Chapter 25
Big Data Analytics and Visualization for Food Health Status Determination Using Bigmart Data

Sumit Arun Hirve
VIT-AP University, India

Pradeep Reddy C. H.
VIT-AP University, India

ABSTRACT

Being premature, the traditional data visualization techniques suffer from several challenges and lack the ability to handle a huge amount of data, particularly in gigabytes and terabytes. In this research, we propose an R-tool and data analytics framework for handling a huge amount of commercial market stored data and discover knowledge patterns from the dataset for conveying the derived conclusion. In this chapter, we elaborate on pre-processing a commercial market dataset using the R tool and its packages for information and visual analytics. We suggest a recommendation system based on the data which identifies if the food entry inserted into the database is hygienic or non-hygienic based on the quality preserved attributes. For a precise recommendation system with strong predictive accuracy, we will put emphasis on Algorithms such as J48 or Naive Bayes and utilize the one who outclasses the comparison based on accuracy. Such a system, when combined with R language, can be potentially used for enhanced decision making.

INTRODUCTION

Classes of 'Big Data'

Big Data could be found in three structures: structured, unstructured, semi-structured.

Structured

Any information that can be put away got to and prepared as the settled configuration is named as a 'Structured' information. Over the timeframe, ability in software engineering has made more noteworthy progress in creating methods (Li, Wang, Lian et al., 2018) for working with such information (where the organization is outstanding ahead of time) and furthermore inferring an incentive out of it. In any case, presently days, we are predicting issues when the size of such information develops to an immense degree, normal sizes are being in the fury of a various zettabyte.

Unstructured

Any information with the obscure frame or the structure is delegated unstructured information. Not-withstanding the size being tremendous, un-organized information represents various difficulties as far as its preparing for inferring an incentive out of it. The regular case of unstructured information is a heterogeneous information source containing a blend of basic content records, pictures, recordings and so on (Anandakumar & Umamaheswari, 2018). Presently, multi-day associations have an abundance of information accessible with them yet sadly they don't realize how to determine an incentive out of it since this information is in its crude shape or unstructured arrangement.

Semi-structured

Semi-organized information can contain both types of information. We can see semi-organized information as a structured in frame however it is not characterized with for example a table definition in social DBMS. A case of semi-organized information is information spoken to in XML document.

Qualities of 'Big Data'

Volume

The name Big Data itself is identified with a size which is tremendous. Size of information assumes the extremely vital job in deciding an incentive out of information (Ahmed, 2019). Likewise, regardless of whether a specific information can be considered as a Big Data or not, it is an endless supply of infor-mation. Henceforth, 'Volume' is one trademark which should be considered while managing Big Data.

Variety

The following part of Big Data is its assortment (Lee, 2019). Assortment alludes to heterogeneous sources and the idea of information, both organized and unstructured. Amid prior days, spreadsheets and

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/big-data-analytics-and-visualization-for-food-health-status-determination-using-bigmart-data/290999

Related Content

Use of Big Data Analytics by Tax Authorities
Brendan Walker-Munro (2022). *Research Anthology on Big Data Analytics, Architectures, and Applications (pp. 1388-1412).*
www.irma-international.org/chapter/use-of-big-data-analytics-by-tax-authorities/291043

Big Data Analytics Using Local Exceptionality Detection
Martin Atzmueller, Dennis Mollenhauerand Andreas Schmidt (2016). *Enterprise Big Data Engineering, Analytics, and Management (pp. 108-125).*
www.irma-international.org/chapter/big-data-analytics-using-local-exceptionality-detection/154557

Big Data Applications in Healthcare Administration
Joseph E. Kasten (2020). *International Journal of Big Data and Analytics in Healthcare (pp. 12-37).*
www.irma-international.org/article/big-data-applications-in-healthcare-administration/259986

COVID-19 Deaths Previsions With Deep Learning Sequence Prediction: Bacille Calmette-Guérin (BCG) and Tuberculosis Track
Heni Bouhamed (2020). *International Journal of Big Data and Analytics in Healthcare (pp. 65-77).*
www.irma-international.org/article/covid-19-deaths-previsions-with-deep-learning-sequence-prediction/259989

A Markov-Chain-Based Model for Group Message Distribution in Connected Networks
Peter Bajorskiand Michael Kurdziel (2020). *International Journal of Data Analytics (pp. 13-29).*
www.irma-international.org/article/a-markov-chain-based-model-for-group-message-distribution-in-connected-networks/258918