International Journal of Information System Modeling and Design
Volume 13 • Issue 7

Improving Big Data Analytics With Interactive Augmented Reality

Sumit Arun Hirve, VIT-AP University, India*
Pradeep Reddy C. H., VIT-AP University, India

ABSTRACT

Since, data is generated every minute by everyone including consumers and/or business worldwide, there is an enormous worth for big data analytics. Big data analytics is a technique for extracting important information from large amounts of a data. Visualization is the best medium to analyze and share information. Visual images help to transmit bid data to the human brain within a few seconds. Visual interpretations help in visualizing data from different angles. Visualization helps to outline problems and understand current trends. Augmented reality enables the user to experience the real world, which is digitally augmented in a way. The main objective of this research work is to find the solution to visualize the analyzed data and show it to the users in a 3D view and to improve the angle of visualization with the help of augmented reality techniques.

KEYWORDS

Android, Information System, Machine Learning, MapReduce, System Modeling, Tableau, Unity, Vuforia

INTRODUCTION

The primary focus is on big data visualizations displayed through augmented reality and the importance of big data analytics using the Tableau business intelligence tool. The visualizations obtained with the same tool are showcased in section 4. Data generation has gained momentum since the outburst of social media platforms. The rate of data generation has become more significant than data processing. This shift caused technology such as big data Analysis to immerge in a short period. The approach helped organizations categorize and organize data so that the results concluded improved market analysis and economic growth. But some contradicting issues were unable to solve just with the perspective of distinguishing the data. That is when the visualization came into the picture. The visualizing techniques introduced a perspective switch that made information much more comprehensive. Somehow, the visualization could not tap into the hyper-dimensional view of the data. To get more knowledge from the datasets, many big data tools and technologies are available in the market for performing computations on massive data. The tools help the data analysts or scientists clean, process and transform the data. The big guys in the market, such as Hadoop, NoSQL databases,

DOI: 10.4018/IJISMD.315124 *Corresponding Author

Map Reduce, YARN, and Spark, are available to perform end to end big data analytics. Apart from the tools and technologies, Tableau is a business intelligence tool popularly known for computing data analytics, extensive data handling and interconnectivity with the databases for data extraction. Tableau is the best for sharing data insights, self-service visual analysis and embedding data insights into web applications. In this paper, the use of Tableau for performing big data analytics on the sample bigmart dataset is done and input of data analytics is fed to the AR application designed in the Unity IDE engine.

Human mind procedures and recalls data better when seen in the graphical arrangement when contrasted with printed groups. Imagined data helps speed with increasing the investigation. What's more, while investigating a lot of information, there is a massive amount of data produced at big organizations and requires accurate analysis visualization (Wang, L et al, 2015; M. A. Defeyter et al, 2009). some innovative approaches are needed to visualize big data due to its complexity (E. Olshannikova, 2015). Along with improvisations and additions in the various business-related curriculums, the employment world looks into every candidate who possesses the right approach to data visualization.

LITERATURE REVIEW

Numerous analysts have smoothed out the definitions for big data to portray a part of big data. Various studies have some other way hinted that data visualization techniques play a vital role in big data analysis. The age of enormous measures of unstructured information expands the need to gather, store, sort out, and process it for removing significant data (I. D. Ltd, 2018). Big data visualization is firmly related as visualizations help discover designs, extricate suspicions, and domain explicit expectations. Visualizing enormous datasets is more diligently when contrasted with conventional datasets. Mathematical demonstration and highlight extractions are likewise utilized to expand customary visualization models while visualizing big data (Wang, 2015). Big Data may be of a wide scope of structures like video, music, text picture, and various associations. The data can be created through cloud applications and informal organizations (Jamiy EL F., 2015). As stated by E. Olshannikova et al., visualizing extracted data is time-consuming, and Sometimes, the evaluated results are not accurate and as estimated. After calculating the results, the limitation was clear: the existing system cannot visualize the information how humans process a large amount of information from a different perspective. When it comes to applying the data visualization for the smart city, there is no platform right now to process data into the VR visualization. So, this paper introduces a custom VR application. The study also compares the recognized work volume with categorized open web platform data. The intense reactions are observed when VR is used on web-based platforms, whereas the proposed VR system notes a more positive response (Broucke, S. V., & Deligiannis, N, 2019). The representations of 2D and 3D spatial data, along with their advantages and disadvantages, are explained in this research paper. These traits concern the technical and perceptual aspects of the system. The comparison proposed in this research helps to formalize the fundamental characteristics that helped comprehend the merits and demerits between existential systems and the proposed one. Overall, after evaluating the results, it is obvious that Systematization is imperative for decision-making and the first phase for effective visualization. It is also inferred that Sophisticated Visualization can be achieved by improving binary categorization. Also, distinguishing the traits with maximum accuracy can improve 2D and 3D visualization (Dubel, S. et al., 2015). There's a familiar framework related to flexible displays and their use in their different interaction methods. Also, this paper scrutinizes the data that can be used in ways to interact (Franke I. S. et al., 2014). This paper explains the tools used for 3d visualizing Shenzhen (Li, X et al., 2015). This study introduces the application of virtual reality in big data visualization. Big data visualization enables us to discover hidden in a comprehendible structure. Along with the virtual reality hyper-dimensional perspective, let's uncover the complex perspective in a more understandable way to make predictions (Donalek C. et al., 2014). Using

