

# Mini-Course on Long Short-Term Memory Recurrent Neural Networks with Keras

by **Jason Brownlee** on [August 16, 2017](#) in **Long Short-Term Memory Networks**

Long Short-Term Memory (LSTM) recurrent neural networks are one of the most interesting types of deep learning at the moment.

They have been used to demonstrate world-class results in complex problem domains such as language translation, automatic image captioning, and text generation.

LSTMs are different to multilayer Perceptrons and convolutional neural networks in that they are designed specifically for sequence prediction problems.

In this mini-course, you will discover how you can quickly bring LSTM models to your own sequence forecasting problems.

After completing this mini-course, you will know:

- What LSTMs are, how they are trained, and how to prepare data for training LSTM models.
- How to develop a suite of LSTM models including stacked, bidirectional, and encoder-decoder models.
- How you can get the most out of your models with hyperparameter optimization, updating, and finalizing models.

Let's get started.

**Note:** This is a big guide; you may want to bookmark it.



Mini-Course on Long Short-Term Memory Recurrent Neural Networks with Keras  
Photo by [Nicholas A. Tonelli](#), some rights reserved.

## Who Is This Mini-Course For?

Before we get started, let's make sure you are in the right place.

This course is for developers that know some applied machine learning and need to get good at LSTMs fast.

Maybe you want or need to start using LSTMs on your project. This guide was written to help you do that quickly and efficiently.

- You know your way around Python.
- You know your way around SciPy.
- You know how to install software on your workstation.

- You know how to wrangle your own data.
- You know how to work through a predictive modeling problem with machine learning.
- You may know a little bit of deep learning.
- You may know a little bit of Keras.

You know how to set up your workstation to use Keras and scikit-learn; if not, you can learn how to here:

- [How to Setup a Python Environment for Machine Learning and Deep Learning with Anaconda](#)

This guide was written in the top-down and results-first machine learning style that you're used to. It will teach you how to get results, but it is not a panacea.

You will develop useful skills by working through this guide.

After completing this course, you will:

- Know how LSTMs work.
- Know how to prepare data for LSTMs.
- Know how to apply a suite of types of LSTMs.
- Know how to tune LSTMs to a problem.
- Know how to save an LSTM model and use it to make predictions.

Next, let's review the lessons.

---

## **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!**

---

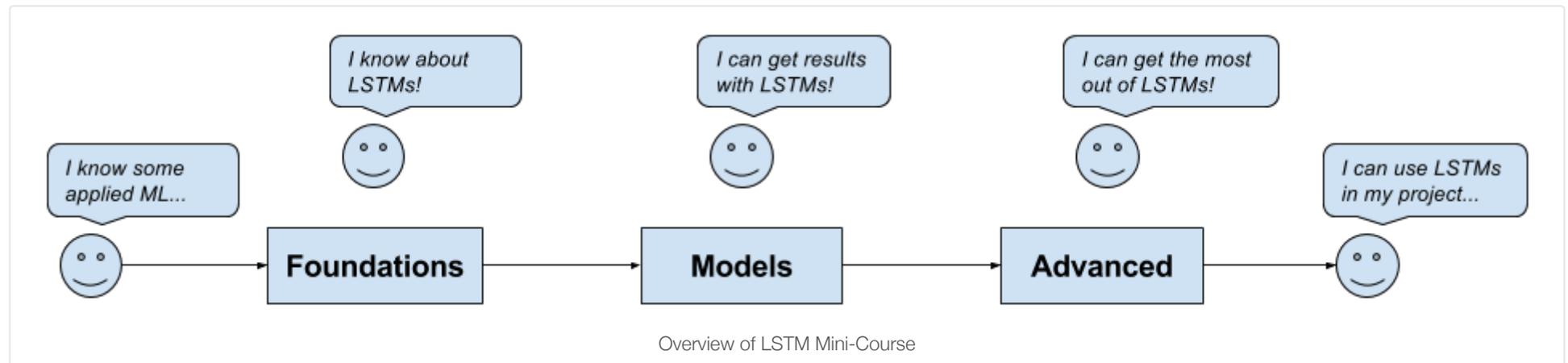
# Mini-Course Overview

This mini-course is broken down into 14 lessons.

