

[Start Here](#)[Blog](#)[Books](#)[About](#)[Contact](#)

Need help with LSTMs in Python? [Take the FREE Mini-Course.](#)

How to Get Good Results Fast with Deep Learning for Time Series Forecasting

by [Jason Brownlee](#) on [July 31, 2017](#) in [Long Short-Term Memory Networks](#)



3 Strategies to Design Experiments and Manage Complexity on Your Predictive Modeling Problem.

It is difficult to get started on a new time series forecasting project.

Given years of data, it can take days or weeks to fit a deep learning model. How do you get started exactly?

For some practitioners, this can lead to paralysis and even procrastination at the very beginning of a trap of only trying and using what has worked before rather than truly exploring the problem.

[Get Your Start in Machine Learning](#)

the

In this post, you will discover practical strategies that you can use to get started when applying deep learning methods like Multilayer Neural Networks and Long Short-Term Memory (LSTM) Recurrent Neural Network models to time series forecasting problems.

The strategies in this post are not foolproof, but they are hard learned rules of thumb that I have discovered while working with large time series datasets.

After reading this post, you will know:

- A strategy to balance the exploration of ideas and the exploitation of what works on your problem.
- A strategy to learn quickly and scale ideas with data to confirm they hold on the broader problem.
- A strategy to navigate the complexity of the framing of your problem and the complexity of the chosen deep learning model.

Let's get started.

1. Strategy for Exploration and Exploitation

It is important to balance exploration and exploitation in your search for a model that performs well on your problem.

I would recommend two different approaches that should be used in tandem:

- Diagnostics.
- Grid Search.

Diagnostics

Diagnostics involve performing a run with one set of hyperparameters and producing a trace of the model's performance over the training epoch.

These plots provide insight into over-learning or under-learning and the potential for specific sets of hyperparameters. They are sanity checks or seeds for deeper investigation of ranges of parameters that can be explored and prevent you from wasting time with more epochs than are reasonably required, or networks that are too large.

They are sanity checks or seeds for deeper investigation of ranges of parameters that can be explored and prevent you from wasting time with more epochs than are reasonably required, or networks that are too large.

