Assignment 1

Computational Intelligence, SS2017

Team Members		
Last name	First name	Matriculation Number
Reeh	Lucas	00630128

Contents

List of Figures

1 Linear Regression

1.1 Derivation of Regularized Linear Regression

1.2 Linear Regression with polynomial features

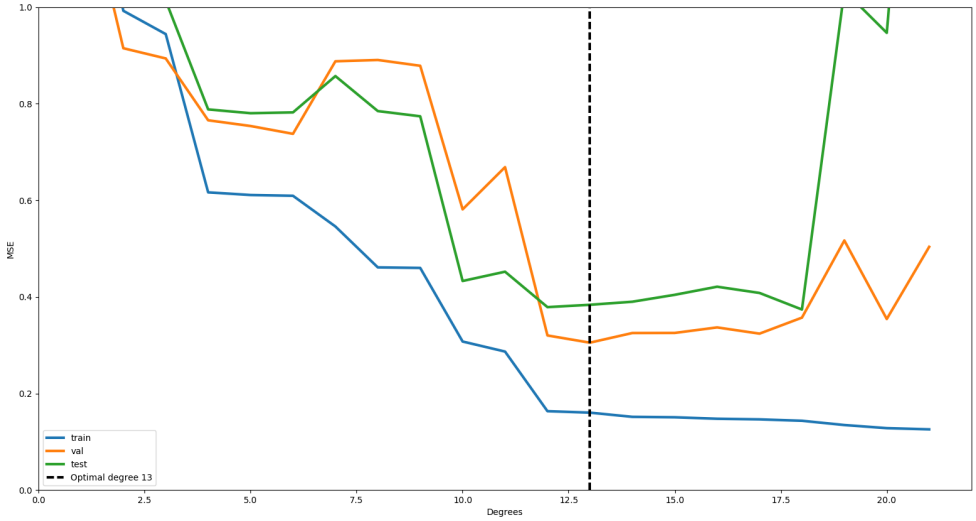

Figure 1: Training, validation and testing errors

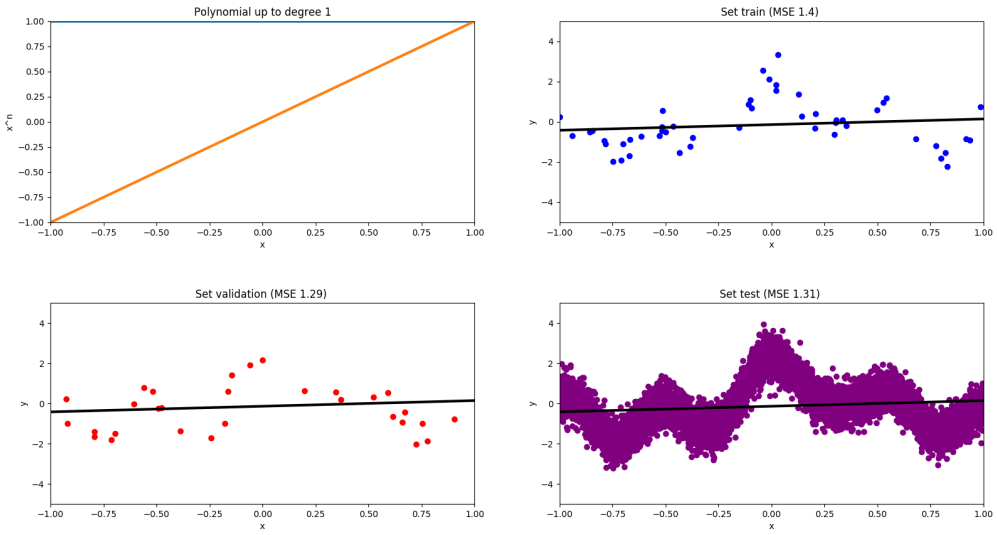

Figure 2: Linear Regression (Polynomial, Degree 1)

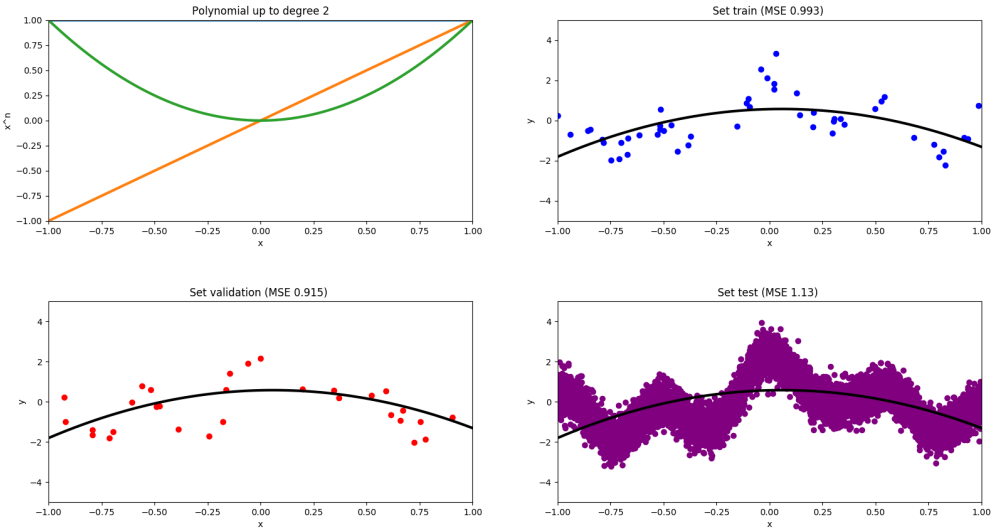

Figure 3: Linear Regression (Polynomial, Degree 2)

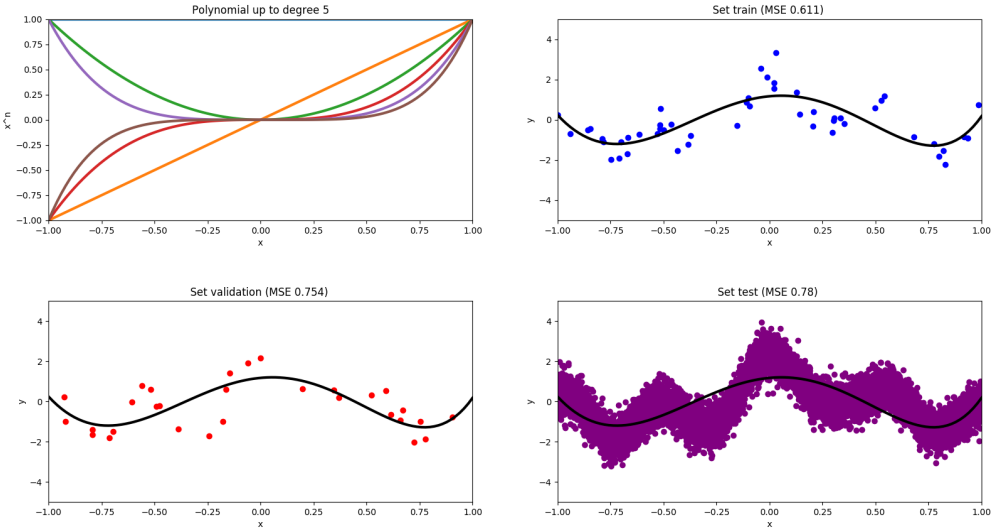

Figure 4: Linear Regression (Polynomial, Degree 5)

