DEEP LEARNING FOR DERIVATIVES PRICING: FROM THEORY TO PRACTICE

Tim Wood, CQF Institute, April 6th, 2021
INTRODUCTION & OBJECTIVES

Who is this talk for?

- To see how and where DL may be applied to derivatives pricing and applications in risk management
- To develop an experimental framework to explore the technology and problem space
- By the end, the audience should have an idea of how-to set-up an experimental framework for themselves
- Additionally, they should have some feeling for the challenges and considerations w.r.t to applying this technique in practice
- As such will cover many topics but not in great depth given the time constraint
NO SMOKE WITHOUT FIRE?
Large and growing body of work

- **Deeply Learning Derivatives**, Ferguson, Green, 2018
- **Deep Hedging**, Buehler et al, 2019
- **A neural network-based framework for financial model calibration**, Oosterlee et al, 2019
- **Neural Networks with Asymptotics Control**, Anatov et al, 2020
- **Deep learning volatility: a deep neural network perspective on pricing and calibration in (rough) volatility models**, Horvath et al, 2020
- **Neural networks for option pricing and hedging: a literature review**, Ruf & Wang, 2020
RISK AWARDS 2021
Industry Confirmation

Risk Awards 2021: new risk engine can run nearly a billion XVA calculations per second

- “calculate, on-the-fly, the impact to its book”
- “has given us greater visibility into our risk”
- “able to better exploit fleeting dislocations in the market”

https://www.risk.net/awards/7736276/technology-innovation-of-the-year-scotiabank
WHAT IS DEEP LEARNING ANYWAY?

Big data and plentiful compute have triggered a resurgence in AI

Deep learning and AI is impacting how we live, whether we realise it or not.

- Recomender systems
- Transport: Autonomous Vehicles
- Smart Cities: Advanced video analytics
- Phara: Drug discovery
- Financial Services …
NEURAL NETWORKS
A quick refresher

\[ y = f \left( \sum w x + b \right) \]

Free Parameters
7 non-linear activations
70 weights
23 biases

+ the magic of backpropagation and stochastic gradient descent or other training methods
WHY ANOTHER METHODOLOGY?

Current Challenges

- Valuation is burdensome for all but the simplest products
- Traditional methods are expensive

Can Deep Learning help?

- Expensive training, but once deployed ANNs are fast and offer very high throughput.
- "We can train a neural network to approximate any function to an arbitrary level of precision"
HOW?

- No magic wonder model that prices all things under all conditions for all people.
- Much of the literature focuses on specific cases.
- Don’t look for a single model but instead train models for specific problems and integrate them.
- We already know what the computationally expensive bits are:
  - Optionality: Investor behavior
  - Market Events: Path dependency
- We shouldn’t throw everything away and re-invent the world
EXPERIMENTAL FRAMEWORK
From data generation through training to benchmarking
EXPERIMENTAL FRAMEWORK
From data generation through training to benchmarking

Pricing Model Implementation*

Training Data Portfolio (X) Valuations (Y)

NN Model Desc.

Benchmark Portfolio

Model Weights

NN Model Desc.

*Same implementation displayed separately for clarity
MODEL ARCHITECTURE
TRAINING DATA
How much and how good?
INTERPRETING RESULTS

How do we know when we have a good model?