Get Your Start in Machine Learning ×

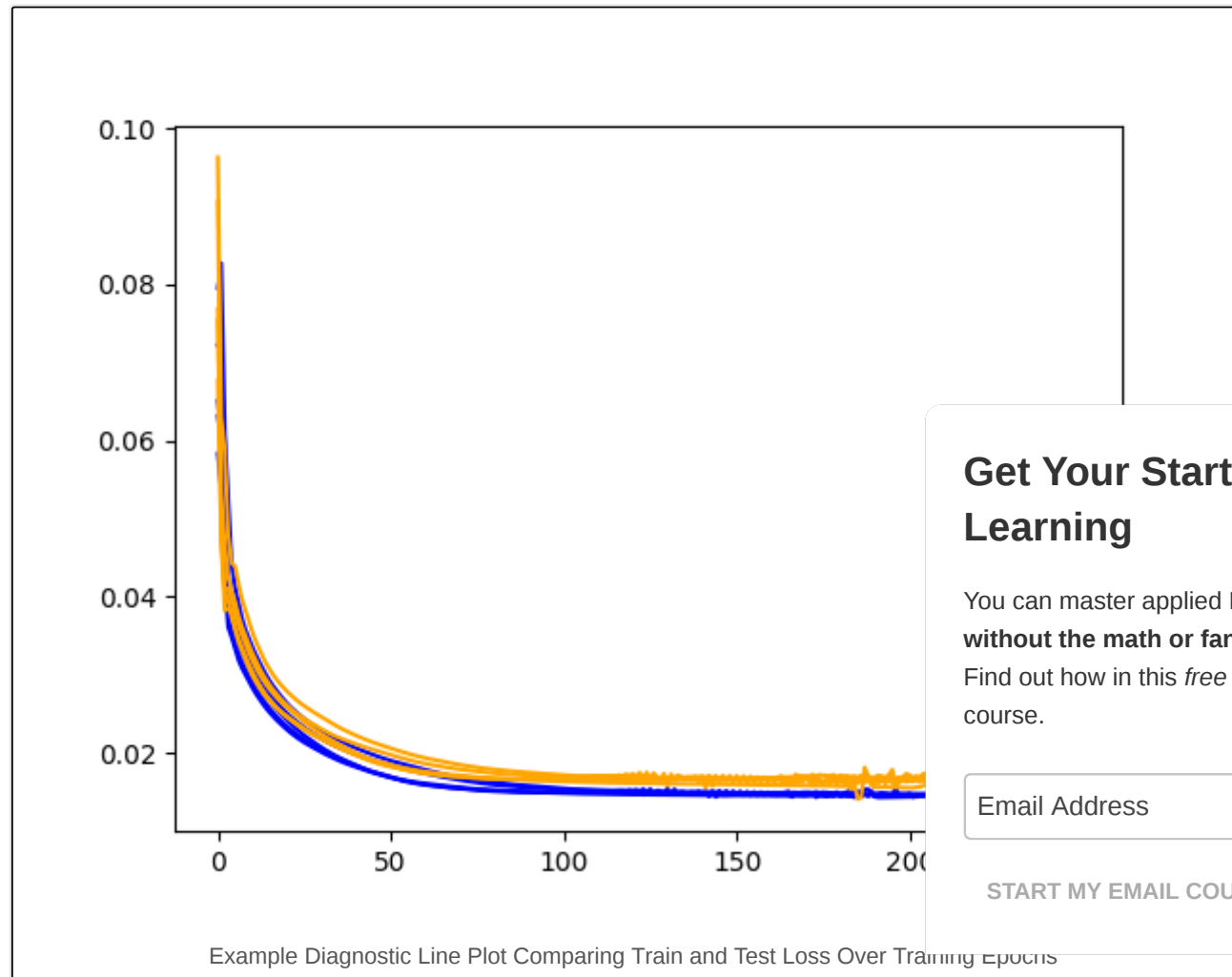
You can master applied Machine Learning **without the math or fancy degree.**

Find out how in this *free* and *practical* email course.

START MY EMAIL COURSE

Get Your Start in Machine Learning

Below is an example of a diagnostic plot from a model run showing training and validation RMSE.



Grid Search

Based on learnings from diagnostic results, the grid search provides a sweep across a suite of values for specific model hyperparameters such as the number of neurons, batch size, and so on.

They allow you to systematically dial in specific hyperparameter values in a piecewise manner.

Get Your Start in Machine Learning

Need help with LSTMs for Sequence Prediction?

Take my free 7-day email course and discover 6 different LSTM architectures (with sample code).

Click to sign-up and also get a free PDF Ebook version of the course.

Start Your FREE Mini-Course Now!

Interleave the Approaches

I would recommend interleaving diagnostic runs and grid search runs.

You can spot check your hypotheses with diagnostics and get the best from promising ideas with grid search.

I would strongly encourage you to test every assumption you have about the model. This includes sampling the data, the choice of model, the choice of activation function, loss function, and more.

Used with the data handling strategy below, you will quickly build a map of what works and what does not.

Below is an example of the results of a grid search of the batch size for a model showing the distribution of times.

