P&I | CONFERENCES

Canadian Pension Risk Strategies

Machine Learning in Quant Finance

Tarik Eldin
Geode

P&I | CONFERENCES

Canadian Pension Risk Strategies

MACHINE LEARNING IN QUANT FINANCE

Tarek Eldin, Head of Research
Geode

P&I | CONFERENCES

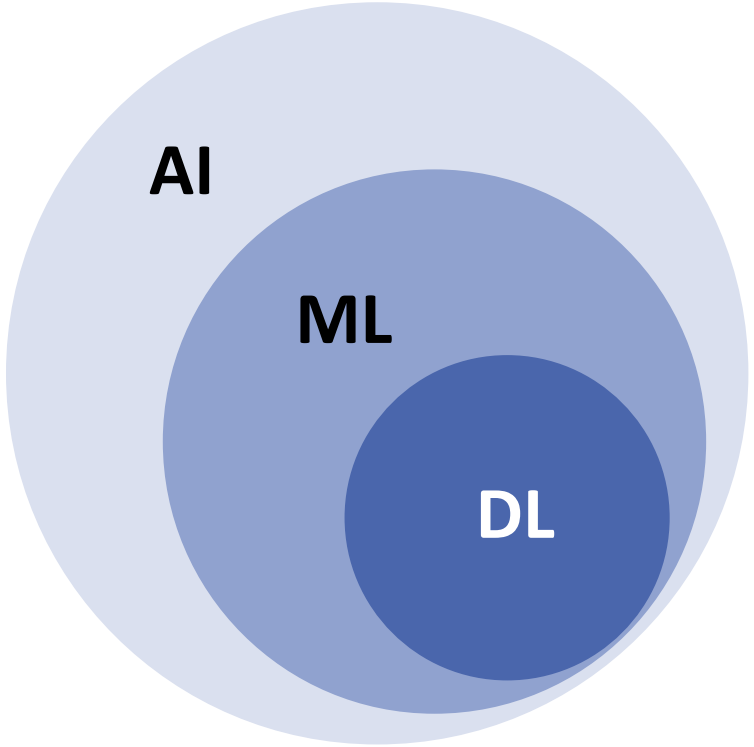
Canadian Pension Risk Strategies

- Definitions, History
- Random Forests
- Neural Networks
- Comparisons; What We've Learned So Far
- A Real Life Example
- Summary

