Data science e Fintech. Audizione in Commissione Finanze Camera dei Deputati, Roma, 25 ottobre 2017

Paolo Giudici

Professore Ordinario di Data Science e Statistica
Dipartimento di Scienze Economiche ed Aziendali, Universitadi Pavia
giudici@unipv.it; Twitter: @stateconomist; Facebook: Data Science Network

Premessa

Una nota simile a questa e' stata sottomessa alla: European Commission public consultation on fintech. E' stata inoltre discussa nel corso del ciclo di incontri organizzati dalla CONSOB, in collaborazione con le principali Fintech e Universita' Italiane.

Nei prossimi giorni, verrÃ inviata a:

- BIS Consultative Document on sound practices on the implications of fintech developments for banks and bank supervisors;
- the EBA Discussion Paper on financial technology consultation;
- the ECB Guide to the assessment of fintech credit institution license applications.

Background

- FinTech denotes companies that combine financial services with (on-line) innovative technologies.
- Advances in big data and big data modelling (data science) have enabled Fintechs to provide **competitive** (e.g. faster and cheaper) banking and finance services: creditech, paytech, wealthtech, insurtech, regtech,...
- More competitive services may bring, however, higher risks to the consumers: e.g. **cyber risk** (the risk of financial losses due to an operational failure in the fintech IT systems) and **scoring risk** : the risk that consumers' choices may be misguided by wrong information, particularly in terms of credit rating. Both are amplified by **systemic risk**, due to the high interconnectdness of Fintech companies.

Proposal

- At @unipv I coordinate a Data Science laboratory since 2001, currently with 3 Faculty, 1 post-doc and 8 Phd students (5 of them from abroad). The laboratory carries out research and training in collaboration with Banks, Fintechs, and Regulators.
- Our current aim is to improve data science methodologies and, specifically, the modelling of information coming from correlation, social and transactional networks, to evaluate and improve scoring and rating of both banks and fintech platforms.
- Examples: credit scoring models in P2P lending; matching risk profiles in roboadvisory; price discovery in cryptocurrency markets; assessment and prioritisation of cyber risks.

Case study: credit scoring in P2P Lending

- So far we have collected from: Lending Club, Mode Finance, Borsa del Credito and N26 (in progress).
- Different statistical and machine learning methods have been applied to the data, with the aim of finding the best predictive model (in terms of default predictive accuracy).
- Here we refer to Mode Finance data, made up of SMEs which have applied for a loan via a P2P lending platform. The proportion of defaults in the sample is equal to 23%.

Results: the network based on the activity indicator - weighted

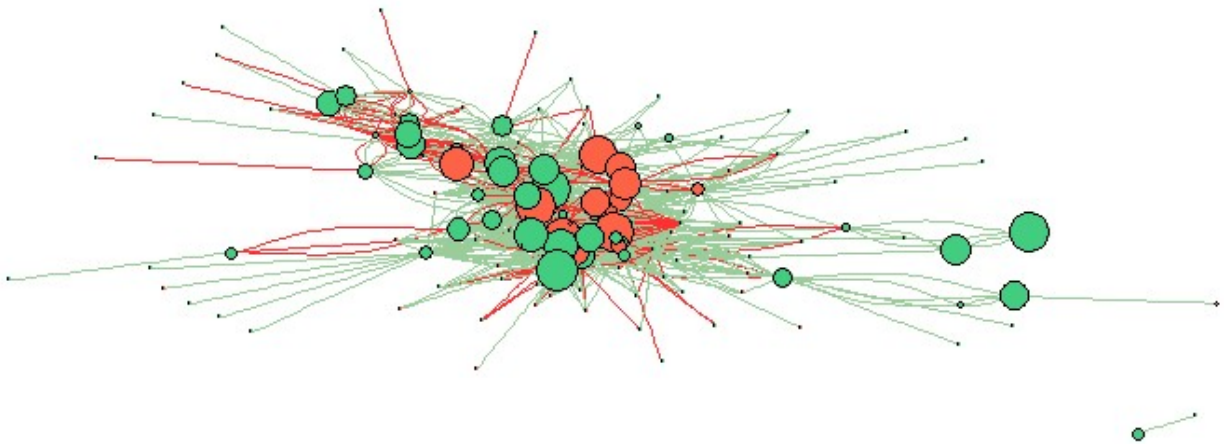

Figure: Correlation network based on the activity indicator.