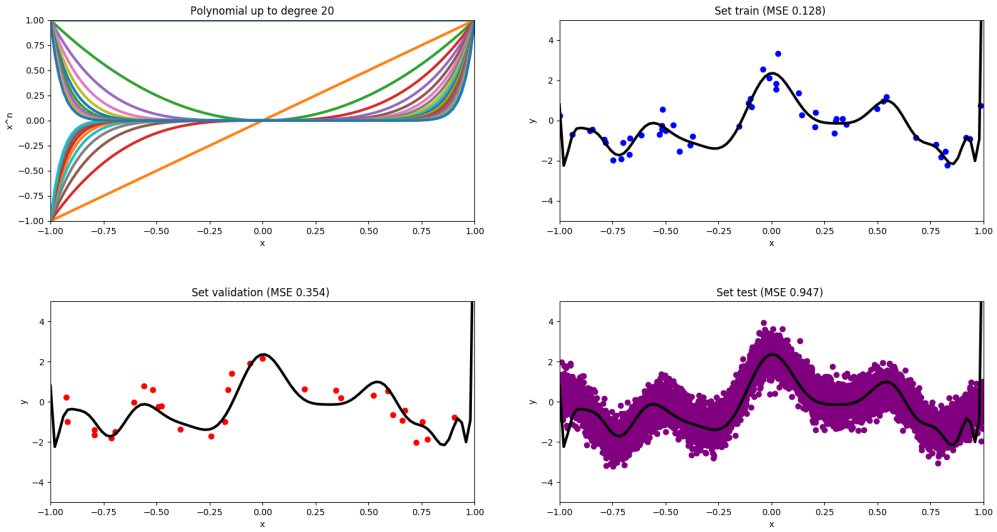

Figure 5: Linear Regression (Polynomial, Degree 20)

- Lowest training error when using degree 21

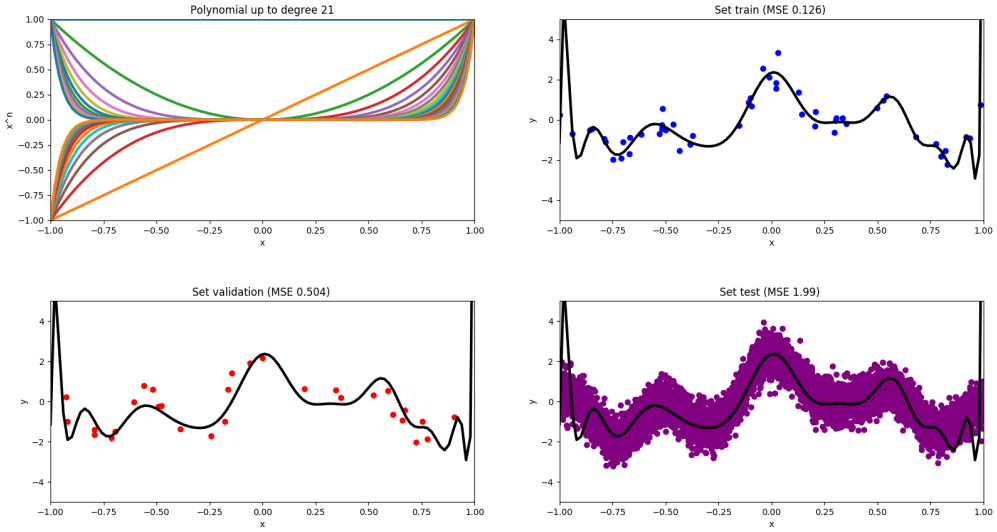

Figure 6: Linear Regression (Polynomial, Degree 21)

- Lowest validation error occurs when using degree 13

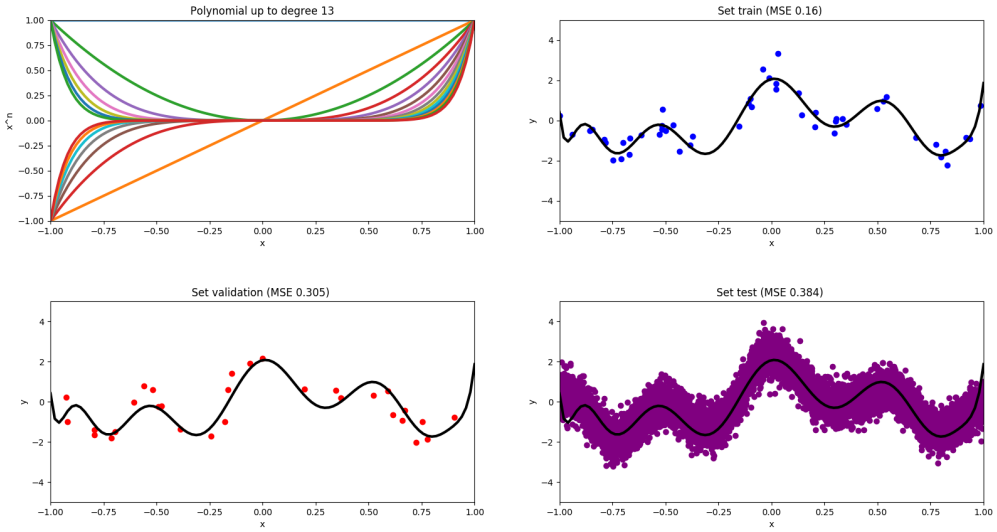

Figure 7: Linear Regression (Polynomial, Degree 13)

- **Discussion**

  Validation sets help to estimate performance of algorithms used for predictions and also to select a hypothesis (lowes error on set data). According to the error in the test set no over-fitting occured up to a degree of 13 (but would on higher degrees as can clearly be seen in Figure for degree 21, outliers and lesser data).

1.3 Linear Regression with radial basis functions

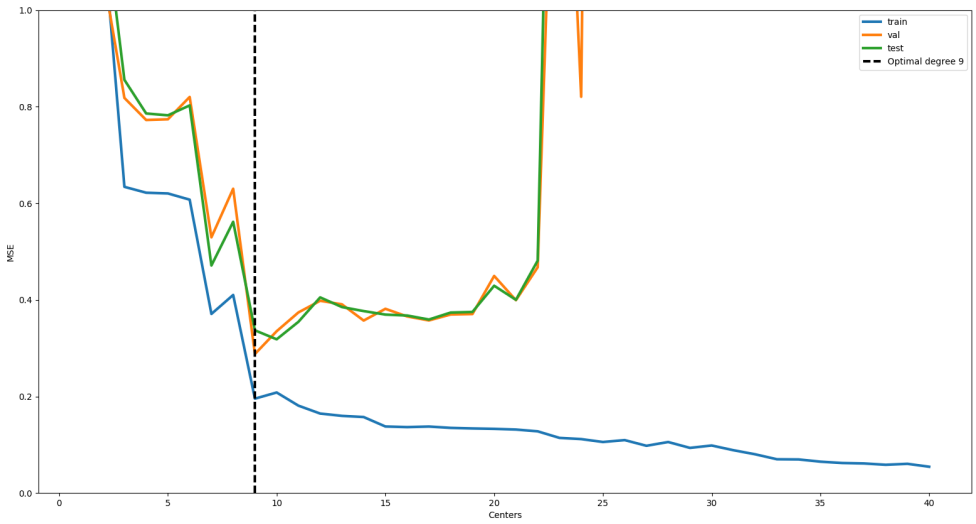

Figure 8: Training, validation and testing errors

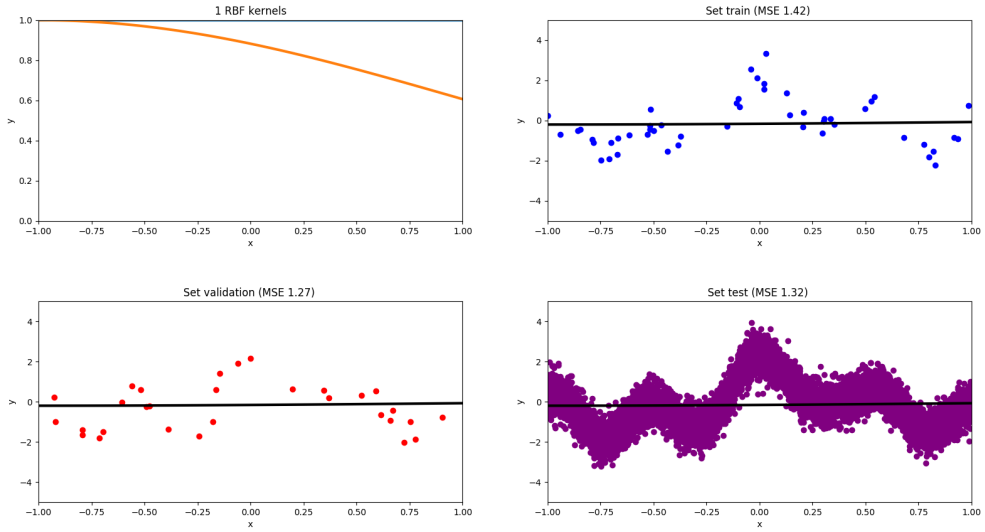

Figure 9: Linear Regression (Bias, Center 1)

