Machine Learning in Finance

Thomas Krabichler and Josef Teichmann

ETH Zürich

January 2021

Exciting times

We see fascinating progress in

- Language and Image recognition,
- Go, Chess,
- Weather forecasting,
- Protein foldings, see https://www.nature.com/articles/d41586-020-03348-4
- economic or social predictions might follow soon ...

Exciting times

We see fascinating progress in

- Language and Image recognition,
- Go, Chess,
- Weather forecasting,
- Protein foldings, see https://www.nature.com/articles/d41586-020-03348-4
- economic or social predictions might follow soon ...

Exciting times

We see fascinating progress in

- Language and Image recognition,
- Go, Chess,
- Weather forecasting,
- Protein foldings, see https://www.nature.com/articles/d41586-020-03348-4
- economic or social predictions might follow soon ...

Exciting times

We see fascinating progress in

- Language and Image recognition,
- Go, Chess,
- Weather forecasting,
- Protein foldings, see
  https://www.nature.com/articles/d41586-020-03348-4
- economic or social predictions might follow soon ...

Exciting times

We see fascinating progress in

- Language and Image recognition,
- Go, Chess,
- Weather forecasting,
- Protein foldings, see
  `https://www.nature.com/articles/d41586-020-03348-4`
- economic or social predictions might follow soon ...

Unique interactions ...

... between academia and financial industry

- since we both see the paradigm changes for modeling,
- and the opportunities to see solutions which we have never seen before in fully realistic frameworks.

Unique interactions ...

... between academia and financial industry

- since we both see the paradigm changes for modeling,
- and the opportunities to see solutions which we have never seen before in fully realistic frameworks.

Neural Networks

Neural networks are nowadays frequently used to approximate functions due to ubiquitous universal approximation properties. A neural network, as for instance graphically represented by the following figure,

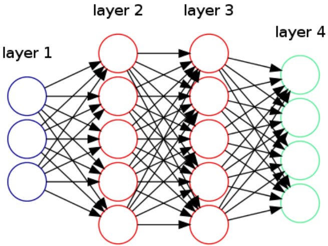

Figure: A 2 hidden layers neural network with 3 input and 4 output dimensions

denoting a concatenation of affine and non-linear functions in a well specified order. This is a feed forward architecture, one among many ...

Neural Networks and Universal Approximation

- Neural networks appeared in the 1943 seminal work by Warren McCulloch and Walter Pitts inspired by certain functionalities of the human brain aiming for artificial intelligence (AI).
- Universal Approximation Theorems (George Cybenko, Kurt Hornik, et al.) show that *one hidden neural layer networks* can already *uniformly approximate* any continuous function on the unit cube.
- Neural networks are an easy to handle (no curse of dimension, just matrix multiplication and one non-linearity are needed, etc), parametric class of functions which is dense in continuous functions.

Neural Networks and Universal Approximation

- Neural networks appear in the 1943 seminal work by Warren McCulloch and Walter Pitts inspired by certain functionalities of the human brain aiming for artificial intelligence (AI).
- Universal Approximation Theorems (George Cybenko, Kurt Hornik, et al.) show that *one hidden neural layer networks* can already *uniformly approximate* any continuous function on the unit cube.
- Neural networks are an easy to handle (no curse of dimension, just matrix multiplication and one non-linearity are needed, etc), parametric class of functions which is dense in continuous functions.

Neural Networks and Universal Approximation

- Neural networks appear in the 1943 seminal work by Warren McCulloch and Walter Pitts inspired by certain functionalities of the human brain aiming for artificial intelligence (AI).
- Universal Approximation Theorems (George Cybenko, Kurt Hornik, et al.) show that *one hidden neural layer networks* can already *uniformly approximate* any continuous function on the unit cube.
- Neural networks are an easy to handle (no curse of dimension, just matrix multiplication and one non-linearity are needed, etc), parametric class of functions which is dense in continuous functions.

A standard perspective on machine learning

- *Learning or Training* is the specification of a neural network which approximates a certain non-linear function on some input space.
- Modern learning technology performs training tasks in a highly accessible and very efficient way (Tensorflow, Theano, Torch).
- "Unreasonable effectiveness" of learning (image and language recognition, classification tasks, etc).

A standard perspective on machine learning

- *Learning or Training* is the specification of a neural network which approximates a certain non-linear function on some input space.
- Modern learning technology performs training tasks in a highly accessible and very efficient way (Tensorflow, Theano, Torch).
- "Unreasonable effectiveness" of learning (image and language recognition, classification tasks, etc).

A standard perspective on machine learning

- *Learning or Training* is the specification of a neural network which approximates a certain non-linear function on some input space.
- Modern learning technology performs training tasks in a highly accessible and very efficient way (Tensorflow, Theano, Torch).
- "Unreasonable effectiveness" of learning (image and language recognition, classification tasks, etc).