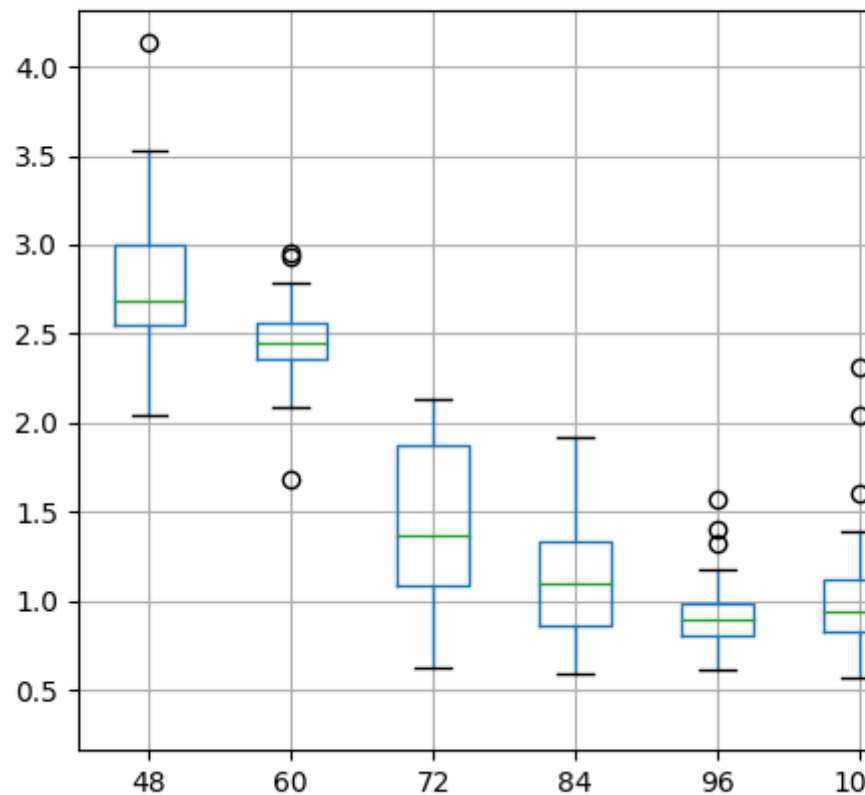
Get Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree.**

Find out how in this *free* and *practical* email course.

START MY EMAIL COURSE

Get Your Start in Machine Learning



Example Box and Whisker Plots Comparing a Model Skill For Different Mode

Get Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree**. Find out how in this *free* and *practical* email course.

START MY EMAIL COURSE

2. Strategy for Handling Data Size

I recommend a strategy of working with smaller samples of data first to test ideas and slowly increasing the amount of data to see if things learned on the small samples hold on larger samples.

For example, if you have multiple years of hourly measurements, you could split your data as follows:

- 1 week sample.

Get Your Start in Machine Learning

- 1 month sample.
- 1 year sample.
- all data.

The alternative is that you fit and explore models on the entire dataset where each model can take days to fit, in turn meaning that your rate of learning is dramatically decreased.

The benefit of this approach is that you can very quickly test ideas, in minutes, with multiple repeats (e.g. statistically significant) and then later scale up only those promising ideas to more and more data.

Generally, with well-framed supervised learning problems, the learnings do scale with the data. Nevertheless, there is a risk that the problems are substantially different at different scales of data and that findings do not hold. You can check for this with simpler models that are faster to train and tease out whether this is an issue early on.

Finally, as you scale models to more data, you can also reduce the number of repeats of experiment

3. Strategy for Model Complexity

Like data size, the complexity of the model is another concern that must be managed and can be scaled.

We can look at this both from the framing of the supervised learning problem and the model itself.

Model Framing Complexity

For example, we may assume a time series forecasting problem that includes exogenous variables (e.g. weather data).

We can scale the complexity of the problem and see what works at one level of complexity (e.g. univariate inputs) holds at more complex levels of complexity (multivariate inputs).

For example, you could work through model complexity as follows:

- Univariate input, single-step output.
- Univariate input, multi-step output.
- Multivariate inputs, single-step output.

Get Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree.**

Find out how in this *free* and *practical* email course.

START MY EMAIL COURSE

- Multivariate inputs, multi-step output.

This too can extend to multivariate forecasts.

At each step, the objective is to demonstrate that the addition of complexity can lift the skill of the model.

For example:

- Can a neural network model outperform a persistence forecast model?
- Can a neural network model outperform a linear forecast model?
- Can exogenous input variables lift the skill of the model over a univariate input?
- Can a direct a multi-step forecast be more skillful than a recursive single-step forecast?

If these questions can't be overcome, or be overcome easily, it can help you quickly settle on a framework.

Complexity in Model Capability

This same approach can be used when working with more sophisticated neural network models like RNNs.

