IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Vol. 10, Issue 4, April 2023

DOI: 10.17148/IARJSET.2023.10416

Data Visualization Tools: A Quick Survey

Dr. S. Sarumathi*[1], T. Udhaya[2], N. Reshma[3], S. Pavithra[4]

Professor, Department of Artificial Intelligence and Data Science, K. S. Rangasamy College of Technology, Tamil Nadu, India*[1]

Department of Information Technology, K. S. Rangasamy College of Technology, Tamil Nadu, India[2-4]

Abstract: With the expansion of information technologies and applications, a large amount of data is created, attracting both the academic community to use this information for knowledge extraction and the industry to build knowledge-based systems. Visualization tool, pattern mining from datasets, and assessing data drift for distinct attributes are three popular machine learning and data science applications. Visualization tool is a critical application in which the right depiction of data may assist in identifying the required features. Data visualization can take numerous graphical forms, such as bar charts and line charts, so that users with non-technical backgrounds can readily utilize the tools to extract information. Furthermore, in addition to visual representations of data, creating drift analysis of data in a certain time period is a standard procedure in several industries such as mHealth, e-commerce, and MOOC education. This paper presents a detailed review of various data visualization tools for different applications. At first, various data visualization tools for different applications are studied in detail. A comparative study is conducted with their merits and demerits for identifying the challenges in those tools and then this paper is concluded with suggestions of solutions for enhancing the efficiency of data visualization tools.

Keywords: Data visualization tools, information technology, visual representation, graphical form.

I. INTRODUCTION

In recent decades, a massive amount of data has been generated from various sectors such as mHealth, education, ecommerce, and Smart City applications [1]. In 2015, for example, 2.5 quintillion bytes of information were produced. Furthermore, not only is volume of information increasing, but so is content creation speed; for example, it is estimated that in 2018, data generation speed will be 50,000 GB/day, up from 100 GB/day in 1992. As a result, managing and extracting knowledge from this massive amount of data is a time-consuming task.

The data science community can advantageously use this information to derive knowledge through using Big Data analytics tools as a result of computation technology and cloud computing services. Preprocessing data is crucial for data mining applications to effectively mine the right knowledge by implementing machine learning techniques. Thus, it is unquestionably important to design a generic data analytics system for the many areas of information technologies in the field of knowledge discovery from Big Data.

Data visualization [2-5] is a crucial activity where the right depiction of the information can assist in identifying the required features. Data visualization can take many different graphical shapes, such a bar chart or a line chart, to make it easier for those without technical backgrounds to use the tools to extract information. Besides that, creating drift analyses of data over a specified time period is a common practice in a variety of industries, including mHealth, e-commerce, and MOOC education. For instance, in e-commerce applications, the impact of location, mobility, and a number of other factors are engaged to forecast product consumption.

In this article the detailed information on different data visualization tools are studied. In addition to this, their performance efficiency and limitations are discussed to further improve the performance of data visualization tools. The remaining part of the article is organized as follows: Section II provides the previous researches related to data visualization tools. Section III compares performance efficiency of those data visualization tools. Section IV concludes an entire discussion and suggests new solution for future enhancement.

II. SURVEY ON DATA VISUALIZATION TOOLS

Data visualization saliency (DVS) model [6] was introduced for estimating the abstract data visualizations. The structural and visual characteristics of visualization which are contrary with the existing visual saliency models were determined and then the enhancement of a modified saliency models were discussed. During the analysis of these models, the

DOI: 10.17148/IARJSET.2023.10416

characteristics of these models were addressed and combined the new data on top-down attention. It can more precisely forecast which areas of visualization will catch the interest of a viewer.

A toolkit called DXR was developed [7] for immersive data visualization. In DXR, a brief declarative visualization grammar which is encouraged by Vega-Lite was used for better specification of visualization designs. Moreover, DXR further provided a Graphical User Interface (GUI) for simple and speedy adjustments and assessments of visualization designs in-situ, or while actually using the virtual environment. DXR also offers configurable graphic markers and reusable templates, providing interesting and captivating visualizations.

