GENERAL ASSEMBLY

More Unsupervised Learning Methods

Patrick Smith

LEARNING OBJECTIVES

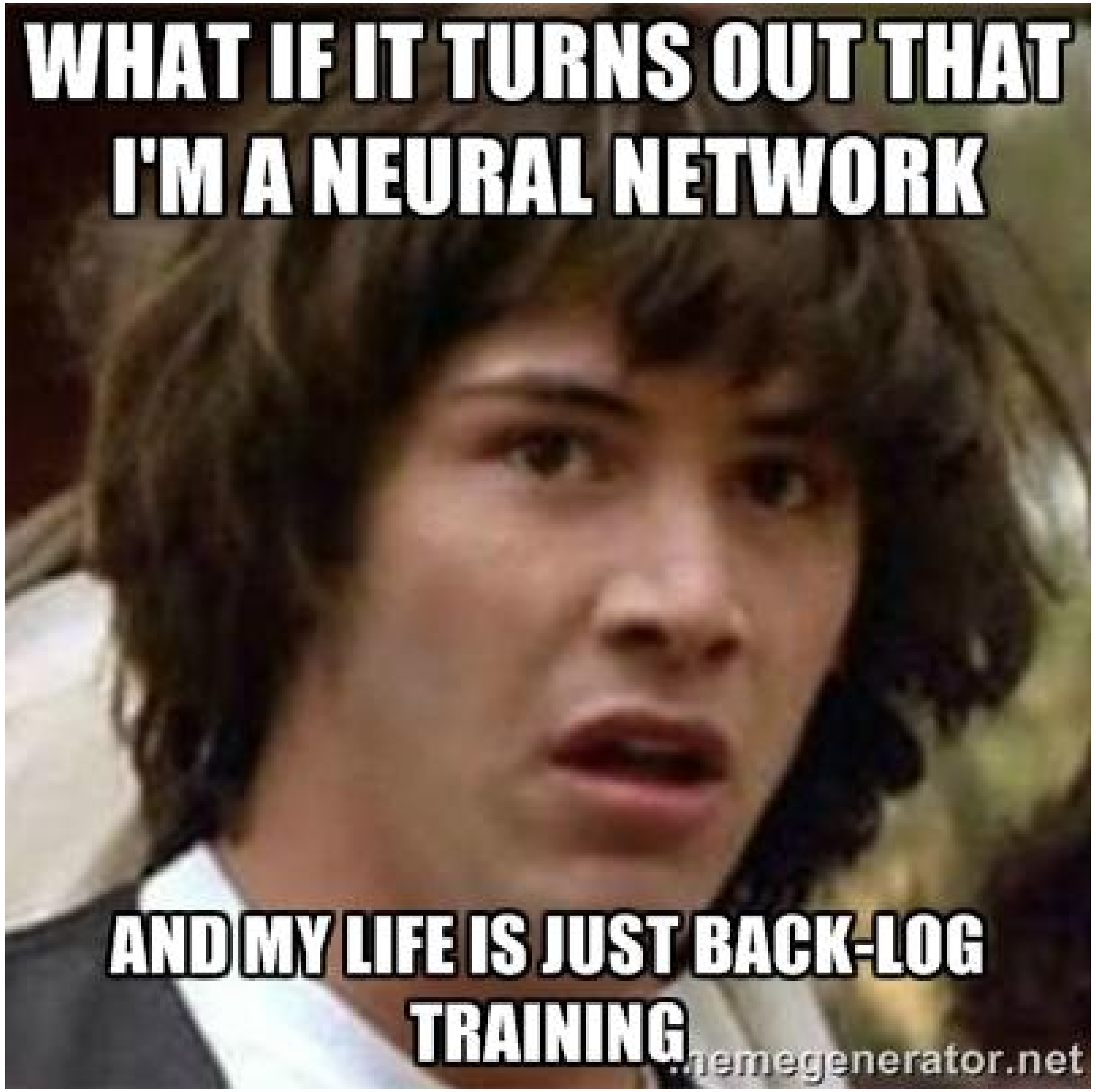
- Identify + Learn Additional Unsupervised Learning Methods
- State when/why each unsupervised learning method is utilized

PRE-WORK

- Have a working understand of scikit learn and numpy
- Be able to create functions from scratch in python
- Have a basic understanding of linear algebra concepts such as matrices, eigenvalues, and eigenvectors

OPENING

More Unsupervised Learning Methods



More Unsupervised Learning Methods

"I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted."

More Unsupervised Learning Methods

"I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted."

Alan Turing, 1947

More Unsupervised Learning Methods

So far this week, we've learned about different forms of clustering - including hierarchical, k-means, and DBSCAN, as well as PCA.

Today, we're going to walk through some other useful unsupervised learning methods in scikit to add to your data science toolbox.

Overview: Other Unsupervised Learning Methods

More Unsupervised Learning Methods

“Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed”

More Unsupervised Learning Methods

“Machine learning is the] field of study that gives computers the ability to learn without being explicitly programmed” (Arthur Samuel, 1959)

Manifold Learning

Manifold Learning

Manifold learning is another form of dimensionality reduction.

This week we learned about dimensionality reduction in a general sense; however, what happens when we want to perform dimensionality reduction on a non-linear dataset? That's where Manifold Learning comes in.

