# Presentations overview – TopQuants Autumn Event 2018

| Speaker | Title | Session |
|---|---|---|
| Svetlana Borovkova (Probability & Partners) | Deep and reinforcement learning with sentiment for stock market trading | 1 |
| Martin Engblom (TriOptima) | Accelerated MVA in the probability matrix method | 1 |
| Charlotte Faber-Werger (Van Lanschot Kempen) | Augmenting discretionary asset management with machine learning | 1 |
| Riccardo Lena (LIST S.p.A.) and Enrico Melchioni (LIST S.p.A.) | The Gn++ model and its implementation | 1 |
| Karim Moussa (ABN AMRO) | Arbitrage-free filtering of option data | 1 |
| Pasquale Cirillo (TU Delft) | A brand new way of looking at recovery risk | 2 |
| Daniel J. Duffy (Datasim) | Efficient and accurate PDE models for one and two factor pricing problems | 2 |
| Daniel Linder (Adyen) | Building a smart, scalable platform that simplifies big data, improves analysis and enhances machine learning | 2 |
| Marcin Rybacki (Cardano) | The cross-currency basis | 2 |
| Joris Tolenaar (Transtrend) | Systematic trading and changed market dynamics | 2 |

**Deep and reinforcement learning with sentiment for stock market trading**

Svetlana Borovkova (Probability & Partners)

In this talk we will review our recent work on combining news and social media sentiment signals with cutting edge Machine Learning techniques to forecast and trade stock markets.

Our first application is developing a sentiment-based intraday trading strategy for a large stock index (EUROSTOXX 50). We utilize an ingenious combination of Kalman Filter for de-noising the news sentiment signal, Multi-Layer Perceptron for features selection and an LSTM network for actual forecasting, to obtain powerful forecasting framework. Out-of-sample forecasts show that downward movements in price can be predicted correctly in more than 75% of cases. Algorithmic trading strategy based of our forecasts shows significant economic gain, even after transaction costs.

Our second application is trading individual stocks that comprise S&P500 index. We use high-frequency stock-specific news sentiment signals, in combination with a reinforcement learning approach. We apply NeuroEvolution of Augmenting Topologies (NEAT), to learn the non-linear functional price response to different input features. The resulting trading strategy not only outperforms the benchmark, but also generates double the average profits made by the benchmark.

Our results demonstrate the great potential of combining external real-time information (such as sentiment) with Machine Learning techniques for algorithmic trading.

**Accelerated MVA in the probability matrix method**

Martin Engblom (TriOptima)

This presentation is focused on Margin Valuation Adjustment (MVA). In particular, it will deal with the following aspects:

- Practitioners' and academics' view of MVA
- Simulating IM using full SIMM and CCP formulas
- Introduction to the Probability Matrix Method
- Practical examples and benchmarks

During the workshop, a live demonstration will be provided.

**Augmenting discretionary asset management with machine learning**

Charlotte Faber-Werger (Van Lanschot Kempen)

You don't typically think about Machine Learning when talking about discretionary fund management. Although the last decade has seen the rise of quant houses in the investment space, most funds are still managed using traditional methods and strategies. The deluge of data and increases in computer processing power are changing this however. In this talk we will highlight select machine learning strategies that support traditional asset management by leveraging commonly used data as well as alternative data. We focus on Van Lanschot Kempen innovation projects that augment stock selection using recommender systems, optimize portfolio construction with non-parametric algorithms and contribute to smart portfolio monitoring by flagging anomalous events using unsupervised machine learning techniques.

**The Gn++ model and its implementation**

Riccardo Lena (LIST S.p.A.) and Enrico Melchioni (LIST S.p.A.)

In this presentation, some theoretical and implementative aspects of the short rate model Gn++ will be discussed. The G++ model, due to its balance between analytical tractability and sophistication, is widely employed for modeling rate curves and pricing interest rate derivatives which are sensitive to the correlations among forward rates. In fact, in a financial environment allowing negative interest rates, a Gaussian model for the risk-free curves is well suited to describe the behaviour of the rate curves. In order to fit as accurately as possible, the behaviour of market curves it is then necessary to find optimal structural parameters for the model. The focus of this presentation will be put, indeed, on the *parameter calibration* process for the Gn++. In particular, the possible cost functions and minimization algorithms will be addressed in order to explain their potential role in the optimization of computational time and accuracy.

