Abstract

Data analysis is prominent in data science researches, but by each day data usage is expanding, and in recent times the usage is becoming indispensable and inseparable in all works of life including engineering profession. This is why data engineering as a discipline sprang up - using the data analysis techniques from statistics, machine learning, pattern recognition or neural networks, together with other technologies such as visualization, optimization, database systems, knowledge discovery etc to produce systems needed in diverse business, science and social science domains. This paper is a novel presentation of data analysis and data engineering discipline, focusing on critical issues that are relevant to both, but divulging more the new trend of moving data science beyond data analysis, to data engineering. Data engineering is a multi-disciplinary field with applications in control, decision theory, and in the emerging areas like bioinformatics. Data engineering is needed in critical activities for business, engineering, and scientific organizations, since service oriented architecture and web services has moved into full swing.
References

5. Chris Olsen, Roxy Peck, Jey L Devore; Introduction to Statistics and Data Analysis; Chegg Books EISBN-13: 9781305445963
7. DJ Patil, Hilary Mason (2015) Data Driven Publisher: O’Reilly Media

Index Terms

Computer Science

Information Sciences

Keywords

Data Science, Data Engineering, Data analysis, Data pipelines, Data infrastructure