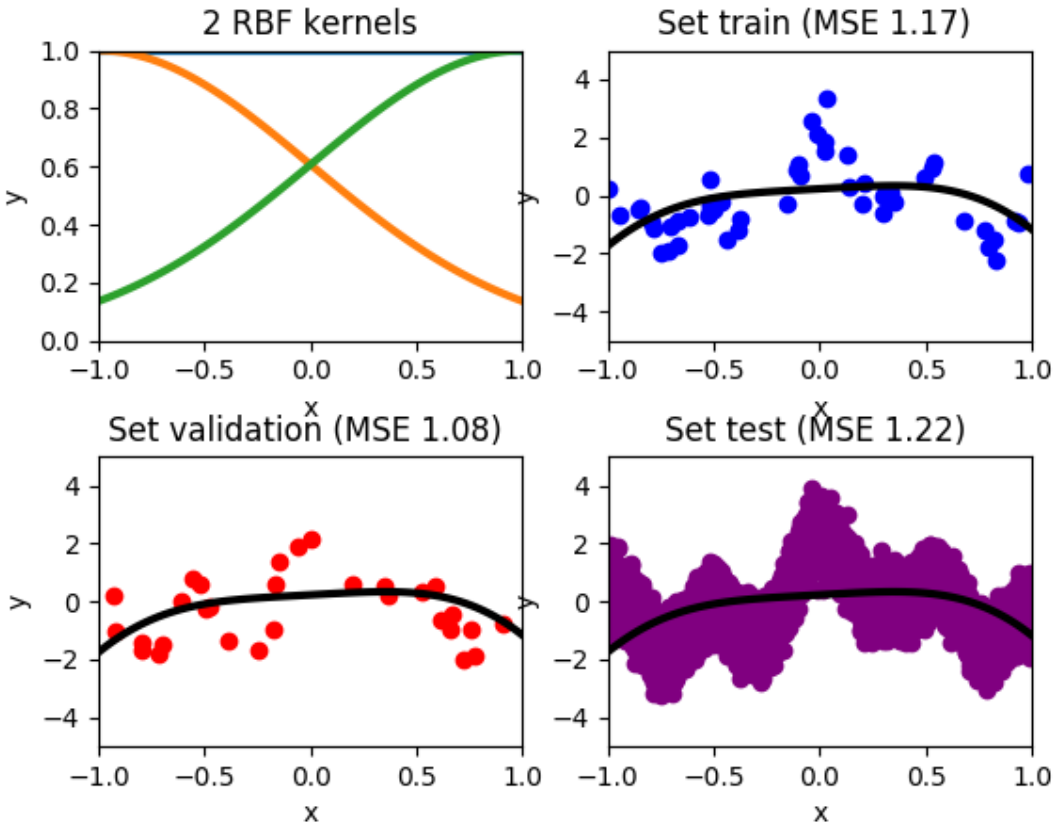

Figure 10: Linear Regression (Bias, Center 2)

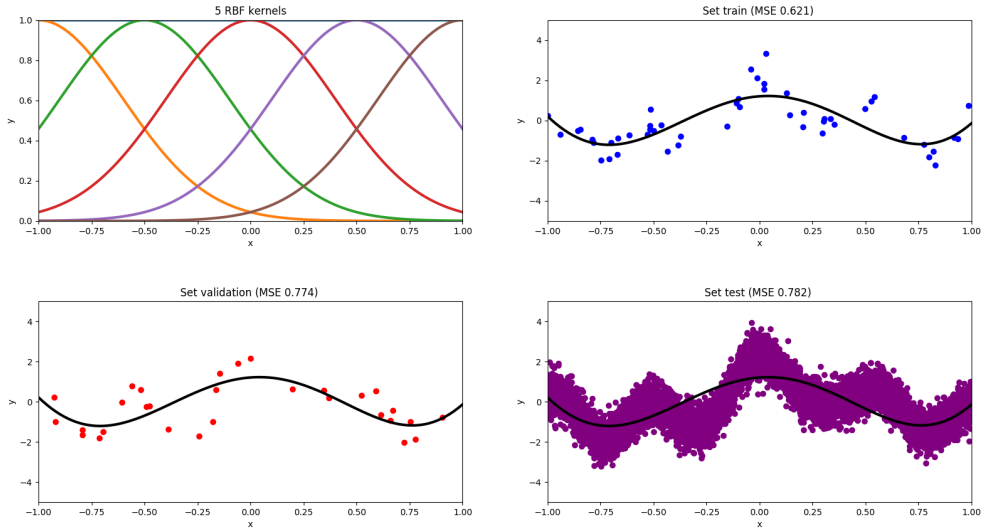

Figure 11: Linear Regression (Bias, Center 5)

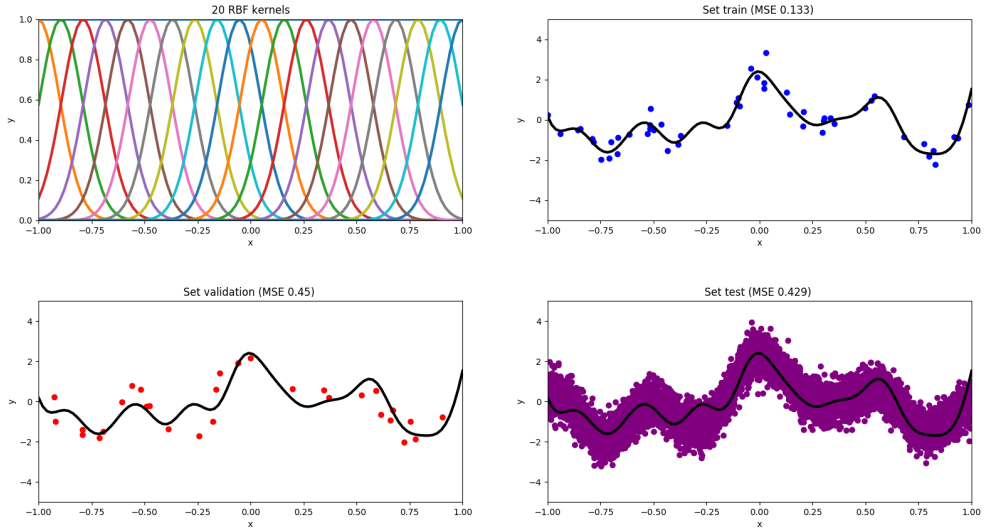

Figure 12: Linear Regression (Bias, Center 20)

- Lowest training error when using center 40

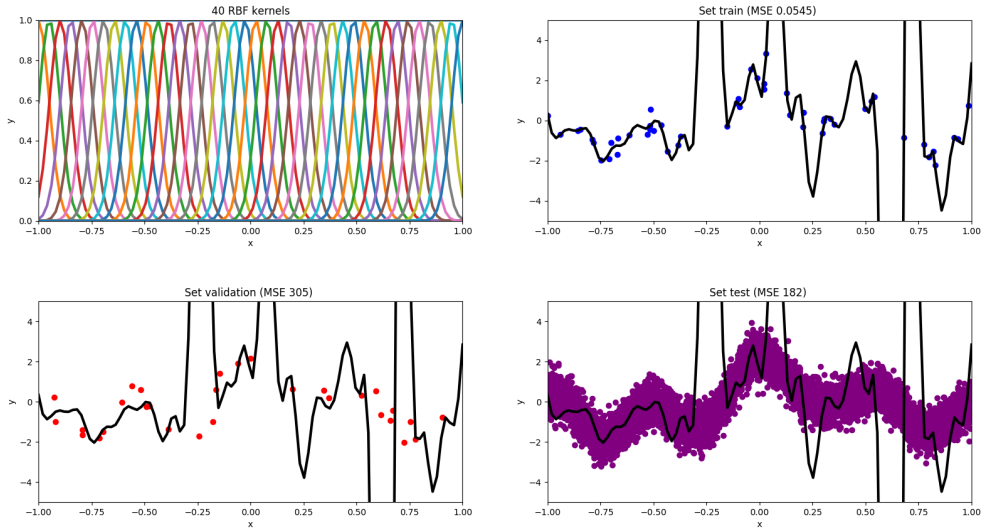

Figure 13: Linear Regression (Bias, Center 40)

