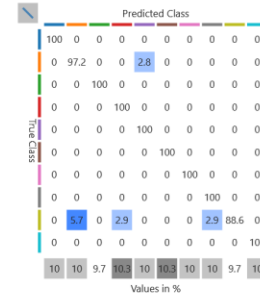
Hyper parameters

MLP 2x200 adam

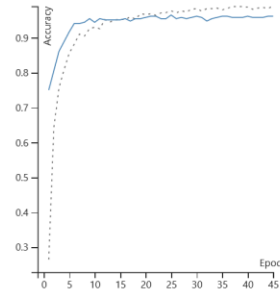
✓ Scores

Score	Train	Test
Accuracy	0.9903	0.9861
Time	3.95s	0.12s

Test confusion matrix



History



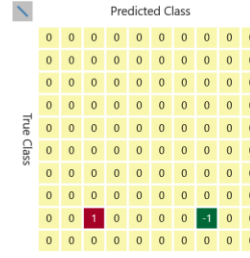
Hyper parameters

Comparison

✓ Scores

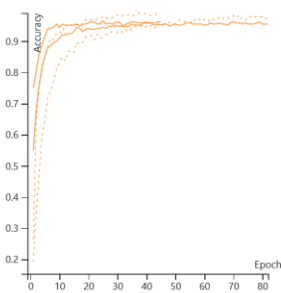
Score	MLP 2x100 adam	MLP 2x200 adam	Difference
Training acc.	99.0	99.0	0.0
Test acc.	98.6	98.6	0.0
Training time	5.15s	3.95s	-1.20s
Test time	0.10s	0.12s	0.02s

Test confusion matrix (difference)



Left is better Both are equal Left is worse

History



MLP 2x100 adam
MLP 2x200 adam

Parameters

Parameter	MLP 2x100 adam	MLP 2x200 adam
method	mlpc_keras	mlpc_keras
activation	relu	relu
batch_size	64	64
dropout	0.5	0.5
early_stopping	true	true
early_stopping_patience	20	20
epochs	300	300
hidden_layer_sizes_short	2x100	2x200
learning_rate_init	0.001	0.001

Juxtaposed Details



Filtering

Selected

★ Optimal
MLP 2x100 adam
MLP 2x200 adam
M CNN b32
MLP 1x100 adam
MLP 1x300 adam
MLP 3x200 adam
M CNN b64
MLP 2x50 adam
M CNN b16
MLP 1x50 adam
SVM linear
SVM rbf
MLP 1x200 adam
RF 400
RF 450
RF 500
SVM poly
KNN 13
MLP 3x300 adam
RF 300
RF 350
KNN 1
KNN 12
KNN 3
KNN 4
KNN 7
KNN 8

All

★ Optimal
MLP 2x100 adam
MLP 2x200 adam
M CNN b32
MLP 1x100 adam
MLP 1x300 adam
MLP 3x200 adam
M CNN b64
MLP 2x50 adam
M CNN b16
MLP 1x50 adam
SVM linear
SVM rbf
MLP 1x200 adam
RF 400
RF 450
RF 500
SVM poly
KNN 13
MLP 3x300 adam
RF 300
RF 350
GP
KNN 1
KNN 12
KNN 3
KNN 4
KNN 7
KNN 8
MLP 2x300 adam
MLP 3x50 adam
RF 250
KNN 11

Filter

Select all

Clear

Invert

Select top* 10

Method

adaboost

decisiontree

gaussiannb

gaussianprocess

kneighbors

mlpc_keras

mnist_cnn

naive

optimal

qda

randomforest

svm

Statistics

Score Extents

Score	Min	Max
Training Accuracy	0.98	1.00
Test Accuracy	0.97	1.00
Training Time	0.00	61.14
Test Time	0.00	0.48

Mean Confusion Matrix

	Predicted Class									
	0	1	2	3	4	5	6	7	8	9
0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
6	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0