<table>
<thead>
<tr>
<th>S</th>
<th>2.00</th>
<th>1.00</th>
<th>0.10</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>95.5</td>
<td>6,273.67</td>
<td>580.12</td>
<td>145.38</td>
<td>6.25</td>
</tr>
<tr>
<td>96</td>
<td>5,549.74</td>
<td>422.24</td>
<td>133.34</td>
<td>6.11</td>
</tr>
<tr>
<td>96.5</td>
<td>4,873.28</td>
<td>284.22</td>
<td>119.47</td>
<td>8.79</td>
</tr>
<tr>
<td>97</td>
<td>4,260.59</td>
<td>169.74</td>
<td>104.15</td>
<td>15.79</td>
</tr>
<tr>
<td>97.5</td>
<td>3,648.24</td>
<td>70.72</td>
<td>85.84</td>
<td>16.92</td>
</tr>
<tr>
<td>98</td>
<td>3,091.33</td>
<td>9.25</td>
<td>70.53</td>
<td>7.43</td>
</tr>
<tr>
<td>98.5</td>
<td>2,569.00</td>
<td>44.13</td>
<td>56.24</td>
<td>2.56</td>
</tr>
<tr>
<td>99</td>
<td>2,076.98</td>
<td>79.28</td>
<td>46.26</td>
<td>0.03</td>
</tr>
<tr>
<td>99.5</td>
<td>1,616.26</td>
<td>98.56</td>
<td>40.18</td>
<td>0.84</td>
</tr>
<tr>
<td>100</td>
<td>1,181.41</td>
<td>105.52</td>
<td>37.62</td>
<td>6.14</td>
</tr>
<tr>
<td>100.5</td>
<td>770.68</td>
<td>101.84</td>
<td>39.73</td>
<td>12.28</td>
</tr>
<tr>
<td>101</td>
<td>385.15</td>
<td>86.18</td>
<td>41.88</td>
<td>12.10</td>
</tr>
<tr>
<td>101.5</td>
<td>20.32</td>
<td>61.47</td>
<td>45.54</td>
<td>7.55</td>
</tr>
<tr>
<td>102</td>
<td>323.90</td>
<td>30.27</td>
<td>51.93</td>
<td>5.40</td>
</tr>
<tr>
<td>102.5</td>
<td>650.20</td>
<td>10.16</td>
<td>59.60</td>
<td>6.49</td>
</tr>
<tr>
<td>103</td>
<td>957.36</td>
<td>56.15</td>
<td>70.91</td>
<td>10.27</td>
</tr>
<tr>
<td>103.5</td>
<td>1,249.02</td>
<td>106.90</td>
<td>82.13</td>
<td>14.08</td>
</tr>
<tr>
<td>104</td>
<td>1,525.05</td>
<td>162.70</td>
<td>97.78</td>
<td>15.56</td>
</tr>
<tr>
<td>104.5</td>
<td>1,787.56</td>
<td>222.79</td>
<td>117.47</td>
<td>14.85</td>
</tr>
</tbody>
</table>

Min     | 29.32    | 9.25     | 37.62    | 0.03     |
| Max    | 6,273.67 | 580.12   | 145.38   | 16.92    |
| Average| 2,252.09 | 142.70   | 76.11    | 8.92     |
NO SINGLE ARCHITECTURE
Each pricing model will require a new round of architecture tuning
Traditional valuation models of the Bermudan Swaption use numerical methods which are very costly to compute. Riskfuel models use machine learning to learn an analytic representation of the target model. The result is fast and accurate valuations and risk sensitivities.

Learn more about the Bermudan Swaption Demo.
EXPERIMENTATION

Accelerated compute yields drastic productivity improvement
TRIVIAL ACCELERATION
A huge performance boost with almost no effort at all

- ANNs map well to parallel architectures
- ANN training and inference workloads exhibit a high degree of trivial parallelism
- All major deep-learning frameworks have been ported to leverage the vast compute available on GPUs
- Pytorch DataParallel further facilitates easy scaling across multiple GPUs
NVIDIA GPU CLOUD

- Low barrier to entry
- Instant productivity
- Continuous functional and performance improvements
- [https://ngc.nvidia.com/](https://ngc.nvidia.com/)
THE ROAD TO PRODUCTION

What challenges or considerations await us?

- Regulators
  - Regulators encourage technology driven innovation including the responsible deployment of machine learning and artificial intelligence.

- Validation
  - To satisfy internal and external validation and audit requirements we should anticipate a strong requirement for reproducibility through the model development cycle.

- Integration
  - Ultimately, our trained model needs to find its way to production in a controlled way. Transition from experimental environment with controlled deployment and versioning.
IN CLOSING

- By recognising the work of Scotiabank and Riskfuel, Risk.net have confirmed the emergence of Deep Learning as an applicable and valid methodology in derivatives valuation.

- Rather than a revolution, Deep Learning provides a valuable addition to the Quant’s toolbox. Existing methods and investments are not discarded but remain highly relevant.

- Apart from operational relief Deep Learning, can open the door not only to better performance but also better modelling.

- Far from being esoteric, Deep Learning has already been commoditized in the form of high quality, freely available frameworks.