# Setting up python, tensorflow, virtualenv etc.

Making a virutalenv in bash

cd projectFolderName

$ virtualenv -p python virtualEnvName

Activate virtualenv in cmd line, after cding into project folder

.\**virtualEnvName**\Scripts\activate

Install packages within a virtual environment without affecting the host system setup. Start by upgrading pip:

pip install --upgrade pip  
  
pip list  # show packages installed within the virtual environment

And to exit virtualenv later:

deactivate  # don't exit until you're done using TensorFlow

Install Tensorflow

pip install --upgrade tensorflow

Install matplotlib

pip install --upgrade matplotlib

<https://matplotlib.org/users/installing.html>

Darknet is a neural net library for c/c++

darkflow is the python port of darknet

# Intro to Machine Learning PDF

## Types of Machine Learning

### Supervised Learning

* trains algo based on example input and output data that was labeled by humans
* the computer is provided with example inputs that are labeled with their desired outputs.
* A common use case of supervised learning is to use historical data to predict statistically likely future events. It may use historical stock market information to anticipate upcoming fluctuations, or be employed to filter out spam emails. In supervised learning, tagged photos of dogs can be used as input data to classify untagged photos of dogs.

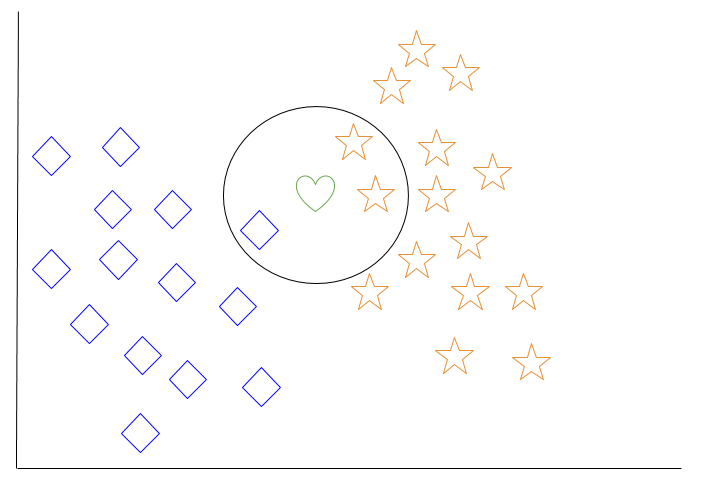
### Unsupervised Learning

* data is unlabeled, so the learning algorithm is left to find commonalities among its input data
* Goal: discovering hidden patterns within a dataset, but it may also have a goal of feature learning, which allows the computational machine to automatically discover the representations that are needed to classify raw data

## Approaches

* machine learning is mostly stats
* **Correlation** is a measure of association between two variables that are not designated as either dependent or independent
* **Regression** at a basic level is used to examine the relationship between one dependent and one independent variable
  + **enables prediction capabilities**

### k-NN: k-nearest neighbor

* pattern recognition model that can be used for classification as well as regression
* the k in k-nn Is a small +’ve int
* In either classification or regression, the input will consist of the k closest training examples within a space.
* In k-NN classification, the output is class membership
* a new object will be assigned to the class of the most common class among it’s k-nearest neighbhors
* If k=3 below, then the heart will be classified as a star
* aka “lazy-learning”

### Decision Tree Learning

* basically sorting the input into an output using a decision tree

### Deep Learning

* attempts to imitate human brain processing of light = sight
* . The output of one layer serves as the input of the successive layer.
* In deep learning, algorithms can be either supervised and serve to classify data, or unsupervised and perform pattern analysis.
* computer vision, speech recognition, IBM Watson

## Programming Languages

* Python
* Java: for security stuff
* R
* C++: for AI in game or robots
  + ML Libraries: mlpack, dlib, shark

# Building a Neural Network for Handwritten Digits

individual neurons are connected in layers with weights assigned to determine how the neuron responds when signals are propagated through the network

*one-hot encoding means the binary vector of [0, 0, 0, 1, 0, 0, 0, 0, 0, 0] represents the number 3*

The archetiecture of a neural network refers to things like the # of layers in the network, number of neurons in each layer, and how they are linked b/w layers

Common Layers:input, hidden, output

A hidden layer means all layers b/w input and output

A *shallow* neural network has only one hidden layer, while a *deep* neural network has multiple

Shallow: often acounter overfitting ( network essentially memorizes the training data that it has seen, and is not able to generalize the knowledge to new data.)

Learning rate: how much the parameters will adjust at each step of the learning process.

* After each pass during training, we tune the weights slightly, to try and reduce loss
* Larger learning rates can converge faster, but also have the potential to overshoot the optimal values as they are updated

# of iterations: how many times we go through the training step,

batch size: how many training examples we are using at each step

A popular choice of loss function in TensorFlow programs is cross-entropy, also known as log-loss, which quantifies the difference between two probability distributions (the predictions and the labels). A perfect classification would result in a cross-entropy of 0, with the loss completely minimized.

A process named gradient descent optimization is a common method for finding the (local) minimum of a function by taking iterative steps along the gradient in a negative (descending) direction