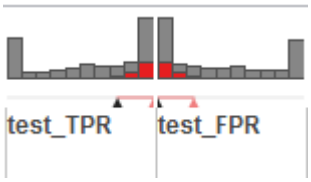
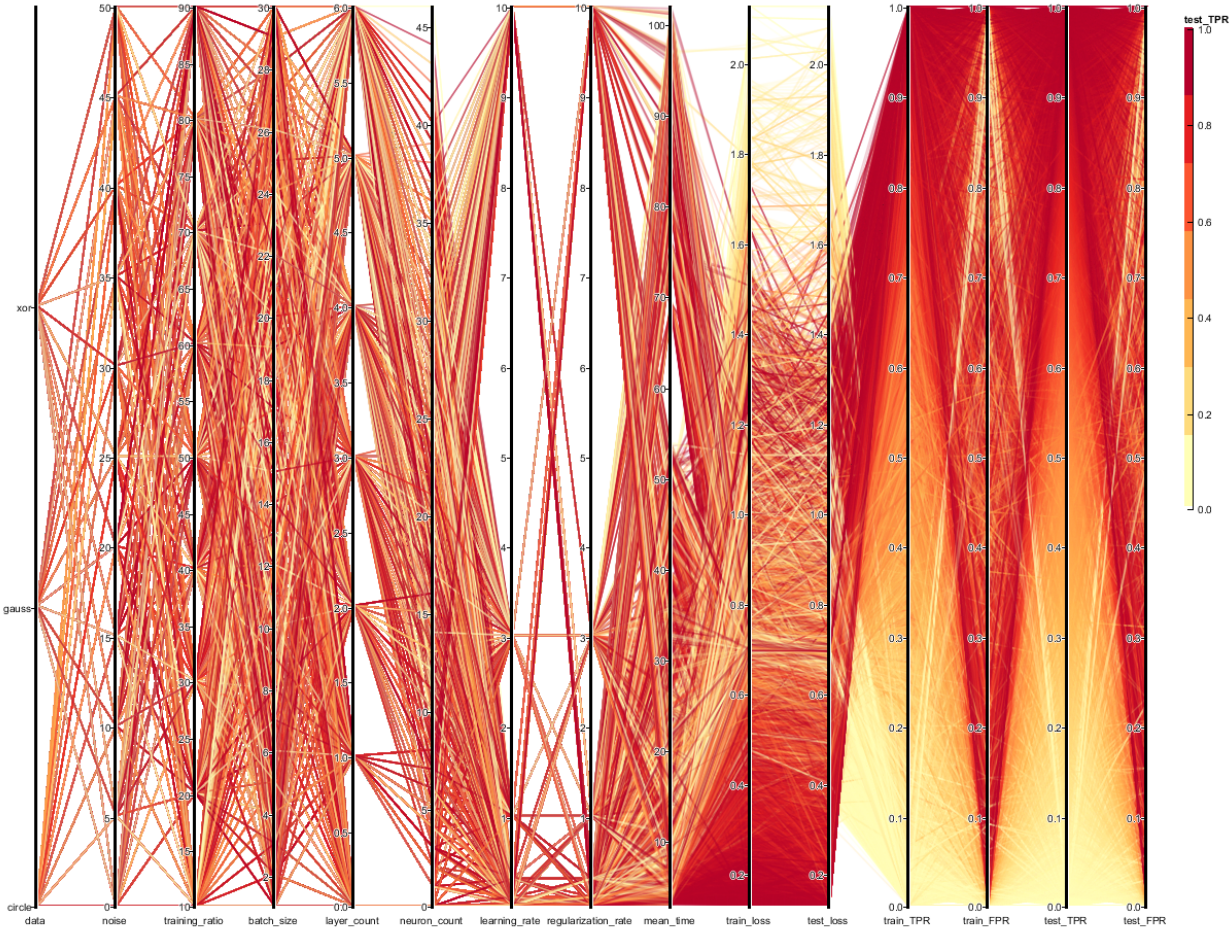
Findings on a first glance