You could complete one lesson per day (recommended) or complete all of the lessons in one day (hardcore!).

It really depends on the time you have available and your level of enthusiasm.

Below are 14 lessons that will get you started and productive with LSTMs in Python. The lessons are divided into three main themes: foundations, models, and advanced.



## Foundations

The focus of these lessons are the things that you need to know before using LSTMs.

- **Lesson 01:** What are LSTMs?
- **Lesson 02:** How LSTMs are trained
- **Lesson 03:** How to prepare data for LSTMs
- **Lesson 04:** How to develop LSTMs in Keras

## Models

- **Lesson 05:** How to develop Vanilla LSTMs
- **Lesson 06:** How to develop Stacked LSTMs

- **Lesson 07:** How to develop CNN LSTMs
- **Lesson 08:** How to develop Encoder-Decoder LSTMs
- **Lesson 09:** How to develop Bi-directional LSTMs
- **Lesson 10:** How to develop LSTMs with Attention
- **Lesson 11:** How to develop Generative LSTMs

## Advanced

- **Lesson 12:** How to tune LSTM hyperparameters
- **Lesson 13:** How to update LSTM models
- **Lesson 14:** How to make predictions with LSTMs

Each lesson could take you 60 seconds or up to 60 minutes. Take your time and complete the lessons at your own pace. Ask questions, and even post results in the comments below.

The lessons expect you to go off and find out how to do things. I will give you hints, but part of the point of each lesson is to force you to learn where to go to look for help (hint, I have all of the answers on this blog; use the search).

I do provide more help in the early lessons because I want you to build up some confidence and inertia.

Hang in there; don't give up!

## Foundations

The lessons in this section are designed to give you an understanding of how LSTMs work and how to implement LSTM models using the Keras library.

### Lesson 1: What are LSTMs?

#### Goal

The goal of this lesson is to understand LSTMs from a high-level sufficiently so that you can explain what they are and how they work to a colleague or manager.

#### Questions

- What is sequence prediction and what are some general examples?
- What are the limitations of traditional neural networks for sequence prediction?
- What is the promise of RNNs for sequence prediction?
- What is the LSTM and what are its constituent parts?
- What are some prominent applications of LSTMs?

## Further Reading

- [Crash Course in Recurrent Neural Networks for Deep Learning](#)
- [Gentle Introduction to Models for Sequence Prediction with Recurrent Neural Networks](#)
- [The Promise of Recurrent Neural Networks for Time Series Forecasting](#)
- [On the Suitability of Long Short-Term Memory Networks for Time Series Forecasting](#)
- [A Gentle Introduction to Long Short-Term Memory Networks by the Experts](#)
- [8 Inspirational Applications of Deep Learning](#)

## Lesson 2: How LSTMs are trained

### Goal

The goal of this lesson is to understand how LSTM models are trained on example sequences.

### Questions

- What common problems afflict the training of traditional RNNs?
- How does the LSTM overcome these problems?
- What algorithm is used to train LSTMs?
- How does Backpropagation Through Time work?
- What is truncated BPTT and what benefit does it offer?
- How is BPTT implemented and configured in Keras?

## Further Reading

- [A Gentle Introduction to Backpropagation Through Time](#)
- [How to Prepare Sequence Prediction for Truncated Backpropagation Through Time in Keras](#)

## Lesson 3: How to prepare data for LSTMs

### Goal

The goal of this lesson is to understand how to prepare sequence prediction data for use with LSTM models.

### Questions

- How do you prepare numeric data for use with LSTMs?
- How do you prepare categorical data for use with LSTMs?
- How do you handle missing values in sequences when using LSTMs?
- How do you frame a sequence as a supervised learning problem?
- How do you handle long sequences when working with LSTMs?
- How do you handle input sequences with different lengths?
- How do you reshape input data for LSTMs in Keras?

### Experiment

Demonstrate how to transform a numerical input sequence into a form suitable for training an LSTM.