Parameters

☒ Group by parameter

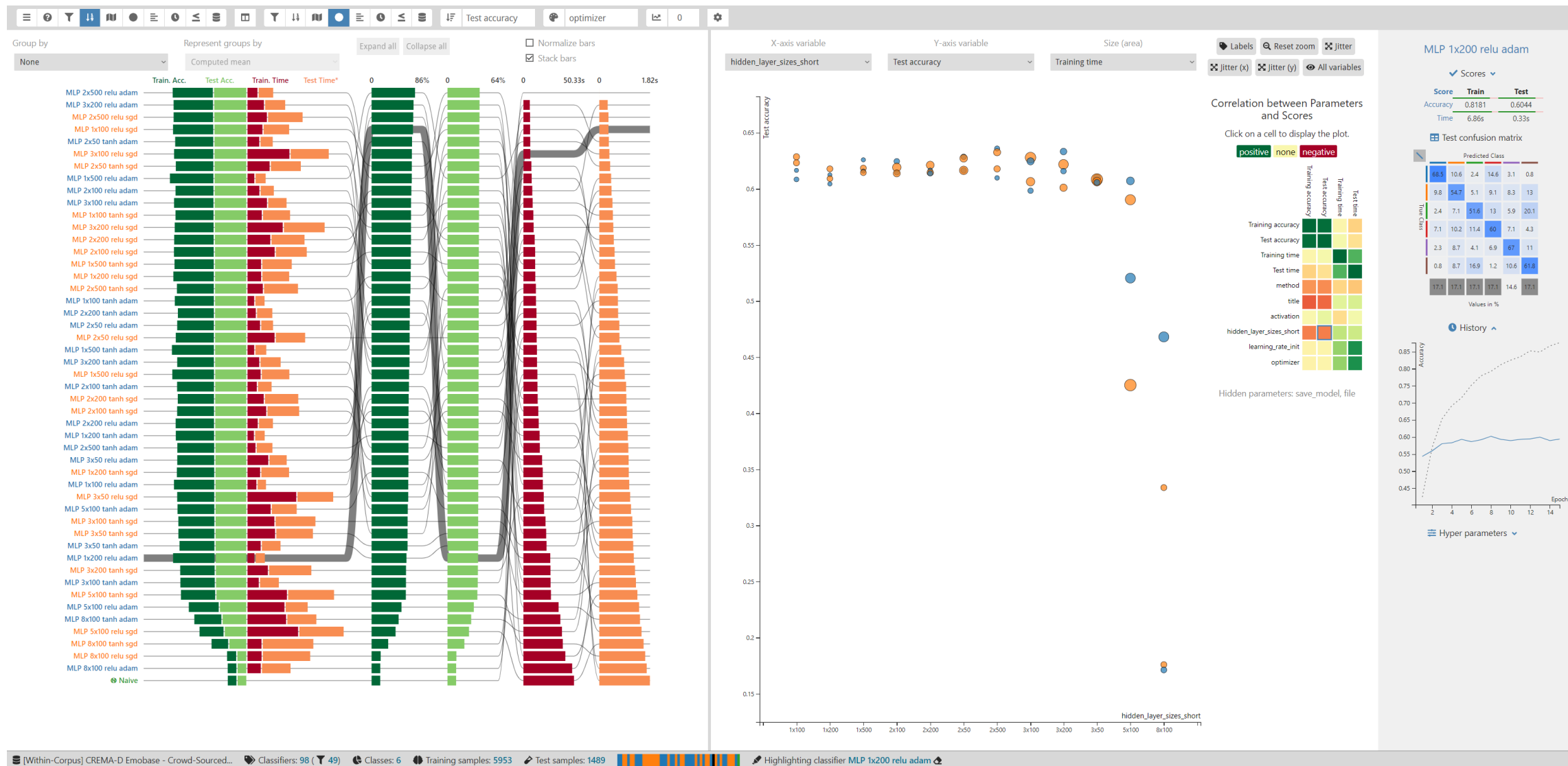
☐ Hide parameters where all values are equal