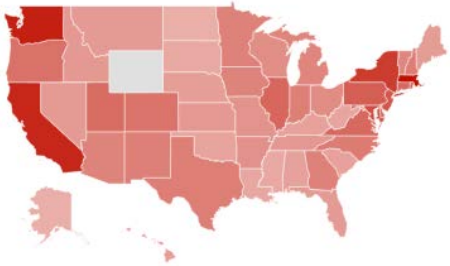
Definitions, History

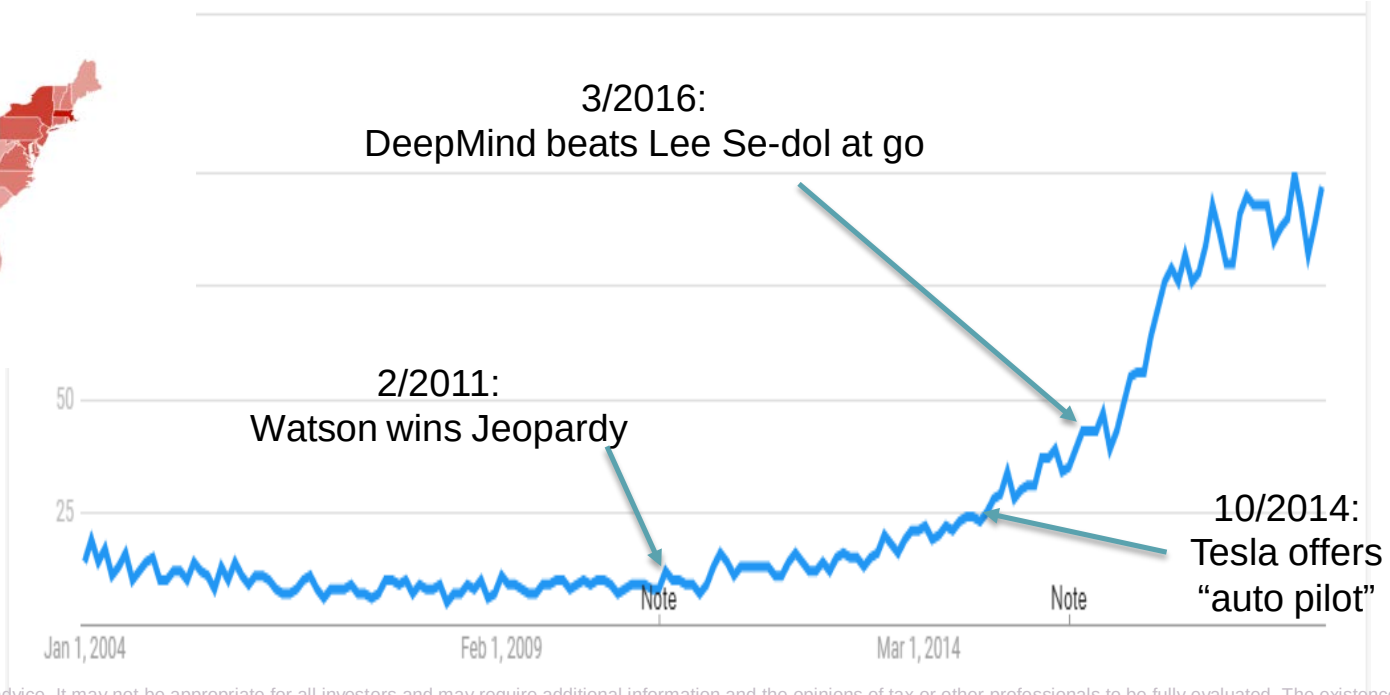
Some Definitions

- **Artificial Intelligence:** Human-like intelligence. The ‘general’ variety is an exciting idea but doesn’t exist yet
- **Machine Learning:** generate classifications, predictions, or relationships between data items in a somewhat self directed way. In recent years established itself in finance.
- **Deep Learning:** as above but also able to identify complex patterns. Modeled after the human brain. Drove many of the (mostly non-finance) breakthroughs in the past decade.



Google Trends “Machine Learning” 1/2004 – 2/2019





A Subjective History of ML in Finance

- **Traditionally, academic finance people have been skeptical**
  - Too easy to over fit
  - Often not driven by structural or at least theoretical models of the world
  - Not easy to interpret
  - Not as well understood as, say linear regressions
- **Practitioners with CS backgrounds led the way**
  - Some (but not many) hedge funds likely began using ML decades ago
  - The typical quant asset manager began sometime in the past 5 years
  - Non-quants aren't there yet but will likely become users
- **ML has now broken through in academic finance as well**
  - See recent papers and conference topics
  - Moritz Zimmermann was a relatively early example
  - Fresh grads now often have a good knowledge base in ML – can be agents of change
- **Progress is closely tied to advances in coding, data, and computational infrastructure**

The History of ML at Geode

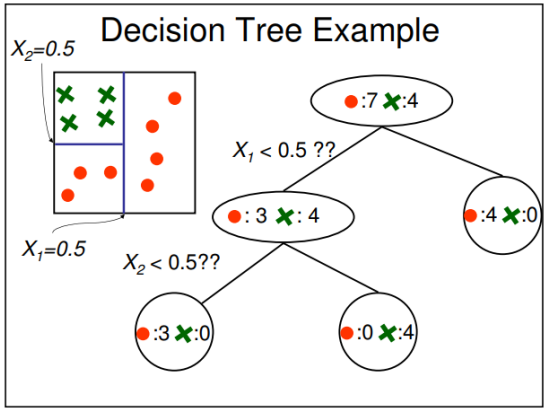
- Started as an experiment led by junior analyst in 2006
- First model, ‘Hal,’ went live in 2007 as an equity market neutral strategy
- Guess what happened in August 2008…
- Subsequently we derisked (bad call) but kept it running (good call) with improvements to risk modelling, risk management and analytics
- A second model was launched in 2014
- Both have been continually improved since then
- Both are among our best performing strategies
- Our current focus is on Neural Networks, uses outside of market neutral equities as well as infrastructure upgrades.

Major Techniques

- **Decision Trees / Random Forests**
- **Neural Networks**
- **SVM: Support Vector Machines**

An Introduction to Random Forests

Key Concepts



- **Decision Tree**: Classifying/Predicting from a set of features
  - Each level splits the data according to different features
  - *Goal*: Achieve good classification with minimal number of decisions

- **Bagging**: (**B**ootstrap **Agg**regation)
  - Replicate the dataset by sampling with replacement
  - Apply a learning method (usually decision tree learning) to each bootstrap sample to produce predictors
  - Use the average of these predictors as the final predictor

- **Random Forests**
  - Do Bagging, but in addition, when growing the tree, select a random sample of features to consider in each step
  - This will lead to very different trees from each bootstrap sample
  - The randomization steps in Bagging and Random Forests will help avoid over-generalization (over-fitting)

