Deep Learning in Trading

www.qplum.co

1620 Plaza 5 Jersey City, NJ 07311

(201) 377-2302 contact@qplum.co (201) 604-5422

Is deep learning used in trading currently?

Absolutely and at scale! Let's look at the few facts we can get hold of.

1. Peter Bown and Robert Mercer, two successful linguists quit their jobs to join RenTech. In the last two years virtually every deep-learning prof and PhD student has received calls from Point72 to join their trading research team.
2. Deep Learning is really young and it's changing a lot. LSTM (Long Short Term Memory), which is the state of the art in contextual learning, wasn't used at scale till two years ago. It will take decades for the methods to be commoditized to an extent that books or WSJ articles are written about it.
3. The basic ideas have been used for years. Deep Learning just formalizes them into a process.

What is the non-math essence of deep learning?

Let's look at how humans learn and we can get better intuition about this. (Appendix A)

Traditional methods work with a lot of data. We need to develop the ability to learn from very few data points.

Let's discuss the structure of the data, before we learn how to predict trades from it.

An example of a deep neural network

Layer 1 is the input layer. Let's feed in daily returns of 5,000 stocks.

Layer 2 has one node.

Layer 3 has 5,000 nodes, that we train to match the input set.

Layer 4 is the output layer. It has 5,000 nodes, in which the ith node has connections with the ith node in the input layer and the ith node in Layer 3. It is trained to fit what will happen the ith stock tomorrow.

Quiz: Have we seen this model of trading before? (Hint: CAPM maybe? Refer Appendix D)

Why go deep and how has that changed everything?

The biggest problem in finance is overfitting. Neural Nets have been around since the sixties but they royally failed. The problem was that we started with very wide networks with less number of hidden layers. Instead when we try to learn a deep neural network, we are being much more "thoughtful". We don't immediately try to dive into the problem. We try to understand the turf. Think of the million rounds of practice that is needed by a soccer player before they become good at free kicks. All this time they are trying to reduce over-fitting. Yup just that: **over-fitting.** We care about future returns. However, we don't have future data. All we have is past data. How do we learn from past data in a way that works well in future ... that's why we need to go deep.
Appendix C illustrates this in the context of capturing sentiment from text.

Let's look at the minds of a few traders and visualize their Deep Neural Networks

- An interest rate swaps trader who has seen the yield curve ten million times.
- Warren Buffett who has seen balance sheets a million times.
- Most good traders prefer to be specialized in one market or one source of alpha rather than covering a wide range of products (deep vs wide).