- Lowest validation error occurs when using center 9

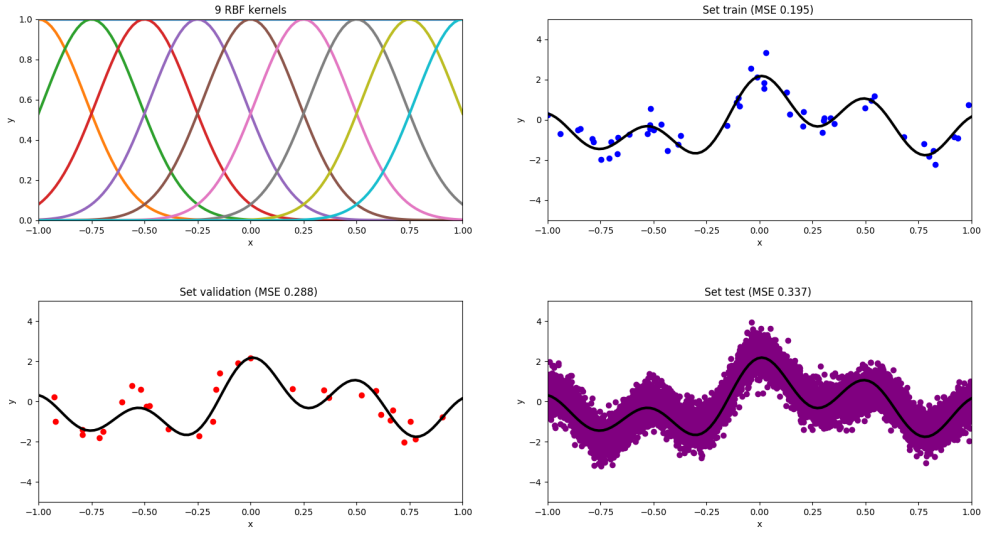

Figure 14: Linear Regression (Polynomial, Degree 9)

- **Discussion**

  Bias function is better because it fits natural phenomen better. Overfitting occurs very early on parameter center 10.

2 Logistic Regression

2.1 Derivation of Gradient

2.2 Logistic Regression training with gradient descent and scipy.optimize

2.2.1 Gradient descent

1. **check_gradient** explaination

   The function check whether the regression functions are really converging at a certain rate. To avoid divergence ;)

2. **gradient descent** degree $l = 1$, 20 and 2000 iterations, learning rate $\eta = 1$

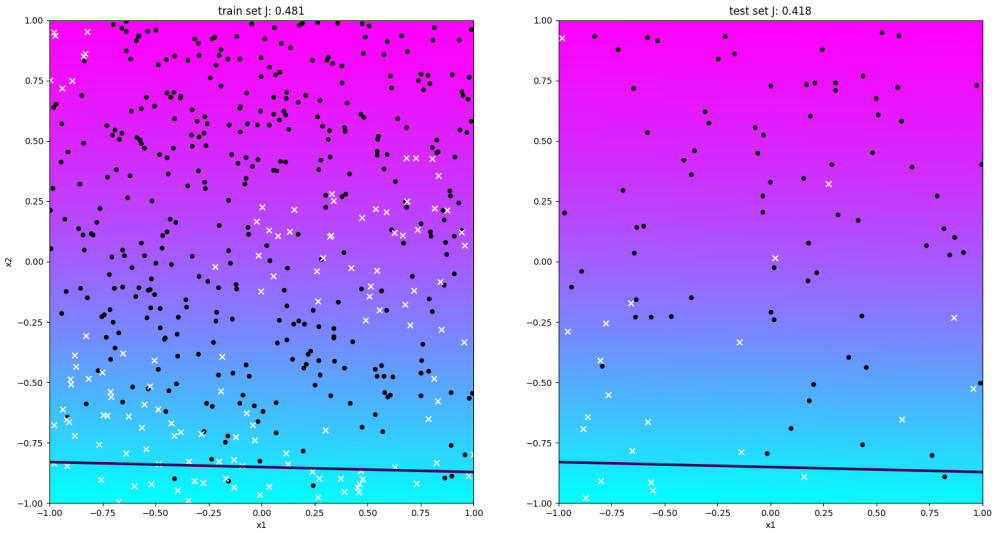

Figure 15: Logistic Regression ($\eta = 1$, $l = 1$, 20 iterations)

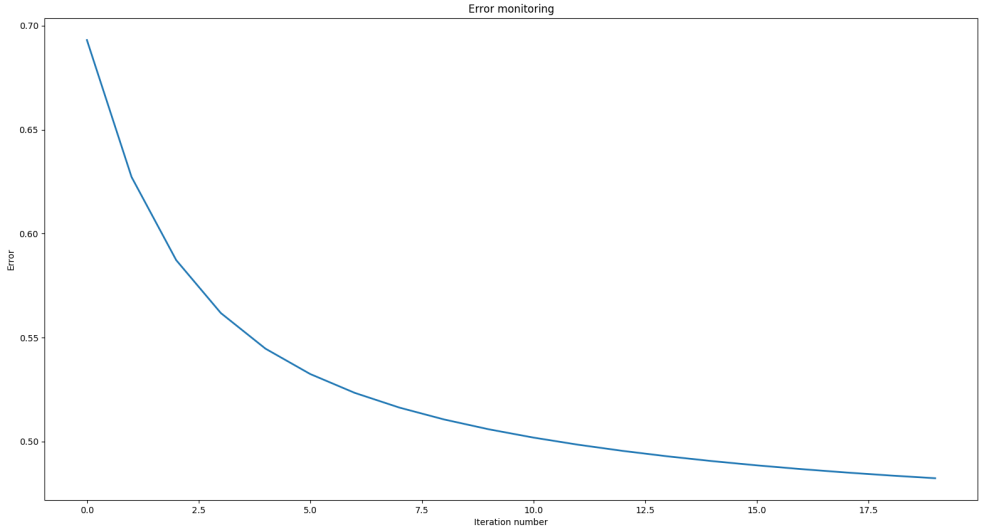

Figure 16: Logistic Regression Errors ($\eta = 1$, $l = 1$, 20 iterations)

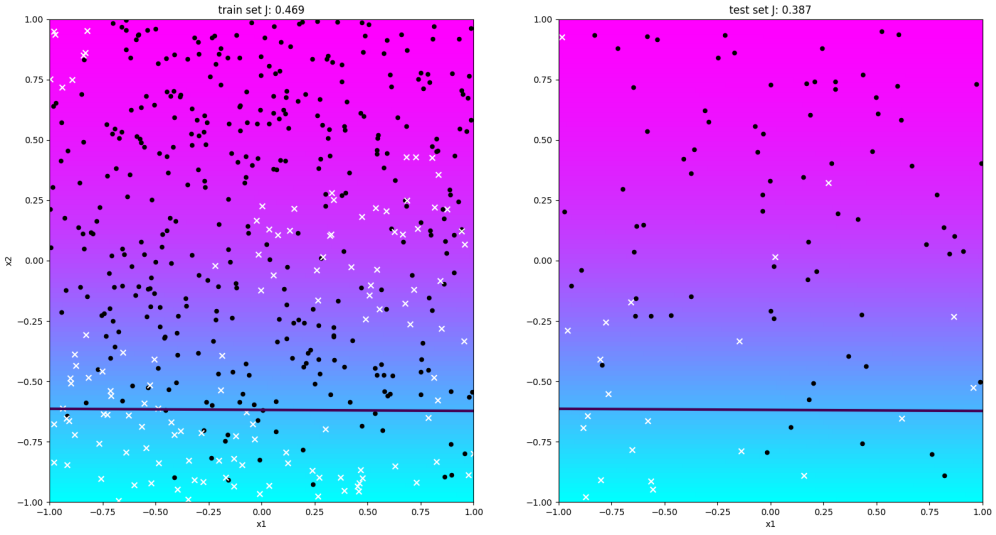

Figure 17: Logistic Regression ($\eta = 1$, $l = 1$, 2000 iterations)

