Reflection Report

My initial neural network regression model was comprised of four layers: one input layer, two hidden layers with 32 and 16 respectively, and an output layer, with a total of 2,209 trainable parameters.

For testing approaches to optimizing the model, I started by adding another hidden layer of 8 neurons. This brought the number of parameters up to 2,337 and increased the training time from 74.95 seconds to 83.22 seconds. The model's performance was greatly improved, with the Root Mean Squared Error (RMSE) decreasing from 86,511.19 to 68,521.76 on the training data and from 86,377.94 to 68,347.84 on test data. The R^2 measure also improved from 0.82 to 0.88 on training data and 0.89 on test data.

On this basis, I further enhanced the initial model by adding two more hidden layers with 64 and 8 neurons, making the total parameter count 6,081. Despite this increased complexity, training time did not change much, at 82.59 seconds. This more complex model performed even better: RMSE fell to 60,976.08 on the training set and 60,830.3 on the test set, with an R^2 score of 0.91 on both. Mean Absolute Error and Mean Absolute Percentage Error improved consistently across all experiments as model complexity was increased.

In summary, adding more hidden layers generally led to better performance by allowing the model to learn more complex relationships in the data. The relatively small increase in training time shows that deeper models can achieve better accuracy. However, I think it's still important to watch out for overfitting as the models become more complex.

Table that summarizes the model results

	With 2 hidden layers		With 3 hidden layers		With 4 hidden layers	
	Train	Test	Train	Test	Train	Test
Root Mean Squared Error	86511.19	86377.94	68521.76	68347.84	60976.08	60830.3
Mean Absolute Error	65021.85	64858.78	50155.65	50088.57	44092.62	43978.12
Mean Absolute Percentage Error	10.82	10.81	8.22	8.23	7.16	7.15
R2 score	0.82	0.82	0.88	0.89	0.91	0.91

Parameters and training time

Initial Model,

Model: "sequential_3"

Layer (type)	Output Shape	Param #
dense_9 (Dense)	(None, 32)	1,664
dense_10 (Dense)	(None, 16)	528
dense_11 (Dense)	(None, 1)	17

Total params: 6,629 (25.90 KB)

Trainable params: 2,209 (8.63 KB)

Non-trainable params: 0 (0.00 B)

Optimizer params: 4,420 (17.27 KB)

```
Training time: 74.95 seconds
```

Modified Model, with 3 hidden layers

Model: "sequential_6"

Layer (type)	Output Shape	Param #
dense_21 (Dense)	(None, 32)	1,664
dense_22 (Dense)	(None, 16)	528
dense_23 (Dense)	(None, 8)	136
dense_24 (Dense)	(None, 1)	9

Total params: 7,013 (27.40 KB)

Trainable params: 2,337 (9.13 KB)

Non-trainable params: 0 (0.00 B)

Optimizer params: 4,676 (18.27 KB)

Training time: 83.22 seconds

Modified Model, with 4 hidden layers

Model: "sequential_4"

Layer (type)	Output Shape	Param #
dense_12 (Dense)	(None, 64)	3,328
dense_13 (Dense)	(None, 32)	2,080
dense_14 (Dense)	(None, 16)	528
dense_15 (Dense)	(None, 8)	136
dense_16 (Dense)	(None, 1)	9

Total params: 18,245 (71.27 KB)

Trainable params: 6,081 (23.75 KB)

Non-trainable params: 0 (0.00 B)

Optimizer params: 12,164 (47.52 KB)

Training time: 82.59 seconds