An interactive browser-based genomics data visualization tool [8] was developed for clinical and transactional laboratory applications. These applications were built for client-side data processing to answer questions regarding information privacy and work with any current web browser on desktop and mobile devices for convenient access. In order to aid in the diagnostic purposes of sequence variants, jsProteinMapper showed the location of an interesting variable in relation to several variant databases and protein domain databases. Using clinicopathologic annotations, jsComut created a highly dynamic and customisable computation map for visual interpretation of genomic data sets, revealing distinctive molecular identities and clinical consequences. The common circular codoneto-amino acid translation table was made interactive in jsCodonWheel, allowing users to instantly map nucleotide changes onto the associated amino acid variances.

A formula-based end user tool called Uvis [9] was developed for data visualization. Uvis is an interactive visualization and user interface tool aimed for end-user developers with knowledge similar to spreadsheet formulas. Designers may drag and drop visual items into Uvis, assigned visual features to formulas, and view the outcome instantly. The formulas resemble spreadsheet formulas and are descriptive. The formulas identify the values of the properties and can make use of database data, visual objects, and end-user input. Here, Uvis was contrasted with well-known visualization programs to support this argument.

A web-based application named as wiz [10] was presented for visualizing large volume of data. By partitioning the associations between the data's relationships using five dimensional visual analytics and conducting multivariate data analysis, such as principal component and linear discriminant analyses, Wiz was enabled scientists and non-scientists to disintegrate the complicated information. The results were presented in bright, publication-ready figures. Wiz was anticipated to be a crucial instrument with a serious influence in our big data era because to the growth of high-throughput techniques for components discovery, knowledge streaming possibilities, and the concentration on artificial intelligence and industrial digitalization.

A data structure [11] was introduced for visualizing the development of technology. The data structure named Dynamic Technology Footprint (DTFootprint) DTFootprint, makes it easier to analyze and visualize the developments and fluctuations of a certain technology and, consequently, the development of a particular technical area, business, or group of individuals. In order to honor the influence of acoustic frequency spectrograms, a pictorial tool on the basis of this data structure was developed and named the Technology Spectrogram (techSpectrogram). Rather of visualizing the fluctuations of an acoustic wave by displaying the progression of its frequency components, this tool displays the movements of a technology by displaying the evolution of its technological components, which were represented by the entire set of IPC-codes.

A web tool [12] ToxAnalyzer was developed for interactive data visualization of chemical compounds. It is an online application that helps individuals in potential toxicogenomic investigations by offering an orderly summary of all the chemical data available in CTD. This program retrieved data from CTD's servers and then processed and presented crucial details about each chemical component detected in CTD using a collection of Python scripts. A user learned more about published research, connected organisms, gene connections, and other related graphs by using the chemical name. User read the source publications via PubMed ID links if they required more information about a certain gene's interactions with chemical compounds.

A Transformed Sigmoid Risk Model (TSRM) [13] was proposed for data visualization on the basis of integrating a heat map and transferred sigmoid function. Using RGB color model, a numerical mapping approach between fire risk ratings and colors was developed. Fire Risk Visualization Tool (FRVT) based on the TSRM was created using HTML 5 and JavaScript. The JavaScript application program connection of Baidu Map was employed as the FRVT's geographic information system engine.

 This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IARJSET.2023.10416

A tool [14] called layerUMAP was presented for visualizing deep learning models in biological sequence classification. To handle learning high-level representations, autoBioSeqpy software and the UMAP library were incorporated. The tool's interactive feature, which enabled users to see the results of hidden layers along the depth of the model, was a massive benefit. The findings demonstrated that layerUMAP may provide illuminating visual feedback regarding models as well as further directed to create better models.

