1 Overview

Artificial Neural Network (ANN) is implemented in this homework. Base ANN class is defined as abstract class used for template of common methods and attributes for derived classes (i.e. uses, implementations) called "ANNRegression" and "ANNClassification". In the base ANN class we can find core of the logic for fit, initialization, predict, cost and optimization. Also, in this class we define abstract method for activation and loss function and their derivatives. Besides mentioned, derived classes have their separated implementations in the structure of output layer, since we are using ANN in regression and multi - classification problems.

2 Housing2r

Housing2r data is extracted and scaled with sklearn standard scaler. Min-max scaler is tested but standard approach yielded better results. For comparison purposes Kernelized ridge regression is used as additional model with RBF kernel ($\sigma = 0.2$, $\lambda = 0.005$). For evaluation of these models chosen loss function is MSE (Mean Squared Error). In
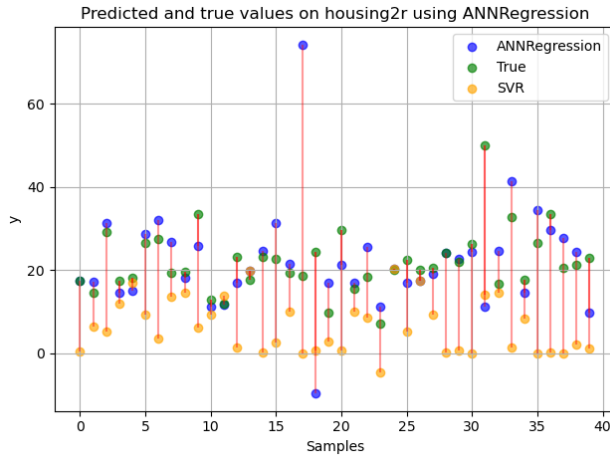

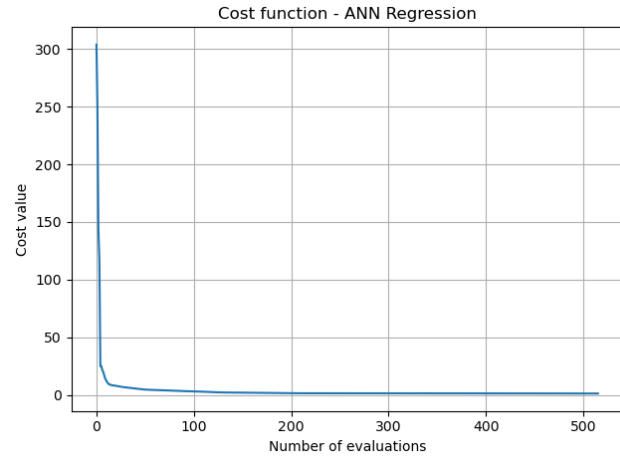Figure 1: Predicted and true values on Housing2r dataset

Figure 2: ANN Regression - Value of cost function through iterations(RMS loss)

Figure 1 predicted values of used models are shown with true values. It is visible that ANN outperforms KRR in most examples. In the Figure 2 cost function is presented. Here we can see that model finds optimum in < 20 with BFGS solver. In the Table 1 ANN model parameters are presented.

ANN regression
$\lambda = 0.0001$
$it = 500$
layers = $5x10x1$

Table 1: Parameters for ANN Regression

3 Housing3r

Housing3r data is extracted and scaled with sklearn standard scaler. Min-max scaler is tested but standard approach yielded better results. For comparison purposes Classification Tree is used as additional model with $min_s amples = 20$. Minimum samples in are left in high value so we can more accurately measure performance of these two models, since default behavior for Trees is to predict finite class label without probability.