### Further Reading

- [How to Scale Data for Long Short-Term Memory Networks in Python](#)
- [How to One Hot Encode Sequence Data in Python](#)
- [How to Handle Missing Timesteps in Sequence Prediction Problems with Python](#)
- [How to Convert a Time Series to a Supervised Learning Problem in Python](#)
- [How to Handle Very Long Sequences with Long Short-Term Memory Recurrent Neural Networks](#)
- [How to Prepare Sequence Prediction for Truncated Backpropagation Through Time in Keras](#)
- [Data Preparation for Variable-Length Input Sequences](#)

## Lesson 4: How to develop LSTMs in Keras

### Goal

The goal of this lesson is to understand how to define, fit, and evaluate LSTM models using the Keras deep learning library in Python.

## Questions

- How do you define an LSTM Model?
- How do you compile an LSTM Model?
- How do you fit an LSTM Model?
- How do you evaluate an LSTM Model?
- How do you make predictions with an LSTM Model?
- How can LSTMs be applied to different types of sequence prediction problems?

## Experiment

Prepare an example that demonstrates the life-cycle of an LSTM model on a sequence prediction problem.

## Further Reading

- [The 5 Step Life-Cycle for Long Short-Term Memory Models in Keras](#)
- [Gentle Introduction to Models for Sequence Prediction with Recurrent Neural Networks](#)

## Models

The lessons in this section are designed to teach you how to get results with LSTM models on sequence prediction problems.

## Lesson 5: How to develop Vanilla LSTMs

### Goal

The goal of this lesson is to learn how to develop and evaluate vanilla LSTM models.

- What is the vanilla LSTM architecture?
- What are some examples where the vanilla LSTM has been applied?

### Experiment

Design and execute an experiment that demonstrates a vanilla LSTM on a sequence prediction problem.

## Further Reading

- [Sequence Classification with LSTM Recurrent Neural Networks in Python with Keras](#)
- [Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras](#)
- [Time Series Forecasting with the Long Short-Term Memory Network in Python](#)

## Lesson 6: How to develop Stacked LSTMs

### Goal

The goal of this lesson is to learn how to develop and evaluate stacked LSTM models.

### Questions

- What are the difficulties in using a vanilla LSTM on a sequence problem with hierarchical structure?
- What are stacked LSTMs?
- What are some examples of where the stacked LSTM has been applied?
- What benefits do stacked LSTMs provide?
- How can a stacked LSTM be implemented in Keras?

### Experiment

Design and execute an experiment that demonstrates a stacked LSTM on a sequence prediction problem with hierarchical input structure.

## Further Reading

- [Sequence Classification with LSTM Recurrent Neural Networks in Python with Keras](#)
- [Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras](#)

## Lesson 7: How to develop CNN LSTMs

### Goal

The goal of this lesson is to learn how to develop LSTM models that use a Convolutional Neural Network on the front end.

## Questions

- What are the difficulties of using a vanilla LSTM with spatial input data?
- What is the CNN LSTM architecture?
- What are some examples of the CNN LSTM?
- What benefits does the CNN LSTM provide?
- How can the CNN LSTM architecture be implemented in Keras?

## Experiment

Design and execute an experiment that demonstrates a CNN LSTM on a sequence prediction problem with spatial input.

## Further Reading

- [Sequence Classification with LSTM Recurrent Neural Networks in Python with Keras](#)

# Lesson 8: How to develop Encoder-Decoder LSTMs

## Goal

The goal of this lesson is to learn how to develop encoder-decoder LSTM models.

## Questions

- What are sequence-to-sequence (seq2seq) prediction problems?
- What are the difficulties of using a vanilla LSTM on seq2seq problems?
- What is the encoder-decoder LSTM architecture?
- What are some examples of encoder-decoder LSTMs?
- What are the benefits of encoder-decoder LSTMs?
- How can encoder-decoder LSTMs be implemented in Keras?

## Experiment

Design and execute an experiment that demonstrates an encoder-decoder LSTM on a sequence-to-sequence prediction problem.