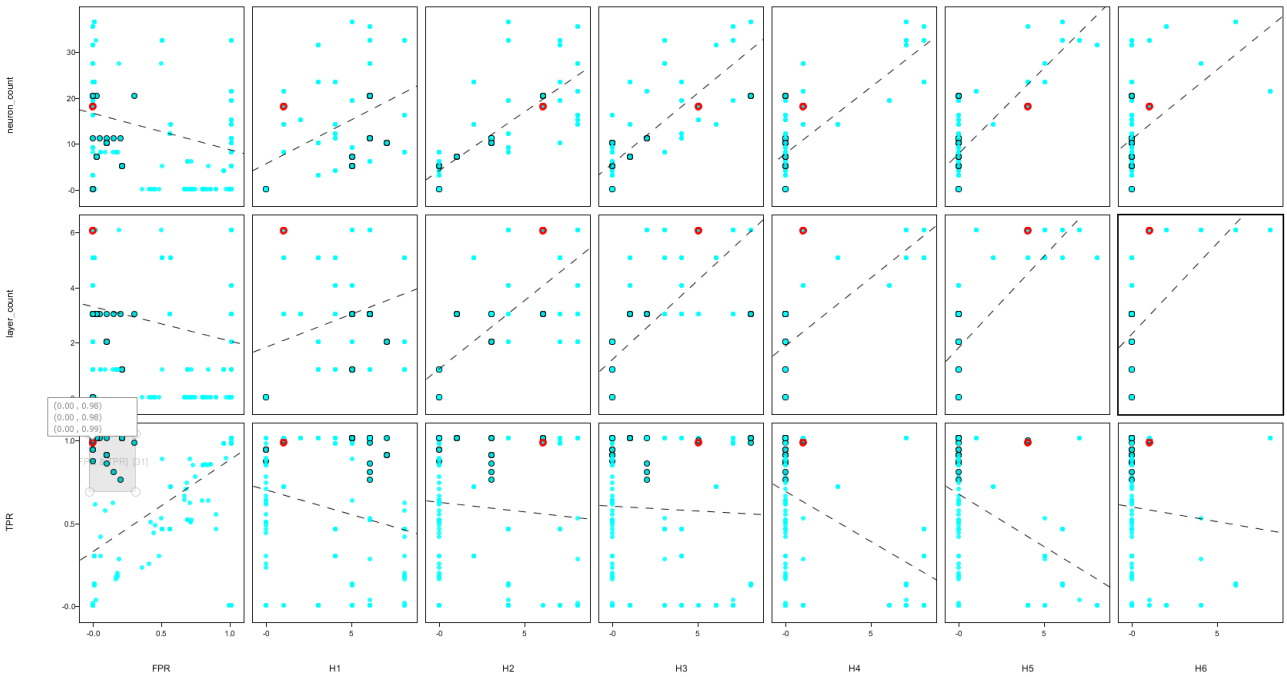


1. It is really difficult to classify the spiral dataset. High TPR and low FPR leave are the measure for a good classification. The spiral dataset has only very few somewhat successful classifications.