immersing technologies with a collaborative approach recognizes a cost-effective advantage (Donalek C. et al., 2014). This paper describes data visualization and its applications (Makela A., 2018). There is a VR system given for high-dimensional data visualization. The procedure relies on the Principal Component Analysis (Sanji S. et al., 2001). An effort is introduced to improve bid data visualization that encompasses the 2-fold analysis and emphasizes visualization and interaction. It uses a piece of geo-tagged information by investigating data sets of different categorizations and using it in the VR platform. First of all, the data will be analyzed and represented in a realistic 3d module so that the analyst can scrutinize it from various perspectives. This can lead to an in-depth analysis of the given datasets. Overall, this study aims to promote the collaboration of today's technology that is gaining popularity for visualizing with interactive abilities to gain comprehensive results of the big data (Moran A. et al., 2015). After analyzing numerous studies on VR immerging as an effective approach for big data visualization, there's an evaluation of users' boundary-less multi-dimensional datasets with traditional 2d visualization represented in this study. Overall, it is recognized that there is no significant extra burden of tasks between traditional and VR visualization rather than correctness and depth of perception (P. Millais et al. 2018). This literature highlights the need, challenges, and ongoing Augmentation-based visualization work (Chandra et al., 2019). As the Big Data visualization performance can be either static or dynamic. The best results are achieved with an interactive approach. This literature gives insights into how the interaction is provided by developing visualization on cross platforms. It also suggests that adding innovative functionality in the visualization, such as interactive brushing freestyle linking on web-based tools, can help scientists process more efficiently. The study also emphasizes Web-based solutions to get more up to date dynamic data for visualization (L. Wang et al., 2015). This handbook discusses the birth of Big data with its complex data structure. It also discusses how the rate of data generated is higher than processed, leading scientists to develop more efficient paradigms. This literature also touches on the security aspect usually ignored in most big data visualization techniques. The research provides the work on various platforms (Buyya R. et al., 2016). This research provides close insights into the need for big data analysis and potential challenges. It's also important that various organizations be aware of the potential of the data that has been generated to anticipate opportunities for competitors and innovation (Jamiy EL et al., 2015). The system evaluates two different types of position assistance Stereoscopic depth perception and 3D scaling cube (F. El Jamiy and R. Marsh, 2019).

DATA ANALYSIS

With the development of computational strategies, the measure of big data has risen dramatically. A quick rate makes it difficult to use such information in the bigmart field without proper pre-handling, which thus prompts more incredible intricacy and integrity issues, making numerous complexities like data analysis, privacy, storage and security. Hence, the bigmart dataset may look simple to deal with as far as its volume. However, it requires a significant convoluted interaction because of its highlights' intricacy, heterogeneity, and hybridity. This interaction is entitled the information revelation measure. The data description helps analysts get an overview of the type of data. It also represents the different fields that contribute to the evaluation process. The data of BIGMART choose for this operation. At first, BIGMART data does not seem huge enough, but it has its challenges, such as customers' personal and transaction details, which raises security issues. The dataset contains 17 distinct data fields such as Item_Weight, Item_Visibility, Item_MRP, etc.

