Anton Grudev

L.M.Korotenko, research supervisor

N.M. Nechai, language adviser

National Mining University, Dnipro, Ukraine

The Importance and Problems of Big Data

In the era of high-tech we can hear the term Big Data more and more often. This fact indicates that the importance of Big Data constantly increases. This term is also used with related concepts such as Business Intelligence or data mining. But what does that mean?

The Big Data is the set of algorithms that are designed to analyze a very large amount of constantly changing data which can be complex or unstructured. Nowadays the amount of data that's being created and stored on a global level is unimaginable and it just keeps growing. That means there's potential to collect key findings from data. The importance of BD is hard to overrate. It can be used in wide range of needs. It is already used by banking, businesses, media, retail, transportation systems etc. For instance, it makes easy to know how to develop infrastructure of a city by summarizing data about the movements of transport or it can make a fairly accurate forecast about exchange rates by analyzing data that contain information about last exchange rates, world political climate, development of industry and so on. Also BD is widely used in Internet, as you may know web stores use the users' information (history of purchases and views) to make their advertising more effective. And all of these are the only top of iceberg. The BD system already starts to be used by different companies. Despite the effectiveness of BD there are some really important problems. Firstly, the users of BD face the lack of memory because BD system consumes lots of it. And this fact requires large investments of money for storage. But there's a solution: users can reduce their expenses by using cloud storages. The second problem is the lack of experience of using BD technologies. There's the short amount of working examples which can be used by companies. So they are afraid of losing money by implanting BD analytics in their structure. Then there is one feature of gathered data for analyzing. It is clear that without good data you cannot get useful findings. As a rule the information that clients have is not useful. So, this problem can be solved by creating special software that will sort information. And the last one is that there's a shortage of data scientists, analysts, engineers and this is a big problem because these specialists create new algorithms or software for solving specific challenges of BD and providing modernization of technologies that get useful information from data. However, there are already some courses for training specialists in this field of study on the Internet and in the near future their number will be increased.

So, the main problems of providing the projects of BD are: security of data, the lack of money and specialists, shortage of data providers, unavailability of companies to introduce the project of BD. To sum up, I'd like to say that the Big Data system pushes the limits of analyzing data and creates new opportunities for making highly useful and correct conclusions.