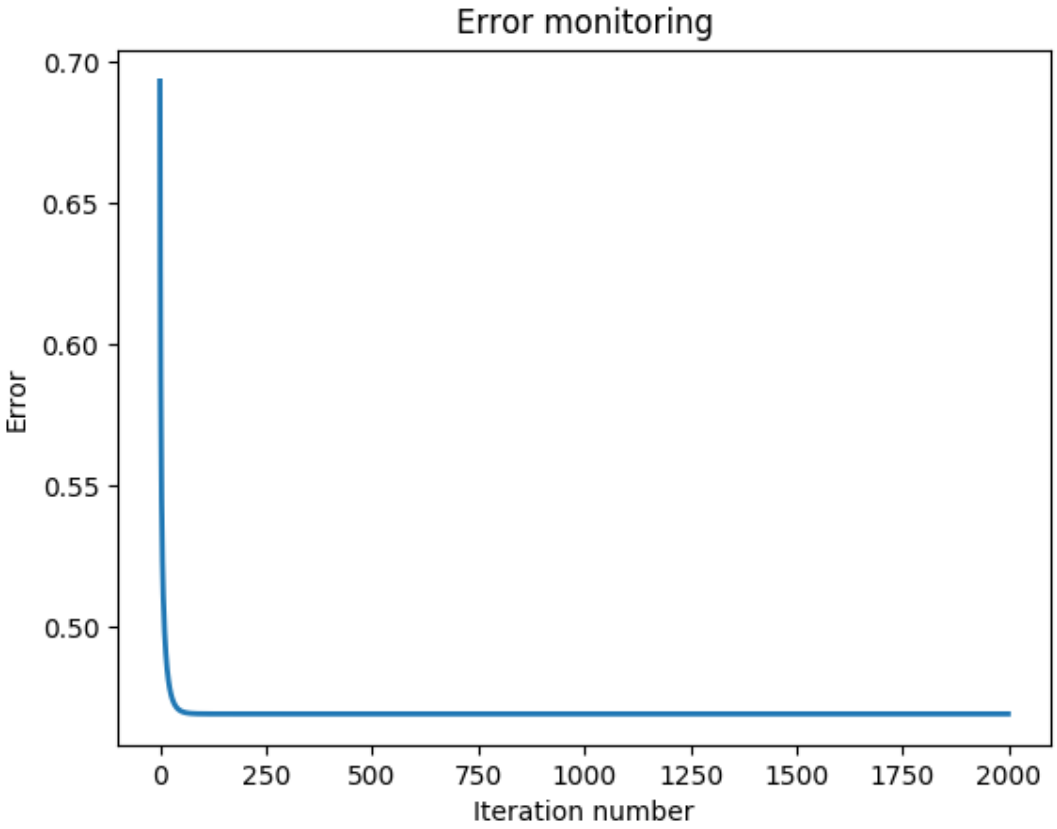

Figure 18: Logistic Regression Errors ($\eta = 1$, $l = 1$, 2000 iterations)

3. **gradient descent** degree $l = 2$, 200 iterations, learning rate $\eta = 0.15, 1.5, 15$

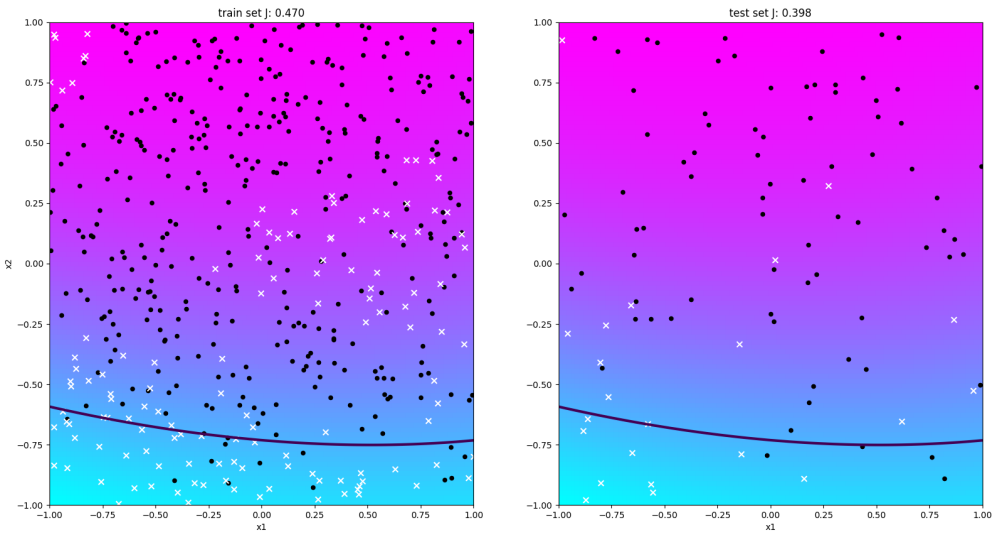

Figure 19: Logistic Regression ($\eta = 0.15$, $l = 1$, 200 iterations)

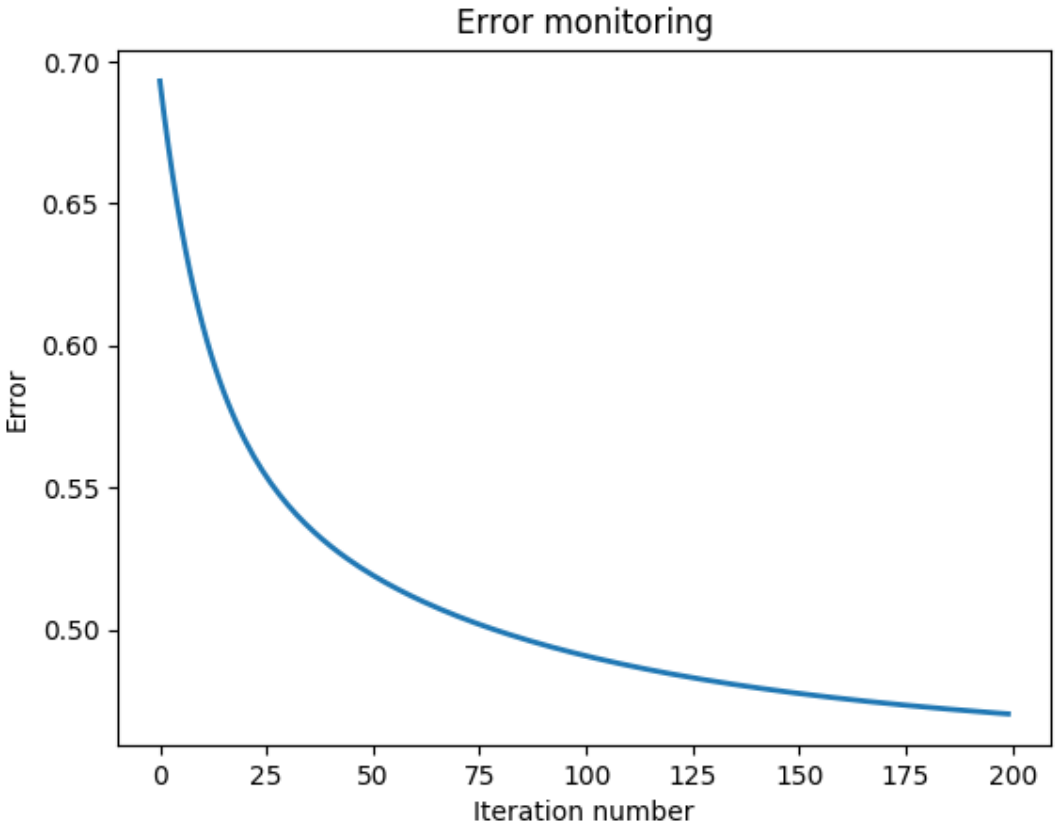

Figure 20: Logistic Regression Errors ($\eta = 0.15$, $l = 1$, 200 iterations)

