Results

Model Diagnostics

Models performed very well on training data. Both the LSTM and AutoML models outperformed the Naïve Bayesian model in fitting the training data. Training scores saw very little variance in performance year-to-year as seen in the small spread in training scores in Figure 1. Testing scores showed very high variance, with some testing years having very high performance and many having very poor performance, as seen in Figure 1. AutoML models in particular showed a very high variance year-to-year in test performance indicating an untrustworthiness in their forecasts.
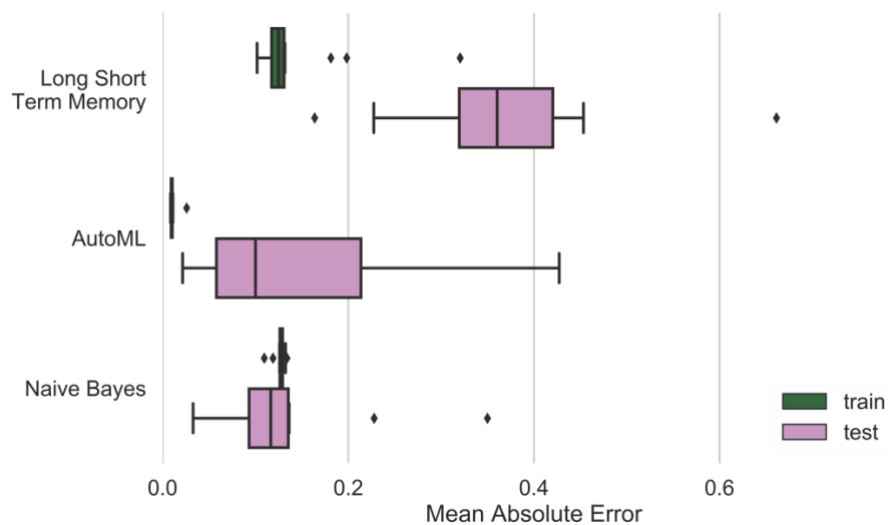

Figure 1. Model training diagnostic results. Individual data points indicate individual training datasets.

Forecasts

Full test data set

Forecasts produced by both the LSTM and AutoML models did not produce sufficiently accurate
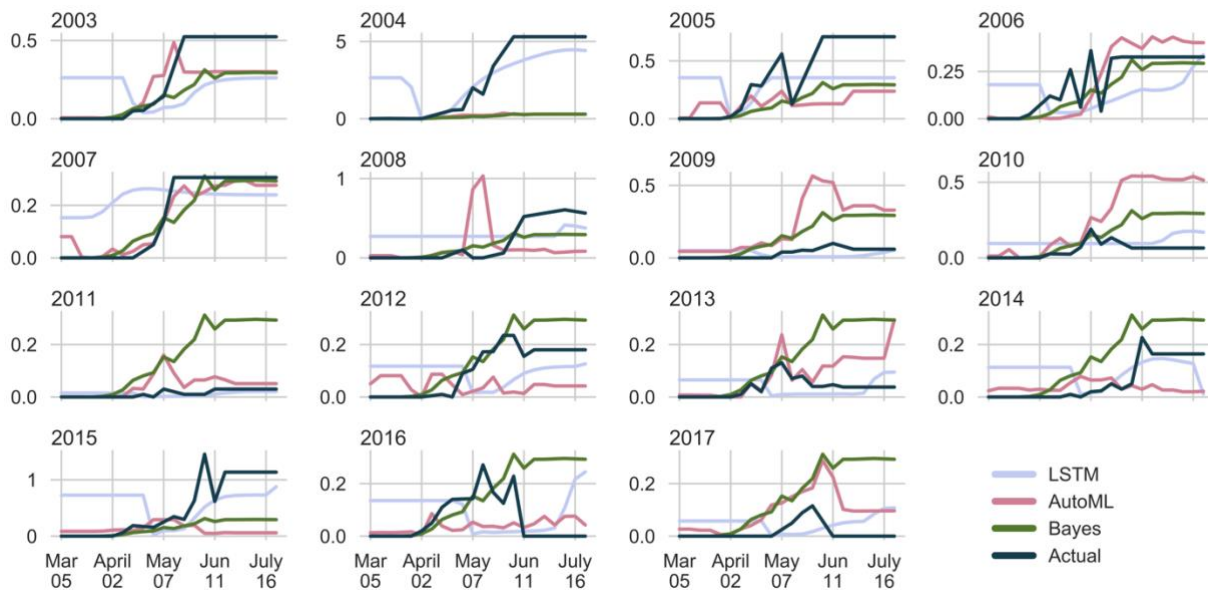
forecasts, as evidenced in Figure 2.

Figure 2. Forecasts from LSTM and AutoML methods when provided with full test data against

the actual values and Naïve Bayes model.

