COLLABORATIVE DOCTORAL PARTNERSHIPS – CALL 2020

THEMATIC FIELD 13: Robustness in complex data analysis and statistical modelling

JRC RESEARCH AREA DESCRIPTION

The "robust statistics" reduces the risk to draw wrong conclusions because of incorrect measurements and observations that seem inconsistent with the rest of the data set: the so-called outliers. Robust statistics also helps to disentangle from the bulk of the data observations clustered around peculiar but interesting pattern. In applications like anti-fraud, outliers and systematic patterns are of great interest. In other application domains, like text and speech translation, which heavily rely on quality data for training deep neural networks or machine learning methods, outliers need to be identified precisely and then eliminated from the training set: this is expected to increase considerably the generalization capacity of the trained models.

The above mentioned application domains, and many others, nowadays dispose of very large amount of data, which often are also very complex, in terms of the number of variables (also called features) that need to be treated simultaneously: in analysing these data it is almost unavoidable to incur in errors, outliers or insignificant or redundant variables/features.

The robust statistical analysis of big and complex data is still an open challenge. This collaborative doctoral partnership proposes to investigate and fill the gap between modern deep learning / machine learning methods with the robustness need.

The working scenarios and case studies will be chosen among concrete policy needs defined by the European Commission, like those mentioned above, or will be proposed by the applicant.

MAIN POLICY FIELDS

Anti-fraud, Customs Union, Single Market, Disinformation.

LINKS / URL WEBSITES

- https://ec.europa.eu/jrc/en/research-topic/antifraud
- https://it.mathworks.com/matlabcentral/fileexchange/72999-fsda

LINKS / REFERENCES TO PUBLICATIONS

- https://orcid.org/0000-0002-0987-3823