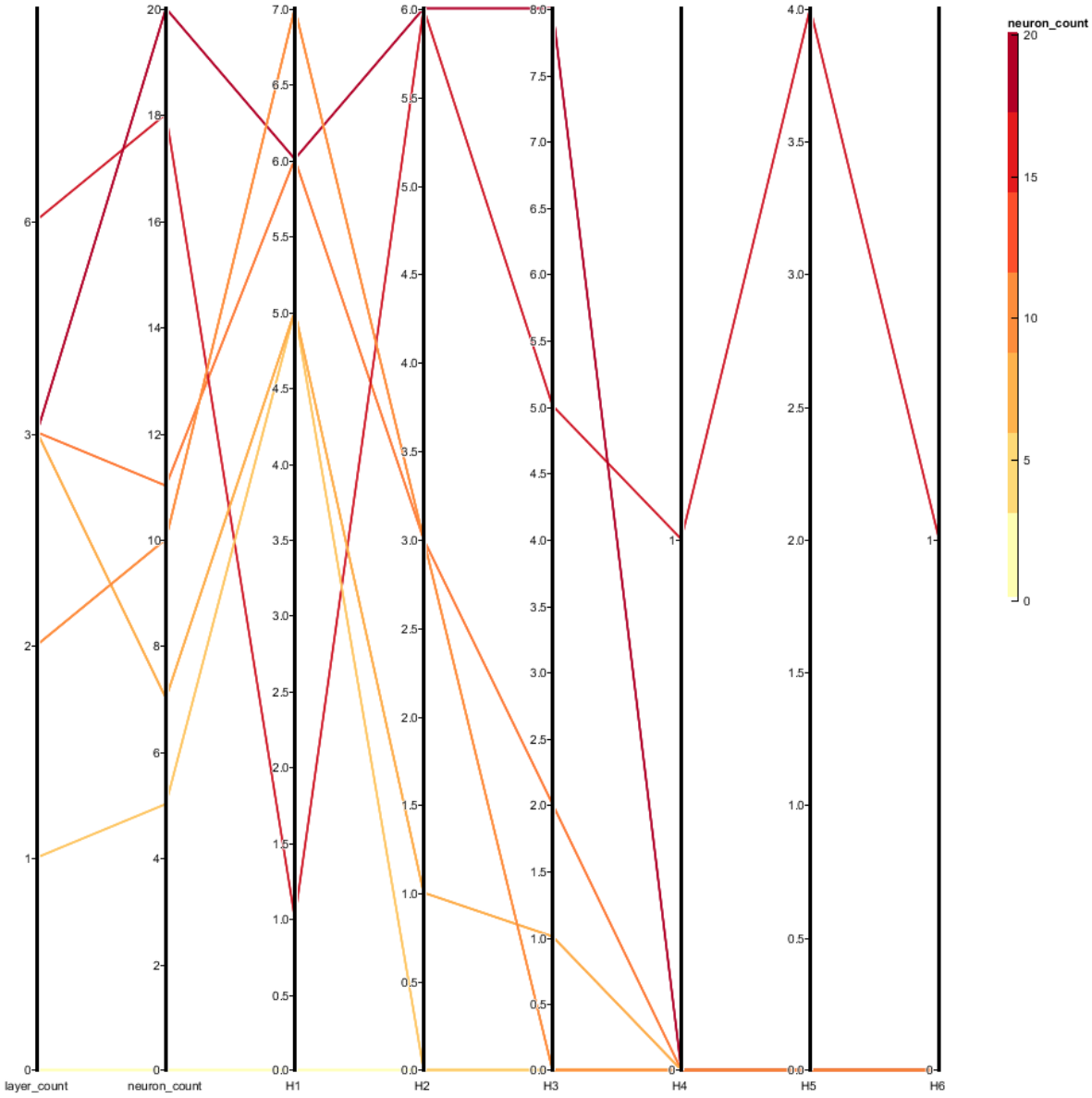


2. SIMPLIFICATION: The PC plot shows that test_TPR and train_TPR are correlated, we'll leave out test_TPR (Analogue for FPR)
train_loss and test_loss are inversely correlated to train_TPR, we'll leave those out too.

