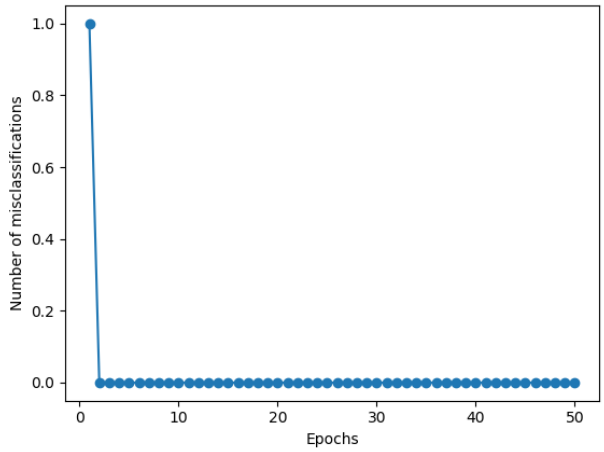
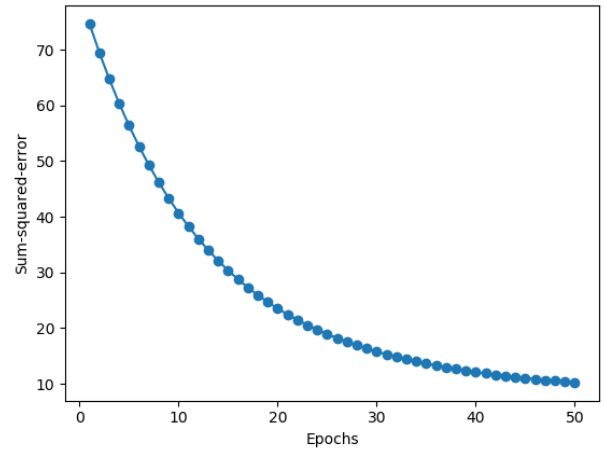
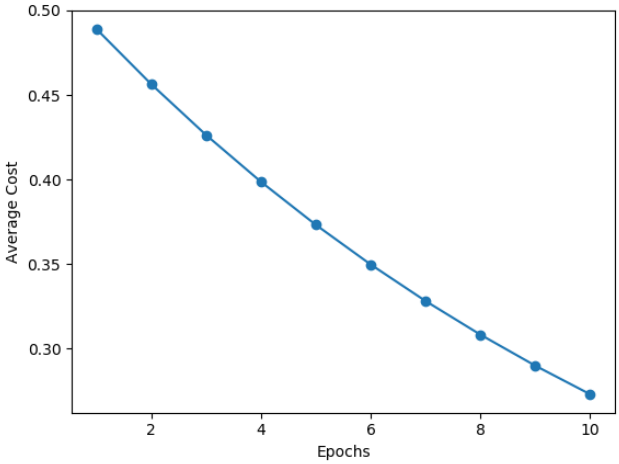
For the iris dataset using perceptron: the final number of errors is: 0



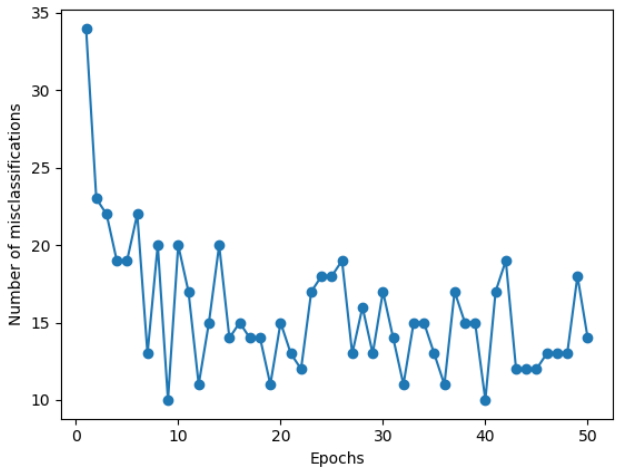
For the iris dataset using adaline: the final sum squared error is: 10.199110233664427



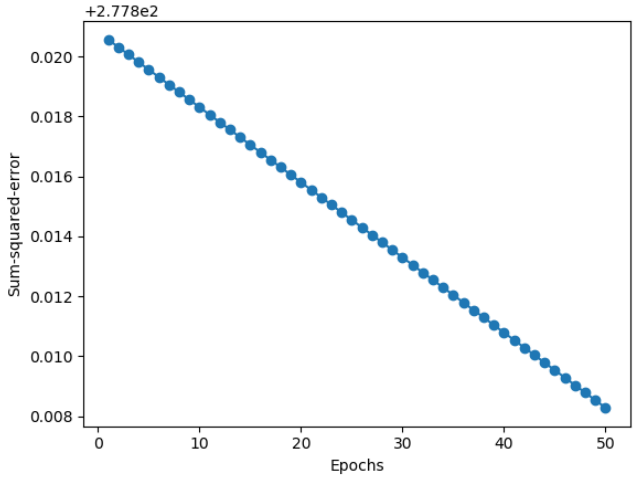
For the iris dataset using sgd: the final Average Cost is: 0.27301753819310526



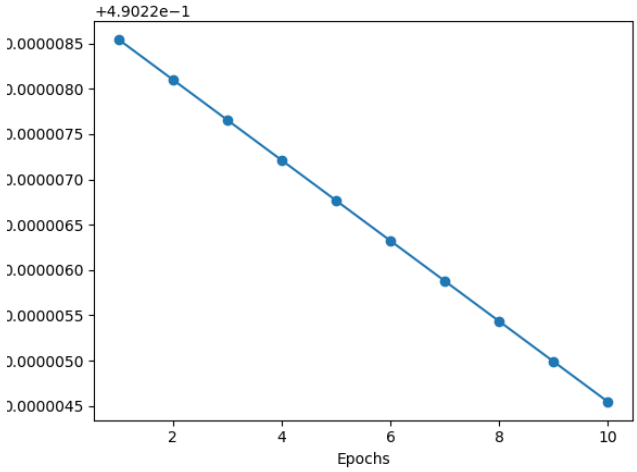
For the BreastCancer dataset using perceptron: the final number of errors is: 14



For the BreastCancer dataset using adaline: the final sum squared error is: 277.80828023320987



For the BreastCancer dataset using sgd: the final Average Cost is: 0.49022454511970454



We see that for the second dataset the Perceptron model doesn't converge but Adaline and SGD do. That's because they compare the true class labels with the continuous valued output to compute the model error and update the weights. In contrast, the perceptron compares the true class labels to the predicted class labels.

Both datasets are normalized and for the perceptron I used 0.01 learning rate, for Adaline and SGD I used 0.0001 and 0.0000000001 learning rates for the iris and Breast Cancer datasets respectively.