Data Science for big bank data

Jorge Martín Arevalillo[1]

[1]Departamento de Estadística, Investigación Operativa y Cálculo Numérico,
Universidad Nacional de Educación a Distancia (U.N.E.D)

Statistical and Machine Learning algorithms appear as part of our ongoing research practice as statisticians and mathematicians, having become powerful tools to address different challenges arising from Big Data scenarios in the banking sector. The research advances in the data science field can complement information database storage and business know-how in the construction of mathematical and statistical models that help to guide the management of commercial campaigns and also highlight business insights for data-driven decision making. Some ideas about how scientific research from academia can be put into action to generate business value are shown as well as several projects carried out for the banking sector and show that the journey from mathematical research to business value is achievable. All these initiatives have arisen as a result of pursuing research in the development of Statistical and Machine Learning algorithms and their applications to solve problems in different fields [1-5]. They all belong to the contracting framework of the Knowledge and Research Transfer Office at University Nacional de Educación a Distancia (U.N.E.D).

Keywords: Big Data. Banking. Machine Learning. Statistical Learning. Business value

REFERENCES

[[1] Jorge M Arevalillo (2019). A machine learning approach to assess price sensitivity with application to automobile loan segmentation. Applied Soft Computing. 76.390-399.

[2] Jorge M Arevalillo, Marcelo B Sztein, Karen L Kotloff, Myron M Levine, Jakub K Simon (2017). Identification of immune correlates of protection in Shigella infection by application of machine learning. Journal of Biomedical Informatics. 74. 1-9.

[3] Jorge M Arevalillo, Hilario Navarro (2013). Exploring correlations in gene expression microarray data for maximum predictive – minimum edundancy biomarker selection and classification. Computers in Biology and Medicine. 43 (10). 1437-1443.

[4] Sara Jurdao, Emilio Chuvieco, Jorge M Arevalillo (2012). Modelling forest fire ignition probability from satellite estimations of life fuel moisture content. Fire Ecology. 8 (1). 77-97.

[5] Jorge M Arevalillo, Hilario Navarro (2011). Uncovering bivariate interactions in high dimensional data using Random Forests with data augmentation. Fundamenta Informaticae. 113 (2). 97-115.