An application of Naïve Bayes Classifier to explore big data using XLSTAT

■ **M.S. Kasture, A.J. Shivagaje, C.G. Shelake and A.J. Nalavade**

See end of the paper for authors' affiliations

Correspondence to :

A.J. Shivagaje
Post Graduate Institute (M.P.K.V.), Rahuri, **Ahmednagar (M.S.) India**
Email : ashokjs123@ gmail.com

ABSTRACT : The present ICT era has changed the scenario of multivariate data or information usage altogether. Organizations treat data as an asset and they try to employ various methodology to come up with organization progress oriented conclusions. A wide range of database tools to manage the huge data and equally number of software's are also developed to visualize, present and analyses the big data. XLSTAT with diversified data analyzing utilities, is one such tool that can be appended to usual Excel software. The present paper gives an application of Naïve Bayes Classifier applied to a data on Global Super Store Orders-2016 (Source: secondary data obtain from data.world platform). This application will give the insight of understanding the concept of Naïve Bayes Classifier. It will also show the effect of continuous data monitoring and maintenance on derived results of Naïve Bayes Classification. The summary of derived output will facilitate the comparison and will also give an idea about the overall trend of the factors under study. The step based analysis of big and diverse data shows that global accuracy of Naïve Bayes Classifier increases with increase in data size.

KEY WORDS : Big data, Confusion matrix, Global accuracy of the model, Posterior probability, Regression