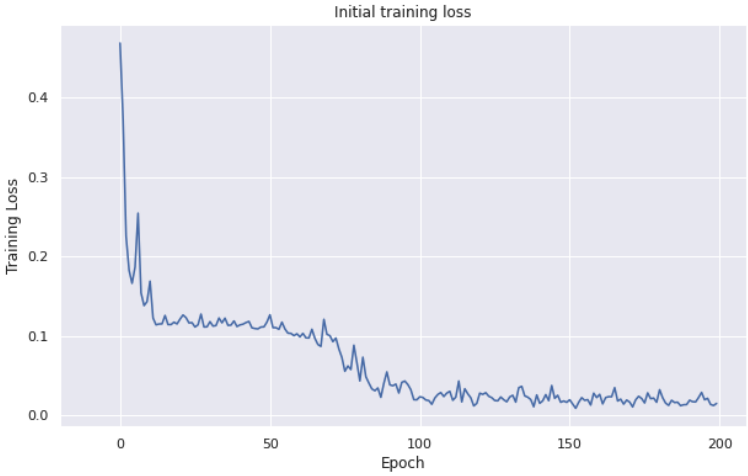
ECS269P – Artificial Intelligence

Name: Mark Orpen
Student Number: 200861742

Coursework: Curve Fitting

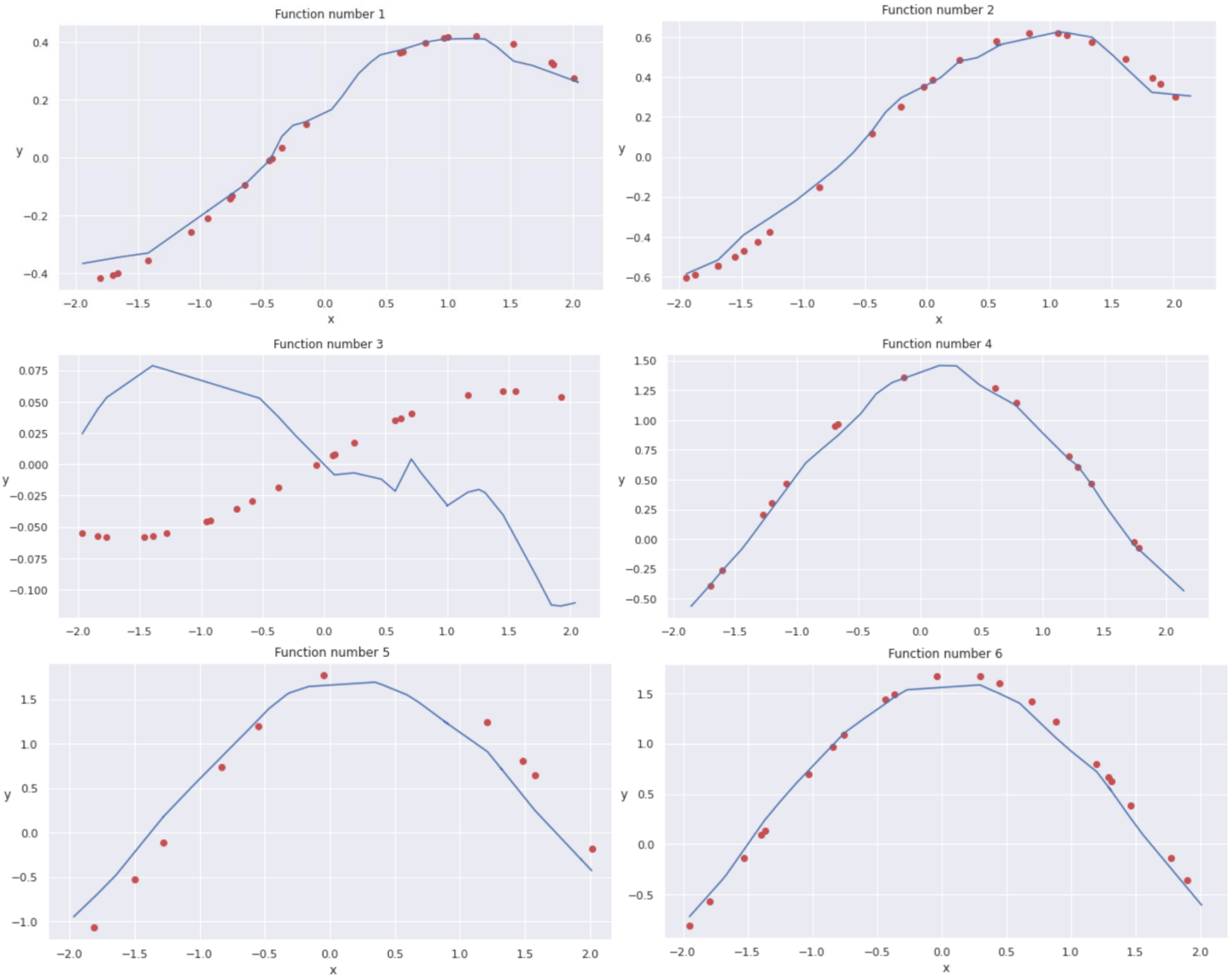
Training Loss Plots



The training loss for each epoch is shown above. The training loss initially decreases significantly in the first 10 epochs. The loss then begins to decrease considerably again between epoch 50 and 100. The training loss then fluctuates for the remaining epochs, however there is a trend showing that the loss continues to fall. This is shown in the plot below, which is a rolling mean of 50 epochs.

Test Functions

The test functions plots are shown below. The model performed well in all functions with the exception of function 3.



Hyperparameter analysis

When exploring the effects of different hyperparameters, the following parameters were used as independent variables.

h_{dim}	Batch size	r_{dim}	Learning Rate	Activation Function	Weight Decay	Optimizer
20	100	2	0.01	ReLU	0	Adam

Results of different hyperparameters:

h_{dim}	Training Loss	Validation Loss
10	0.0394	0.0364
14	0.0243	0.0195
20	0.0105	0.0119
40	0.0085	0.0101
80	0.0123	0.0392

Batch size	Training Loss	Validation Loss
24	0.0086	0.0137
50	0.0084	0.0071
100	0.0105	0.0119
200	0.0208	0.0201
400	0.0264	0.0301

Learning Rate	Training Loss	Validation Loss
0.1	0.0955	0.1066
0.05	0.2627	0.2549
0.01	0.0105	0.0119
0.005	0.0252	0.0239
0.001	0.1087	0.1146

Optimizer	Training Loss	Validation Loss
Adam	0.0394	0.0364
SGD	0.2608	0.2611

Activation Function	Training Loss	Validation Loss
ReLU	0.0394	0.0364
PReLU (0.25)	0.0098	0.0109
PReLU (0.1)	0.0072	0.0067
PReLU (0.05)	0.0081	0.0070

Weight Decay	Training Loss	Validation Loss
0.01	0.2627	0.2628
0.005	0.2618	0.2614
0.001	0.0153	0.0251
0.0005	0.0116	0.0086
0.0001	0.0102	0.0076

r_dim	Training Loss	Validation Loss
4	0.0115	0.0082
6	0.0073	0.0046
10	0.0072	0.0047
20	0.0085	0.0057