Predictive Analytics On Political Data

Abstract:

Data in every organization are increasing in a rapid speed along with the volume of it increasing at a large extend through various domains such as education, social network, meteorology, government and much more. Accordingly, Big Data refers to data as traditional data, Machine-generated Sensor data, and Social data which are both structured and unstructured. Apache’s Hadoop has proven to provide salient properties such as scalability, ease of use, and most notable robustness to node failures. Processing K-Nearest Neighbor queries in high-dimensional data have received a lot of attention by researchers in recent a year, which provides a way for predicting in a better way as done in this analysis. This research analysis is to find the means to enhance the classification accuracy, utilizing logistic regression and K-Nearest Neighbor a fast process technique in a machine learning technique for classification. The work aims at enhancing the rate of accuracy, true positive, false positive, precision, recall, sensitivity, specificity rate of classification using the Hadoop platform.