From Trees to Forests

Bootstrap sampling

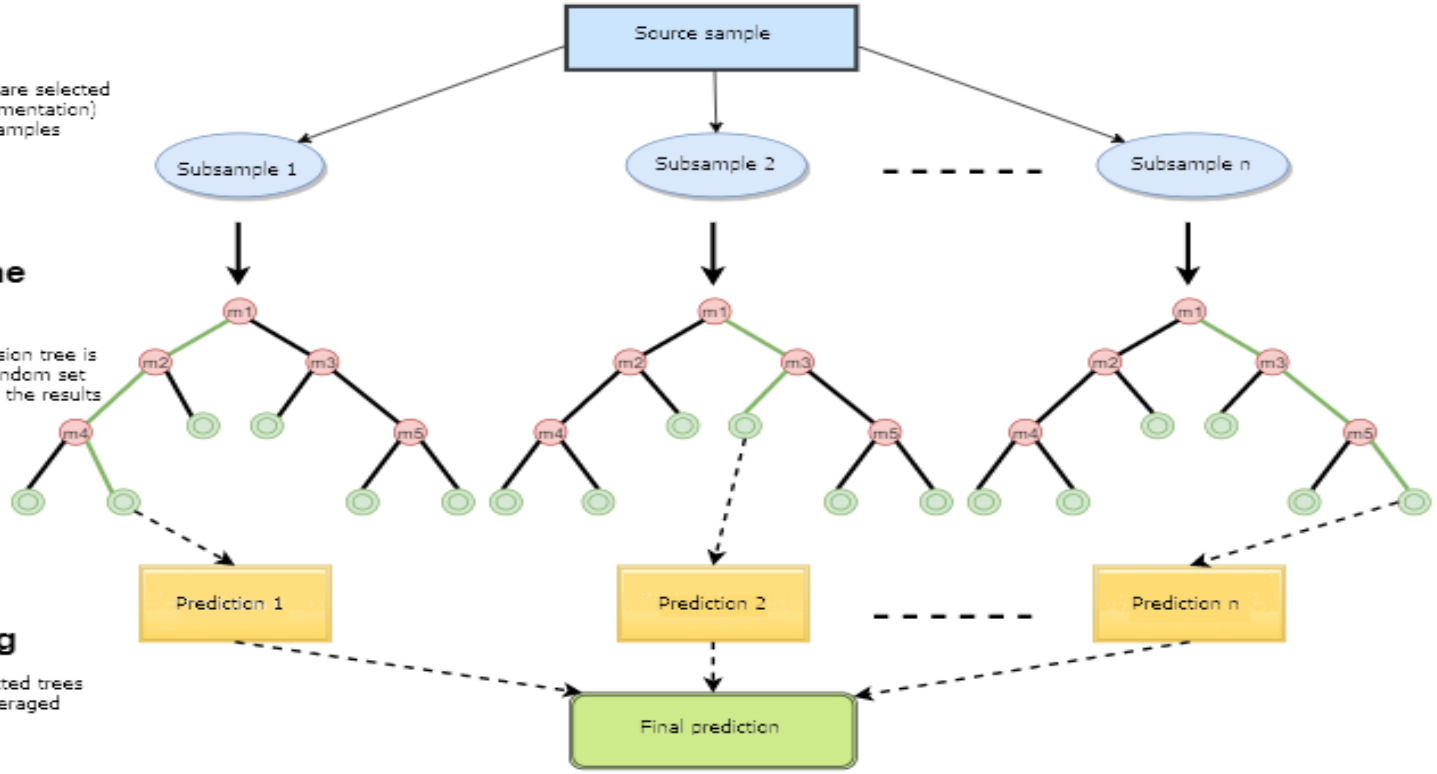
r (percentage) examples are selected (0.63 in classical implementation) in n random subsamples

Building the models

for each subsample, a decision tree is constructed based on a random set of m features (covariants), the results fall into leaves

Bootstrap aggregating

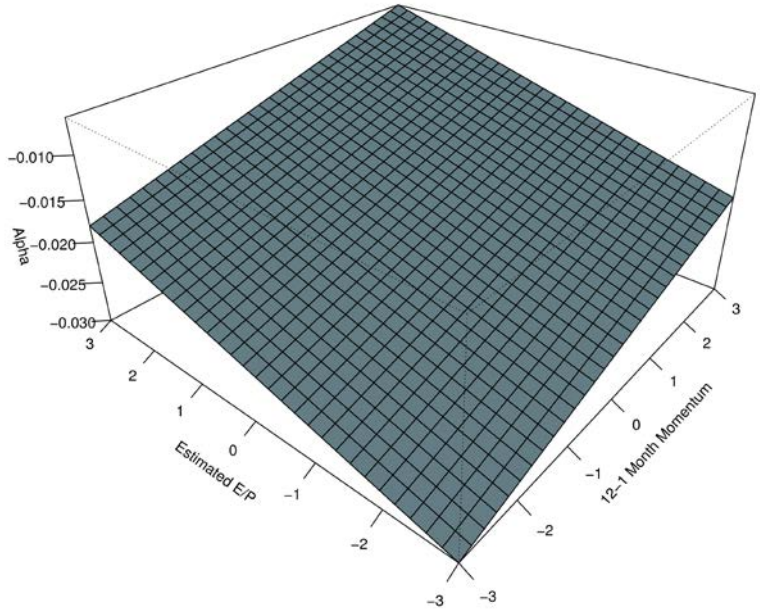
results from all constructed trees are gathered and averaged