For example:

- Model the problem as a mapping of inputs to outputs (e.g. no internal state or BPTT).
- Model the problem as a mapping problem with internal state across input sequences only (no BPTT).
- Model the problem as a mapping problem with internal state and BPTT.

At each step, the increased model complexity must demonstrate skill at or above the prior level of complexity. If not, the added complexity must be justified by a commensurate increase in model skill or capability.

For example:

- Can an LSTM outperform an MLP with a window?
- Can an LSTM with internal state and no BPTT outperform an LSTM where the state is reset after each sample?
- Can an LSTM with BPTT over input sequences outperform an LSTM that is updated after each time step?

Further Reading

Get Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree.**

Find out how in this *free* and *practical* email course.

START MY EMAIL COURSE

This section provides more resources on the topic if you are looking go deeper.

- [How to Work Through a Time Series Forecast Project](#)
- [How to Tune LSTM Hyperparameters with Keras for Time Series Forecasting](#)
- [How To Improve Deep Learning Performance](#)

Summary

In this tutorial, you discovered how you can overcome the paralysis that can strike at the beginning of a deep learning project.

Specifically, you learned how you can systematically break down complexity and strategies that you can use to get results fast:

- A strategy to balance the exploration of ideas and the exploitation of what works on your problem
- A strategy to learn quickly and scale ideas with data to confirm they hold on the broader problem
- A strategy to navigate the complexity of the framing of your problem and the complexity of your problem

Do you have any questions?

Ask your questions in the comments below and I will do my best to answer.

Develop LSTMs for Sequence Prediction

Develop Your Own LSTM models in Minutes

...with just a few lines of python code

Discover how in my new Ebook:
[Long Short-Term Memory Networks with Python](#)

It provides **self-study tutorials** on topics like:
CNN LSTMs, Encoder-Decoder LSTMs, generative models, data preparation, making predictions and much more...

Finally Bring LSTM Recurrent Neural Networks to Your Sequence Predictions Projects

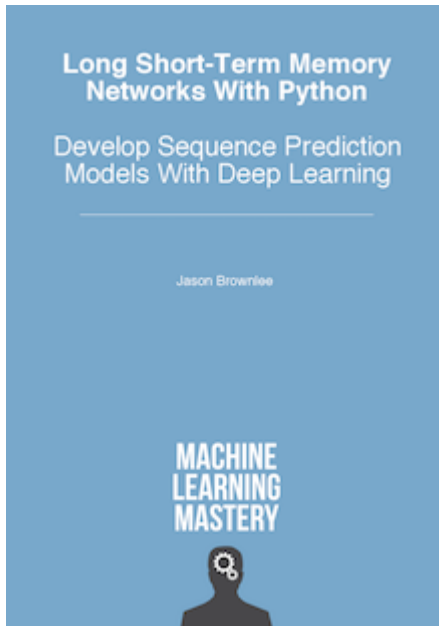
Get Your Start in Machine Learning

You can master applied Machine Learning
without the math or fancy degree.

Find out how in this *free* and *practical* email
course.

START MY EMAIL COURSE

Skip the Academics. Just Results.

[Click to learn more.](#)

About Jason Brownlee

Dr. Jason Brownlee is a husband, proud father, academic researcher, author, professional developer, and entrepreneur. He is passionate about helping developers get started and get good at applied machine learning. [Learn more.](#)

[View all posts by Jason Brownlee](#) →

Get Your Start in Machine Learning

You can master applied Machine Learning **without the math or fancy degree.** Find out how in this *free* and *practical* email course.

[START MY EMAIL COURSE](#)

[◀ Why One-Hot Encode Data in Machine Learning?](#)

[9 Ways to Get Help with Deep Learning in Keras ▶](#)

Get Your Start in Machine Learning

4 Responses to *How to Get Good Results Fast with Deep Learning for Time Series Forecasting*



hal9000 August 4, 2017 at 4:59 am #

REPLY ↩

Hi,

again a excellent write-up!

In “Complexity in Model Capability” you write about using LSTMs with and without BPTT and keeping the internal states. I’m not sure what you mean with BPTT. Isn’t BPTT always used to when training LSTMs?