Manifold Learning: Recall PCA

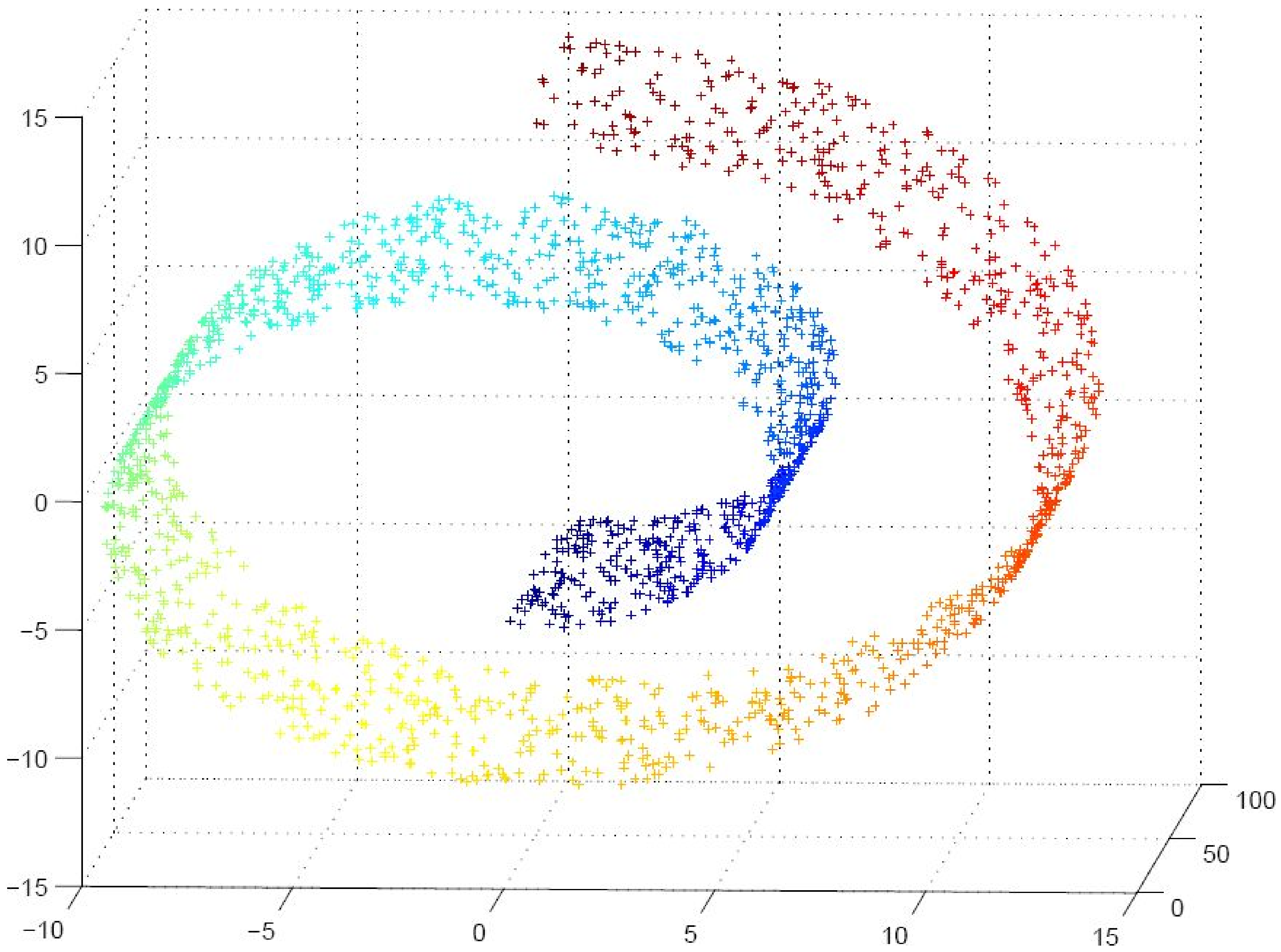
There are 3 assumptions that PCA makes:

- ***Linearity:*** Our data does not hold nonlinear relationships.
- ***Large variances define importance:*** our dimensions are constructed to maximize remaining variance.
- ***Principal components are orthogonal:*** each component (columns of Z) is completely un-correlated with the other

Manifold Learning:

- Manifold learning is a way to generalize a dimensionality reduction technique like PCA to take into account non-linear structures in the data.
- When we set up Manifold Learning, it learns the structure of a highly dimensional data from the data itself, and then performs the reduction.