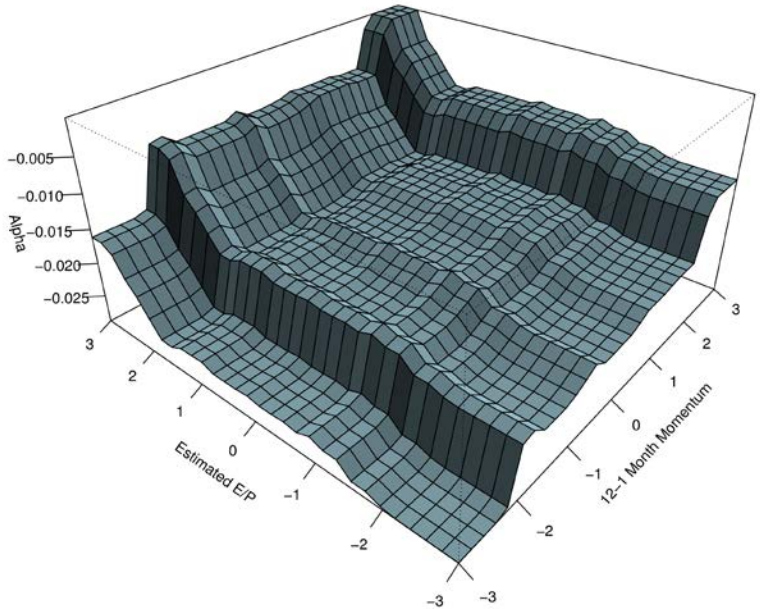
Source: https://www.mql5.com/en/articles/3856