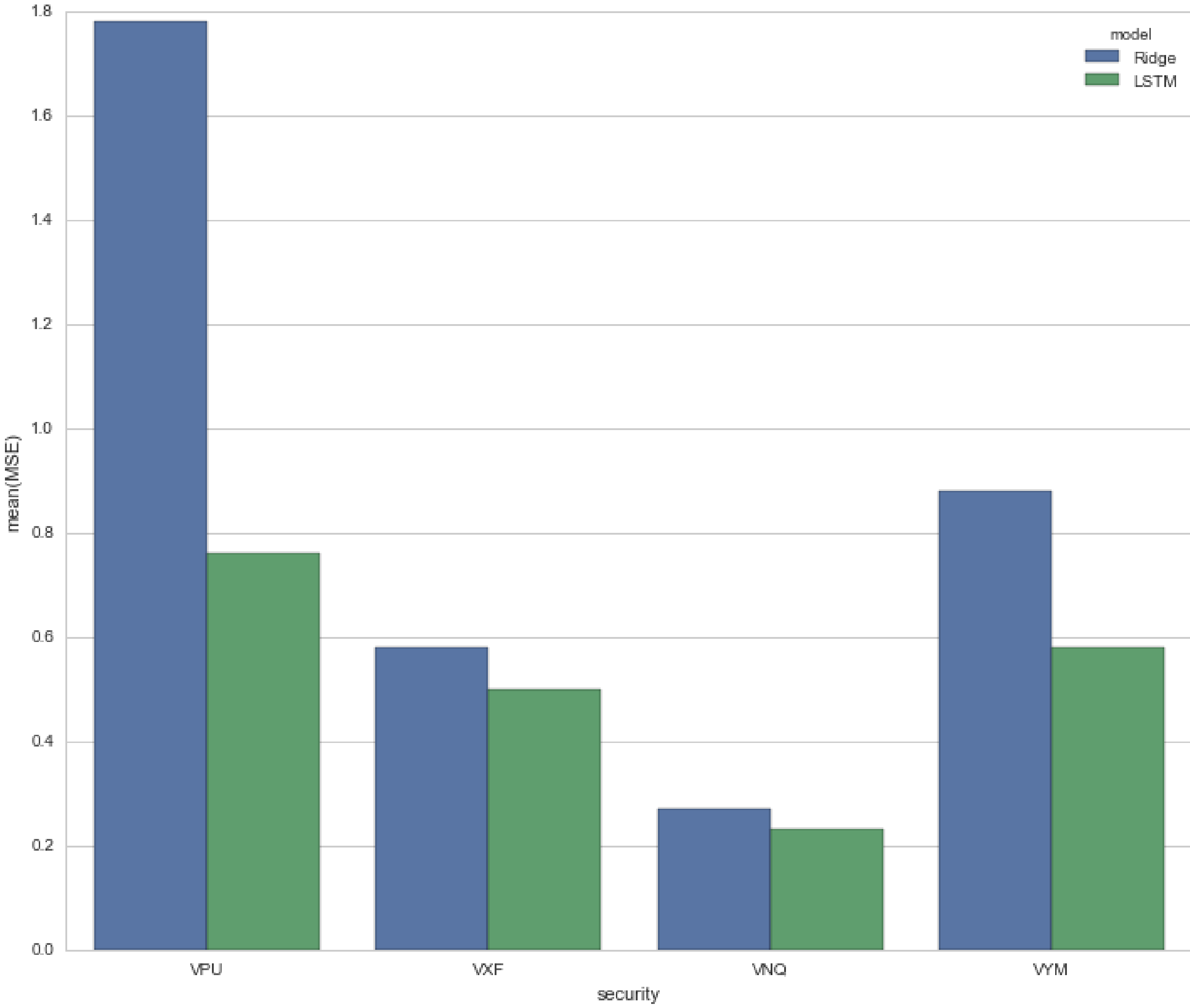
Why should a portfolio manager care?

Deep Learning is at a point where the skill set is substantially widely available to make it possible to be used. I admit that It is costly to use it but not extremely cost prohibitive. The rewards are likely to be there since not everyone knows about it. It can save costs by reducing the part of the team we dedicate to learning alphas. We can learn patterns, we did not identify previously. Also all this is very new. It is not found in traditional finance textbooks. **There is an opportunity.**

Do we use deep learning at qplum or is all this just hot air?

We use it, in multiple frequencies. These are all the areas where we use it now :

- Autoencoders to learn better summaries of the data. We build models on top of that.
- High frequency models for prediction.
- Medium frequency models, which do much better than Linear Regression models.



Our real results where we show that Deep Learning (blue) has higher accuracy than Linear Regression(green) in predicting future returns of equity ETFs.

Current state of the art

1. LSTM is the holy grail of sequence predictions.
A major part of the financial modelling is sequence prediction - whether that's volatility models or volume models or the toughest one of all - return prediction models.
This is the underlying task in such problems - Given a sequence of values, can we predict the next number in the sequence?
LSTM models naturally fit this criteria, because of its recursive nature. Additionally, the hidden state and the memory cell tremendously help retain the useful features of the sequence.

2. Feature engineering is the thing of the past in the era of neural networks.
Neural networks are really good at coming up with features on their own. A number of people in finance work day-in-day-out in coming up with features. Neural nets are poised to take over this segment of the market.

3. Neural networks provide an easy way to combine market data and other data sources.
Since neural nets work in the latent space, it's super easy to combined your market data input with other data sources you might have. That can be anything from sentiment analysis, summary of SEC filings to visual or audio inputs.

4. Additionally, neural networks make it easy to do multivariate modelling where there are a lot of relationships between inputs, and there is a time-varying nature to it.