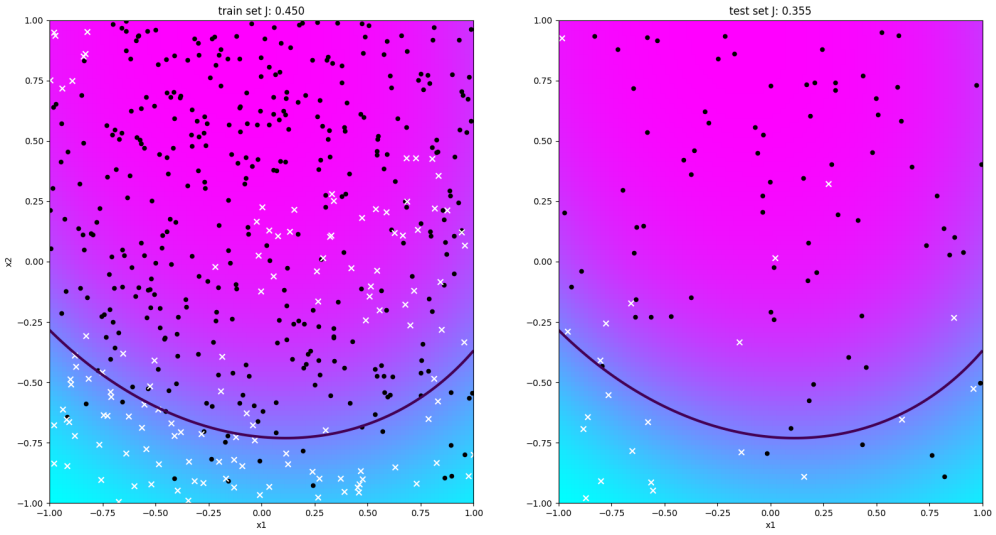

Figure 21: Logistic Regression ($\eta = 1.5$, $l = 1$, 200 iterations)

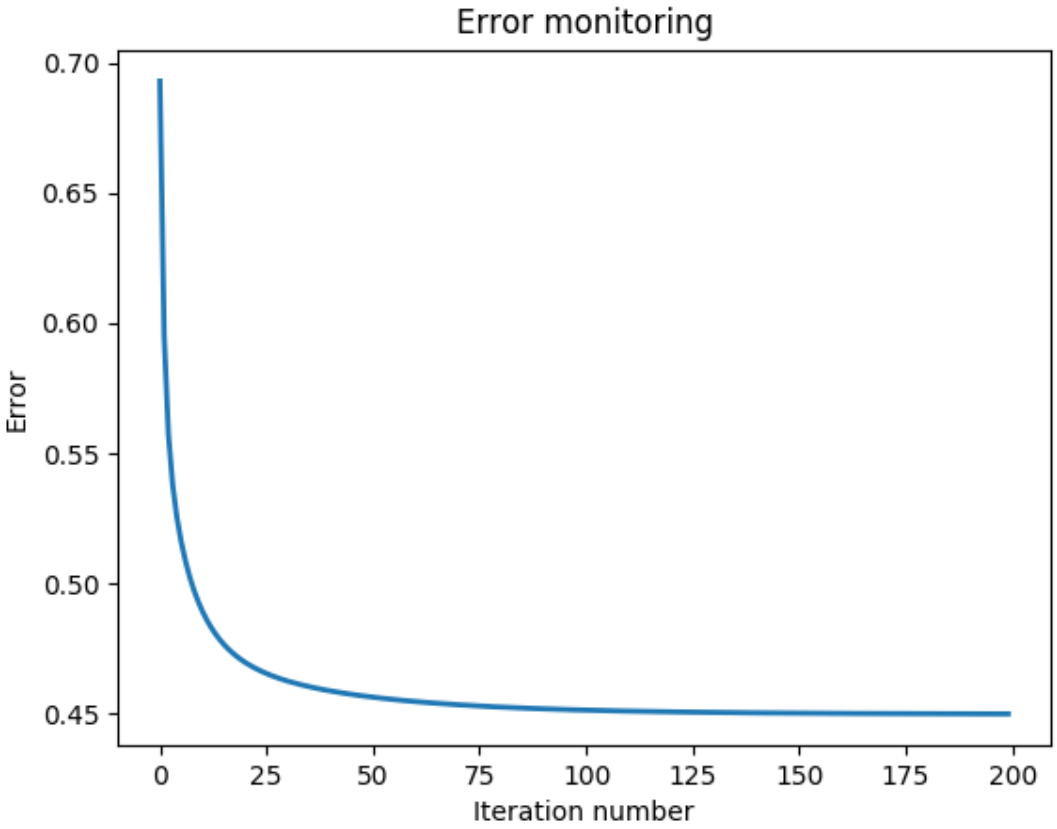

Figure 22: Logistic Regression Errors ($\eta = 1.5$, $l = 1$, 200 iterations)

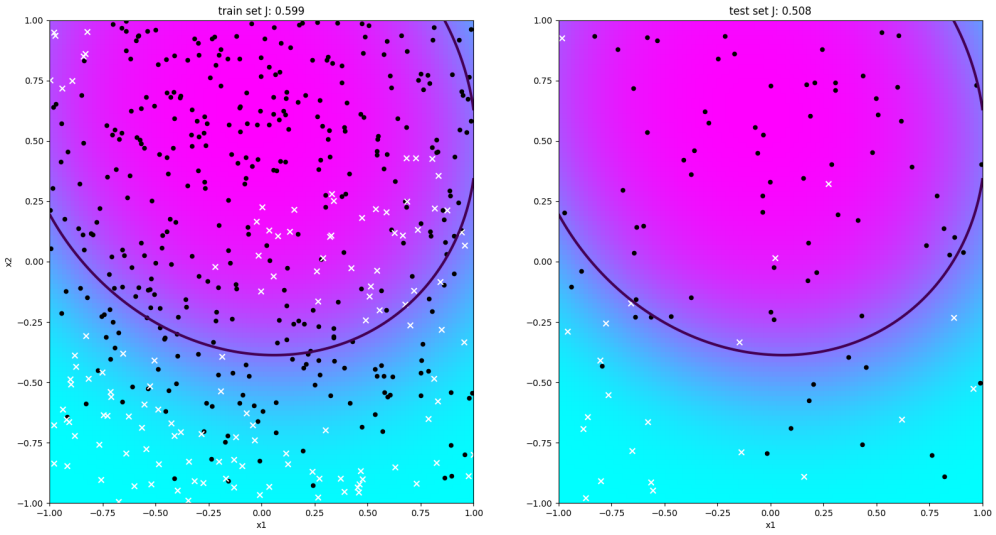

Figure 23: Logistic Regression ($\eta = 15$, $l = 1$, 200 iterations)

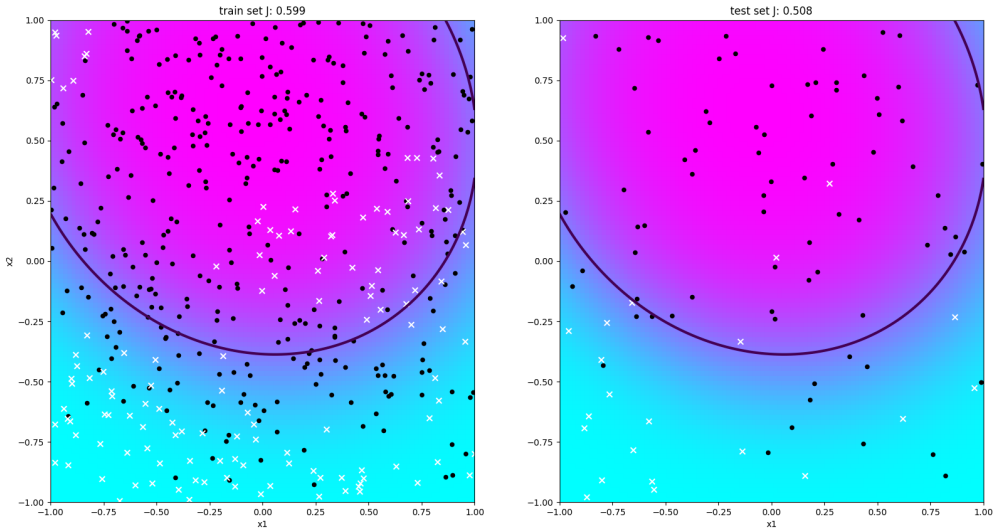

Figure 24: Logistic Regression Errors ($\eta = 15$, $l = 1$, 200 iterations)

Discussion: Too low or too hight learning rates lead to divergence or spinning between lower and hight cost (oscillates).

