Assignment 7

Jenni Simon
09-116-005

Exercise 1

a) I assumed that the regression models are supposed to be linear, so I chose the single linear regression model $PRP \sim MMAX$ as in assignment 4 and the multiple linear regression model $PRP \sim MMAX + CACH + MMIN + CHMAX + MYCT$ as found in assignment 5. Figure 1 shows the MSR per fold of the 10-fold cross validation. Note that the folds are identical for both models.

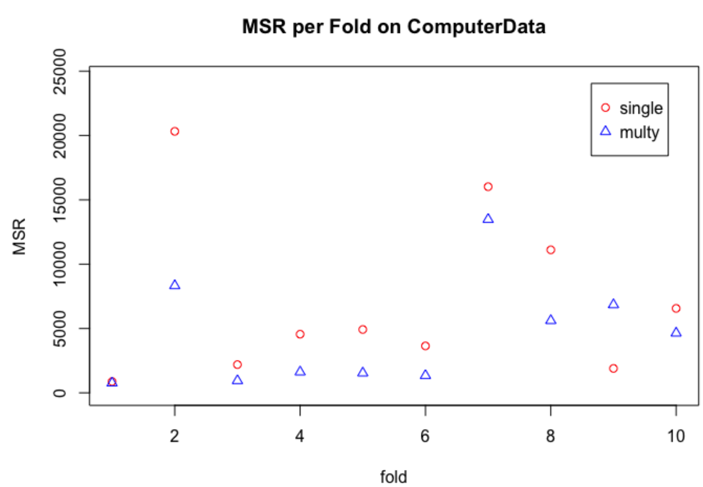

Fig. 1: Mean squared residuals on the ten cross-validation test-sets from the Computer dataset .

To compare the models the two-sided paired t-test is performed. The results shown in Listing 1 suggest that multiple regression performs better given the positive mean difference and relatively low p-value. However, at a significance level of 95% the Null-hypothesis that there is no difference in means can not be rejected.

Listing 1: Results of paired t-test single vs. multiple regression.

```
        Paired t-test

data:  err.single and err.multy
t = 2.0074, df = 9, p-value = 0.07564
alternative hypothesis: difference in means is not equal to 0
95 percent confidence interval:
 -342.6602 5743.2296
sample estimates:
mean of the differences
               2700.285
```

b) To find the best value for k, 10-fold cross-validation has been performed. The resulting average MSR for different k's are shown in Figure 2. A value of $k = 1$ clearly gives the best performance in this setting.

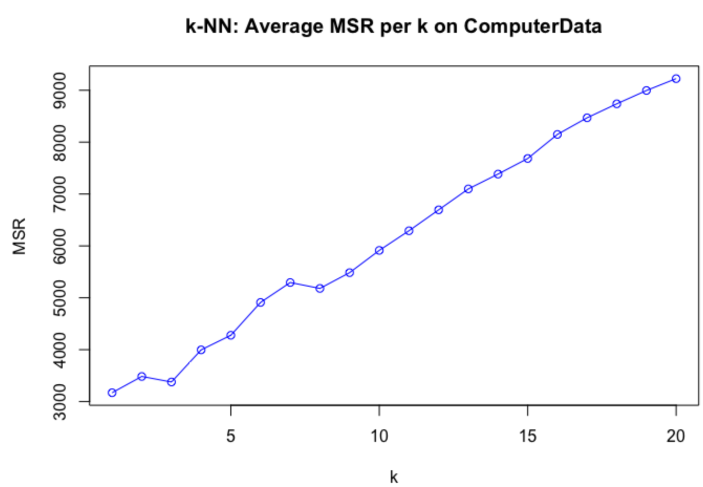

Fig. 2: Average mean squared residuals on the ten cross-validation test-sets from the Computer dataset for different values of k.

c) Figure 3 compares the MSR for 1-NN and the multiple regression model $PRP \sim MMAX + CACH + MMIN + CHMAX + MYCT$ on identical cross-validation folds.

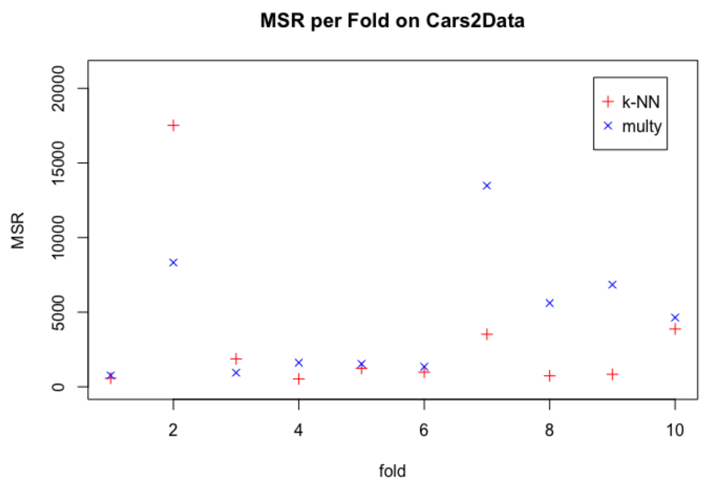

Fig. 3: Mean squared residuals on the ten cross-validation test-sets from the Computer dataset.

The results of the two-sided paired t-test are shown in Listing 2. Given the high p-value the Null-hypothesis can not be rejected and we can not asses that one model performs better than the other.

In this case I would favour the multiple linear regression model for its simplicity and better interpretability. Linear models also generalise better with little training data and have a much lower bias.