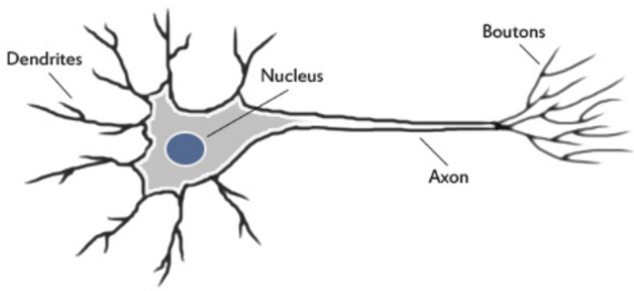
5. It's important to understand when neural networks do not work - they don't work if you don't have enough data.
Small data sets are bottlenecks when it comes to convergence. Large datasets come with computation problems.

Appendix A

How do humans learn?

- Basic unit is a neuron.
- Think of the difference between System-1 and System-2. We start with simple logic.
- We learn from mistakes. (“loss and reward build new neural pathways and weights”)
- System 2 thinking (“logical thinking”) is nothing but a deep neural networks in the brain.

Neuron



Electrical signals are input to a neuron via dendrites

If the sum of these signals exceeds a threshold the neuron fires

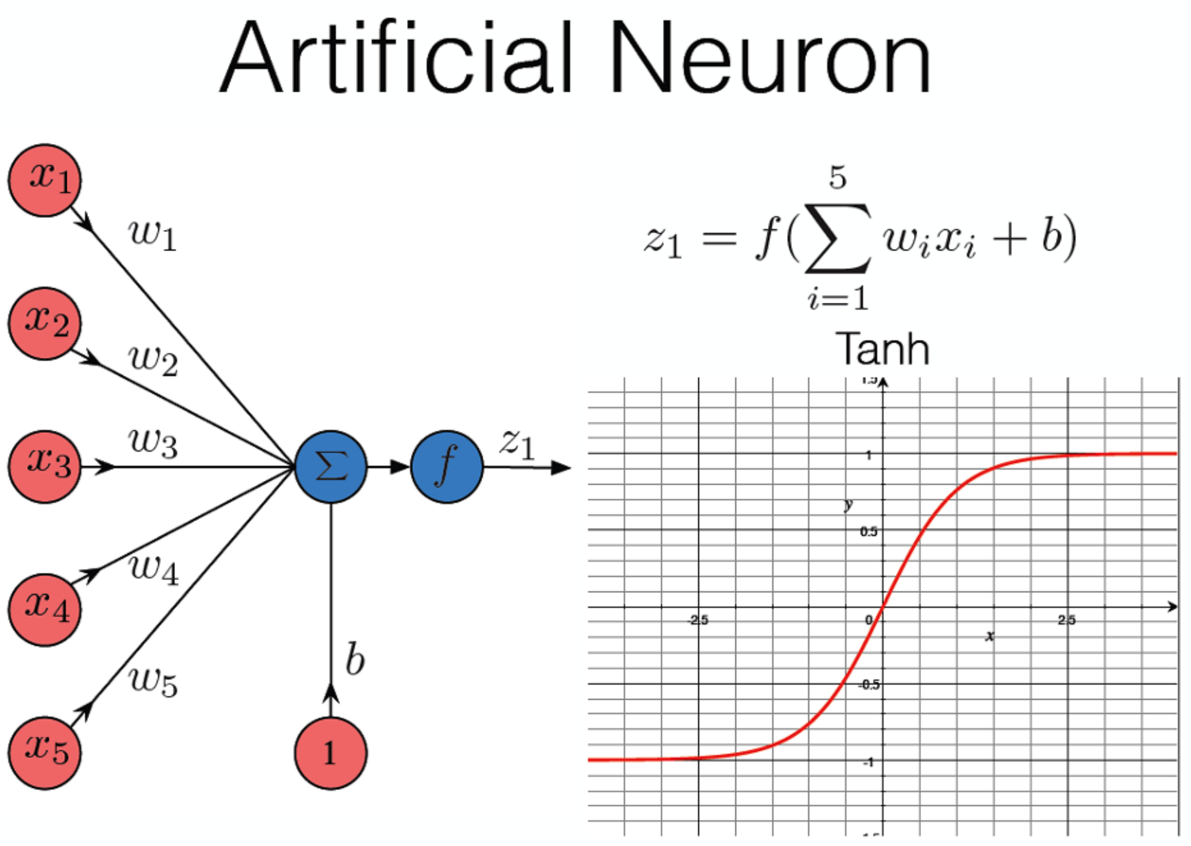
Resulting electrical signal is passed to other neurons via Axons