Example: Linear vs Non-Linear (Random Forest Based) Alpha Modeling

Linear Regression

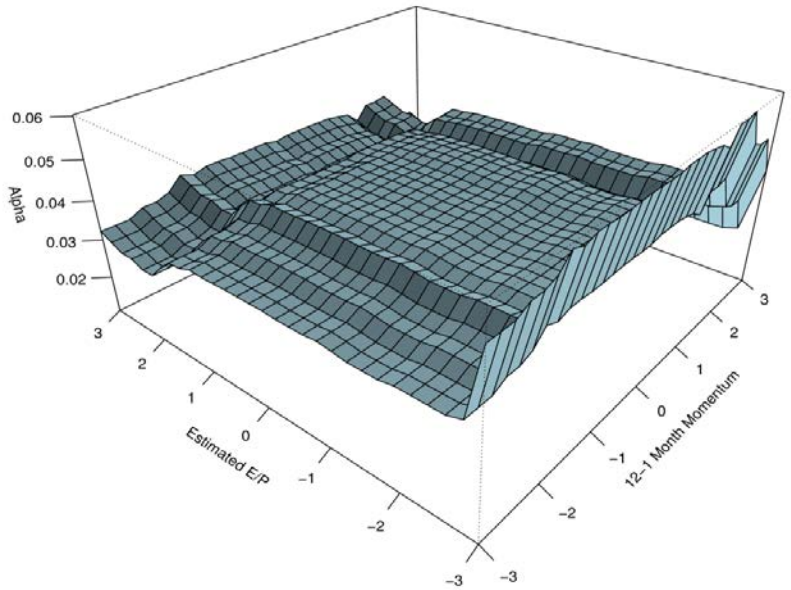


Random Forrest



For illustrative purposes only

Another Example: Dynamic Non-Linear Alpha Modeling

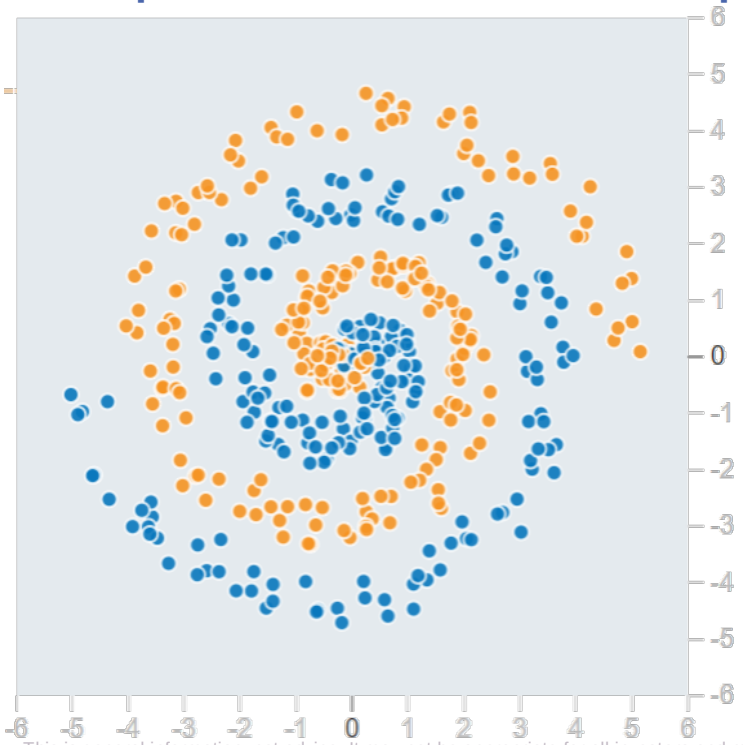


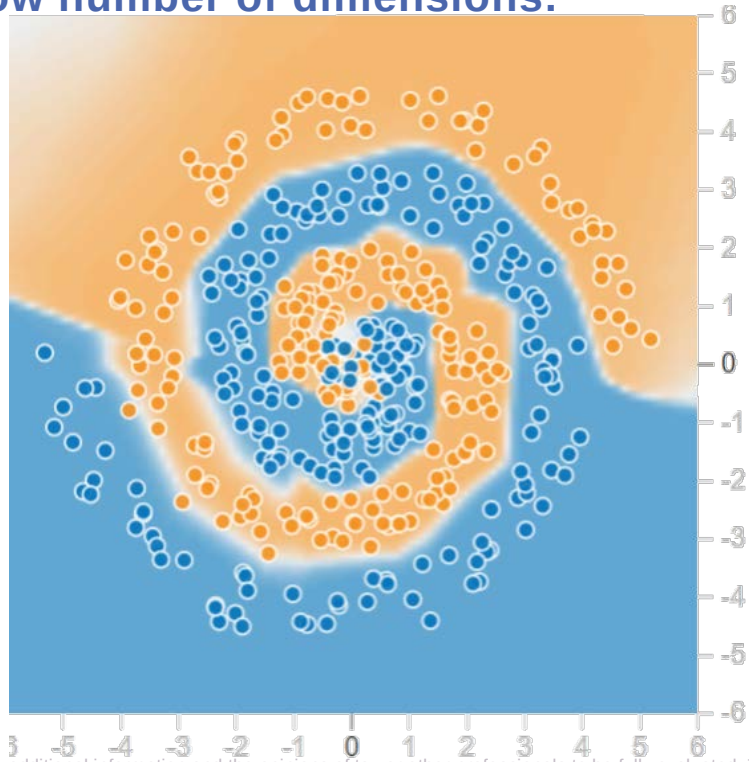
For illustrative purposes only

An Introduction to Neural Networks

Neural Networks: Illustration

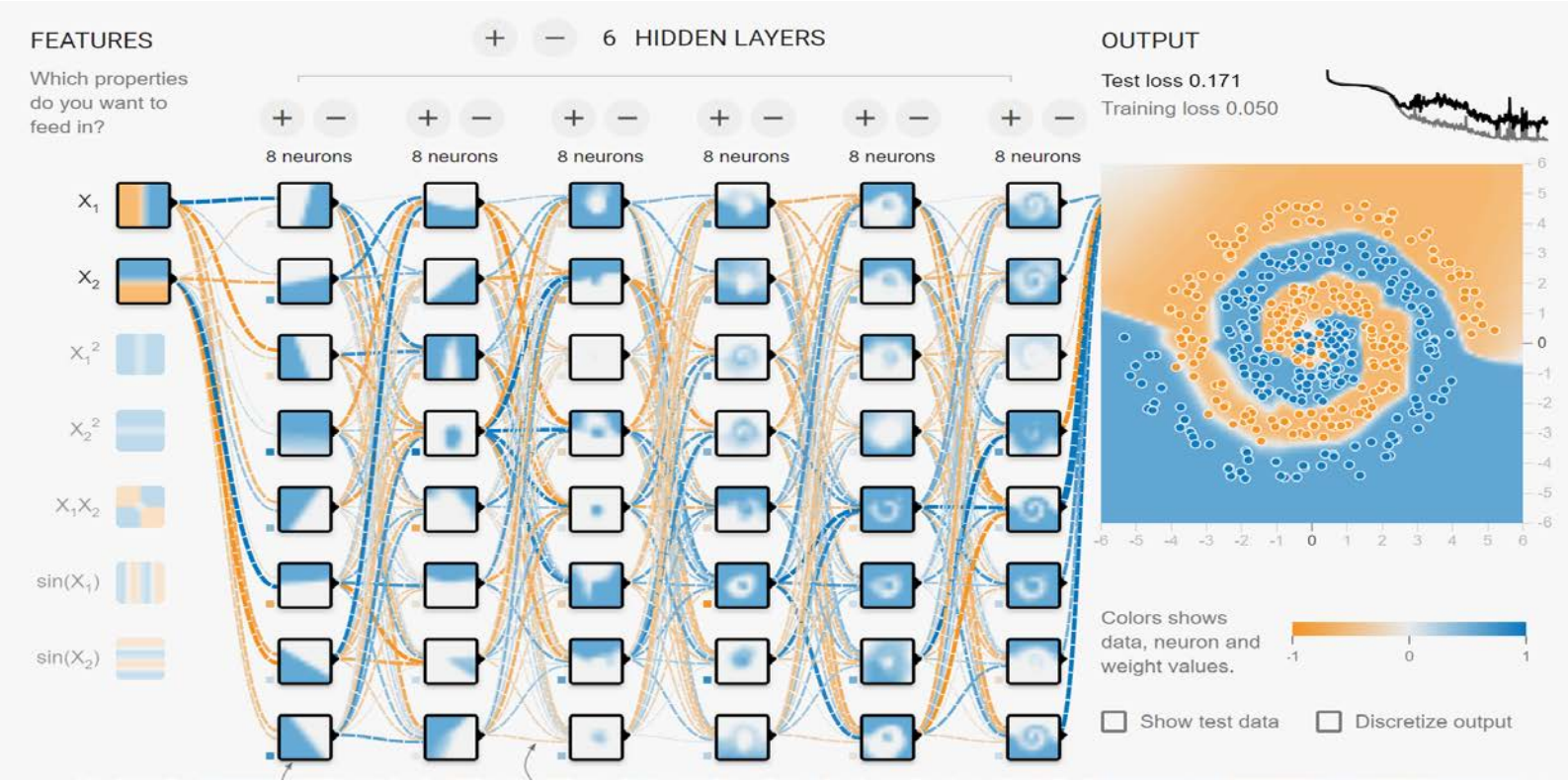
- **Example of nonlinear relationship – low number of dimensions:**





