

5th International Conference on

Big Data Analysis and Data Mining

June 20-21, 2018 | Rome, Italy

Machine learning on source code

Vadim Markovtsev
Madrid , Spain

Machine learning on source code (MLoSC) is an emerging and exciting domain of research which stands at the sweet spot between deep learning, natural language processing, social science and programming. We've accumulated petabytes of source code data which is open, yet there have been few attempts to fully leverage the knowledge that is sealed inside. This talk gives an introduction into the current trends in MLoSC, presents the tools and some of the applications, such as deep code suggestions and structural embeddings for fuzzy deduplication. There will be an additional emphasis on mining the "big code".

vadim@sourced.tech