**Arbitrage-free filtering of option data**

Karim Moussa (ABN AMRO)

Although option price data play an essential role in numerous applications in economics, little attention has been devoted to their pre-processing. The importance of option price data in applications suggests, however, that it may be worthwhile to perform automated checks to establish whether these data are of sufficient quality, and to make appropriate corrections if needed. To this end, this presentation discusses a simple filtering procedure for option price data based on the no-arbitrage principle, in which the data are adjusted only if basic no-arbitrage conditions are violated. In such case, quotes are adjusted by using information from other

option prices in the data set that are considered more trustworthy until the violations are resolved.

**A brand new way of looking at recovery risk**

Pasquale Cirillo (TU Delft)

Answering a major demand in modern credit risk management, we propose a nonparametric survival approach for the modeling of the recovery rate and the recovery time of a defaulted counterparty, by introducing what we call the Recovery Reinforced Urn Process, a special type of combinatorial stochastic process.

The new model allows for the elicitation and exploitation of prior knowledge and experts' judgements, and for the constant update of this information over time, as soon as new data become available. We show how to use it to perform Bayesian nonparametric prediction about the recovered amounts and the (total) recovery time of a series of defaulted exposures.

An application to real data is provided using the Single Family Loan-Level Dataset by Freddie Mac.

Joint work with Dan Cheng (TU Delft).

**Efficient and accurate PDE models for one and two factor pricing problems**

Daniel J. Duffy (Datasim)

In this talk we discuss the application of several finite difference methods (FDM) to the pricing of derivative products whose behaviour is described by linear and nonlinear partial differential equations (PDE). We employ a number of mathematical techniques to transform the PDE in some way in order to make it suitable for approximation by FDM. For example, we model a number of features such as UVM (uncertain volatility model), early exercise, stochastic short rates and PDE domain transformation. We then introduce three main finite difference methods, namely Alternating Direction Explicit (ADE), Crank Nicolson and Method of Lines (MOL) and we compare them based on efficiency, accuracy and applicability considerations.

We conclude the talk with a summary of using ADE and MOL for two-factor PDEs and we compare them with the Fractional Steps and Alternating Direction Implicit (ADI) methods.

**Building a smart, scalable platform that simplifies big data, improves analysis and enhances machine learning**

Daniel Linder (Adyen)

With all the hype around big data, I'm going to give a practical talk that will aim to answer the following questions.

When is a big data platform necessary?

What are the important pieces of a big data platform?

What are some recommended technologies for big data processing?

How can we continually test the data and code to make sure nothing breaks behind the scenes?

How can we make an environment where data scientists can experiment and break things while still deploying stable products?

How can we get our data scientists building products together in a more reproducible, understandable way with less mistakes instead of working alone and keeping ideas isolated?

How can Data Science be leveraged in FinTech to bring real value to the industry?

I will try to take the vantage point of a Data Scientist, Data Engineer and Project manager, so that it's technical enough for techies, but interesting enough for non-techies.

I'll give a glimpse at the architecture of the big data platform built at Adyen along with a deeper dive in the current and future data science products being built on this platform.

Then I will give my personal take on lessons learned and what I would have done differently along the path of building this big data infrastructure.

**The cross-currency basis**

Marcin Rybacki (Cardano)

The crisis of 2008 had dramatic effects on financial markets. From an interest rates perspective, the markets have been left fragmented. More realistic pricing of credit risk by market participants has, in fact, caused a tenor basis to arise within the same family of rates. From a foreign-exchange perspective, the markets have been left dislocated. The arbitraging mechanism that kept foreign exchange term structures at the level implied by the covered interest rate parity has in fact stopped working. Result of the dislocation has been the rise of a non-negligible cross-currency basis.

A direct impact of the cross-currency basis on the balance sheet of a pension fund is noticeable for trades that are collateralized in different currencies. For long term maturities the difference in computed present values, with or without basis, can be significant. Hence, the analytics must capture the basis to reflect market valuations. Modifying an existing analytical framework for this purpose can present several challenges.

The presentation, based on joint work with Giampietro Carpentieri, provides an overview of the cross-currency basis and the challenges associated with its treatment.

**Systematic trading and changed market dynamics**

Joris Tolenaar (Transtrend)

At Transtrend we have systematically traded hundreds of futures markets around the world for more than 25 years. In this presentation I will illustrate why systematic trading is perhaps much less systematic than one might think. The world around us keeps on changing and this requires continuous adaptation. Any strategy that does not adapt will experience diminishing performance. In particular, I will illustrate why and how market dynamics have fundamentally changed post financial crisis, the implications this has for a trend following strategy and how Transtrend has adapted. In other words, a practitioner's view on trend following strategies, markets and modelling.