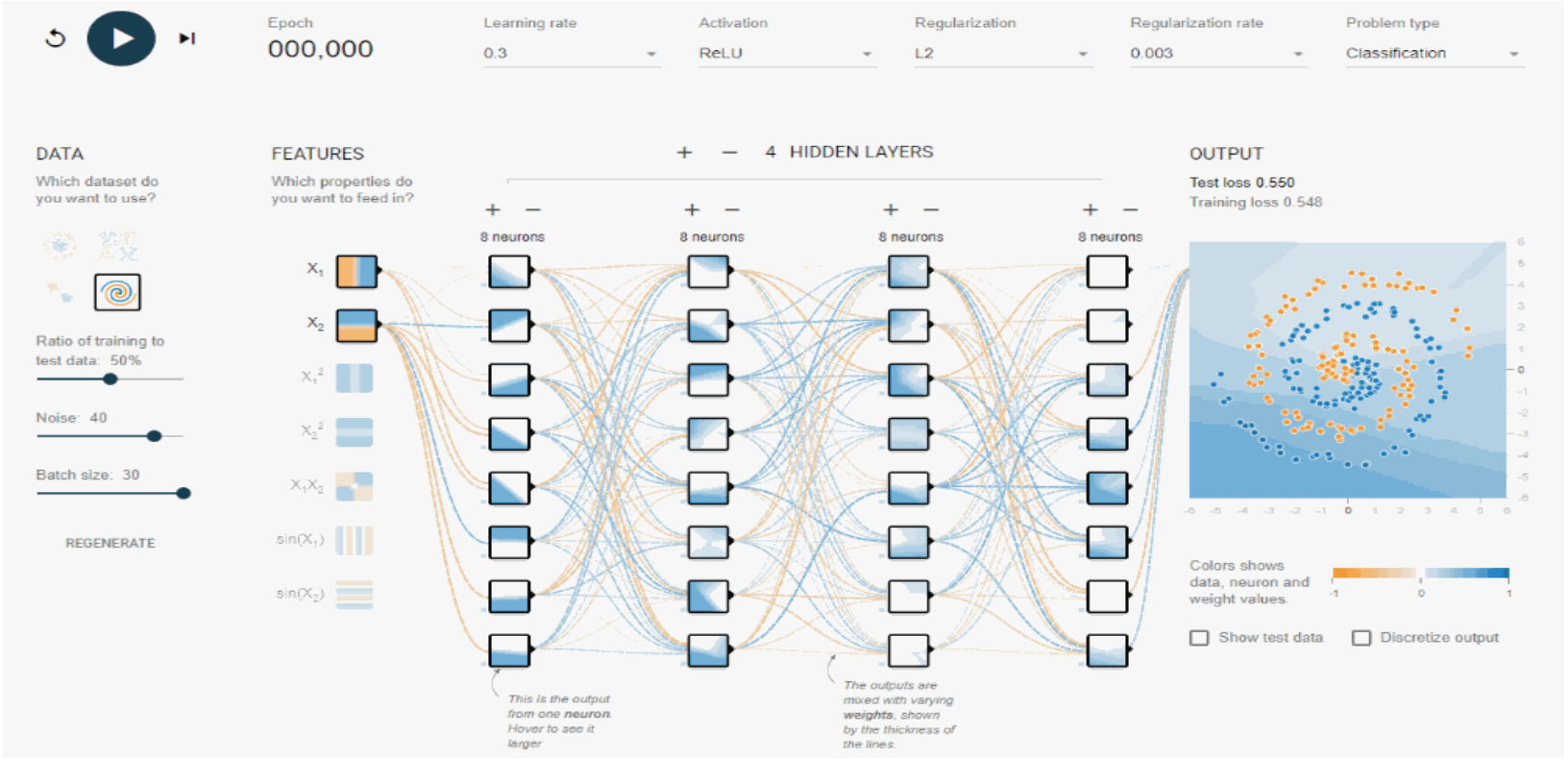
Source: playground.tensorflow.org

Neural Networks: ‘Neurons’ and ‘Layers’



Source: playground.tensorflow.org

An Animated Example: The Learning Process



Source: playground.tensorflow.org

Comparisons;
Technology Needs;
What We've Learned So Far

Linear Regressions RF vs NN

	Linear Regressions	RF	NN
What they Model	Linear additive relationships	Non linear and interactive relationships	Complex Patterns in high dimensionality data
Data Needed	A fair amount	A lot	A ton
Computing Needs	Little	A lot	A ton
Transparency	Well understood, easy to interpret	While trees are intuitive, forests are pretty opaque	Hard to analyze and interpret
Collective Experience	~150 years	Decades	Years