A handy web tool called StarmapVis [15] was developed to provide a dynamic posterior evaluation of spatial transcriptomic or single-cell expression data. Modern web browsers are used to investigate the plethora of viewing angles that two dimensional media cannot examined because to the simple user interface. Knowledge on clustering was displayed interactively in scatter plots, and connectivity networks show the progression and cross-comparison of data from various coordinate systems. This tool distinctive characteristic was automated camera view animation. A helpful animation transition from two-dimensional spatial omic data to three-dimensional single cell coordinates was also provided by StarmapVis.

III. COMPARATIVE ANALYSIS

A comparative analysis is presented in terms of merits and demerits of different data visualization tools whose operational details are studied in the above section. From the following Table 1, both merits and demerits in the above studied tools used for visualization of data are investigated and the best solution is suggested to overcome those drawbacks in data visualization tools to obtain better accuracy.

TABLE I COMPARISON OF DIFFERENT DATA VISUALIZATION TOOLS

Ref No.	Tools or Models	Merits	Demerits	Performance Metrics
[1]	Data visualization saliency (DVS) model	It offered a substantial improvement in performance over the other models.	It currently applies only to static images. It cannot capture the interactive component itself.	Percentage Improvement for the DVS Model Relative to Text only models: Area Under ROC Curve-Judd (AUC-J) = 2% Shuffled AUC (sAUC) = 5% AUC-Borji (AUC-B) = 4%
[2]	DXR	Reduces the need for tedious manual visual encoding and low-level programming to create immersive data-driven content.	Hope to leverage advanced graphics processing unit shader programs to improve frame rates for large data sizes.	For 10000 data items: Construction time = 11.59 secs Frame Rate = 170
[3]	Interactive browser-based genomics data visualization tool	Significantly improve the analysis of genomics data in clinical and translational settings.	Doesn't allows visual exploration of custom Data.	Average time to load the data and render the initial graphics was <2 seconds per gene.
[4]	UVIS	Relatively fast, and can customize visualizations.	Inability to create visual objects recursively.	Showing temperature greater than 25.
[5]	Wiz	Can handle datasets exceeding 50,000 instances by utilizing WebGL plot elements for large datasets.	Different sheets do not have to have the same feature variables which affect the performance.	Variance = 0.6%

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⋇ Impact Factor 8.066 ⋇ Vol. 10, Issue 4, April 2023

DOI: 10.17148/IARJSET.2023.10416

[6]	DTFootprint, techSpectrogram	Better efficiency.	It is not applicable for heterogeneous data.	-Nil-
[7]	ToxAnalyzer	User friendly.	Lack of custom configurations.	Number of genes sorted by organism for mus musculus = 12000
[8]	TSRM, FRVT	Conducive to ensuring regional process safety.	The fire risks posed by hazardous materials transportation (e.g., pipelines and trucks) were not considered in the TSRM and FRVT.	When generalization radius = 1024 m, 1536 m, and 2048 m, the fire risk distribution in the area can be appropriately expressed.
[9]	layerUMAP	Standard tool for the visual analysis of deep learning models.	Cannot capture the hidden outputs from the layers in a branch of the built model.	Accuracy = 78% (for 13 kernel size and 5o filter number)
[10]	StarmapVis	Higher frame rate.	Can further utilize low-cost VR headsets, such as Google Cardboard and Daydream, to enable more immersive visualization.	Availability = open source Smartphone support = Yes

IV. CONCLUSION

Data visualization tools are widely used in various fields, including business, science, education, and government. They are particularly useful for data analysts, researchers, and decision-makers who need to make sense of large amounts of data quickly and efficiently. Overall, data visualization tools are essential for transforming complex data sets into accessible and actionable insights that can drive decision-making and improve outcomes. In this paper, a survey on data visualization tools is presented in detail. Also, merits and demerits of these tools are discussed to suggest the future directions towards increase the performance of data visualization tools. From the comparative analysis, it is concluded that the StarmapVis works effectively to visualize the data. In future, StarmapVis can further utilize low-cost VR headsets, such as Google Cardboard and Daydream, to enable more immersive visualization.

