

Networked data-science for research, academic communities and beyond

Wednesday 24 October 2018 18:00 (1 hour)

There is an exceptional way of doing data-driven research employing networked community. The following examples can illustrate the approach: Galaxy Zoo or Tim Gower's blog. However many cases of collaboration with the data-science community within competitions organised on Kaggle or Coda Lab platforms usually get limited by restrictions on those platforms. Common Machine Learning quality metrics do not necessarily correspond to the original research goal. Constraints imposed by the problem statement typically look artificial for ML-community. Preparing a perfect competition takes a considerable amount of efforts. On the contrary research process requires a lot of flexibility and ability to look at the problem from different angles. I'll describe the alternative research collaboration process can bridge the gap between domain-specific research and data science community. Notably, it can involve academic researchers, younger practitioners and all enthusiasts who are willing to contribute. Such research process can be supported by an open computational platform that will be described along with essential examples for the audience of the workshop.

Presenter: USTYUZHANIN, Andrey