Introduction to Machine Learning and Network Analytics in Finance Minitrack

József Mezei
School of Business and Management
Lappeenranta University of Technology
Risklab Finland, Arcada University of Applied Sciences
Jozsef.Mezei@lut.fi

Peter Sarlin
Department of Economics
Hanken School of Economics
Risklab Finland, Arcada University of Applied Sciences
peter@risklab.fi

We are experiencing enormous growth in the interest of application of various computational methods in finance, which is the consequence of various developments in the last 15 years. As a result, the number and importance of contributions utilizing various machine learning techniques and network analytics has increased significantly in many areas of finance. The presence of previously unprecedented structured and unstructured data in interconnected financial markets calls for the development of new models capable of addressing the complex underlying phenomenon. The papers accepted to the minitrack offer a representative sample of recent advances in the application of big data, machine learning and network analytics in a financial context. A common theme in most of the papers is the combination of structured and unstructured (textual) data when developing new applications.

We accepted four papers to the minitrack to be presented in a single session. The articles offer a nice combination of methods and application areas dominating the domain in recent years, ranging mainly from approaches for text mining to network analysis. The papers were selected from eight high quality submissions.

The first paper "*What do they mean? Using Media Richness as an Indicator for the Information Value of Stock Analyst Opinion regarding post-earnings Firm Performance*" is authored by Matthias Eickhoff (University of Göttingen). The paper proposes to apply media-richness theory as the basis of understanding firm performance. The model combines the use of high-richness (text of analyst reports and earnings conference calls) and low-richness (stock price and balance sheet related variables) data sources. The results of the data analysis demonstrate the benefits of using high-richness data sources as they are shown to include more investment relevant information than traditionally used low-richness data.

The second paper "*Bayesian network models for systemic risk estimation*" by Paolo Giudici (University of Pavia), Paola Cerchiello (University of Pavia) and Giancarlo Nicola (University of Pavia) presents a new Bayesian network model to estimate systemic risk. The model incorporates both structured and unstructured data in the form of financial market data and financial tweets, respectively. The proposed graphical Gaussian model is exemplified through the analysis of the Italian banking system.

The third paper "*How The Market Can Detect Its Own Mispricing - A Sentiment Index To Detect Irrational Exuberance*" is authored by Jonas Krinitz (Albert-Ludwigs-Universität Freiburg), Simon Alfano (Albert-Ludwigs-Universität Freiburg) and Dirk Neumann (Albert-Ludwigs-Universität Freiburg). The authors propose a new sentiment index-based model to capture the relationship between news sentiment and stock prices. The introduced "Sentiment to CDAX index'" can serve as an early-warning indicator by identifying potential over- or undervaluation of stock prices. One novelty of the paper is to utilize corporate announcements as the basis of the sentiment index instead of traditionally used journalistic news.

The fourth paper "*Possibilistic clustering for crisis prediction: Systemic risk states and membership degrees*" by József Mezei (Lappeenranta University of Technology) and Peter Sarlin (RiskLab Finland at Arcada and Hanken School of Economics) proposes to use possibilistic clustering for systemic risk analysis. In the proposal, various vulnerability indicators are used as the basis of a generalized objective function-based mixed fuzzy-possibilistic clustering model. The resulting clusters are used as the basis of a classification procedure to pre-crisis states. The developed model is used to identify vulnerability to systemic banking crises in Europe during the last decades.

URI: http://hdl.handle.net/10125/41322
ISBN: 978-0-9981331-0-2