In the Figure 3 cost function is presented. Here we can see that model optimization path is stable and reaches new stationary state after approx. 200 iterations. For evaluation of these models chosen loss function is cross entropy loss. In the Table 2 ANN model parameters are presented.
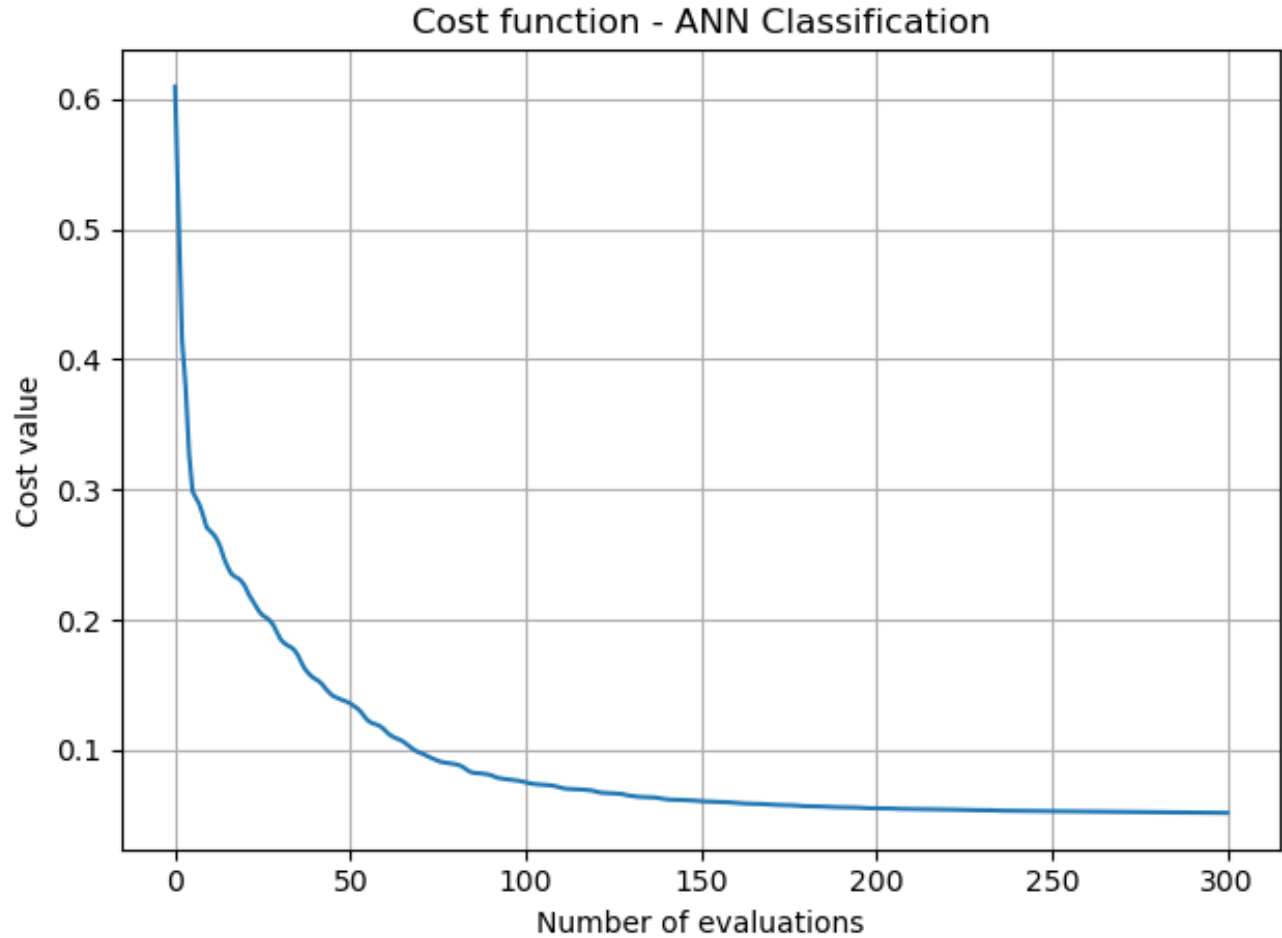
Figure 3: ANN Classification - Value of cost function through iterations(Cross Entropy loss)

ANN classification
$\lambda = 0.02$
$it = 300$
layers $= 9x10x2$

Table 2: Parameters for ANN Classification