

Data Science, Data Science, Data Science...

Are you a data scientist?

M. Rosário Oliveira

*CEMAT and Dep. Matemática, Instituto Superior Técnico, Univ. Lisboa,
rosario.oliveira@tecnico.ulisboa.pt*

Palavras-chave: data science, statistics, machine learning, big data.

Resumo: Data Science is the area that studies problems associated with data acquisition, storage, and organization, as well as the extraction of useful information (or knowledge) from that data. Usually, these problems involve large volumes and complex datasets, called big data. The abundance of data can be justified by several factors. One is the expansion of the Internet. According to the Internet Live Stats [1], in 2016 around 40% of the world's population had access to the Internet (in Portugal this number grows to 67.3%). The Internet activities of these users produce vast amounts of data about their professional interests, social preferences, and habits. Another factor is technological advances in computing, data storage, and telecommunications that, among other things, has changed the way human beings interact and live. To deal with this immense data volume and complexity several theoretical and methodological tools were developed that supported technological growth. Side-by-side with more traditional disciplines like mathematics and statistics or informatics, more recent scientific areas like machine learning and artificial intelligence, visualisation of data and information, among others, gave their contribution.

The ultimate challenges for Data Science are to transform the enormous abundance of data into useful information for the solution of a specific problem. The applications of Data Science are vast, namely: telecommunications, biology, astronomy, economics, finance, law, sociology, climatology, health sciences, and in general all areas of industry.

The interest of companies and organizations in Data Science has recently exploded. The private and public sectors are increasingly interested in the value of information and its use as a differentiating factor among competitors. Moreover, it is unquestionably the value of Data Science as a tool that enables organisations to extract valuable insights from data, and use them for substantial societal benefit [2].

In this talk, we discuss the links between Data Science, Statistics, Machine Learning, Artificial Intelligence, Data Mining, and other scientific fields that put data as its main concern. We dedicate special attention to the role of Statistics in an area of data abundance. Topics like career opportunities, education, and the importance of collaborations between industry, universities, and government sectors are also addressed.

Referências

- [1] <https://www.internetlivestats.com/internet-users/>, October 2019.
- [2] Royal Statistical Society. *Dynamics of data science skills: How can all sectors benefit from data science talent?* 2019. This report can be viewed online at: royalsociety.org/dynamics-of-data-science-skills