

# How to get started with Deep Learning on your own:

##Quick-start:

This is a quick-start for advanced users who just want to get going. If you'd like a more detailed step-by-step, keep reading.

- 1) Use the nvidia AML on AWS ( 10 minutes):  
<https://github.com/NVIDIA/nvidia-docker/wiki/Deploy-on-Amazon-EC2>
- 2) Get started with nvidia-docker (5 minutes):  
<https://github.com/NVIDIA/nvidia-docker>
- 3) Get started with the Nvidia DIGITS container ( 5 minutes):  
<https://github.com/NVIDIA/nvidia-docker/wiki/DIGITS>
- 4) Download a dataset (10+ minutes):  
<http://image-net.org/download-images>
- 5) Do Deep Learning ( 10+ minutes):  
<https://www.nvidia.com/dli>

## Introduction

Congratulations, you are only a few steps away from building your very own AI. This may be your first time training a neural network or you may have just taken one of the many courses offered by the Deep Learning Institute. Either way you are interested in turning some “problem” into an “AI problem”.

Before progressing further you'll want to make sure you understand the basics workflows of deep learning. Feel free to run through the free fundamentals of deep learning course. In this course you will be working in a development environment identical to the one we are about to set up.

## Overview

To get started with your AI problem we are going to need to run through a few steps.

- 1) Get setup with a Linux machine with up-to-date Nvidia drivers
- 2) Install nvidia-docker and learn the docker basics
- 3) Start a pre-made Deep Learning Container with the framework of your choice.
- 4) Download a dataset
- 5) Do deep learning

### Setting up a Linux machine

#### Using NGC

The fastest way to get started in the cloud is to sign up for NGC, the Nvidia GPU Cloud. If you would like to use this service, go here to sign up and follow the walkthrough on their site:

#### #### In the cloud

The easiest way to setup a Linux machine with Nvidia drivers is to setup a cloud account on Amazon and deploy the pre-built Nvidia AML.

You can follow these directions to do this:

<https://github.com/NVIDIA/nvidia-docker/wiki/Deploy-on-Amazon-EC2>

You may also be interested in signing up for the Nvidia GPU Cloud, which provides a light wrapper around Amazon and other Cloud Service Providers to ease deployment:

<https://www.nvidia.com/en-us/gpu-cloud/>

#### #### On your own machine

If you are looking to use your own training you will have to manually install the latest Nvidia drivers along with the Nvidia convolutional network library cudnn.

Drivory install:

<http://www.nvidia.com/Download/index.aspx>

Cudnn binary:

<https://developer.nvidia.com/rdp/form/cudnn-download-survey>

### ## Installing nvidia-docker

#### #### In the cloud

If you are using a Nvidia cloud image, nvidia-docker should already be installed and up-to date.

#### #### On your own machine

You can follow the installation instructions on this page to get the latest version of nvidia-docker installed:

<https://github.com/NVIDIA/nvidia-docker>

### ## Understanding Docker and nvidia-docker

This blog post will give you a good understanding of nvidia-docker.;

<https://devblogs.nvidia.com/parallelfforall/nvidia-docker-gpu-server-application-deployment-made-easy/>

### ## Starting a Deep Learning container

If this is your first time doing deep learning, we recommend getting started with a tool called DIGITs. DIGITs is an application that simplifies the deep learning process and lets you use do deep learning in Tensorflow, Pytorch, or Caffe (to learn more see some related DLI courses).

This wiki page will show you how to use a container for deep learning (it's really only a single docker command!):

<https://github.com/NVIDIA/nvidia-docker/wiki/DIGITS>

If you are interested in learning more about DIGITS it is completely free, open source, and additional details can be found on GitHub here:

<https://github.com/NVIDIA/DIGITS>

At this point you should have DIGITS running, the setup is over.

Now you're ready to download a dataset and get started!

### ## Downloading data

If you have your own dataset for the problem you are trying to solve that's great. If not you can download popular image, text, segmentation, and signals datasets from one of these websites:

<https://www.kaggle.com/datasets>

<https://research.google.com/youtube8m/>

<http://www.image-net.org/>

### ### Passing data between environments

Manipulating data prior to training is not a trivial task. While some datasets come tailored to learning, most data is much messier. No matter what *state* your data is in, it is of no use unless it is in the same environment you plan to train it. While there are many ways to ensure data is in the right environment, here are a few:

- 1) Downloading a dataset directly to a training environment using the [wget](#) command.
- 2) Moving between machines using [secure copy and/or ssh](#).

### ## Deep Learning

At this point I'd hope you know what problem you want to solve, if however you are still lost, fear not. The Deep Learning Institute is here to help. Take a look at our course catalog and find some lab material that is relevant to your problem. Are you trying to determine what an image is? Are you trying to locate something in an image? Are you trying to generate text? Translate or classify text? Are you trying to track things in real-time video? Do you have a lot of data and need to detect anomalous patterns? All totally doable with Deep Learning and we can show you how.

<http://www.nvidia.com/dli>