Link prediction in knowledge networks using exogenous and endogenous attributes: a machine learning approach

by Giovanni Cerulli | Antonio Zinilli | IRCRES-CNR | IRCRES-CNR

Abstract ID: 51
Inviato: 07/04/2022
Evento: XX Workshop Annuale SIEPI
Argomento: 12. Reti, infrastrutture, energia
Parole chiave: Knowledge networks, Machine learning, Optimal prediction

We propose a supervised machine learning approach to predict partnership formation between universities. We focus on successful joint R&D projects funded by Horizon 2020 programme in three research domains: Social Sciences and Humanities, Physical and Engineering Sciences, and Life Sciences.
We perform two connected analyses: link formation prediction, and feature importance detection.
As for link prediction, using out-of-sample cross-validated accuracy and a set of network endogenous and exogenous attributes, we obtain 90% prediction accuracy when both types of attributes are used, and around 65% when using only the exogenous ones. This proves that partnership predictive power is on average 25% larger for universities already incumbent in the programme than for newcomers.
As for feature importance, by computing super-learner average partial effects and elasticities, we find that the endogenous attributes are the most relevant in affecting the probability to generate a link, and observe a largely negative elasticity of the link probability to feature changes, fairly uniform across attributes and domains.Inserisci l'Abstract (max 200 parole).