HUB-NUS FinTech Workshop

Date: Thursday, 21 March 2019
Time: 8:50 am to 4:45 pm
Venue : 21 Heng Mui Keng Terrace, I^3 Building, National University of Singapore, 119613

Program Agenda:

Time	Speaker	Title
08:50 ~ 09:00	Assoc Prof Yi-Chun CHEN (NUS RMI) Prof. Dr. Ulrich HORST (HUB)	*Welcome address*
09:00 ~ 09:30	Prof. Dr. Stefan LESSMANN (HUB)	From accuracy to value: a multi-objective approach for value driven feature selection
09:30 ~ 10:00	Hitoshi IWASAKI (NUS)	Topic sentiment asset pricing with DNN supervised learning
10:00 ~ 10:30	Dr. Simon TRIMBORN (NUS)	Investing with cryptocurrencies - topic modelling dependent information retrieval from cryptocurrency experts
10:30 ~ 11:00		*Tea break*
11:00 ~ 11:30	Prof. Dr. Vladimir SPOKOINY (HUB)	Optimal stopping via reinforced regression
11:30 ~ 12:00	Dr. Richard PETERSON (CEO of MarketPsych)	Sentimental markets: how information flow drives asset prices
12:00 ~ 14:00		*Lunch break*
14:00 ~ 14:30	Prof. Mike K. P. SO (HKUST)	Efficient estimation of high-dimensional dynamic covariance by risk factor mapping: applications for financial risk management
14:30 ~ 15:00	Niels WESSELHÖFFT (HUB)	Utilizing high-dimensional high-frequency data for lower sampling frequencies
15:00 ~ 15:30		*Tea break*
15:30 ~ 16:00	Paolo PAGNOTTONI (Pavia University)	Vector error correction models to measure connectedness of bitcoin exchange markets
16:00 ~ 16:30	Prof. Jin-Chuan DUAN (NUS)	Variable selection with big data based on Zero Norm and via sequential Monte Carlo
1800		*Conference Dinner (by invitation only)* Carnivore CHIJMES, 30 Victoria Street, #01-30 Singapore 187996

Program Details:

Title:	From accuracy to value: a multi-objective approach for value driven feature selection
Speaker:	Prof. Dr. Stefan Lessmann
Email:	stefan.lessmann@hu-berlin.de
Abstract:	In credit scoring, feature selection aims at removing irrelevant data to improve the performance of the scorecard and its interpretability. Standard techniques treat feature selection as a single-objective task and rely on statistical criteria such as correlation. Recent studies suggest that using profit-based indicators may improve the quality of scoring models. We extend the use of profit measures to feature selection and develop a multi-objective wrapper framework based on the NSGA-II genetic algorithm with two fitness functions: the Expected Maximum Profit (EMP) and the number of features. Experiments on multiple credit scoring data sets demonstrate that the proposed approach develops scorecards that can yield a higher expected profit using fewer features than conventional feature selection strategies.

Title:	Topic sentiment asset pricing with DNN supervised learning
Speaker:	Hitoshi Iwasaki
Email:	iwasaki@u.nus.edu
Abstract:	We develop an innovative deep neural network (DNN) supervised learning approach to extracting insightful topic sentiments from analyst reports at the sentence level and incorporating this qualitative knowledge in asset pricing and portfolio construction. The topic sentiment analysis is performed on 113,043 Japanese analyst reports and the topic sentiment asset pricing model delivers superior predictive power on stock returns with adjusted R2 increasing from 1.6% (benchmark model without sentiment) to 14.0% (in-sample) and 13.4% (out-of-sample). We find that topics reflecting the subjective opinions of analysts have greater impact than topics of objective facts and justification of the quantitative measures.

Title:	Investing with cryptocurrencies - topic modelling dependent information retrieval from cryptocurrency experts
Speaker:	Dr. Simon Trimborn
Email:	simon.trimborn@nus.edu.sg
Abstract:	Cryptocurrency (CC) market is a new market emerged in 2009. Sentiment is believed to have influence on CC investment, hence sentiment information can have effect in portfolio investment. We retrieve topic modelling dependent information from CC experts' tweets and investigate the connection to the market state.

Title:	Optimal stopping via reinforced regression
Speaker:	Prof. Dr. Vladimir Spokoiny
Email:	spokoiny@wias-berlin.de
Abstract:	In this note, we propose a new approach towards solving numerically optimal stopping problems via reinforced regression based Monte Carlo algorithms. The main idea of the method is to reinforce standard linear regression algorithms in each backward induction step by adding new basis functions based on previously estimated continuation values. The proposed methodology is illustrated by a numerical example from mathematical finance.

