
Forecasting with parametric models and neural networks: a comparison (exposé)

Niklas Paulig^{1,2,3}

¹ Technical University of Dresden, Dresden, Germany

² Mat.Nr: 4064033

³ Master VWL, Fak. Wirtschaftswissenschaften

January 13, 2021

For this master thesis, I choose 14 stock indices from around the world to compare the forecasting abilities of different parametric statistical models against the forecasting abilities of artificial neural networks (ANNs). The focus will lie on the inner workings and parameterizations of those two tools as well as accuracy and goodness-of-fit measures. The research question asks whether black-box approaches like neural networks can outperform profound statistical forecasting models.

1 Idea

When performing time-series forecasting the usual goal is not only to predict future values from a given set of data but also to get an understanding of the data generating process. Both forecasting branches I cover in this thesis do this, but very differently.

For classical time-series analysis, I will set up an exponential smoothing model based on the works of Brown (1962), Holt (1957) and Winters (1960) that will act as a baseline. Various error metrics to measure out-of-sample performance will be provided, such as RMSE, MAE and MAPE.

In a next step, the more advanced Box and Jen-

kins (1976) methodology is used to fit an ARIMA model to the data and compare it to the naive exponential smoothing in terms of accuracy and robustness. As the 14 time-series are stock indices with fluctuating variances I will also specify and fit a generalized ARCH model to the series in order to account for its volatility clustering (Bollerslev, Engle, and Nelson (1994) based on the work of Engle (1982)).

The classic models will be compared using relative entropy estimators like the Akaike information criterion and other related estimators such as Hannan-Quinn or the Schwarz information criterion. Unfortunately due to the nature of artificial neural networks not using likelihood functions, those cannot be compared to the classical models in this way.

The second part of the thesis will introduce artificial neural networks, that substantially differ from the classical models in the way they work and produce forecasts. The development of neural networks will be traced back to the Perceptron algorithm (Rosenblatt (1957)) and the idea of a neuron will be explained.

As a benchmark for feed-forward neural networks a multi-layer perceptron (MLP) with different parametrizations is trained on the data and out-of-sample performance in comparison to classical models is discussed. In a final step, a more

complex recurrent neural network (RNN) (see for example Williams and Zipser (1989) or Werbos (1990)), a Long Short-Term Memory (LSTM) model after Hochreiter and Schmidhuber (1997) is fitted and compared in order to get a more holistic overview over the forecasting abilities of simple and more complex artificial neural networks.

2 Structure

The thesis structure is planned to contain three chapters. The first chapter introduces the different inference and forecast models and explains their origin, use cases and inner workings. I will thereby start with the classical models (exp. smoothing, ARIMA, GARCH) and continue with the neural networks. The last part of the first chapter will introduce the various error and performance metrics and explain their behavior.

Chapter two introduces the dataset and firstly describes the pre-fit transformations and statistical tests and secondly the fitting process itself. The models are then compared to each other with the above-introduced out-of-sample performance metrics.

The last chapter draws a conclusion to the question whether the rather complicated and opaque neural networks can outperform classical statistical analysis, and if they do, to what magnitude.

References

- Bollerslev, Tim, Robert F Engle, and Daniel B Nelson (1994). "ARCH models". In: *Handbook of econometrics* 4, pp. 2959–3038.
- Box, George EP and Gwilym M Jenkins (1976). "Time series analysis: Forecasting and control San Francisco". In: *Calif: Holden-Day*.
- Brown, Robert Goodell (1962). "Smoothing, forecasting and prediction of discrete time series". In:
- Engle, Robert F (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". In: *Econometrica: Journal of the Econometric Society*, pp. 987–1007.
- Hochreiter, Sepp and Jürgen Schmidhuber (1997). "Long short-term memory". In: *Neural computation* 9.8, pp. 1735–1780.
- Holt, Charles C (1957). "Forecasting trends and seasonals by exponentially weighted averages. carnegie institute of technology". In: *Pittsburgh ONR memorandum*.
- Rosenblatt, Frank (1957). *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory.
- Werbos, Paul J (1990). "Backpropagation through time: what it does and how to do it". In: *Proceedings of the IEEE* 78.10, pp. 1550–1560.
- Williams, Ronald J and David Zipser (1989). "A learning algorithm for continually running fully recurrent neural networks". In: *Neural computation* 1.2, pp. 270–280.
- Winters, Peter R (1960). "Forecasting sales by exponentially weighted moving averages". In: *Management science* 6.3, pp. 324–342.