4. **Adaptative gradient descent (GDad)** degree $l = 1, 2, 5, 15$, 1000 iterations, learning rate $\eta = 1$

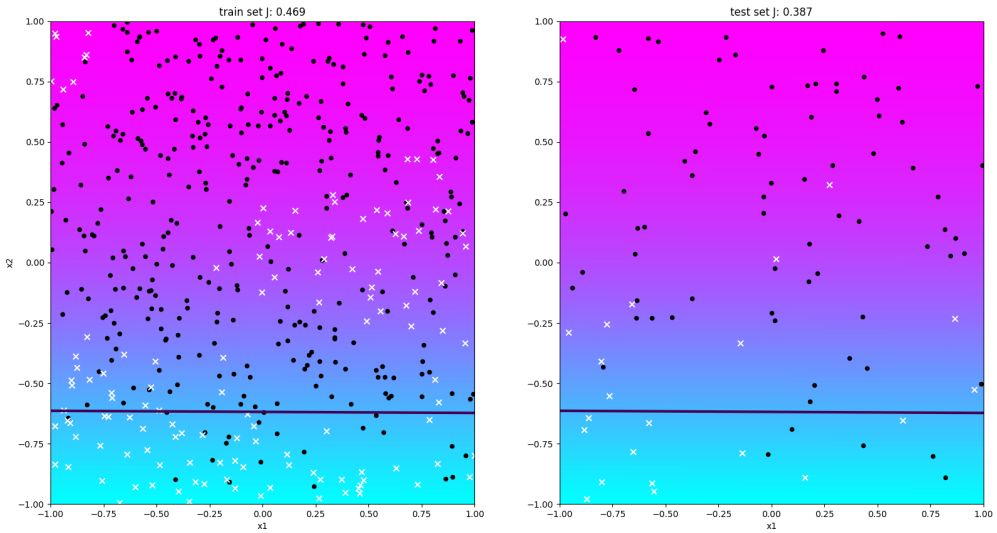

Figure 25: Logistic Regression (adaptive) ($\eta = 1$, $l = 1$, 1000 iterations)

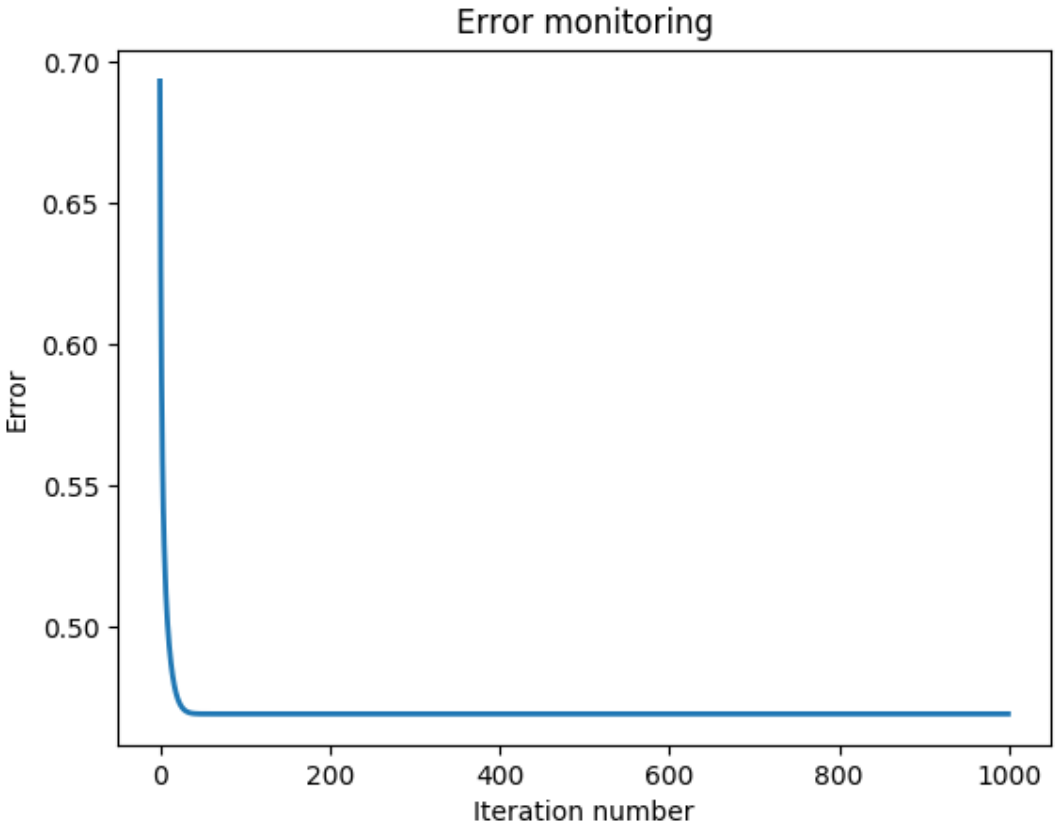

Figure 26: Logistic Regression (adaptive) Errors ($\eta = 1$, $l = 1$, 1000 iterations)

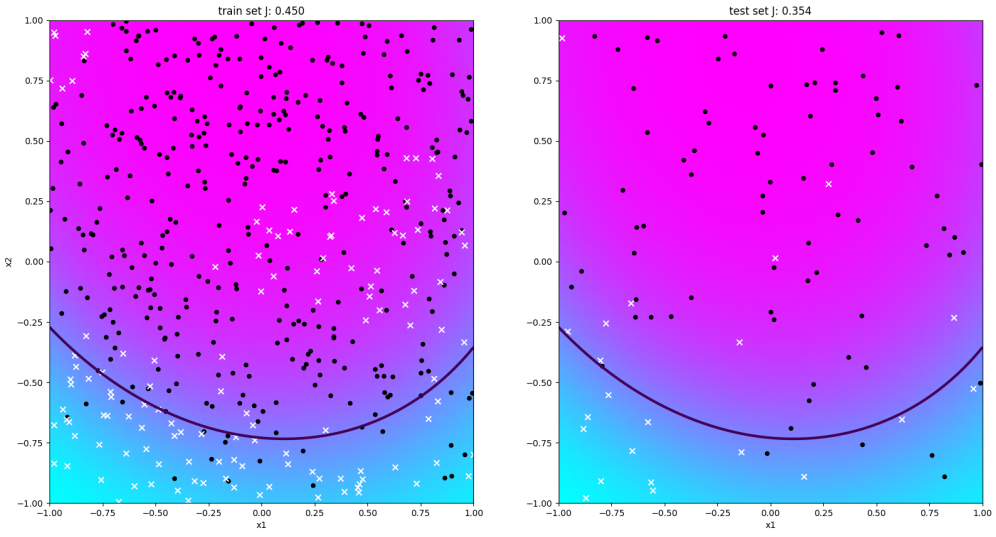

Figure 27: Logistic Regression (adaptive) ($\eta = 1$, $l = 2$, 1000 iterations)

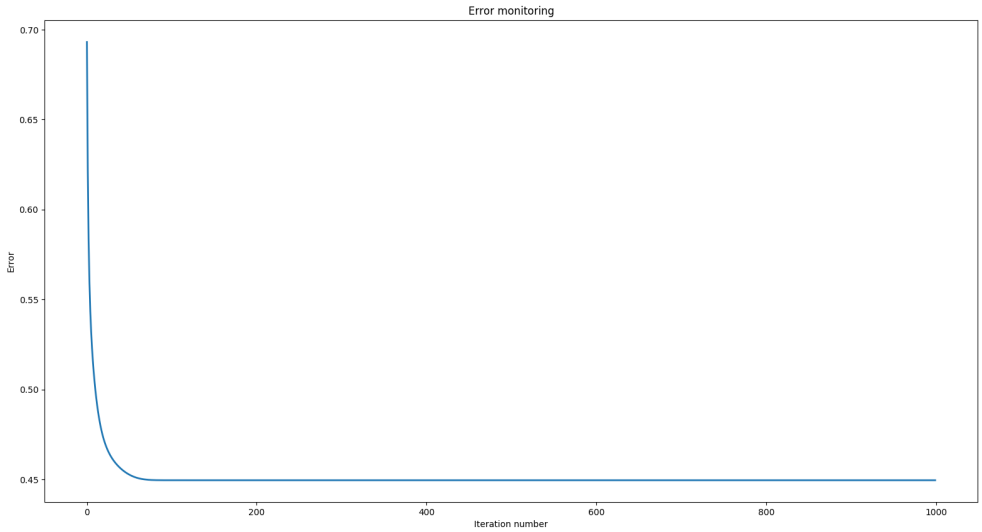

Figure 28: Logistic Regression (adaptive) Errors ($\eta = 1$, $l = 2$, 1000 iterations)

