CS235 Fall'23 Project Implementation Correctness Report: Use Deep Learning to Predict Car Sales Price

LIAM HSIEH (NETID:LHSIE013)

1 DEEP NEURAL NETWORKS

1.1 Cross Validation

This implemented method is multi-layer Neural Network and network models with different layers and setting are evaluated.

model	layers	neurons	activation
1	2	5-5-15-1	linear
2	2	5-5-15-1	relu
3	3	5-5-10-5-1	linear
4	3	5-5-10-5-1	relu
5	4	5-5-15-15-5-1	linear
6	4	5-5-15-15-5-1	relu
7	5	5-5-15-20-15-5-1	linear
8	5	5-5-15-20-15-5-1	relu

Table 1. All prediction models candidates

Table 1 provides an overview of candidate prediction models with different architectures.

- Layers: indicates the number of hidden layers in each model.
- Neurons: specifies the number of neurons for each layer, following the format that denotes the configuration of hidden layers.
- Activation: indicates the activation function used in each model, which can be either Linear or ReLU.
- The input and output dimensions are consistent across all models, with the input representing features and the output being a single response.

We use mean absolute error (MAE) for the cross validation that the absolute value of the error between the predicted values and the actual values is averaged over all testing data set as the performance measure. The number of replications for each ANN is 10. To assess the impact of adding more hidden layers to the Artificial Neural Network (ANN) architecture, we conducted an analysis of variance (ANOVA) test followed by Tukey's Honest Significant Difference (HSD) test. The primary objective was to determine whether there were statistically significant differences in the performance metrics as we varied the number of hidden layers. Basically, the ANOVA test was employed to evaluate the overall differences in mean performance across different configurations of hidden layers. MAE was used to quantify the effectiveness of each ANN configuration. To further investigate specific differences between pairs of hidden layer configurations identified by the ANOVA test, we applied Tukey's HSD test. This post hoc test allows for a rigorous pairwise comparison while accounting for the multiple comparisons problem.

1.2 Experimental results

The results of ANOVA is that we got 0.0319 as p-value, and the result of Tukey HSD is as follows:

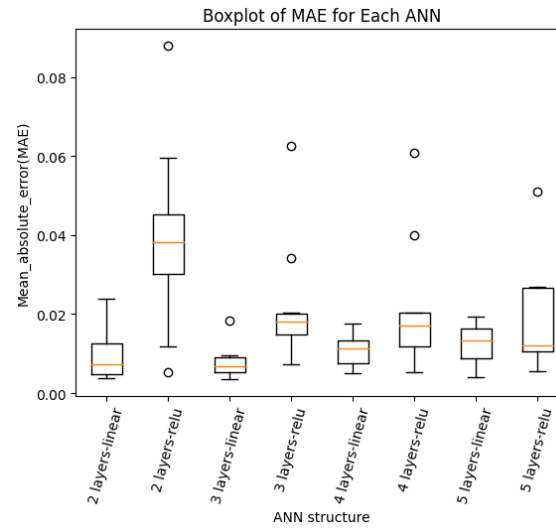| group1	group2	meandiff	p-adj	lower	upper	reject
2 layers-linear	2 layers-relu	-0.0099	0.97	-0.051	0.0312	False
2 layers-linear	3 layers-linear	-0.0085	0.9869	-0.0496	0.0326	False
2 layers-linear	3 layers-relu	-0.009	0.9819	-0.0501	0.0321	False
2 layers-linear	4 layers-linear	-0.01	0.9676	-0.0511	0.031	False
2 layers-linear	4 layers-relu	-0.0152	0.809	-0.0562	0.0259	False
2 layers-linear	5 layers-linear	-0.0147	0.8267	-0.0558	0.0263	False
2 layers-linear	5 layers-relu	-0.0156	0.7881	-0.0567	0.0255	False
2 layers-relu	3 layers-linear	0.0014	1.0	-0.0397	0.0425	False
2 layers-relu	3 layers-relu	0.0009	1.0	-0.0402	0.042	False
2 layers-relu	4 layers-linear	-0.0002	1.0	-0.0412	0.0409	False
2 layers-relu	4 layers-relu	-0.0053	0.9992	-0.0464	0.0358	False
2 layers-relu	5 layers-linear	-0.0049	0.9995	-0.0459	0.0362	False
2 layers-relu	5 layers-relu	-0.0057	0.9987	-0.0468	0.0354	False
3 layers-linear	3 layers-relu	-0.0005	1.0	-0.0416	0.0406	False
3 layers-linear	4 layers-linear	-0.0016	1.0	-0.0427	0.0395	False
3 layers-linear	4 layers-relu	-0.0067	0.9966	-0.0478	0.0344	False
3 layers-linear	5 layers-linear	-0.0063	0.9977	-0.0474	0.0348	False
3 layers-linear	5 layers-relu	-0.0072	0.9949	-0.0482	0.0339	False
3 layers-relu	4 layers-linear	-0.0011	1.0	-0.0422	0.04	False
3 layers-relu	4 layers-relu	-0.0062	0.9979	-0.0473	0.0349	False
3 layers-relu	5 layers-linear	-0.0058	0.9986	-0.0469	0.0353	False
3 layers-relu	5 layers-relu	-0.0066	0.9967	-0.0477	0.0344	False
4 layers-linear	4 layers-relu	-0.0051	0.9994	-0.0462	0.036	False
4 layers-linear	5 layers-linear	-0.0047	0.9996	-0.0458	0.0364	False
4 layers-linear	5 layers-relu	-0.0056	0.9989	-0.0467	0.0355	False
4 layers-relu	5 layers-linear	0.0004	1.0	-0.0407	0.0415	False
4 layers-relu	5 layers-relu	-0.0005	1.0	-0.0416	0.0406	False
5 layers-linear	5 layers-relu	-0.0009	1.0	-0.042	0.0402	False

