Full Name: Lee Taylor, ST Number: 190211479

CSC3831 Predictive Analytics – Machine Learning House Price Prediction

First Model – Linear Regression

A grid search is performed to find the best hyperparameters for the linear regression model on the dataset. The grid search takes in the model type, a dictionary of hyperparameters to search over, and the datasets. The grid search returns a list of results for each set of hyperparameters.

Fit Intercept	Copy X	N jobs	MSE	R^2
TRUE	TRUE	1	0.37206	0.621966
TRUE	TRUE	2	0.37206	0.621966
TRUE	TRUE	3	0.37206	0.621966
TRUE	TRUE	4	0.37206	0.621966
TRUE	FALSE	-1	0.37206	0.621966
TRUE	FALSE	1	0.372047	0.621979
TRUE	FALSE	2	0.372047	0.621979
TRUE	FALSE	3	0.372047	0.621979
TRUE	FALSE	4	0.372047	0.621979
FALSE	TRUE	-1	0.372046	0.621981
FALSE	TRUE	1	0.372046	0.621981
FALSE	TRUE	2	0.372046	0.621981
FALSE	TRUE	3	0.372046	0.621981
FALSE	TRUE	4	0.372046	0.621981
FALSE	FALSE	-1	0.372046	0.621981
FALSE	FALSE	1	0.372046	0.621981
FALSE	FALSE	2	0.372046	0.621981
FALSE	FALSE	3	0.372046	0.621981
FALSE	FALSE	4	0.372046	0.621981

These results suggest that the choice of hyperparameters has a minimal impact on the model's performance. The MSE scores are relatively close together and the R^2 scores are all above 0.6, indicating that the model is performing relatively well, but there is still room for improvement.

Second Model – Multi-Layer Perceptron

Using the TensorFlow Keras library I performed multiple grid searches & cross-validation to optimize the model's hyperparameters. The model is a multi-layer perceptron with the number of units, activation function, number of layers, as hyperparameters. The grid search considers all combinations of these hyperparameters and trains and evaluates the model for each combination.
First grid-search results:

Units Per Layer	Activation Function	No. of Layers	Optimizer	Epochs	Batch Size	MSE	MAE
64	relu	2	adam	10	8	0.224044	0.316482
64	relu	2	adam	10	16	0.212253	0.312662
64	relu	2	adam	10	32	0.221621	0.321429
64	relu	2	adam	10	64	0.228725	0.346533
64	relu	2	adam	10	128	0.219787	0.322156
64	relu	2	adam	10	256	0.22963	0.330435
64	relu	3	adam	10	8	0.22722	0.33222
64	relu	3	adam	10	16	0.21132	0.320985
64	relu	3	adam	10	32	0.220123	0.319673
64	relu	3	adam	10	64	0.212145	0.31642
64	relu	3	adam	10	128	0.227734	0.320977
64	relu	3	adam	10	256	0.225559	0.324945