Best regards!



Jason Brownlee August 4, 2017 at 7:03 am #

BPTT is always used, but it has no “time” to work on if you give it a sequence of one element.



Ali August 7, 2017 at 2:20 am #

How long a time series should be for using deep learning?



Jason Brownlee August 7, 2017 at 8:47 am #

REPLY ↩

Great question.

When using RNNs, the limit that researchers suggest is 200-400 time steps per sequence, but you can have multiple sequences and manage the resetting of internal state manually.

Some basically, infinite!

Get Your Start in Machine Learning



You can master applied Machine Learning **without the math or fancy degree.**

Find out how in this *free* and *practical* email course.

START MY EMAIL COURSE

Get Your Start in Machine Learning

But test for your problem and discover the point of diminishing returns in terms of sequence length and length before resetting state.

I have some posts on BPTT that I would recommend reading, use the blog search.

Leave a Reply

Name (required)

Email (will not be published) (required)

Website

SUBMIT COMMENT

Get Your Start in Machine Learning



You can master applied Machine Learning **without the math or fancy degree.**

Find out how in this *free* and *practical* email course.

START MY EMAIL COURSE

Welcome to Machine Learning Mastery

Hi, I'm Dr. Jason Brownlee.

My goal is to make practitioners like YOU awesome at applied machine learning.

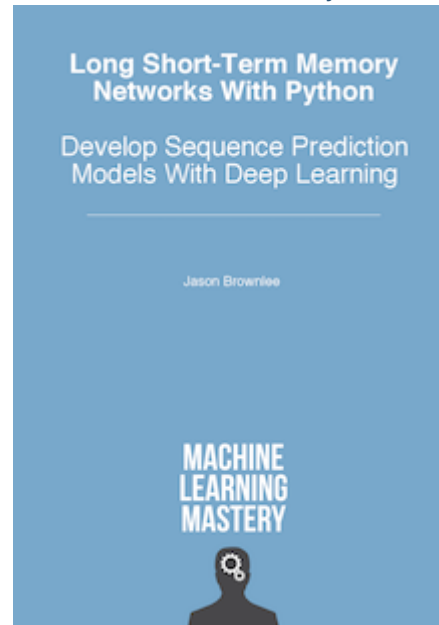
Get Your Start in Machine Learning

[Read More](#)

Deep Learning for Sequence Prediction

Cut through the math and research papers.
Discover 4 Models, 6 Architectures, and 14 Tutorials.

Get Started With LSTMs in Python Today!



Get Your Start in Machine Learning ×

You can master applied Machine Learning **without the math or fancy degree.** Find out how in this *free* and *practical* email course.

START MY EMAIL COURSE

POPULAR



Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras

JULY 21, 2016

Get Your Start in Machine Learning

**Your First Machine Learning Project in Python Step-By-Step**

JUNE 10, 2016

**Develop Your First Neural Network in Python With Keras Step-By-Step**

MAY 24, 2016

**Sequence Classification with LSTM Recurrent Neural Networks in Python with Keras**

JULY 26, 2016

**How to Setup a Python Environment for Machine Learning and Deep Learning with Anaconda**

MARCH 13, 2017

**Time Series Forecasting with the Long Short-Term Memory Network in Python**

APRIL 7, 2017

**Multi-Class Classification Tutorial with the Keras Deep Learning Library**

JUNE 2, 2016

**Regression Tutorial with the Keras Deep Learning Library in Python**

JUNE 9, 2016

**Multivariate Time Series Forecasting with LSTMs in Keras**

AUGUST 14, 2017

**How to Implement the Backpropagation Algorithm From Scratch In Python**

NOVEMBER 7, 2016

Get Your Start in Machine Learning



You can master applied Machine Learning **without the math or fancy degree.** Find out how in this *free* and *practical* email course.

[START MY EMAIL COURSE](#)

Get Your Start in Machine Learning



You can master applied Machine Learning **without the math or fancy degree.**

Find out how in this *free* and *practical* email course.

START MY EMAIL COURSE