REFERENCES

[1]. S. S. I. Arafat,M. S. Hossain,M. M. Hasan, S. A. H. Imam, M. M.Islam, S. Saha,... and T. I. Juthi, “VIM: A Big Data Analytics Tool for Data Visualization and Knowledge Mining”, *IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, pp. 224-227, 2017.
[2]. R. Akase and Y. Okada, “Web based time-tunnel: An interactive multidimensional data visualization tool using genetic algorithm”, 1*9th IEEE International Conference on Information Visualisation*, pp. 95-100, 2015. IEEE.
[3]. R. K. Barik, R. K. Lenka, S. M. Ali,N.Gupta, A. Satpathy, and A. Raj, “Investigation into the efficacy of geospatial big data visualization tools”, *International Conference on Computing, Communication and Automation (ICCCA)*, pp. 88-93, 2017.
[4]. M. J. Samonte, H. J. Bacer., R. PRamirez, and J. J. Tuballa, “An Interactive Data Visualization Tool for Observing History of Fire Incident”, *International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 402-406, 2019.
[5]. A. Lousa, I. Pedrosa, and J. Bernardino, “Evaluation and analysis of business intelligence data visualization tools”, *14th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1-6, 2019.
[6]. L. E. Matzen, M.J. Haass, K.M. Divis, Z. Wang, and A. T. Wilson, “Data visualization saliency model: A tool for evaluating abstract data visualization”, *IEEE transactions on visualization and computer graphics*, vol. 24, pp. 563-573.
[7]. R. Li. J.Sicat, J.Choi, M.Cordeil, W. K. Jeong, B. Bach, and H. Pfister, “DXR: A toolkit for building immersive data visualizations”, *IEEE transactions on visualization and computer graphics*, vol. 25, pp. 715-725, 2018.
[8]. T. M. Pearce, M. N. Nikiforova, and S. Roy, “Interactive browser-based genomics data visualization tools for translational and clinical laboratory applications”, *The Journal of Molecular Diagnostics*, vol. 21, pp. 985-993, 2019.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ჯ Impact Factor 8.066 ჯ Vol. 10, Issue 4, April 2023

DOI: 10.17148/IARJSET.2023.10416

[9]. M. A. Kuhail, and S. Lauesen, "Uvis: A formula-based end-user tool for data visualization". *IEEE Access*, vol. 8, pp. 110264-110278.

[10]. C. Balzer, R. Oktavian, M. Zandi, D. Fairen-Jimenez, and P. Z. Moghadam, "Wiz: A web-based tool for interactive visualization of big data", *Patterns*, vol. 1, pp. 100107, 2020.

[11]. E. Perez-Molina, &F. Loizides, "Novel data structure and visualization tool for studying technology evolution based on patent information: The DTFootprint and the TechSpectrogram", *World Patent Information*, vol. 64, 102009.

[12]. D. R. Rodrigues, D. C. B. Mariano, L.H.S. Santos, and C. A. Tagliati, "ToxAnalyzer: A user-friendly web tool for interactive data analysis and visualization of chemical compounds from the Comparative Toxicogenomics Database (CTD)™", *Computational Toxicology*, vol.19, p. 100170, 2021.

[13]. D. Liu, Z. Xu, C. Fan, and Y. Zhou, "Development of fire risk visualization tool based on heat map", *Journal of Loss Prevention in the Process Industries*, vol. 71, p. 104505, 2021.

[14]. R. Jing, LLi, M. Xue, L.Yu. and J. Luo, "layerUMAP: A tool for visualizing and understanding deep learning models in biological sequence classification using UMAP", *Iscience*, vol. 25, p. 105530, 2022.

[15]. S. Ma. X. Fang, Y. Li. J. Yao, D.C. Morgan, Y. Xia, and J. W. Ho, "StarmapVis: An interactive and narrative visualization tool for single-cell and spatial data", *Computational and Structural Biotechnology Journal*, vol. 21, 1598-1605, 2023.

 This work is licensed under a Creative Commons Attribution 4.0 International License