Test data set up to and including April

Forecasts produced by both the LSTM and AutoML models did not produce sufficiently accurate

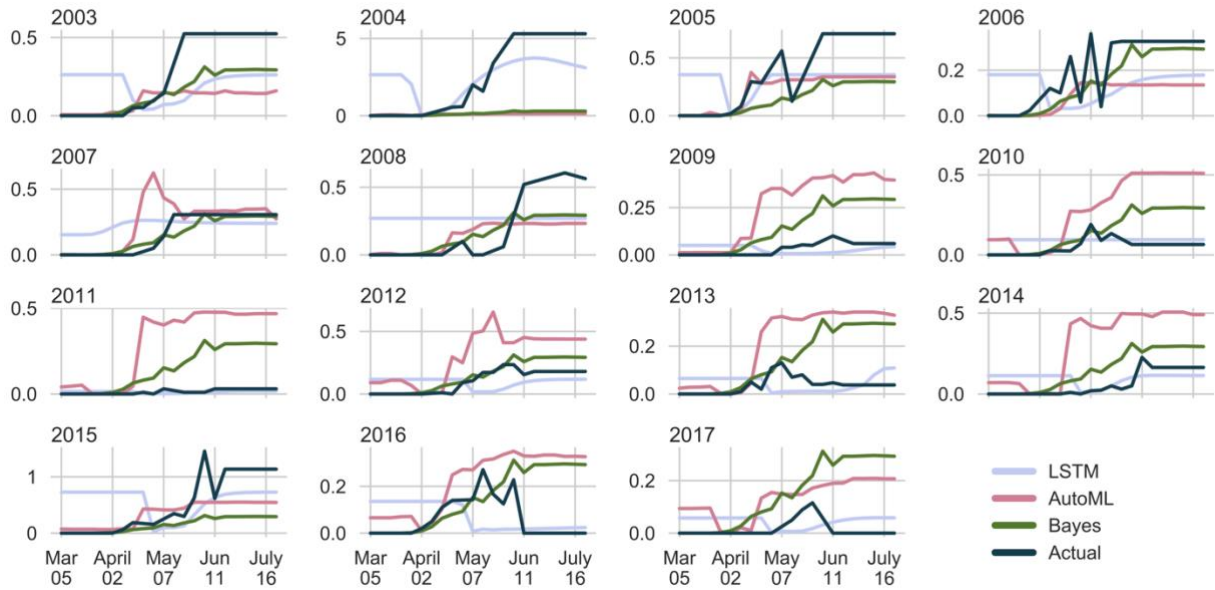forecasts, as evidenced in Figure 3.

Figure 3: Forecasts from LSTM and AutoML, methods when provided with data from January to April against the actual values and Naïve Bayes model.

Test data set up to and including June

Forecasts produced by both the LSTM and AutoML models did not produce sufficiently accurate forecasts, as evidenced in Figure 4.
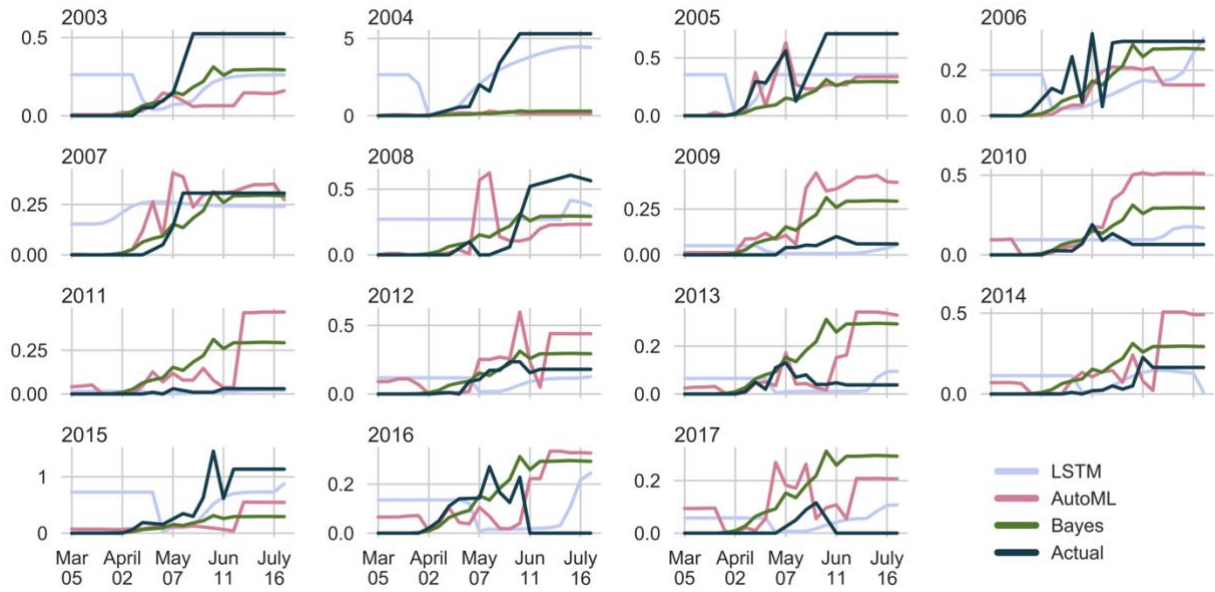
Figure 4: Forecasts from LSTM and AutoML methods when provided with data from January to June against the actual values and Naïve Bayes model.