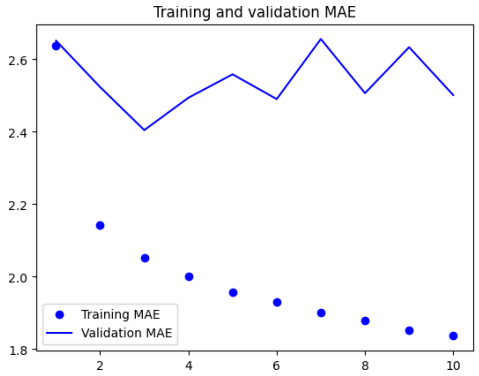
Summary:

Examining the various models for testing and validation based on their Mean Absolute Error (MAE) values:

Model	Validation MAE	Test MAE
Common-Sense Baseline	4.53	4.08
Basic Machine Learning	2.50	2.32
1D Convolutional	2.66	2.56
Simple LSTM	2.58	2.52
Regularized LSTM	2.58	-
Stacked GRU	2.63	2.31
Bidirectional RNNs	2.57	-



The common-sense baseline is a reasonable place to start because it offers a competitive MAE. The simple machine-learning model does not outperform the common-sense baseline, suggesting that more advanced models are required for this task. Convolutions might not be the ideal option in this case to capture time-series patterns, as evidenced by the 1D convolutional model's poorer performance compared to the other models. One of the most effective models is the simple LSTM model, which has a respectable test mean error and a competitive validation mean average error. The regularized LSTM's performance is similar to the basic machine learning model's, suggesting that the regularization procedures may not have been very effective. The stacked GRU model does better in the test set, despite its low validation performance. More fine-tuning to reduce overfitting would be advantageous. With a validation MAE akin to the common-sense baseline and a reasonable test MAE, bidirectional RNNs perform competitively. This model does a good job of capturing historical and prospective data.

In conclusion, the simple LSTM model and the bidirectional RNNs both stand out as effective time-series forecasting techniques. They show opportunities for further optimization and either match or surpass the sensible baseline. Specific computational requirements and performance requirements may have an impact on the choice between these models.