More Unsupervised Learning Methods



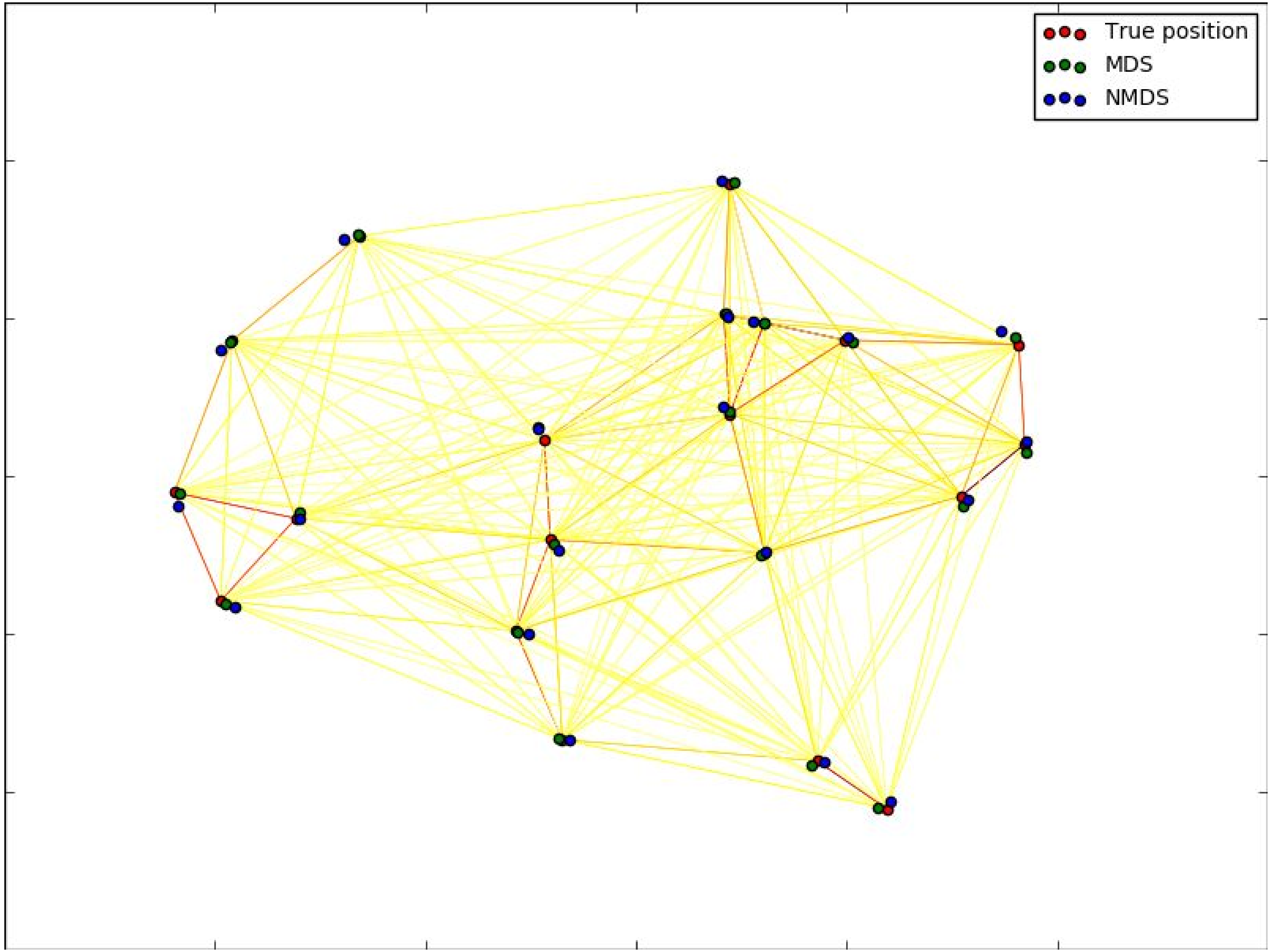
Manifold Learning: Isomapping

- Isomapping is one of the oldest and simplest implementations of manifold learning - it seeks to reduce the dimensionality while mainting the distance between points.
- When we set up Manifold Learning, it learns the structure of a highly dimensional data from the data itself, and then performs the reduction.

Manifold Learning: Multidimensional Scaling

- In a basic sense, multidimensional scaling is a method to provide a visual representation of similarities among a group of data.
- With multidimensional scaling, we're taking a highly dimensional data set and representing it in as little as two dimensions.