The way these neuron are connected and the threshold of each neuron governs how we learn

About 100 billion neurons in human brain each with about 1000 synaptic connections

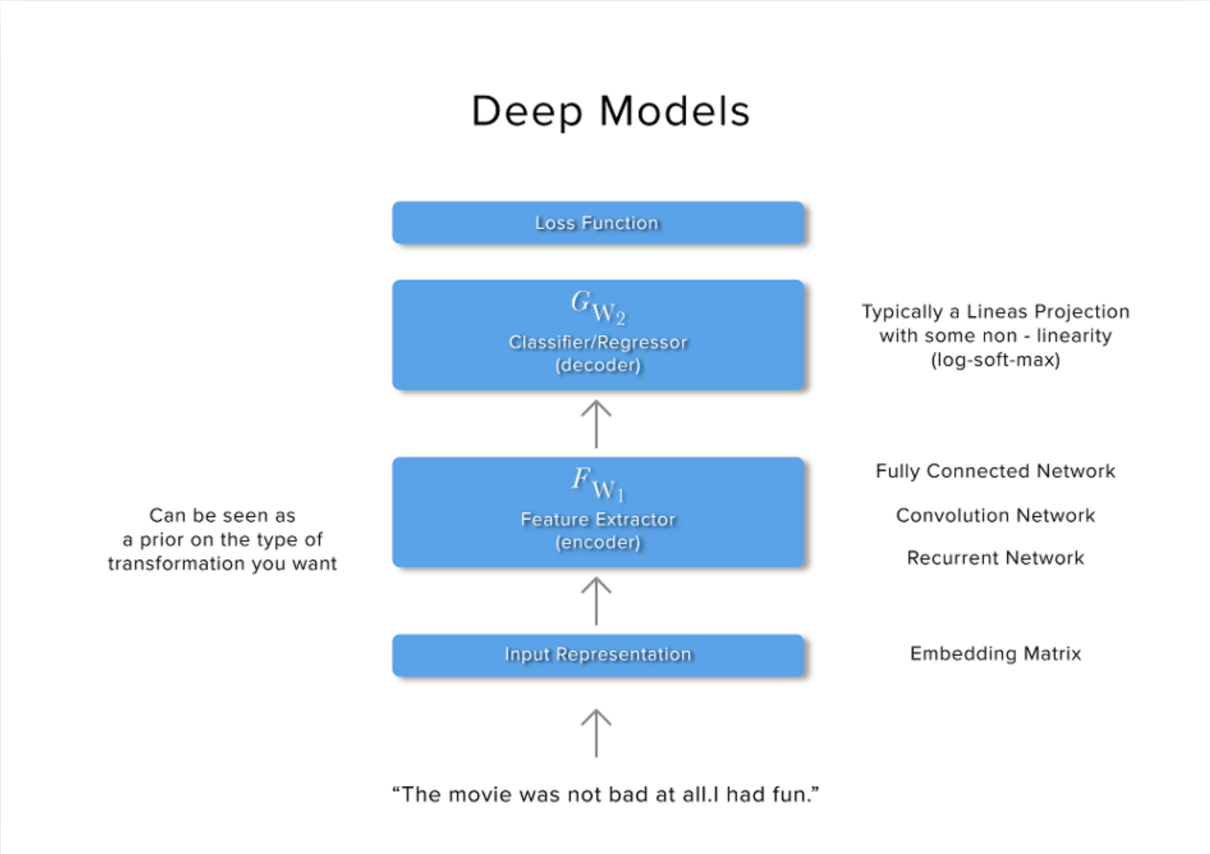
Appendix B

Artificial Neuron



Appendix C

Why go deep?