Full Name: Lee Taylor, ST Number: 190211479

64	relu	4	adam	10	8	0.212145	0.327391
64	relu	4	adam	10	16	0.209875	0.316931
64	relu	4	adam	10	32	0.215377	0.321995
64	relu	4	adam	10	64	0.226254	0.313717
64	relu	4	adam	10	128	0.219476	0.319936
64	relu	4	adam	10	256	0.226489	0.338178
64	relu	5	adam	10	8	0.237961	0.338494
64	relu	5	adam	10	16	0.215081	0.319046
64	relu	5	adam	10	32	0.21513	0.324387
64	relu	5	adam	10	64	0.209051	0.318896
64	relu	5	adam	10	128	0.215715	0.322011
64	relu	5	adam	10	256	0.248842	0.360523
64	tanh	2	adam	10	8	0.215241	0.317689
64	tanh	2	adam	10	16	0.238018	0.341173
64	tanh	2	adam	10	32	0.232614	0.325855
64	tanh	2	adam	10	64	0.248103	0.338166
64	tanh	2	adam	10	128	0.251979	0.346708
64	tanh	2	adam	10	256	0.273365	0.365192
64	tanh	3	adam	10	8	0.225148	0.318866
64	tanh	3	adam	10	16	0.226702	0.332043
64	tanh	3	adam	10	32	0.235575	0.329446
64	tanh	3	adam	10	64	0.221118	0.325245
64	tanh	3	adam	10	128	0.235043	0.342608
64	tanh	3	adam	10	256	0.247696	0.353213
64	tanh	4	adam	10	8	0.243508	0.355629
64	tanh	4	adam	10	16	0.233445	0.340998
64	tanh	4	adam	10	32	0.224969	0.327489
64	tanh	4	adam	10	64	0.225109	0.322122
64	tanh	4	adam	10	128	0.238279	0.326479
64	tanh	4	adam	10	256	0.256057	0.355652
64	tanh	5	adam	10	8	0.239673	0.329573
64	tanh	5	adam	10	16	0.235766	0.34191
64	tanh	5	adam	10	32	0.221314	0.31752
64	tanh	5	adam	10	64	0.223995	0.318587
64	tanh	5	adam	10	128	0.234465	0.330605
64	tanh	5	adam	10	256	0.243001	0.338845
128	relu	2	adam	10	8	0.208366	0.305084
128	relu	2	adam	10	16	0.209042	0.310396
128	relu	2	adam	10	32	0.210125	0.307196
128	relu	2	adam	10	64	0.217805	0.311788
128	relu	2	adam	10	128	0.216145	0.326357
128	relu	2	adam	10	256	0.21918	0.326501
128	relu	3	adam	10	8	0.223651	0.335616
128	relu	3	adam	10	16	0.207523	0.308899
128	relu	3	adam	10	32	0.225145	0.326913
128	relu	3	adam	10	64	0.203492	0.30959
128	relu	3	adam	10	128	0.212474	0.321558
128	relu	3	adam	10	256	0.223571	0.328741
128	relu	4	adam	10	8	0.234412	0.319751
128	relu	4	adam	10	16	0.206783	0.306865
128	relu	4	adam	10	32	0.213773	0.31956
128	relu	4	adam	10	64	0.214953	0.323949
128	relu	4	adam	10	128	0.213828	0.326555
128	relu	4	adam	10	256	0.231709	0.322284
128	relu	5	adam	10	8	0.224689	0.319643
128	relu	5	adam	10	16	0.221376	0.320123
128	relu	5	adam	10	32	0.216928	0.316637
128	relu	5	adam	10	64	0.219234	0.328439
128	relu	5	adam	10	128	0.210652	0.31016
128	relu	5	adam	10	256	0.217441	0.32655
128	tanh	2	adam	10	8	0.255132	0.347482
128	tanh	2	adam	10	16	0.237666	0.342861
128	tanh	2	adam	10	32	0.224524	0.32177
128	tanh	2	adam	10	64	0.223519	0.319779
128	tanh	2	adam	10	128	0.244833	0.356113
128	tanh	2	adam	10	256	0.260176	0.366086
128	tanh	3	adam	10	8	0.23689	0.334883
128	tanh	3	adam	10	16	0.236081	0.336716
128	tanh	3	adam	10	32	0.218373	0.317835
128	tanh	3	adam	10	64	0.228984	0.341201

Full Name: Lee Taylor, ST Number: 190211479

128	tanh	3	adam	10	128	0.246746	0.35016
128	tanh	3	adam	10	256	0.238227	0.330724
128	tanh	4	adam	10	8	0.240521	0.359553
128	tanh	4	adam	10	16	0.231095	0.335379
128	tanh	4	adam	10	32	0.239318	0.343657
128	tanh	4	adam	10	64	0.227629	0.336927
128	tanh	4	adam	10	128	0.234796	0.331781
128	tanh	4	adam	10	256	0.233418	0.332583
128	tanh	5	adam	10	8	0.250075	0.347792
128	tanh	5	adam	10	16	0.256327	0.341721
128	tanh	5	adam	10	32	0.238161	0.339108
128	tanh	5	adam	10	64	0.246192	0.337979
128	tanh	5	adam	10	128	0.25335	0.346402
128	tanh	5	adam	10	256	0.241074	0.348754
256	relu	2	adam	10	8	0.22492	0.330053
256	relu	2	adam	10	16	0.21829	0.319995
256	relu	2	adam	10	32	0.206546	0.311331
256	relu	2	adam	10	64	0.222013	0.333738
256	relu	2	adam	10	128	0.226403	0.331925
256	relu	2	adam	10	256	0.21609	0.32517
256	relu	3	adam	10	8	0.208955	0.313305
256	relu	3	adam	10	16	0.213413	0.322922
256	relu	3	adam	10	32	0.216532	0.309685
256	relu	3	adam	10	64	0.233748	0.323179
256	relu	3	adam	10	128	0.220572	0.316795
256	relu	3	adam	10	256	0.207801	0.315225
256	relu	4	adam	10	8	0.227595	0.328941
256	relu	4	adam	10	16	0.222006	0.310008
256	relu	4	adam	10	32	0.21559	0.320607
256	relu	4	adam	10	64	0.228188	0.339168
256	relu	4	adam	10	128	0.214371	0.324241
256	relu	4	adam	10	256	0.218668	0.319749
256	relu	5	adam	10	8	0.204824	0.306975
256	relu	5	adam	10	16	0.2072	0.312006
256	relu	5	adam	10	32	0.211578	0.313068
256	relu	5	adam	10	64	0.207819	0.31816
256	relu	5	adam	10	128	0.218923	0.318841
256	relu	5	adam	10	256	0.209012	0.322161
256	tanh	2	adam	10	8	0.241161	0.332925
256	tanh	2	adam	10	16	0.236437	0.352759
256	tanh	2	adam	10	32	0.235537	0.345777
256	tanh	2	adam	10	64	0.222765	0.319905
256	tanh	2	adam	10	128	0.276093	0.384288
256	tanh	2	adam	10	256	0.257078	0.348434
256	tanh	3	adam	10	8	0.244419	0.347092
256	tanh	3	adam	10	16	0.240922	0.345308
256	tanh	3	adam	10	32	0.234029	0.32941
256	tanh	3	adam	10	64	0.253037	0.358586
256	tanh	3	adam	10	128	0.247425	0.341759
256	tanh	3	adam	10	256	0.238934	0.339356
256	tanh	4	adam	10	8	0.246199	0.345001
256	tanh	4	adam	10	16	0.285021	0.369102
256	tanh	4	adam	10	32	0.250844	0.354546
256	tanh	4	adam	10	64	0.248201	0.343142
256	tanh	4	adam	10	128	0.24275	0.352775
256	tanh	4	adam	10	256	0.27193	0.353073
256	tanh	5	adam	10	8	0.253979	0.347408
256	tanh	5	adam	10	16	0.289563	0.393264
256	tanh	5	adam	10	32	0.255788	0.36383
256	tanh	5	adam	10	64	0.262953	0.355329
256	tanh	5	adam	10	128	0.335417	0.427542
256	tanh	5	adam	10	256	0.250764	0.340192
512	relu	2	adam	10	8	0.22466	0.326564
512	relu	2	adam	10	16	0.20764	0.308138
512	relu	2	adam	10	32	0.207466	0.306897
512	relu	2	adam	10	64	0.204488	0.308437
512	relu	2	adam	10	128	0.220727	0.334703
512	relu	2	adam	10	256	0.225631	0.326718
512	relu	3	adam	10	8	0.212763	0.316141
512	relu	3	adam	10	16	0.216349	0.313481