More Unsupervised Learning Methods



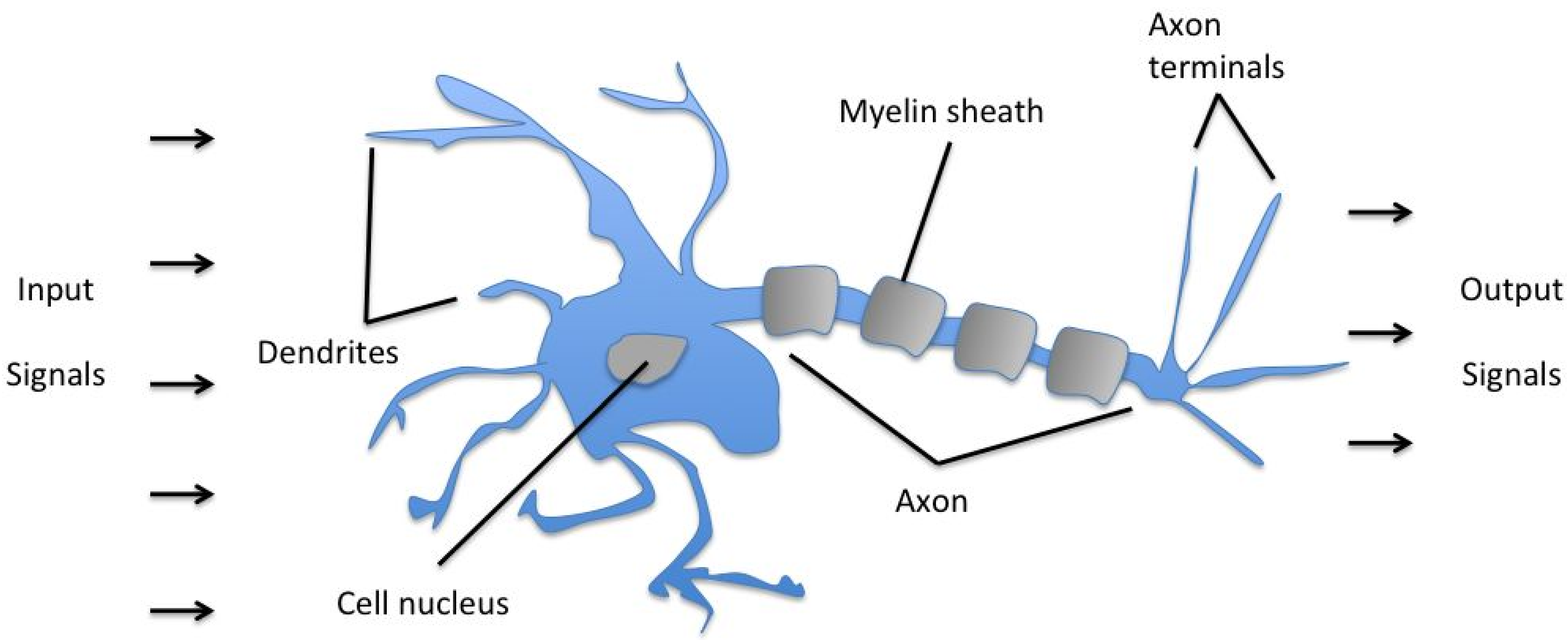
Artificial Neural Networks

More Unsupervised Learning Methods

Artificial Neural Networks are simulated models of neural networks in the brain.

In a practical sense, neural networks have become most useful to tackle problems that are simple for the human brain but difficult for machines, IE: pattern recognition.

More Unsupervised Learning Methods

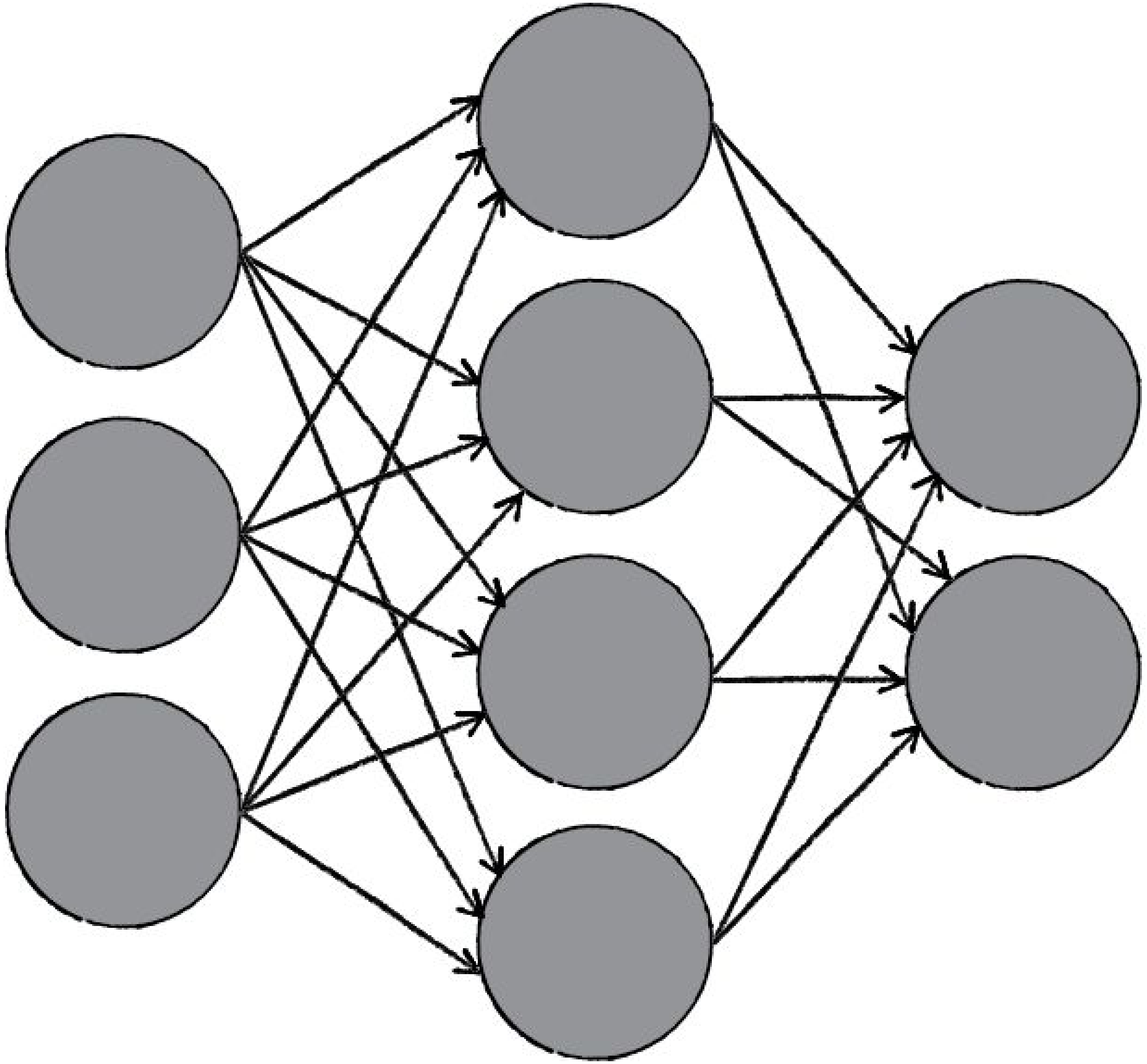


Schematic of a biological neuron.