Listing 2: Results of paired t-test 1-NN vs. multiple regression.

```
Paired t-test

data:  err.knn and err.multy
t = -0.8421, df = 9, p-value = 0.4215
alternative hypothesis: difference in means is not equal to 0
95 percent confidence interval:
 -4944.528  2261.889
sample estimates:
mean of the differences
             -1341.32
```

Exercise 2

a) I chose the single linear regression model $mpg \sim weight$ as in assignment 4 and the multiple linear regression model $mpg \sim weight + year$ as found in assignment 5. Figure 4 shows the MSR per fold of the 10-fold cross validation.

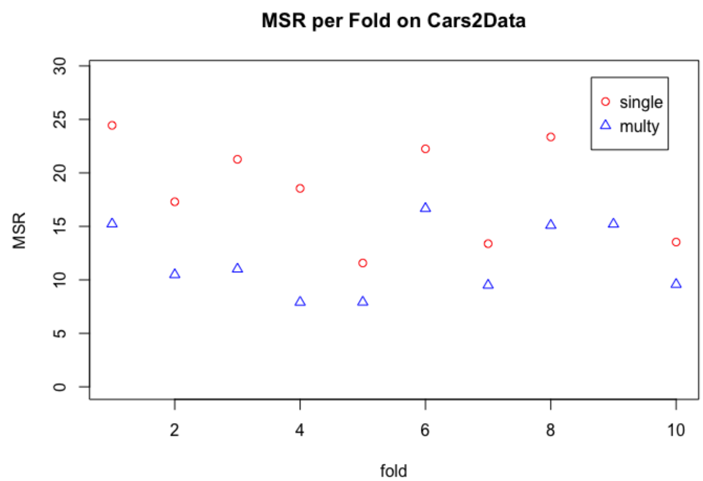

Fig. 4: Mean squared residuals on the ten cross-validation test-sets from the Cars dataset .

To compare the models the two-sided paired t-test is performed. The results shown in Listing 3 suggest that multiple regression performs significantly better given the very low p-value.

Listing 3: Results of paired t-test single vs. multiple regression.

```
        Paired t-test

data:  err.single and err.multy
t = 8.3581, df = 9, p-value = 1.558e-05
alternative hypothesis: difference in means is not equal to 0
95 percent confidence interval:
 5.134794 8.945783
sample estimates:
mean of the differences
               7.040289
```

b) To find the best value for k, 10-fold cross-validation has been performed. The resulting average MSR for different k's are shown in Figure 5. A value of $k = 2$ clearly gives the best performance in this setting.

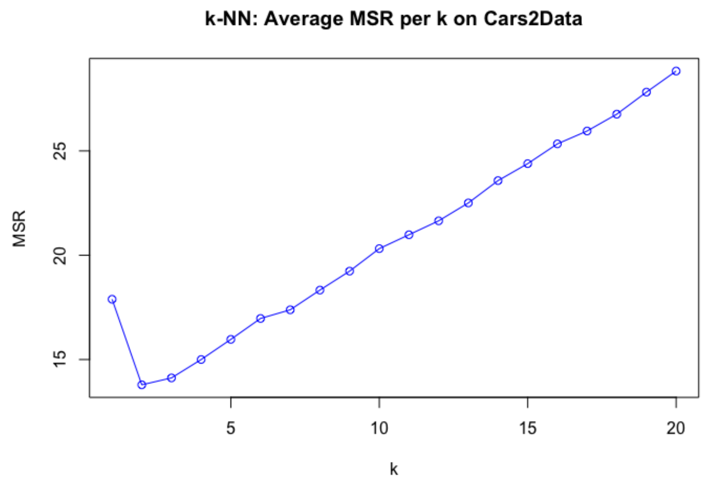

Fig. 5: Average mean squared residuals on the ten cross-validation test-sets from the Cars dataset for different values of k.

c) Figure 6 compares the MSR for 2-NN and the multiple regression model $mpg \sim weight + year$ on identical cross-validation folds.

The results of the two-sided paired t-test are shown in Listing 4. Given the small p-value I would favour the 2-NN model in this case.

Listing 4: Results of paired t-test 1-NN vs. multiple regression.

```
        Paired t-test

data:  err.knn and err.multy
t = 1.9265, df = 9, p-value = 0.08616
alternative hypothesis: difference in means is not equal to 0
95 percent confidence interval:
 -550.3705 6867.0413
sample estimates:
mean of the differences
               3158.335
```

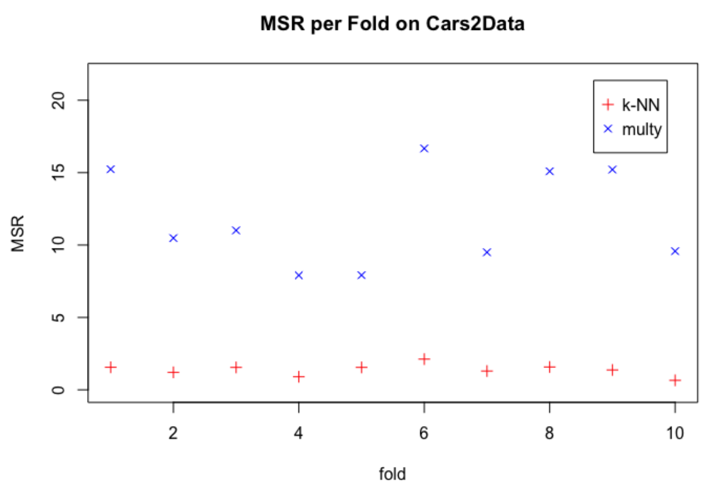

Fig. 6: Mean squared residuals on the ten cross-validation test-sets from the Cars dataset.