Running *a* Machine Learning Model is Trivial

Linear Regression

R

```r
model <- lm(y ~ ., training_data)
```

Python

```python
from sklearn.linear_model import LinearRegression

model = LinearRegression()
model.fit(x, y)
```

Random Forest

R

```r
library(randomForest)

model <- randomForest(y ~ ., training_data)
```

Python

```python
from sklearn.ensemble import RandomForestClassifier

model = RandomForestClassifier()
model.fit(x, y)
```

Simple Neural Network *(Tensorflow, Keras, PyTorch, etc. are better-suited for deep learning models)*

R

```r
library(neuralnet)

model <- neuralnet(y ~ ., training_data)
```

Python

```python
from sklearn.neural_network import MLPRegressor

model = MLPRegressor()
model.fit(x, y)
```

…building a good one is harder though

ML and Technology

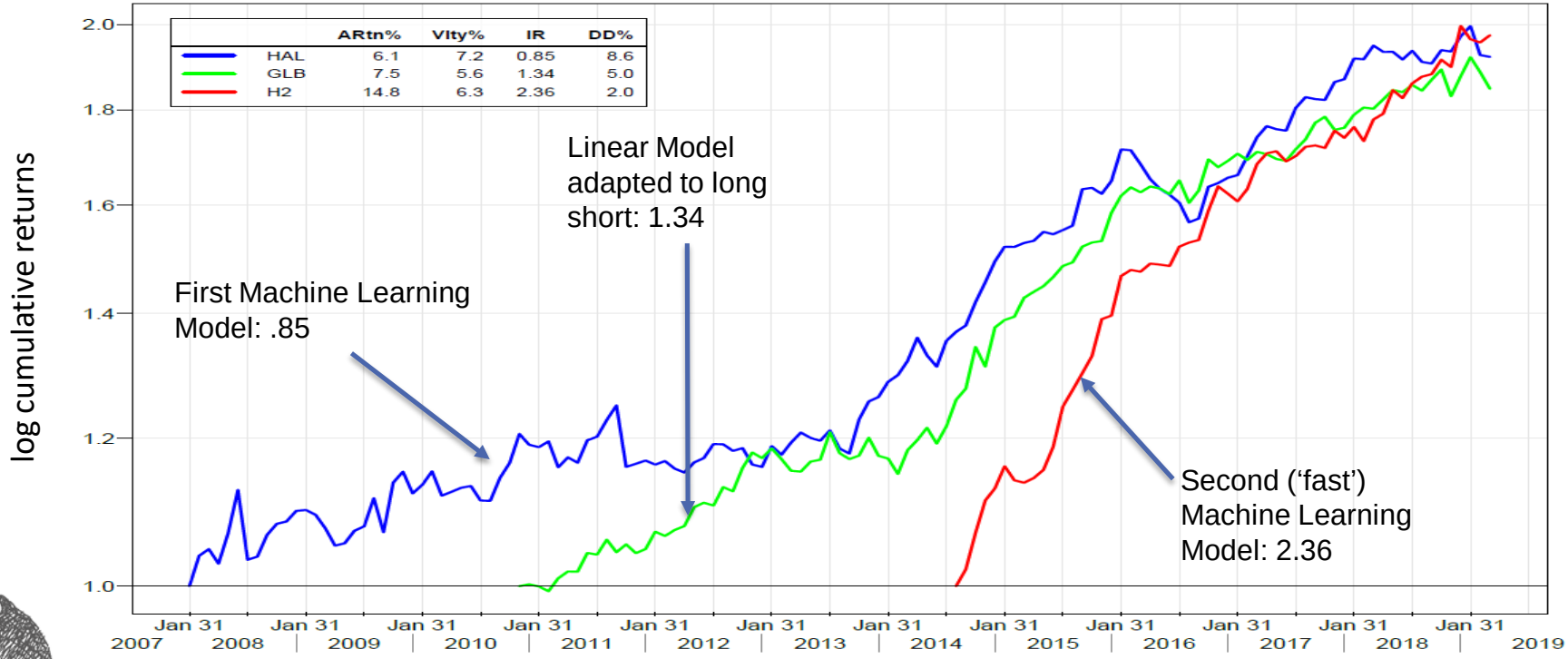
- **ML is enabled by advances in technology**
- **This also means the industry has to keep up and upgrade**
  - Cloud computing, more capacity, more elasticity
  - GPU servers
  - More powerful programming languages. See the rise of Python/TensorFlow/PyTorch
  - Data Management
  - Version control and collaboration tools
  - Visualization tools

Some Key Questions and Selected Learnings

- **Contrary to popular belief, the task of designing and structuring models does not go away. It only becomes more complex**
  - Exactly what are you predicting?
  - Using what as inputs?
  - Time series or X-section?
  - Are relationships static or allowed to change over time?
- **You need to think about if/how to transform the data going in**
- **Tuning!**
- **How to counteract fitting issues!!**
- **ML models can usually handle more inputs than traditional regressions**
- **How to handle Risk Modeling, and the related topics of risk decomposition and attribution is more tricky than in traditional models**