More Unsupervised Learning Methods

- When a neural network receives information, it processes it through a system of nodes
- While this may seem simplistic, a great, intertwined network of these nodes are utilizing parallel computations can be exceptionally powerful.
- Neural networks learn and adapt just the way a human brain would to changing information.

More Unsupervised Learning Methods



More Unsupervised Learning Methods

We can implement artificial neural networks as a supervised learning method or an unsupervised learning method.

More Unsupervised Learning Methods

In the former, we train the network to learn.

Let's take the example of a segment of farm animals - a cow, horse, pig, and goat. We can tell train the network to recognize each animal and then have it provide us it's best guesses, and in return we give it the correct answers, and the network then compares its answers to the correct answers and learns the correction classifications.

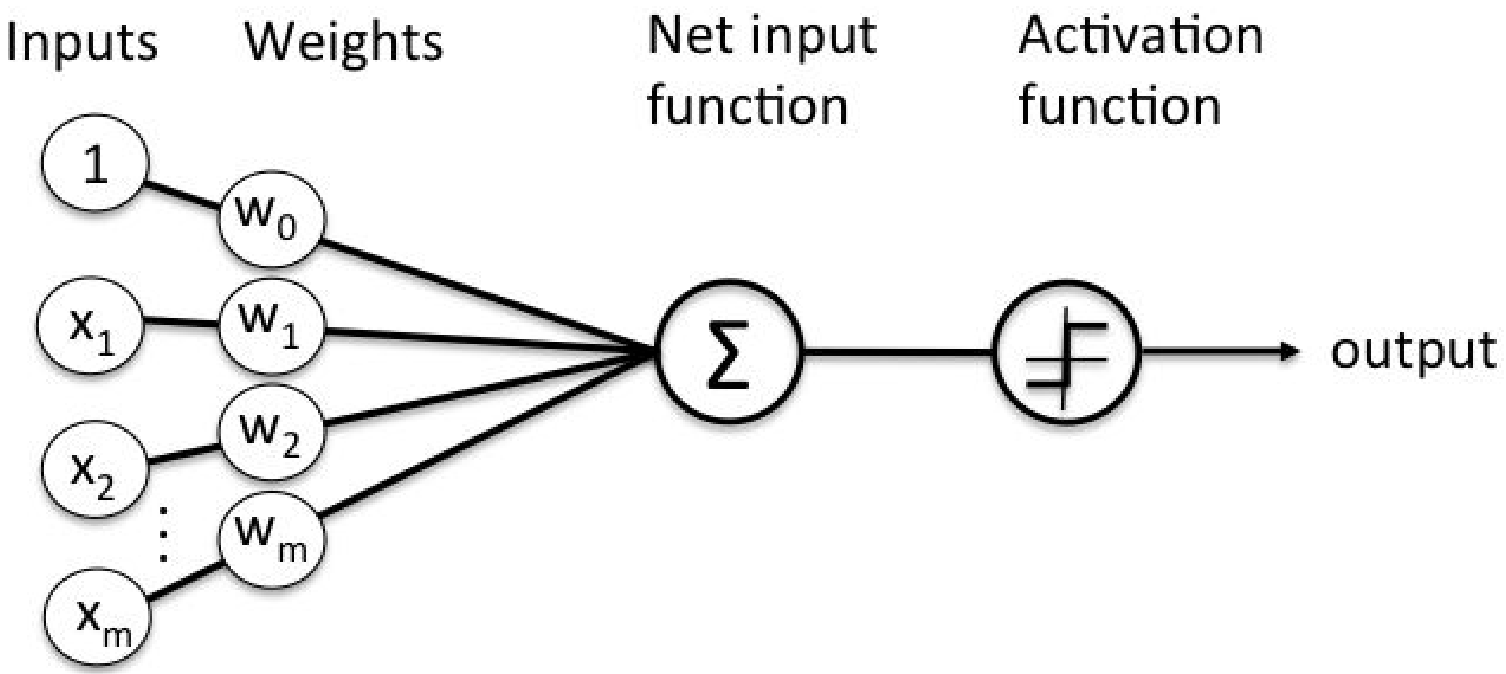
More Unsupervised Learning Methods

With unsupervised neural networks, we train the neural network in a situation where there is not a set of known answers.