## Further Reading

- [How to Use the TimeDistributed Layer for Long Short-Term Memory Networks in Python](#)
- [How to Learn to Add Numbers with seq2seq Recurrent Neural Networks](#)
- [How to use an Encoder-Decoder LSTM to Echo Sequences of Random Integers](#)

## Lesson 9: How to develop Bi-directional LSTMs

### Goal

The goal of this lesson is to learn how to develop Bidirectional LSTM models.

### Questions

- What is a bidirectional LSTM?
- What are some examples where bidirectional LSTMs have been used?
- What benefit does a bidirectional LSTM offer over a vanilla LSTM?
- What concerns regarding time steps does a bidirectional architecture raise?
- How can bidirectional LSTMs be implemented in Keras?

### Experiment

Design and execute an experiment that compares forward, backward, and bidirectional LSTM models on a sequence prediction problem.

## Further Reading

- [How to Develop a Bidirectional LSTM For Sequence Classification in Python with Keras](#)

## Lesson 10: How to develop LSTMs with Attention

### Goal

The goal of this lesson is to learn how to develop LSTM models with attention.

### Questions

- What impact do long sequences with neutral information have on LSTMs?
- What is attention in LSTM models?
- What are some examples where attention has been used in LSTMs?
- What benefit does attention provide to sequence prediction?
- How can an attention architecture be implemented in Keras?

## Experiment

Design and execute an experiment that applies attention to a sequence prediction problem with long sequences of neutral information.

## Further Reading

- [Attention in Long Short-Term Memory Recurrent Neural Networks](#)

# Lesson 11: How to develop Generative LSTMs

## Goal

The goal of this lesson is to learn how to develop LSTMs for use in generative models.

- What are generative models?
- How can LSTMs be used as generative models?
- What are some examples of LSTMs as generative models?
- What benefits do LSTMs have as generative models?

## Experiment

Design and execute an experiment to learn a corpus of text and generate new samples of text with the same syntax, grammar, and style.

## Further Reading

- [Text Generation With LSTM Recurrent Neural Networks in Python with Keras](#)

## Advanced

The lessons in this section are designed to teach you how to get the most from your LSTM models on your own sequence prediction problems.

# Lesson 12: How to tune LSTM hyperparameters

## Goal

The goal of this lesson is to learn how to tune LSTM hyperparameters.

## Questions

- How can we diagnose over-learning or under-learning of an LSTM model?
- What are two schemes for tuning model hyperparameters?
- How can model skill be reliably estimated given LSTMs are stochastic algorithms?
- List LSTM hyperparameters that can be tuned, with examples of values that could be evaluated for:
  - Model initialization and behavior.
  - Model architecture and structure.
  - Learning behavior.

## Experiment

Design and execute an experiment to tune one hyperparameter of an LSTM and select the best configuration.

## Further Reading

- [How to Evaluate the Skill of Deep Learning Models](#)
- [How to Tune LSTM Hyperparameters with Keras for Time Series Forecasting](#)
- [How to Grid Search Hyperparameters for Deep Learning Models in Python With Keras](#)
- [How To Improve Deep Learning Performance](#)

# Lesson 13: How to update LSTM models

## Goal

The goal of this lesson is to learn how to update LSTM models after new data becomes available.

## Questions

- What are the benefits of updating LSTM models in response to new data?
- What are some schemes for updating an LSTM model with new data?

## Experiment

Design and execute an experiment to fit an LSTM model to a sequence prediction problem that contrasts the effect on the model skill of different model update schemes.

## Further Reading

- [How to Update LSTM Networks During Training for Time Series Forecasting](#)

# Lesson 14: How to make predictions with LSTMs

## Goal

The goal of this lesson is to learn how to finalize an LSTM model and use it to make predictions on new data.

## Questions

- How do you save model structure and weights in Keras?
- How do you fit a final LSTM model?
- How do you make a prediction with a finalized model?

## Experiment

Design and execute an experiment to fit a final LSTM model, save it to file, then later load it and make a prediction on a held back validation dataset.

## Further Reading

- [Save and Load Your Keras Deep Learning Models](#)
- [How to Train a Final Machine Learning Model](#)

## The End!

***(Look How Far You Have Come)***