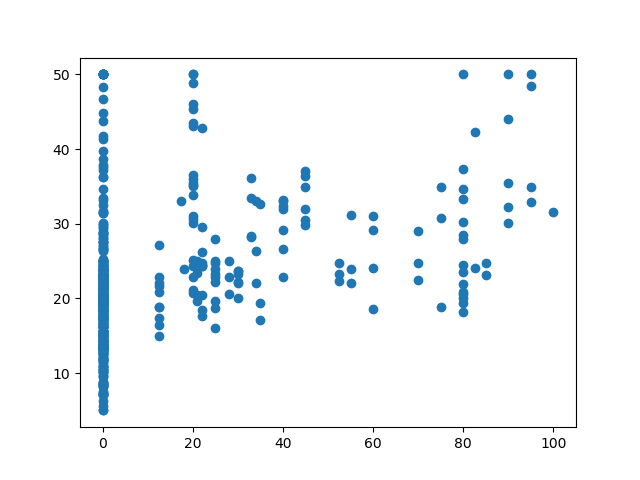
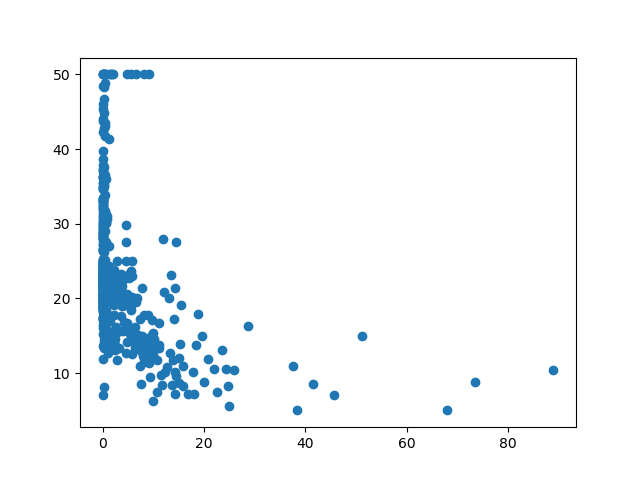
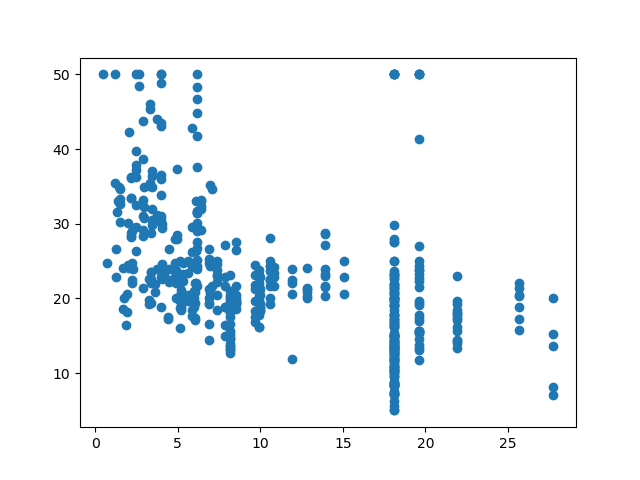
EDA:

Boston dataset has 13 consecutive features and a continuous variable “house price”. Below are the plots of house price on 13 features and the corresponding value shown in order, respectively. The first plot shows the relationship between the first feature and the house price.



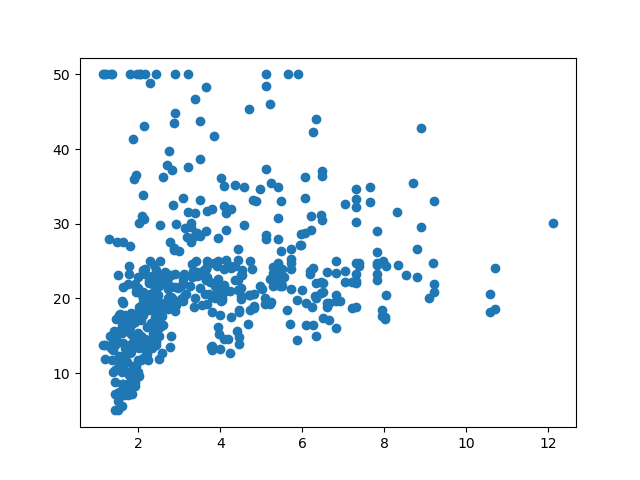
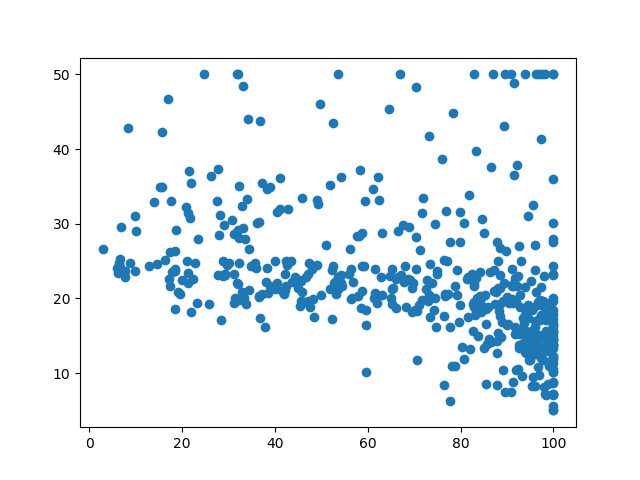
图片包含 屏幕截图