In scikit, we can implement a form of unsupervised neural networks by utilizing Restricted Boltzmann Machines, which are models that learn a probability distribution

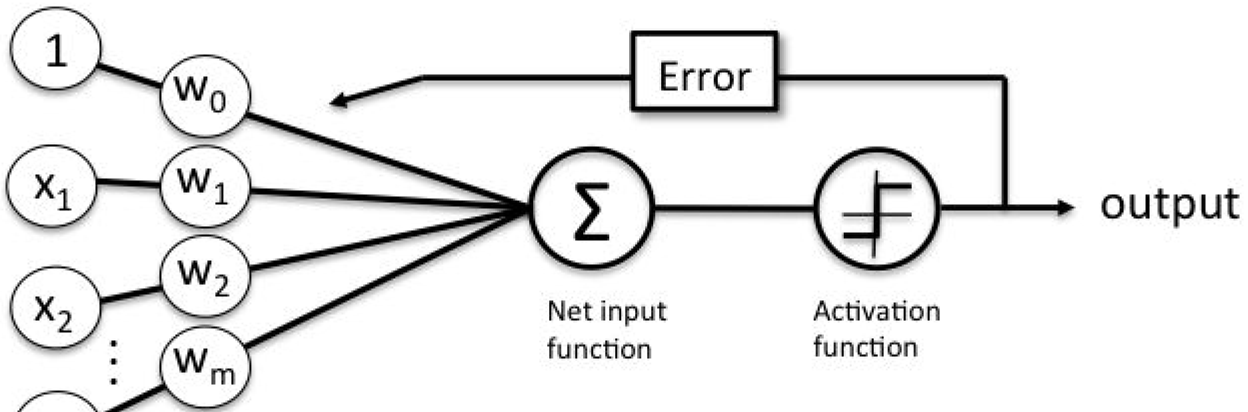
Guided Practice: Constructing a Neural Network

More Unsupervised Learning Methods

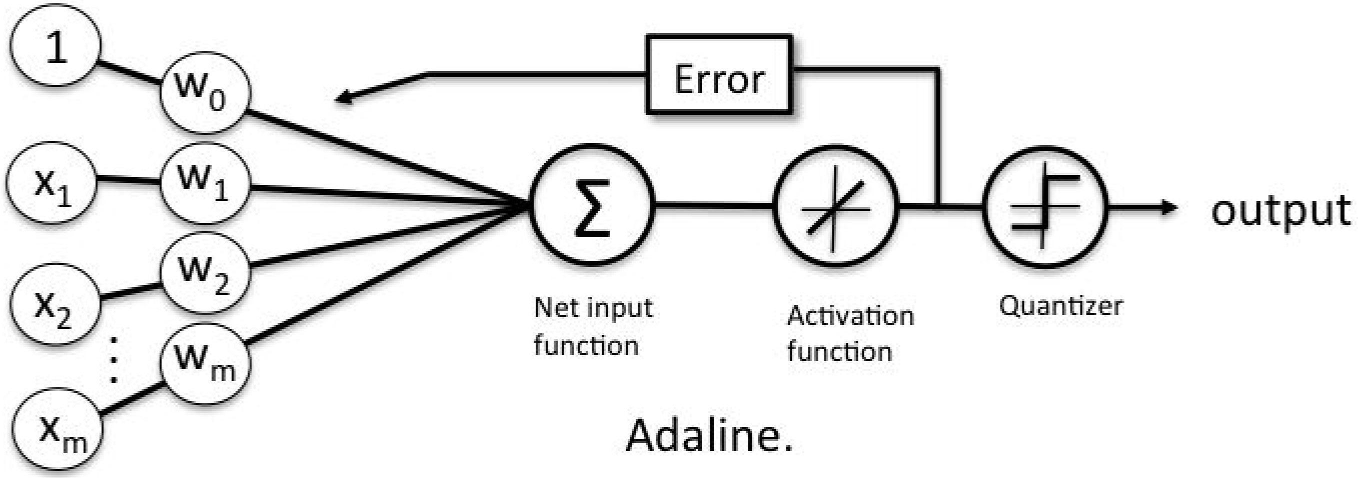


Schematic of Rosenblatt's perceptron.