Nicolas Swoboda

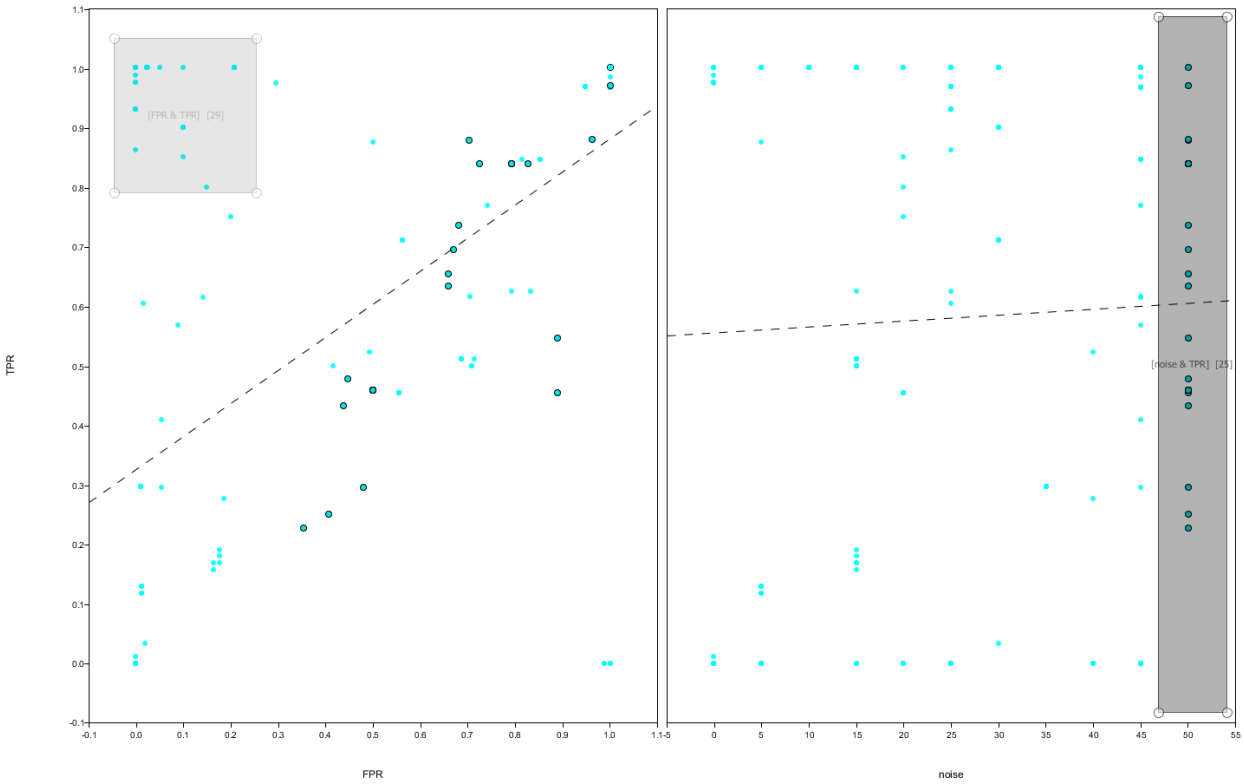


3. The highlighted point is the only run with more than three layers that provided useful results. I'll consider it an outlier (It's an xor run with only one neuron in H6). Good runs have