The network based on the solvency indicator

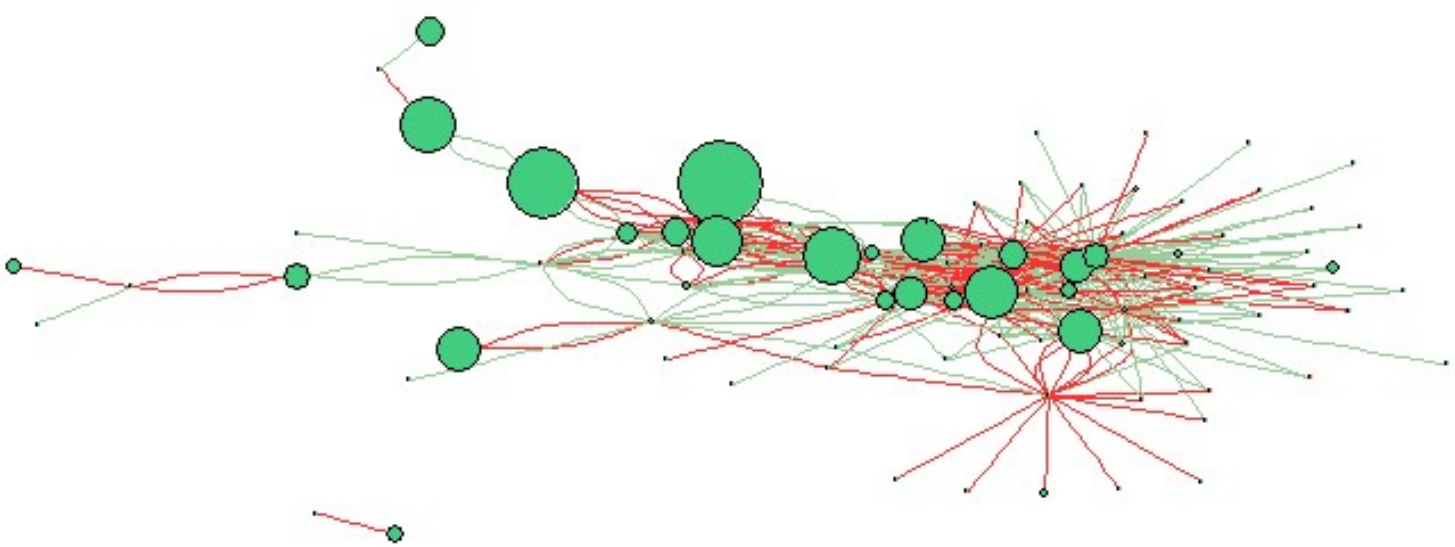

Figure: Correlation network based on the solvency indicator

Predictions: Standard Model

Variable	Estimate	P-value	Significance
Intercept	-3.39	0.011	*
Solvency ratio	0.01	0.539	
Debt to equity ratio	-0.07	0.517	
Current ratio	0.21	0.032	*
Cash over total assets	-2.51	0.579	
Return on equity	-0.08	0.008	**
Return on assets	0.01	0.963	
Return on Capital Employed	0.09	0.044	*
Coverage	-0.01	0.875	
Activity ratio	-1.92	0.001	***
Predictive accuracy (AUROC)			**0.71**

Table: The estimated baseline regression model

Predictions: Network based model

Variable	Estimate	P-value	Significance
Intercept	-1.53	0.033	*
Solvency ratio	-0.02	0.012	*
Debt to equity ratio	-0.00	0.576	
Current ratio	0.24	0.072	*
Cash over total assets	1.08	0.443	
Return on equity	-0.11	0.000	***
Return on assets	0.02	0.876	
Return on capital employed	0.01	0.212	
Coverage	0.02	0.248	
Activity ratio			
Degree Centrality	0.01	0.026	*
Closeness	1.05	0.002	**
Predictive accuracy (AUROC)			0.82

Main findings

- Network models can improve default predictions and, therefore, better protect P2P lending users. In addition, they provide useful information that can used to monitor companies that may trigger and spread contagion.
- We expect that further network information (e.g. social networks and transactional networks) can further improve model performance.
- Similar results are emerging in roboadvisory risk profile matching, crypto price discovery and in the prioritisation of cyber risks.

Policy implications

- Data Science methods can be very helpful to protect investors and to preserve financial stability, encouraging the diffusion of Fintech innovations.
- This requires investing in financial data science education, a mix between Finance, Statistics and Computer Programming, and especially in "collaborative" education, through cooperation between Universities, the Financial industry, and Regulators.
- Such cooperation can be boosted by "light budget" initiatives, such as: a) "outsourcing" the activity of data science laboratories in Fintech districts, innovation hubs and sandboxes; b) issuing "dedicated" research projects, that include funding of Phd and Post-doc positions; c) create a "pooled database" of network data that can augment existing databases (such as Credit Bureaus), for the benefit of an

Some Recent papers

- (2017) Pejman Abedifar, Paolo Giudici, Shatha Hashem. Heterogeneous market structure and systemic risk: evidence from dual banking systems. **Journal of Financial Stability.**
- (2017) Paolo Giudici, Peter Sarlin, Alessandro Spelta. The interconnected nature of financial systems: direct and common exposures. **Journal of Banking and Finance.**
- (2017) Paolo Giudici, Laura Parisi. Sovereign risk in the Euro area: a multivariate stochastic process approach. **Quantitative Finance.**
- (2016) Paolo Giudici, Alessandro Spelta. Graphical network models for international financial flows. **Journal of Business and Economic Statistics**, 34 (1), pp. 126-138.
- (2017) Paolo Giudici, Laura Parisi. Correlation networks to measure the systemic implications of bank resolution. Submitted to Review of Finance
- (2017) Paolo Giudici, Laura Parisi. Corisk: measuring systemic risk through default probability contagion. Submitted to Quantitative Finance
- (2017) Stefan Avdjiev, Paolo Giudici, Alessandro Spelta. Measuring contagion risk in international banking. Bank for International Settlements research paper. Submitted to Journal of Financial Stability
- (2017) Paolo Giudici, Branka Hadji Misheva. Scoring models for P2P lending platforms: a network approach. Submitted to Journal of Banking and Finance