Title:	Sentimental markets: how information flow drives asset prices
Speaker:	Dr. Richard Peterson
Email:	rpeterson@marketpsych.com
Abstract:	There is substantial and growing quantitative evidence that information flow drives asset prices across equities, currencies, commodities, and cryptocurrencies. Innovations in natural language processing, cloud computing, and the proliferation of online news and social media have yielded vast new datasets on which to study the role of information (themes and sentiments) in driving human risk taking behavior, and by proxy, asset prices. Dr. Richard Peterson will review academic literature as well as industry best practices in examining the role of media (news, social media, and search) in driving price patterns. Dr. Peterson will also review machine learning and statistical techniques that appear to capture significant relationships in media data. Dr. Peterson has an Electrical Engineering undergraduate degree, a medical doctorate, and completed residency training in psychiatry. He performed postdoctoral research in Neuroeconomics at Stanford and is a board-certified psychiatrist. As an Associate Editor of the Journal of Behavioral Finance, a former fund manager, and a data vendor of the Thomson Reuters MarketPsych Indices, Dr. Peterson presents a hybrid academic and industry understanding of asset pricing phenomena.

Title:	Efficient estimation of high-dimensional dynamic covariance by risk factor mapping: applications for financial risk management
Speaker:	Prof. Mike K P SO
Email:	immkpso@ust.hk
Abstract:	This paper aims to explore a modified method of high-dimensional dynamic variance-covariance matrix estimation via risk factor mapping, which can yield a dependence estimation of asset returns within a large portfolio with high computational efficiency. The essence of our methodology is to express the time-varying dependence of high-dimensional return variables using the co-movement concept of returns with respect to risk factors. A novelty of the proposed methodology is to allow mapping matrices, which govern the co-movement of returns, to be time-varying. We also consider the flexible modeling of risk factors by a copula multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) model. Through the proposed risk factor mapping model, the number of parameters and the time complexity are functions of a small number of risk factors instead of the number of stocks in the portfolio, making our proposed methodology highly scalable. We adopt Bayesian methods to estimate unknown parameters and various risk measures in the proposed model. The proposed risk mapping method and financial applications are demonstrated by an empirical study of the Hong Kong stock market. The assessment of the effectiveness of the mapping via risk measure estimation is also discussed.

Title:	Utilizing high-dimensional high-frequency data for lower sampling frequencies
Speaker:	Niels Wesselhöfft
Email:	wesselhn@hu-berlin.de
Abstract:	Weekly, quarterly and yearly risk distributions are crucial for risk reporting according to Basel III and Solvency II. For the respective data frequencies, the authors motivate that available data series are not sufficient in order to estimate Value at Risk and Expected Shortfall of high confidence levels. Accordingly, this paper presents a semi-parametric estimation method, rescaling data from high- to low-frequency which allows to obtain significantly more data points for the estimation of the respective distributions. The presented methodology in the α-stable framework, which is able to mimic multifractal behaviour in asset returns, provides tail events which never occurred in the original low-frequency dataset.

Title:	Vector error correction models to measure connectedness of bitcoin exchange markets
Speaker:	Paolo Pagnottoni
Email:	paolo.pagnottoni01@universitadipavia.it
Abstract:	Bitcoins are traded on various exchange platforms and, therefore, prices may differ across trading venues. We aim to investigate return connectedness across eight of the major exchanges of Bitcoin, both from a static and a dynamic viewpoint. To this end, we extend the order-invariant forecast error variance decomposition proposed by Diebold and Yilmaz (2012) to a generalized vector error correction framework. Our results suggest that there is strong connectedness among the exchanges, as expected, although some of them behave dissimilarly. We identify Bit nex and Coinbase as leading exchanges during the considered period, while Kraken as a follower exchange. We also obtain that connectedness across exchanges is strongly dynamic, as it evolves over time.

Title:	Variable selection with big data based on Zero Norm and via sequential Monte Carlo
Speaker:	Prof. Jin-Chuan Duan
Email:	bizdjc@nus.edu.sg
Abstract:	Selecting a subset from many potential explanatory variables in linear regressions has long been the subject of research interest, and the matter is made more important in the era of big data when many more variables become available/accessible. Of late, the l_1-norm penalty based techniques such as Lasso of Tibshirani (1996) have become very popular. However, the variable selection problem in its natural setting is a zero-norm penalty problem, i.e., a penalty on the number of variables as opposed to the l_1-norm of the regression coefficients. The popularity of the l_1-norm penalty or its variants has more to do with computational considerations, because selection with the zero-norm penalty is a highly demanding combinatory optimization problem when the number of potential variables becomes large. In this paper, we devise a sequential Monte Carlo (SMC) method as a practical tool for zero-norm variable selection problems, and the selection task can, for example, be completed under half-an-hour using a typical multi-core desktop computer for a problem of selecting out of 1,000 potential variables regardless of the number of observations involved. The essence of our SMC method is to understand that the selection problem is equivalent to the task of sampling from a discrete probability function defined over all possible combinations comprising, say, k regressors out of p >= k potential variables, where the peak of this function corresponds to the optimal combination. The solution technique sets out to sequentially generate samples, and after a while the final sample represents the target probability function. We demonstrate through a simulation study the method's superiority in selecting right variables vis-a-vis the adaptive Lasso.

Workshop organizers: Ying Chen, Min Dai, Ulrich Horst, Simon Trimborn, Chao Zhou
Workshop coordinator: Ge Zhang
Workshop is sponsored by NUS-RMI and HUB-NUS joint grant “Augmented Intelligence in Digital Society”.