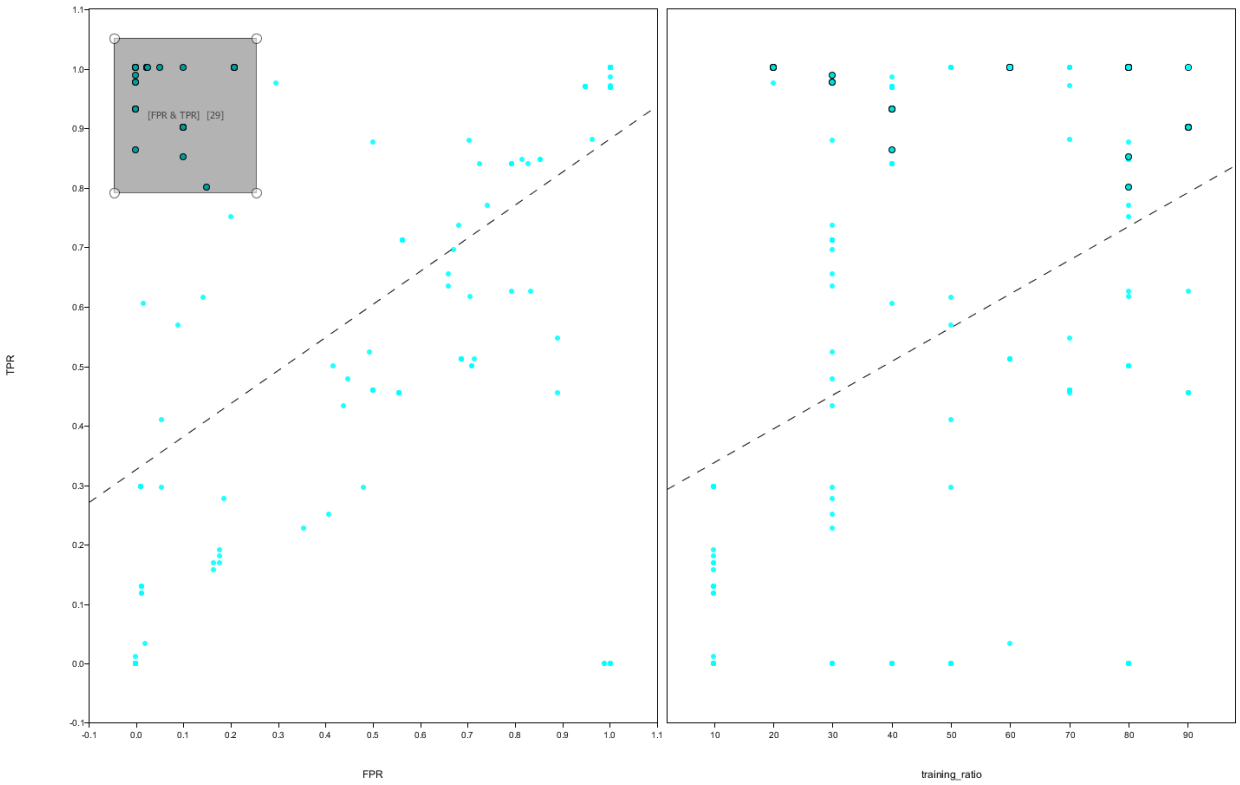


4. The more neurons a good run has the more layers it has. (red lines have high values in H3 and up)