Minimum occurrence 0

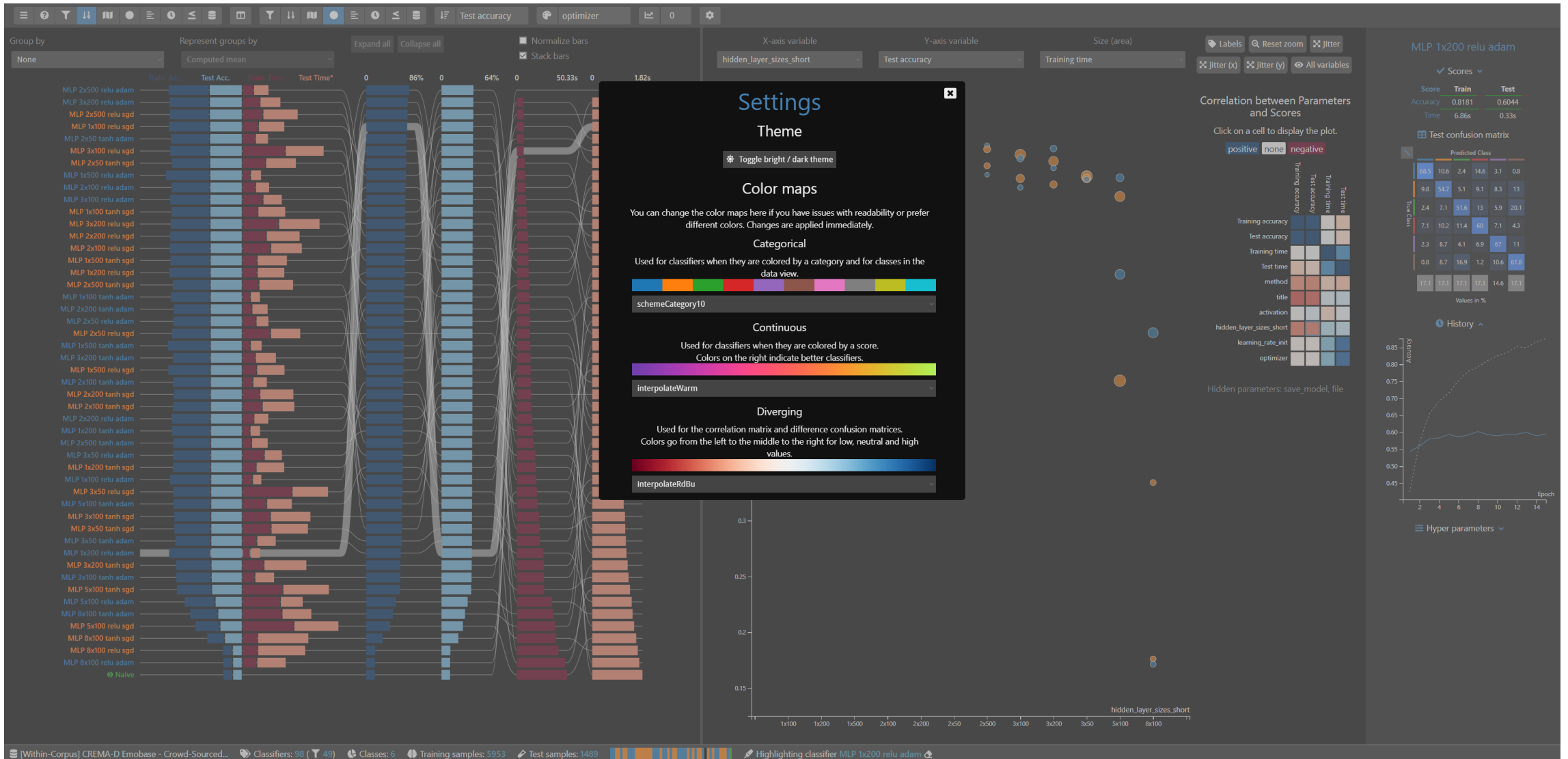
Occurrence of parameter values

Parameter	Value	Occurr.	Total		
steps_per_epoch	1000	3	3	—	+
save_model	false	12	12	—	+
random_state	42	25	28	—	+
random_state	0	3	28	—	+
optimizer	adam	9	9	—	+
optimizer	sgd	0	9	—	+
n_neighbors	13	1	8	—	+