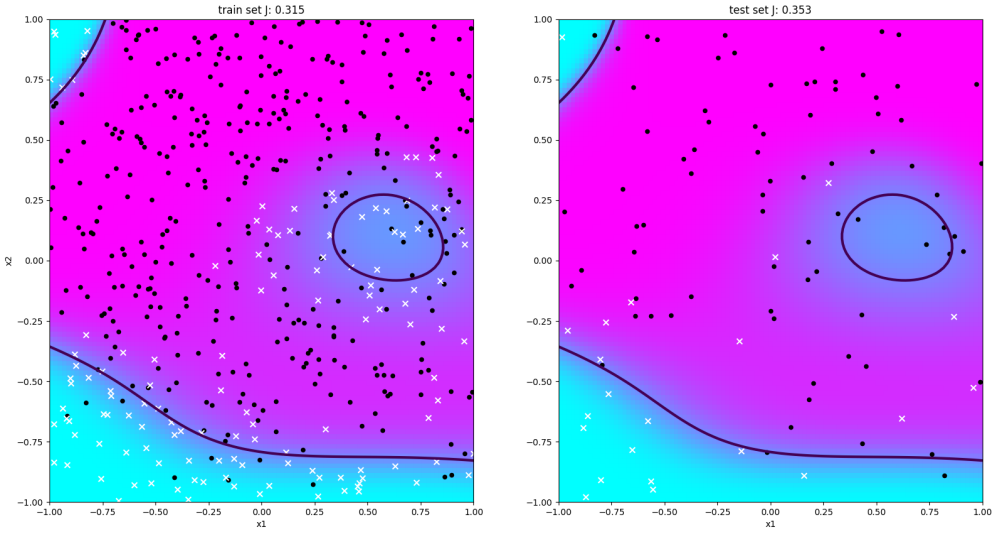

Figure 29: Logistic Regression (adaptive) ($\eta = 1$, $l = 5$, 1000 iterations)

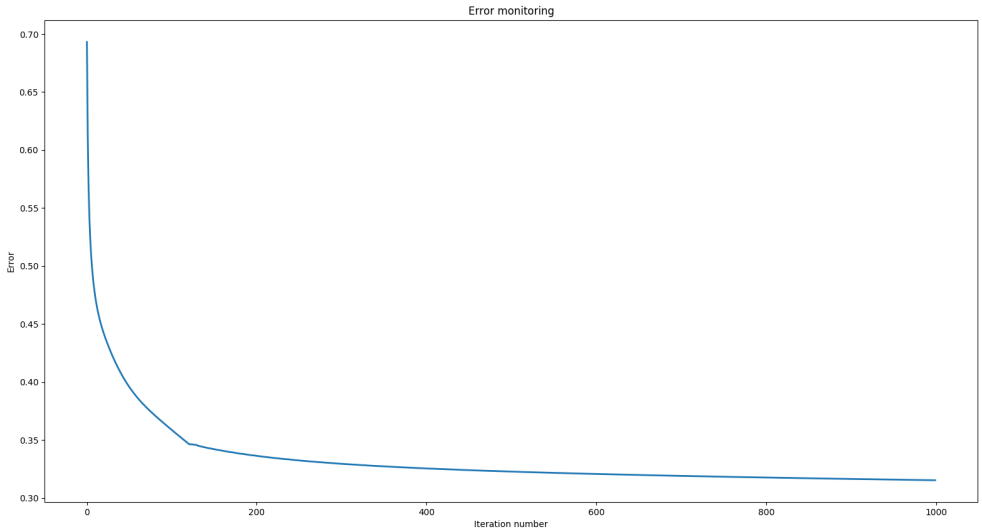

Figure 30: Logistic Regression (adaptive) Errors ($\eta = 1$, $l = 5$, 1000 iterations)

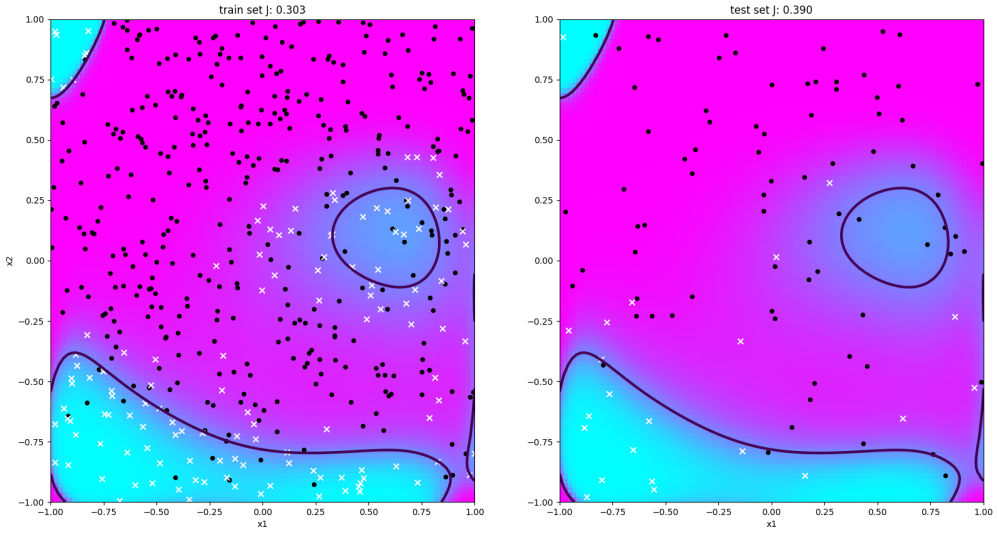

Figure 31: Logistic Regression (adaptive) ($\eta = 1$, $l = 15$, 1000 iterations)

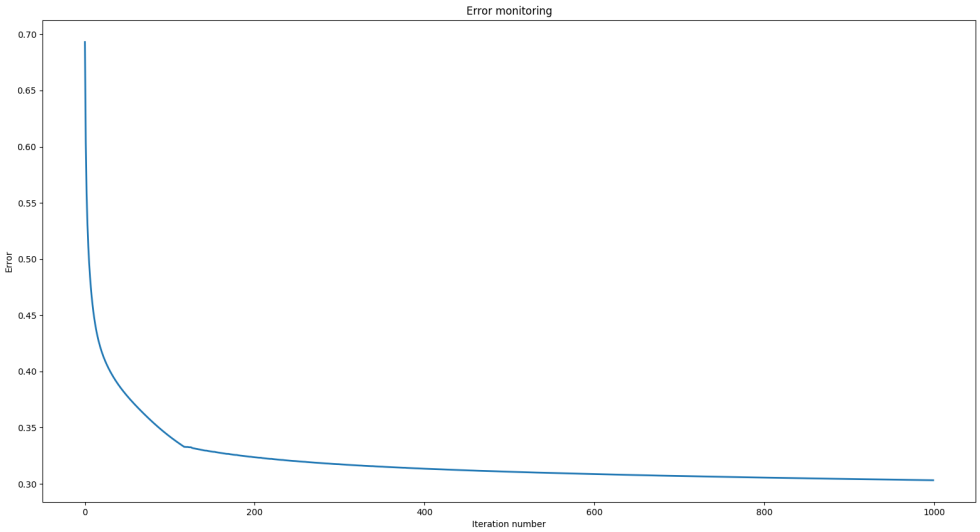

Figure 32: Logistic Regression (adaptive) Errors ($\eta = 1$, $l = 15$, 1000 iterations)

5. Stopping When error between iteration becomes too low, threshold regression should be stopped.

2.2.2 Scipy optimizer