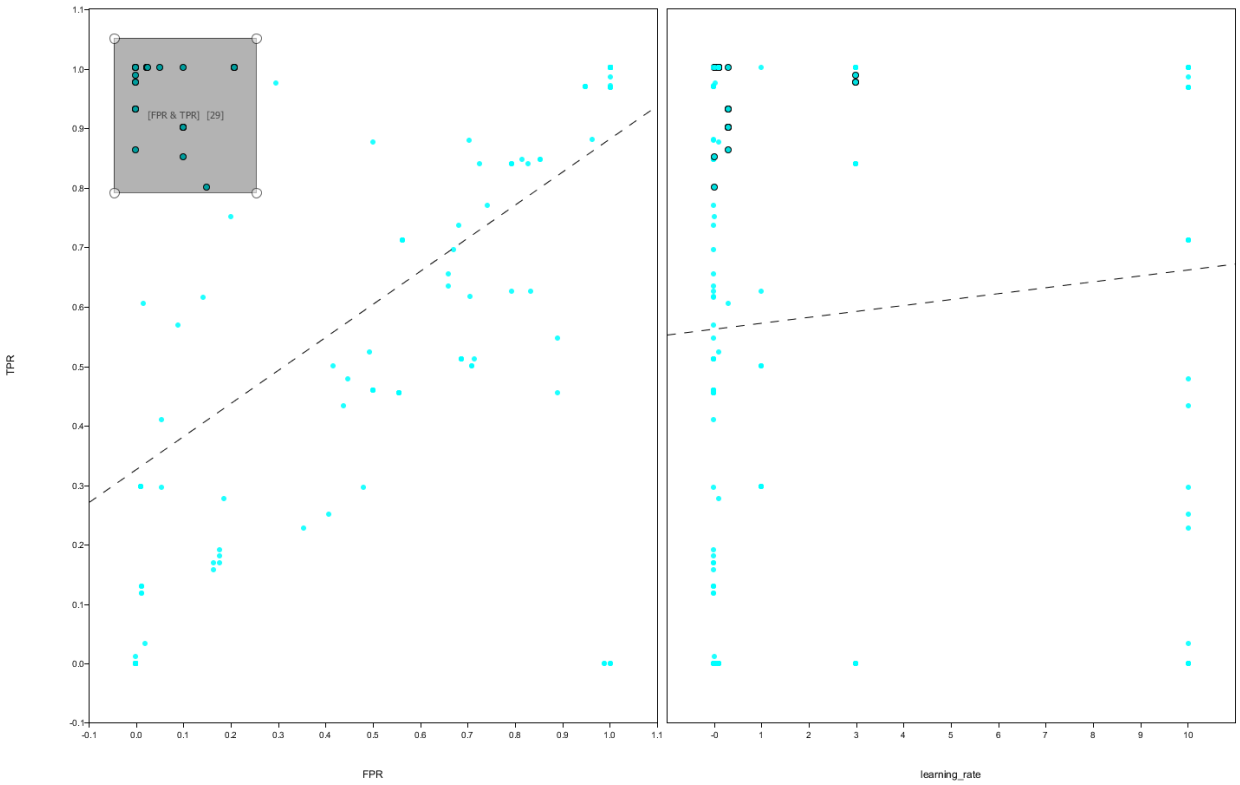
Nicolas Swoboda



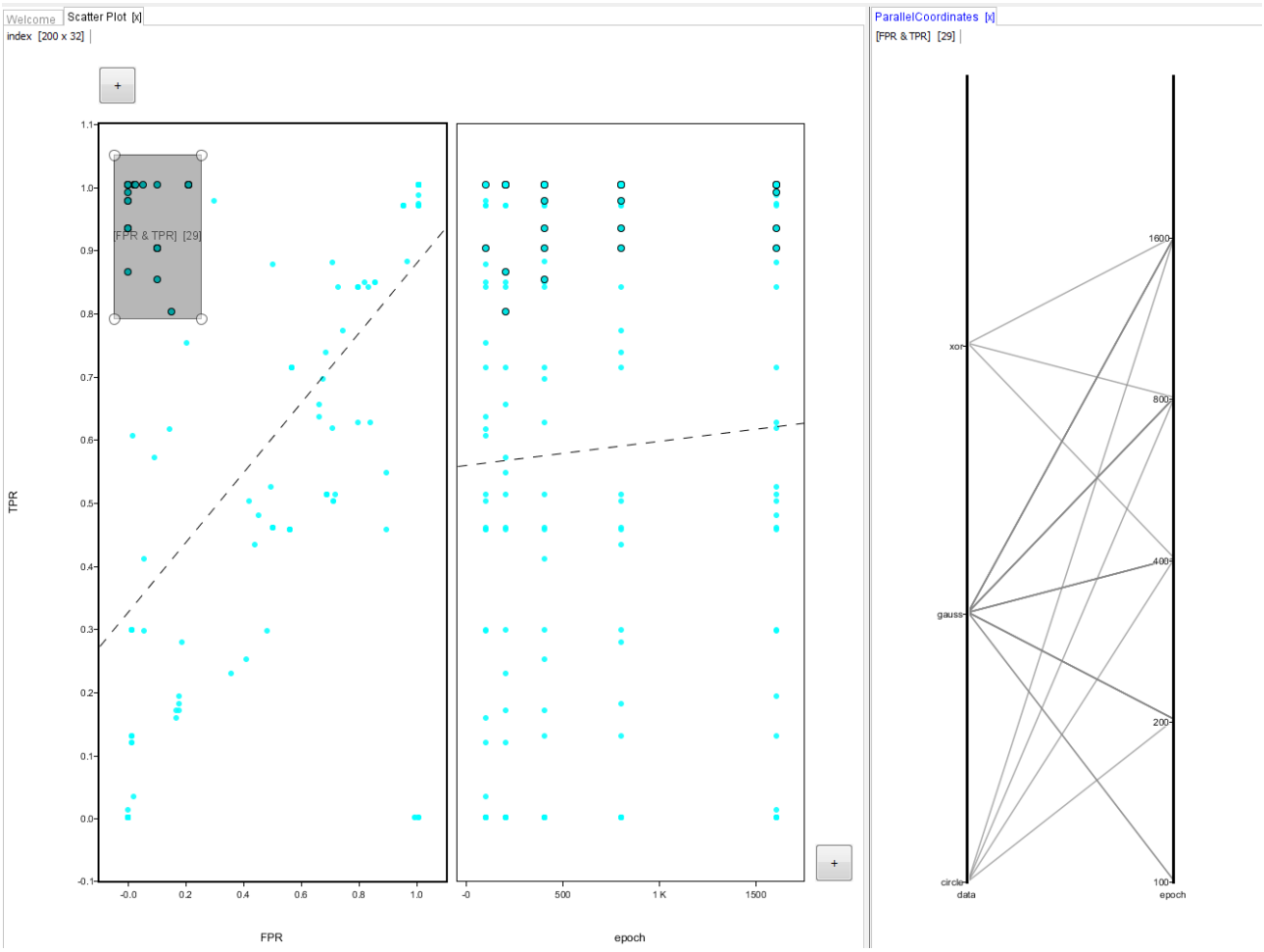
5. Unsurprisingly, high noise results in high FPR.