n_neighbors	1	1	8	—	+
n_neighbors	12	1	8	—	+
n_neighbors	2	1	8	—	+

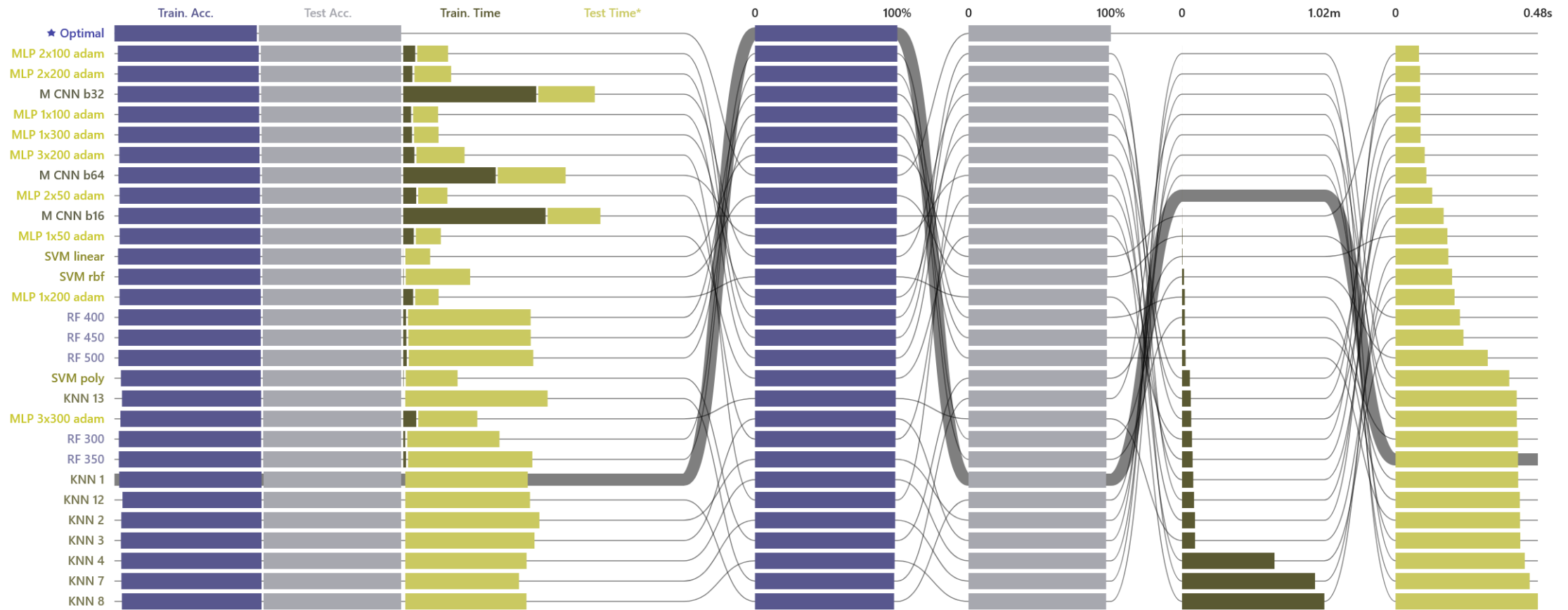
Multiple Views



Color Settings



Simulated Colorblindness



The ranking with an alternative color map as seen with protanopia, simulated with the [Firefox Accessibility Inspector](#)