NEC'2019

XXVII International Symposium on Nuclear Electronics & Computing

Montenegro, Budva, Becici, 30 September - 4 October 2019

Contribution ID: **245**   Type: **Sectional**

DYNAMIC APACHE SPARK CLUSTER FOR ECONOMIC MODELING

Thursday 3 October 2019 12:45 (15 minutes)

Modern econometric modeling of macroeconomic processes usually meets certain challenges due to the incompleteness and heterogeneity of the initial information, as well as huge data volumes involved. In the work, on the example of modeling the level of employment in the regions of the Russian Federation was shown the effectiveness of joint using Big Data technologies and automated deployment of a dynamic virtual computing cluster for solving such problems. There were constructed several models of the regional labor market, taking into account such basic macroeconomic indicators as per capita income, the volume of paid services to the population per capita, the industrial production index and others. The classification of the subjects of the Russian Federation according to the level of employment was obtained, it is stable against different methods (single linkage, complete linkage, Ward's method). For the analysis, it was used a dynamic Apache Spark cluster deployed by the means of the SIMPLE environment developed at CERN.

Primary author: Ms GAVRILENKO, IULIIA (Research Assistant, Plekhanov Russian University of Economics, Moscow, Russia)

Co-authors: Dr LITMAATH, Maarten (CERN); SHARMA, Mayank (CERN); TATYANA, Tikhomirova (Plekhanov Russian University of Economic)

Presenter: Ms GAVRILENKO, IULIIA (Research Assistant, Plekhanov Russian University of Economics, Moscow, Russia)

Session Classification: Machine Learning Algorithms and Big Data Analytics

Track Classification: Machine Learning Algorithms and Big Data Analytics