512	relu	3	adam	10	32	0.217768	0.330387
512	relu	3	adam	10	64	0.209336	0.309705
512	relu	3	adam	10	128	0.218821	0.321909
512	relu	3	adam	10	256	0.209158	0.320715
512	relu	4	adam	10	8	0.216645	0.30532
512	relu	4	adam	10	16	0.218291	0.322156
512	relu	4	adam	10	32	0.219542	0.3255
512	relu	4	adam	10	64	0.210059	0.324888
512	relu	4	adam	10	128	0.240048	0.330429
512	relu	4	adam	10	256	0.213656	0.320748
512	relu	5	adam	10	8	0.21028	0.315247
512	relu	5	adam	10	16	0.215092	0.333958
512	relu	5	adam	10	32	0.205468	0.317355
512	relu	5	adam	10	64	0.215524	0.326526
512	relu	5	adam	10	128	0.218851	0.331722
512	relu	5	adam	10	256	0.211248	0.317664
512	tanh	2	adam	10	8	0.26916	0.371122
512	tanh	2	adam	10	16	0.249523	0.359523
512	tanh	2	adam	10	32	0.251282	0.347249
512	tanh	2	adam	10	64	0.249699	0.33647
512	tanh	2	adam	10	128	0.252707	0.335468
512	tanh	2	adam	10	256	0.24265	0.338714
512	tanh	3	adam	10	8	0.24744	0.345029
512	tanh	3	adam	10	16	0.260181	0.351524
512	tanh	3	adam	10	32	0.24503	0.352272
512	tanh	3	adam	10	64	0.239263	0.332768
512	tanh	3	adam	10	128	0.26393	0.350336
512	tanh	3	adam	10	256	0.240569	0.344233
512	tanh	4	adam	10	8	0.312941	0.399384
512	tanh	4	adam	10	16	0.276091	0.358734
512	tanh	4	adam	10	32	0.26419	0.363237
512	tanh	4	adam	10	64	0.272449	0.370399
512	tanh	4	adam	10	128	0.271018	0.368783
512	tanh	4	adam	10	256	0.261508	0.356
512	tanh	5	adam	10	8	0.438602	0.483205
512	tanh	5	adam	10	16	0.292322	0.408722
512	tanh	5	adam	10	32	0.264102	0.377943
512	tanh	5	adam	10	64	0.260966	0.357417
512	tanh	5	adam	10	128	0.271158	0.383269
512	tanh	5	adam	10	256	0.252574	0.367451

Figure 1 - Grid Search 1 Results – Green indicates near best result – Orange indicates best result

The majority of the best scoring models according to MSE, used 512 units per layer, with a relatively low batch size. I decided to carry on testing with a smaller space of these attributes, but with different epochs and optimizers.

Units/Layer	Act	Layers	Optimizer	Epochs	Batch_Size	MSE	MAE
512	relu	2	adam	25	64	0.200694	0.307518
512	relu	2	adam	50	64	0.207431	0.294119
512	relu	2	rmsprop	10	64	0.237384	0.335353
512	relu	2	rmsprop	25	64	0.271016	0.353882
512	relu	2	rmsprop	50	64	0.248458	0.327234
512	relu	2	adagrad	10	64	0.297886	0.389166
512	relu	2	adagrad	25	64	0.266866	0.364524
512	relu	2	adagrad	50	64	0.254512	0.353608
512	relu	2	adamax	10	64	0.212671	0.314056
512	relu	2	adamax	25	64	0.203614	0.299226

Figure 2 - Grid Search 2 Results – Green indicates best result

These results indicate changing the optimizer and epochs created minimal differences in the MSE score.

Full Name: Lee Taylor, ST Number: 190211479