6. Training ration above 20 (up to 90) works out fine. Below 20 it has a good FPR but the TPR suffers.

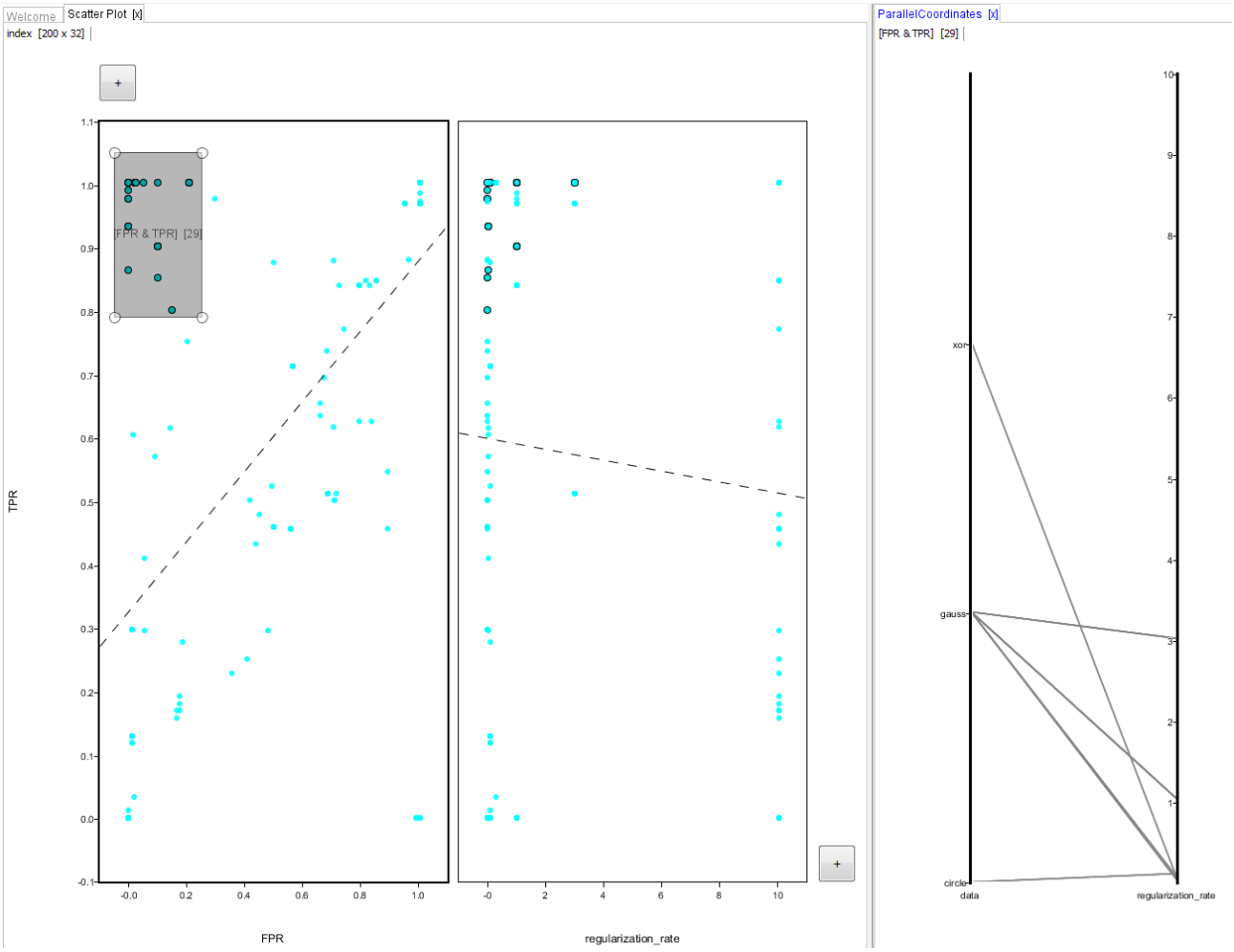


7. Learning rate above 3 is not recommended.



8. The number of epochs is quite irrelevant. 100+ is fine.

Nicolas Swoboda



9. A very low regularization rate is important, especially for XOR and CIRCLE. GAUSS can work with a regRate up to 3. No information in the „tiny“ data set for SPIRAL.

Nicolas Swoboda