描述已自动生成

图片包含 屏幕截图

描述已自动生成图片包含 文字, 地图

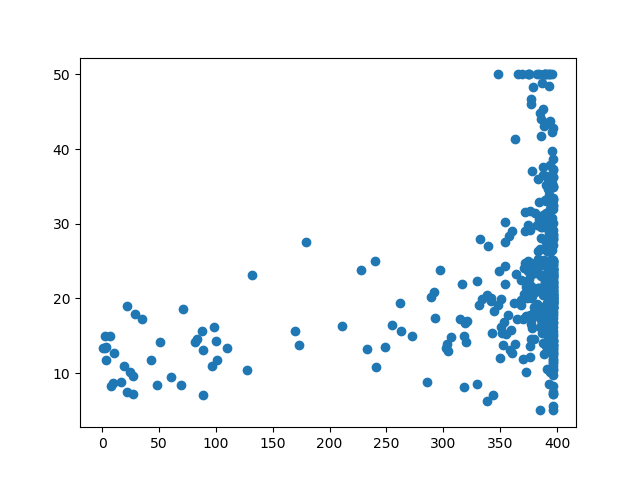
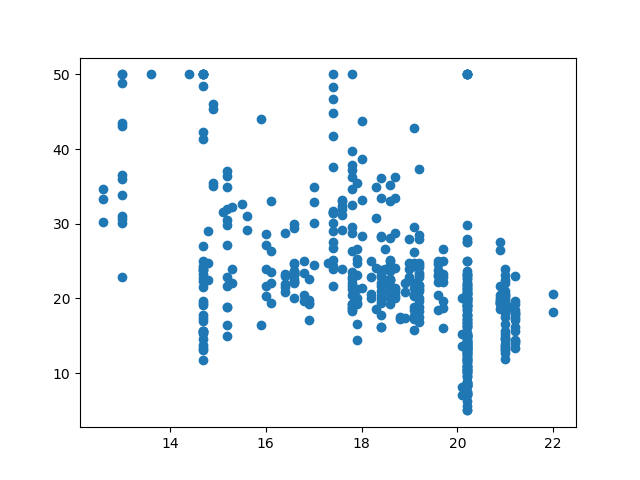
描述已自动生成



图片包含 屏幕截图

描述已自动生成图片包含 屏幕截图

描述已自动生成



图片包含 文字, 地图

描述已自动生成

From the above plots, it is clear that feature 6 has strong linearly positive relationship with the value, while feature 13 has strong linearly negative relationship with the value. Also, feature 3, 5 and 7 may have weak negative relationship with the value, and feature 8 has positive but not linear relationship with the value.

Network Structure:

The network has 3 layers, each of which is a full connected layer.

The first layer directly takes data saved in array as input and maps each 13-dimensional data to the hidden layer of dimension 30, followed by a ReLU layer as activation function.

The second layer is a hidden layer, whose input size is 30 and output size is 10. All data mapped from the first layer to this hidden layer will be transformed to a 10-dimensional data. Finally, same with the first layer, a ReLU activation function is added at the end.

The third layer is the output layer, which takes data from the hidden layer as input, and use a linear function with bias to predict the final value.

The loss function is L2 loss, the optimizer is the stochastic gradient descent optimizer, and the learning rate is 0.005.

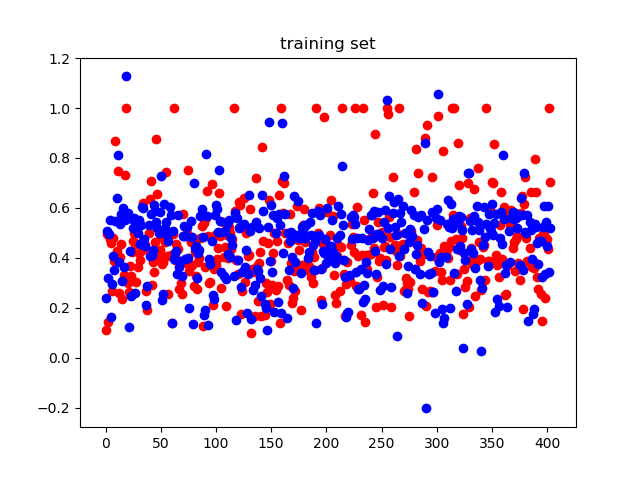
Evaluation:

After 100 iterations, the loss function is almost convergent. The initial loss on the validation set is about 0.577 and the final loss is about 0.061 (shown in pycharm console).

图片包含 屏幕截图

描述已自动生成

Below is the scatter plot of true values and prediction on the training set. Red points are the true values and the blue points are the predicted values. From the scatter plot, we can conclude the model fits well on the training set.



Below is the scatter plot of true values and prediction on the validation set. Same with above, red points are the true values and blue points are the prediction. From the scatter plot, we can also conclude the model fits well on the validation set.