Applications in Finance

- instead of using established numerical algorithms each time one *learns* solutions by specifying their properties.
- Successfully implemented in several areas: deep hedging (Bühler-Gonon-Teichmann-Wood 2017), deep calibration (Cuchiero-Khosrawi-Teichmann), deep simulation (Cuchiero-Gonon-Grigoryeva-Ortega-Teichmann), deep asset liability management (Krabichler-Teichmann 2020).
- we see solutions of problems which we have never seen before and we are at the edge of fully realistic modelling: high dimensional risk management tasks with trading constraints and transaction costs can be treated, high dimensional calibration problems can be solved in real time. Networks with *tens of thousands of weights* are trained to perform this work. *Universality* guarantees that any strategy can be approximated.

Applications in Finance

- instead of using established numerical algorithms each time one *learns* solutions by specifying their properties.
- Successfully implemented in several areas: deep hedging (Bühler-Gonon-Teichmann-Wood 2017), deep calibration (Cuchiero-Khosrawi-Teichmann), deep simulation (Cuchiero-Gonon-Grigoryeva-Ortega-Teichmann), deep asset liability management (Krabichler-Teichmann 2020).
- we see solutions of problems which we have never seen before and we are at the edge of fully realistic modelling: high dimensional risk management tasks with trading constraints and transaction costs can be treated, high dimensional calibration problems can be solved in real time. Networks with *tens of thousands of weights* are trained to perform this work. *Universality* guarantees that any strategy can be approximated.

Applications in Finance

- instead of using established numerical algorithms each time one *learns* solutions by specifying their properties.
- Successfully implemented in several areas: deep hedging (Bühler-Gonon-Teichmann-Wood 2017), deep calibration (Cuchiero-Khosrawi-Teichmann), deep simulation (Cuchiero-Gonon-Grigoryeva-Ortega-Teichmann), deep asset liability management (Krabichler-Teichmann 2020).
- we see solutions of problems which we have never seen before and we are at the edge of fully realistic modelling: high dimensional risk management tasks with trading constraints and transaction costs can be treated, high dimensional calibration problems can be solved in real time. Networks with *tens of thousands of weights* are trained to perform this work. *Universality* guarantees that any strategy can be approximated.

An intelligent solution?

- We set up a market environment given by a fixed (large) number of possible scenarios of future market developments, e.g. an exponential Brownian motion.
- We introduce a method to measure risk and a liability depending on future market developments, e.g. mean squared distance from a European Call contract with respect to the risk neutral measure.
- We have a premium at hand to invest in the market.
- We set up an artificial trader, i.e. a trader who stores experience in an neural network: she can invest in the market to improve ...

An intelligent solution?

- We set up a market environment given by a fixed (large) number of possible scenarios of future market developments, e.g. an exponential Brownian motion.
- We introduce a method to measure risk and a liability depending on future market developments, e.g. mean squared distance from a European Call contract with respect to the risk neutral measure.
- We have a premium at hand to invest in the market.
- We set up an artificial trader, i.e. a trader who stores experience in an neural network: she can invest in the market to improve ...

An intelligent solution?

- We set up a market environment given by a fixed (large) number of possible scenarios of future market developments, e.g. an exponential Brownian motion.
- We introduce a method to measure risk and a liability depending on future market developments, e.g. mean squared distance from a European Call contract with respect to the risk neutral measure.
- We have a premium at hand to invest in the market.
- We set up an artificial trader, i.e. a trader who stores experience in an neural network: she can invest in the market to improve ...

An intelligent solution?

- We set up a market environment given by a fixed (large) number of possible scenarios of future market developments, e.g. an exponential Brownian motion.
- We introduce a method to measure risk and a liability depending on future market developments, e.g. mean squared distance from a European Call contract with respect to the risk neutral measure.
- We have a premium at hand to invest in the market.
- We set up an artificial trader, i.e. a trader who stores experience in an neural network: she can invest in the market to improve ...

What is a model?

- **Classical approach:** model has few interpretable parameters, which can be adapted to represent features of reality and predict future, e.g. Black-Scholes with few clearly interpretable parameters. Price and hedging strategy for a European Call contract can be calculated and assessed.
- **Occam's razor:** the fewer parameters for the same model quality, the better.
- **Machine learning models:** use neural networks with millions of parameters, which by themselves have no direct meaning. Only model performance guides model selection.

What is a model?

- **Classical approach:** model has few interpretable parameters, which can be adapted to represent features of reality and predict future, e.g. Black-Scholes with few clearly interpretable parameters. Price and hedging strategy for a European Call contract can be calculated and assessed.
- **Occam's razor:** the fewer parameters for the same model quality, the better.
- **Machine learning models:** use neural networks with millions of parameters, which by themselves have no direct meaning. Only model performance guides model selection.

What is a model?

- **Classical approach:** model has few interpretable parameters, which can be adapted to represent features of reality and predict future, e.g. Black-Scholes with few clearly interpretable parameters. Price and hedging strategy for a European Call contract can be calculated and assessed.
- **Occam's razor:** the fewer parameters for the same model quality, the better.
- **Machine learning models:** use neural networks with millions of parameters, which by themselves have no direct meaning. Only model performance guides model selection.

See how that looks in reality ...

https://people.math.ethz.ch/~jteichma/index.php?content=teach_mlf2019