A Real Life Example

Some Live Results*



*HAL: From 2008/01/31 to 2019/03/31, GLB: From 2010/12/31 to 2019/03/31. H2: From 2014/09/30 to 2019/03/31. After t-costs and fixed fees but gross of performance fees

Key Properties

Correlation Coefficients*

		HAL	GLB	H2
Machine Learning #1	HAL	1.00	0.33	0.47
Linear	GLB	**0.33**	1.00	**0.26**
Linear Learning #2	H2	0.47	0.26	1.00

Largely similar inputs but fairly different returns from traditional model

Performance and Betas*

		Return	Vol	IR	Beta
Machine Learning #1	HAL	7.6%	5.6%	1.36	-0.15
Linear	GLB	8.6%	5.8%	1.48	-0.13
Linear Learning #2	H2	14.8%	6.3%	2.36	-0.13

Slightly negative market betas (more for ML models)

The machine is not the only one who is learning

*Measured from 2014/09/30 to 2019/03/31. Returns After t-costs and fixed fees but gross of performance fees

Summary

Summary

- **After a delayed start in main stream finance, Machine Learning has now broken through with a vengeance**
- **Machines are less self-organizing than one might think. Structuring and modeling is still key to good design. Training in traditional statistics helps**
- **The tools and infrastructure are at a point where getting started is quite straightforward – maybe deceptively so**
- **Based on over a decade of live results we are pretty confident Machine Learning – if applied well and to the right problems – works and has further potential**

Thank You

Important Information

FIC has prepared this presentation for, and only intends to provide it to, institutional and sophisticated investors in one-on-one or comparable presentations. Do not distribute or reproduce this report.

This is general information, not advice. It may not be appropriate for all investors and may require additional information and the opinions of tax or other professionals to be fully evaluated. The existence of risk management processes does not imply low risk. Past performance is no guide to future returns. A portfolio could fail to meet its objective, and you could lose money. There is no guarantee that a particular portfolio will meet its investment objectives. Please see important Legal Disclosure at the end of this document.

Risks

Past performance is no guarantee of future results. An investment may be risky and may not be suitable for an investor's goals, objectives and risk tolerance. Investors should be aware that an investment's value may be volatile and any investment involves the risk that you may lose money.

Performance results for individual accounts will differ from performance results for composites and representative accounts due to factors such as portfolio size, especially if currently only funded with affiliated fee paying seed capital, timing of investments, market conditions, account objectives and restrictions, and factors specific to a particular investment structure.

The value of a strategy's investments will vary day to day in response to many factors, including in response to adverse issuer, political, regulatory, market or economic developments. The value of an individual security or a particular type of security can be more volatile than the market as a whole and can perform differently from the value of the market as a whole. Nearly all accounts are subject to volatility in foreign exchange markets.

Derivatives may be volatile and involve significant risk, such as, credit risk, currency risk, leverage risk, counterparty risk and liquidity risk. Using derivatives can disproportionately increase losses and reduce opportunities for gains in certain circumstances. Derivatives may have limited liquidity and may be harder to value, especially in declining markets. Derivatives involve leverage because they can provide investment exposure in an amount exceeding the initial investment. Leverage can magnify investment risks and cause losses to be realized more quickly. A small change in the value of an underlying asset, instrument, or index can lead to a significant loss. Assets segregated to cover these transactions may decline in value and are not available to meet redemptions. Government legislation or regulation could affect the use of these transactions and could limit the ability to pursue such investment strategies.

Important Information (cont'd)

Risks (cont'd)

These materials may contain statements that are "forward-looking statements," which are based on certain assumptions of future events. Forward-looking statements are based on information available on the date hereof, and Fidelity Investments Canada ULC ("FIC") does not assume any duty to update any forward-looking statement. Actual events may differ from those assumed by FIC when developing forward-looking statements. There can be no assurance that forward-looking statements, including any projected returns, will materialize or that actual market conditions and/or performance results will not be materially different or worse than those presented.

Performance Data

Performance data is generally presented gross of any fees and expenses, including advisory fees, which when deducted will reduce returns.

* * *

Certain data and other information in this presentation have been supplied by outside sources and are believed to be reliable as of the date of this document. Data and information from third-party databases, are self-reported by investment management firms that generally pay a subscription fee to use such databases, and the database sponsors do not guarantee or audit the accuracy, timeliness or completeness of the data and information provided including any rankings. Rankings or similar data reflect information at the time rankings were retrieved from a third-party database, and such rankings may vary significantly as additional data from managers is reported. FIC has not verified and cannot verify the accuracy of information from outside sources, and potential investors should be aware that such information is subject to change without notice. Information is current as of the date noted.

 FIC does not provide legal or tax advice and we encourage you to consult your own lawyer, accountant or other advisor before making an investment.

P&I | CONFERENCES and AON present

THANK YOU TO OUR SPONSORS

BMO

CIBC Asset Management

Fidelity INVESTMENTS
FIDELITY CANADA INSTITUTIONAL™

NISA

Sun Life Financial

ASSOCIATE SPONSOR:

RiskFirst
Technology for
Asset Owners | Consultants | Investment Managers

Canadian Pension Risk Strategies