More Unsupervised Learning Methods

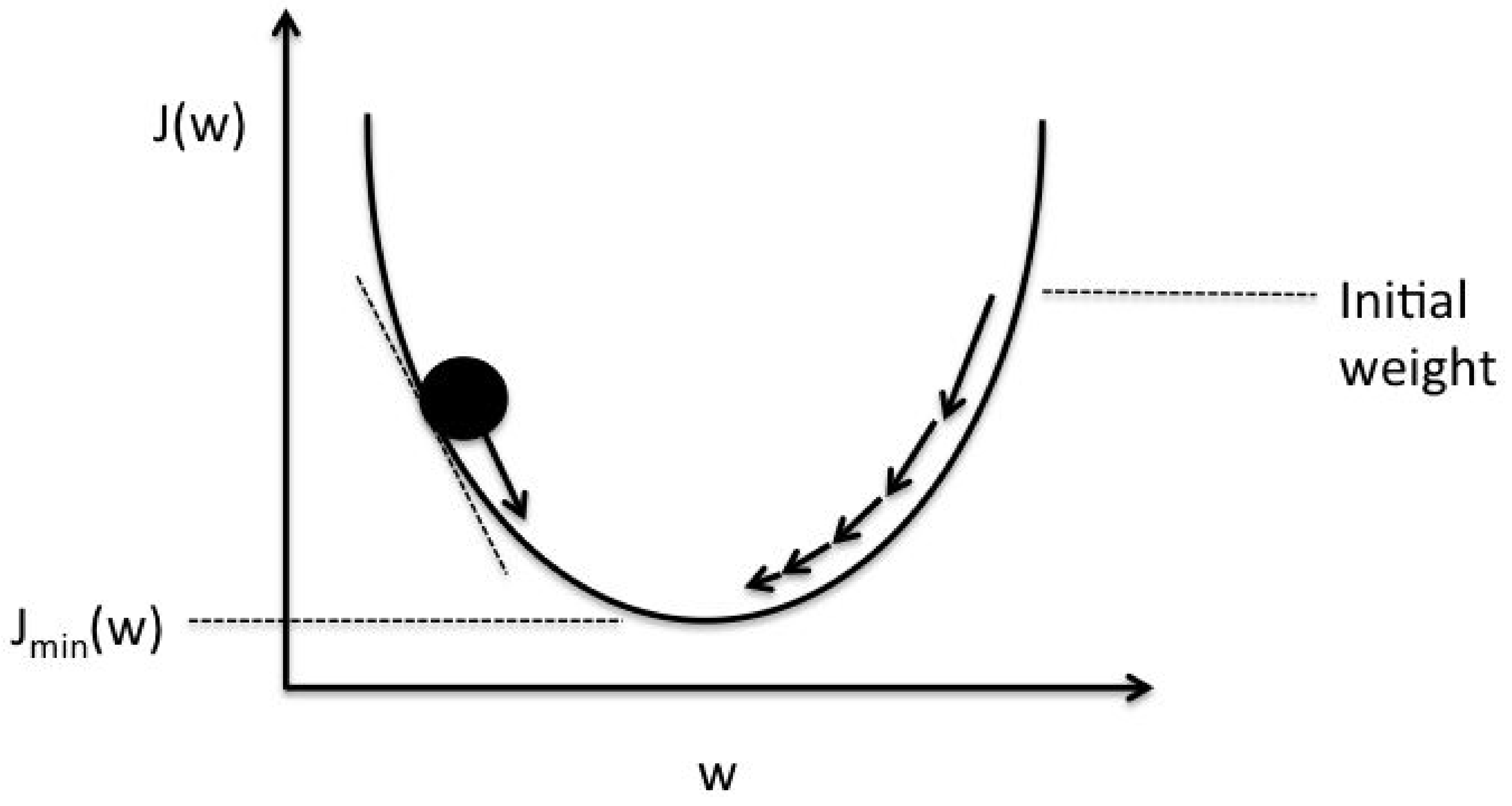


Perceptron rule.



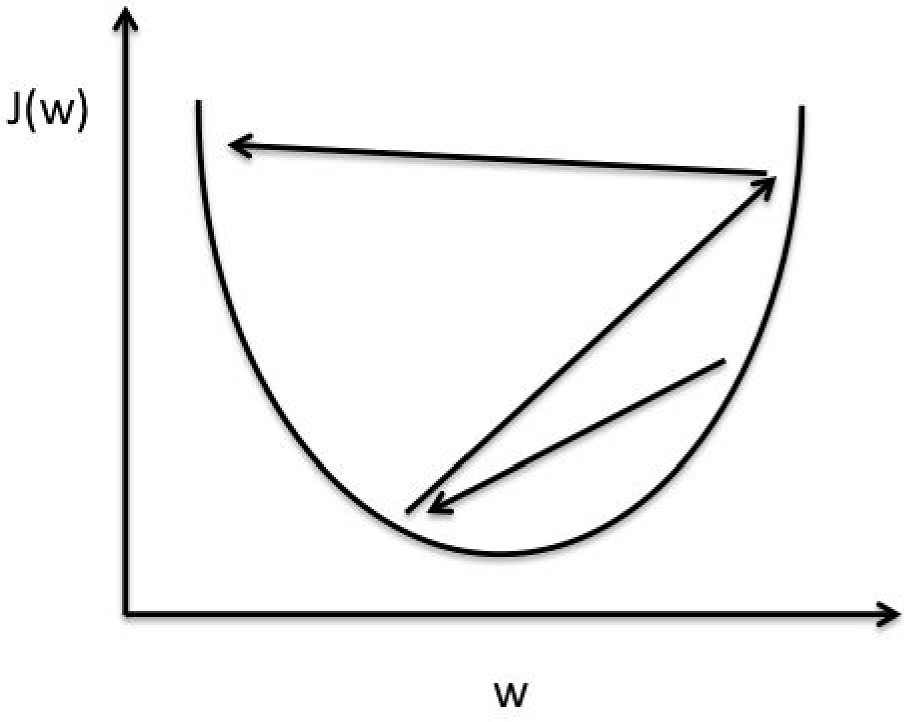
Adaline.