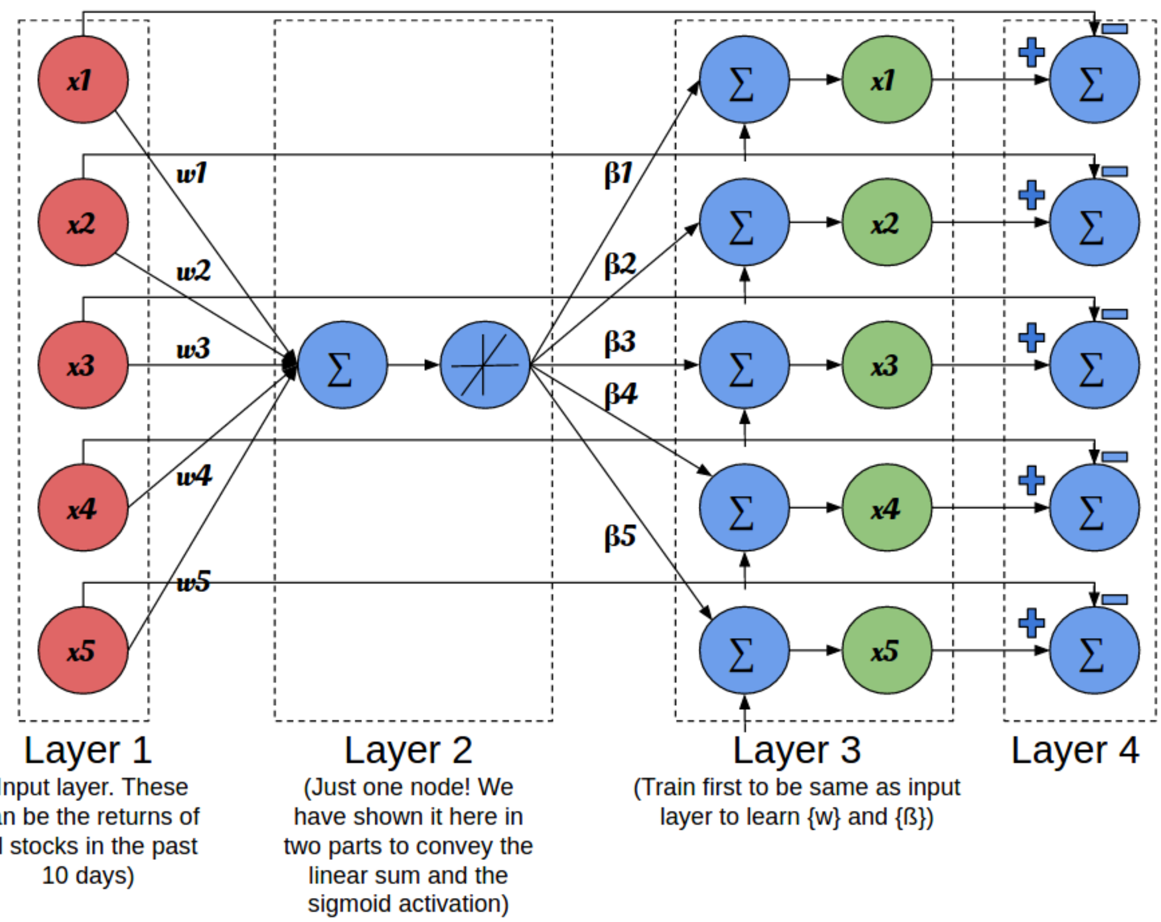
So is the movie good or bad?

In this example, we are trying to learn the context of the conversation. If we implement a one layer network and try to infer from "bad" or "not bad", the results we will get will be very noisy.

If we include the entire context of the sentence and reinforce it with the sentence after that, then the picture is much more clear. A shallow and wide network will find it very hard to train for this. This is where deep neural networks, and specially memory networks like LSTM excel.

Appendix D

CAPM model in Neural Networks



As you can see a natural representation of CAPM in a neural network setting is naturally deep and not one layer.

References

1. Deep Learning Book (2016), Ian Goodfellow, Y. Bengio, A. Courville [e-book]
2. Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
3. Coates, Adam, Honglak Lee, and Andrew Y. Ng. "An analysis of single-layer networks in unsupervised feature learning." Ann Arbor 1001.48109 (2010): 2. [pdf]
4. Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." Journal of Machine Learning Research 15.1 (2014): 1929-1958. [pdf]
5. Deep learning (2015), Y. LeCun, Y. Bengio and G. Hinton [pdf]

Disclosures

The factual information set forth herein has been obtained or derived from sources believed to be reliable but it is not necessarily all-inclusive and is not guaranteed as to its accuracy and is not to be regarded as a representation or warranty, express or implied, as to the information's accuracy or completeness, nor should the attached information serve as the basis of any investment decision.

This document is intended exclusively for the use of the person to whom it has been delivered and it is not to be reproduced or redistributed to any other person.