Table 2. Multiple Comparison of Means - Tukey HSD, FWER=0.05

In this case, the p-value is less than the conventional significance level of 0.05, suggesting that there is some evidence of a difference in means among the groups. However, based on the Tukey HSD results, no pairs of groups show a significant difference in means. All the "reject" values are marked as "False." Therefore, despite the statistically significant ANOVA p-value, the post-hoc Tukey HSD test suggests that the observed differences in means may not be specifically attributed to differences between individual pairs of groups.
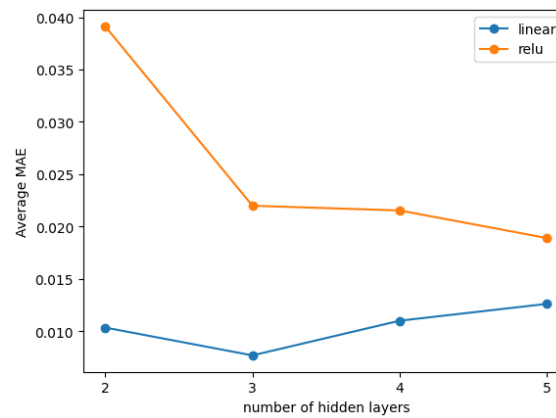
	mean	std. dev
2 layers-linear	0.010360	0.007321
2 layers-relu	0.039120	0.022072
3 layers-linear	0.007703	0.004066
3 layers-relu	0.021993	0.015098
4 layers-linear	0.011014	0.004181
4 layers-relu	0.021542	0.015800
5 layers-linear	0.012623	0.004956
5 layers-relu	0.018909	0.013284

Table 3. Mean and Std. deviation of MAE

Therefore, we summarize the results in Table 3 and also illustrate the results in the Fig. 1a. Although none of them significantly outperforms the others, we finally select 3-layers-linear as our prediction model because it not only has the lowest average MAE (see Fig. 1b) but also has smallest standard deviation of MAE.

157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208

(a) Boxplot of different ANN models

(b) Mean MAE of ANN Models with Varying Number of Hidden Layers

Fig. 1. Comparison of ANN models.

4