More Unsupervised Learning Methods

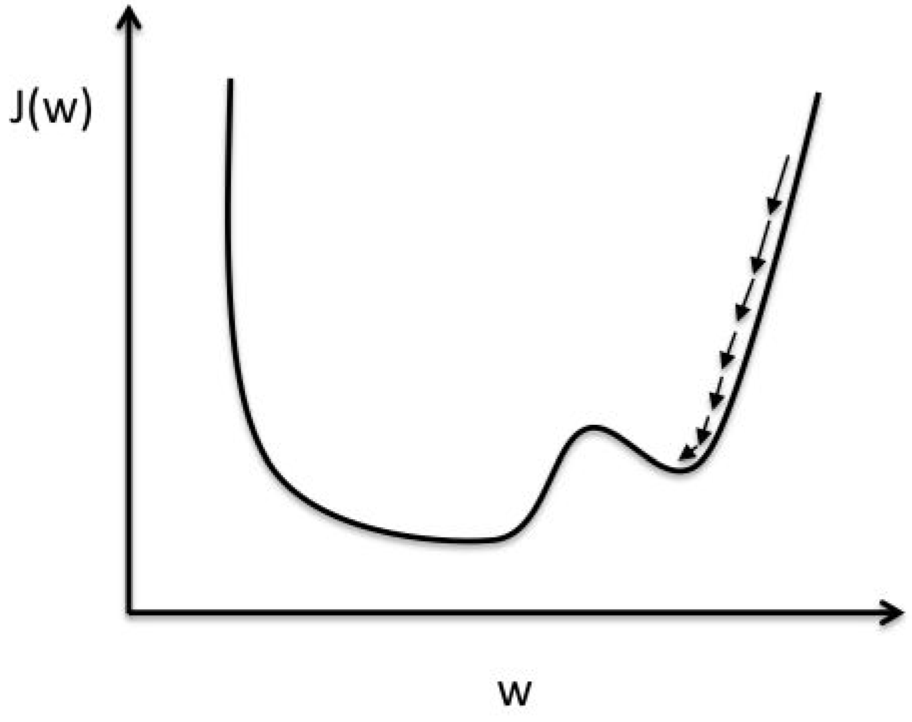


Schematic of gradient descent.

More Unsupervised Learning Methods

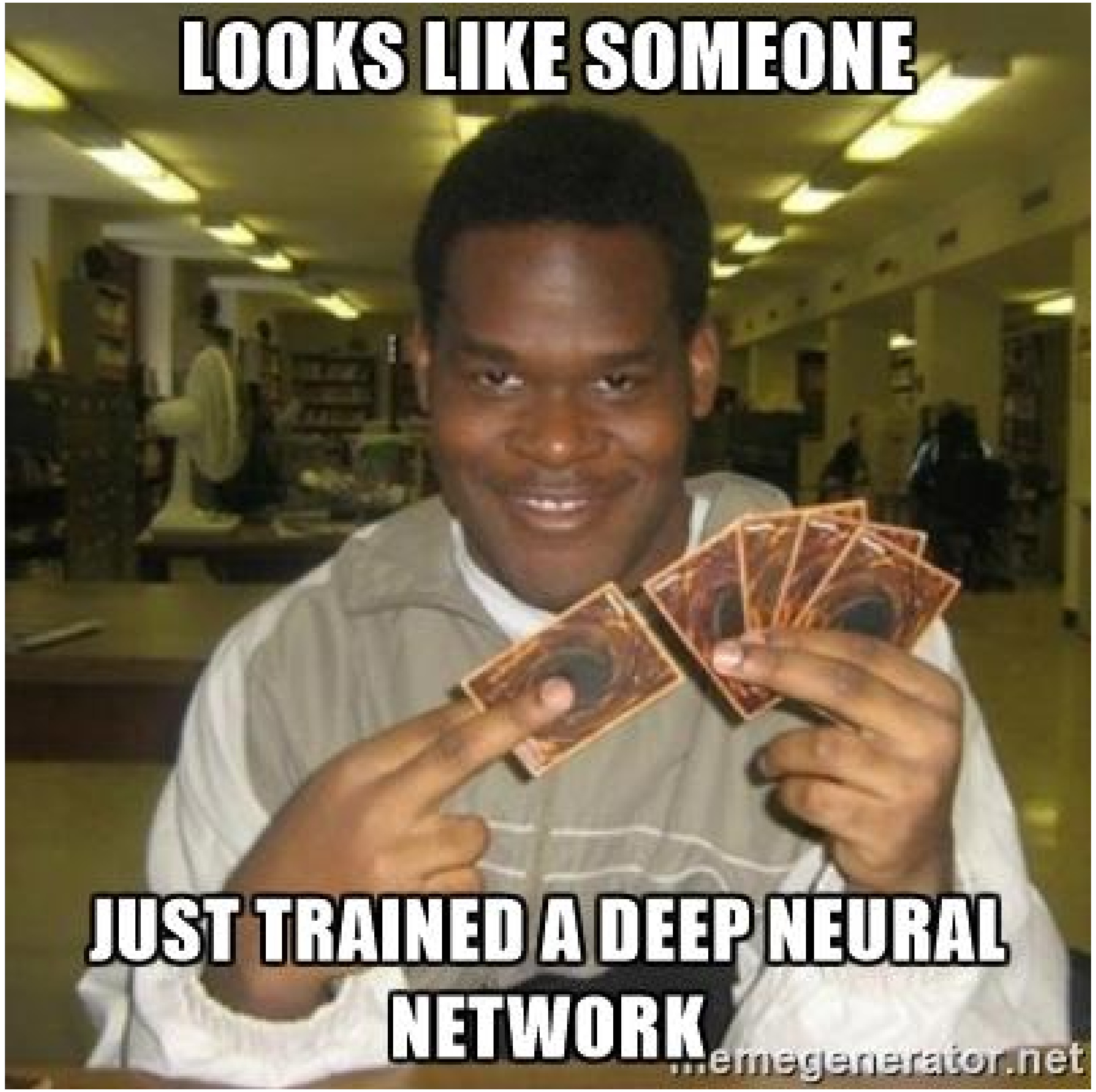


Large learning rate: Overshooting.



Small learning rate: Many iterations until convergence and trapping in local minima.

More Unsupervised Learning Methods



Conclusion

Q & A