The data processing is carried out through standardized processing stages. Initially, intended data is gathered and prepared from various verified sources. The collected information is then categorized for the extraction process to begin. After that, the filtered data is transformed into a more meaningful form that will ease the data aggregation process. The analyst scrutinizes the aggregated data by using statistical tools. Once the data analysis is complete, the visualizing IDE such as Tableau, PowerBI and QlikView is used to put the analysis into perspective. These Visualizations carry insights into

hyper-dimensions as well. Data Processing - Extracting or cleaning unnecessary information is called data processing. The data is processed, and the digital information is stored for the computer to understand the language.

Big data analytics is an unpredictable cycle of getting information about market trends and critical business decisions. Analytics on data provide a new way for advancement and pave a way to remove new datasets. Queries to the business knowledge have essential answers regarding tasks and execution. Big data investigation is a front line assessment that incorporates complex applications with parts such as judicious models, authentic figuring, and structures. Tremendous information examination applications routinely join data from inside systems and external sources, like environment data or fragment data on clients aggregated by pariah data organizations providers. Moreover, streaming examination applications are getting ordinary in gigantic data conditions. Extensive information has reached progressively valuable in the production of network investigation. Massive store network investigation uses important information and quantitative strategies to upgrade dynamic cycles across the inventory network. Specifically, immense store network examination broadens datasets for the extended assessment that goes past the customary inside data found on enormous business resource masterminding (ERP) and creation network the board (SCM) systems. Also, huge stock organization examination executes significantly incredible verifiable methods on new and existing data sources. The pieces of information amassed empower better, taught and additional convincing decisions that benefit and improve the reserve chain. Likely snares of tremendous data assessment exercises join a shortfall of inward examination capacities and the huge cost of utilizing experienced data analysts and data experts to fill the openings. Tableau Tool Analytics: The Bigmart data has been given input to the well-known big data analytics tool, such as Tableau. The analytical computation results in diagrammatic representations, which display the statistics of the structured dataset. Below given are the data analytics techniques performed with the public tableau. tool.

Figure 3 shows the total sales for each item type listed in the dataset, such as fruits and vegetables, snacks, households, dairy, frozen foods, bread, hard drinks, breakfast and seafood etc. This graph's analytical conclusion is the maximum sales of a particular item type in the retail store. This finding or knowledge would help the stocking manager or store manager decide, resulting in higher sales with high demand. Figure 4 shows the outlet size with the establishment years and inclining with the item types. The data analytics obtained in this module are forwarded to the unity engine as inputs, and the visualization is produced in 3D format on smart devices. The following section gives the details on the AR application developed for augmenting the statistical informationMarker Based Augmented Reality

AR is an advanced technology that expands actual conditions on a cell phone screen by overlaying them with computerized content. AR applications can utilize various techniques like associating PC created substances to 'markers' or deciding where to add information with GPS to superimpose

Figure 1. Bigmart Dataset Description

Item_Weight	Item_Visibility	Item_Type	Item_MRP	Outlet_Identifier	Outlet_Establishme	Outlet_Size	Outlet_Loc	Outlet_Type	Item_Outle	Protiens	Carbohydrates	Fats_Conte	Health_Sta	Selected	Item_Iden	Selected (1
continuous	continuous	discrete	continuous	discrete	continuous	discrete	discrete	discrete	continuous	discrete	discrete	discrete	discrete	discrete	string	discrete
														class	meta	meta
9.3	0.016047301	Dairy	249.8092	OUT049	1999	Medium	Tier 1	Supermarket Type1	3735.138	Medium	Medium	Very High	Both	No	FDA15	No
5.92	0.019278216	Soft Drinks	48.2692	OUT018	2009	Medium	Tier 3	Supermarket Type2	443.4228	Less	Less	High	Unhealthy	No	DRC01	No
17.5	0.016760075	Meat	141.618	OUT049	1999	Medium	Tier 1	Supermarket Type1	2097.27	Normal	Normal	Normal	Healthy	No	FDN15	No
19.2	0	Fruits and Vegetables	182.095	OUT010	1998	Small	Tier 3	Grocery Store	732.38	Normal	Normal	Normal	Healthy	No	FDX07	No
8.93	0	Household	53.8614	OUT013	1987	High	Tier 3	Supermarket Type1	994.7052	Normal	Normal	Normal	Healthy	No	NCD19	No
10.395	0	Baking Goods	51.4008	OUT018	2009	Medium	Tier 3	Supermarket Type2	556.6088	Medium	Medium	Very High	Both	No	FDP36	No
13.65	0.012741089	Snack Foods	57.6588	OUT013	1987	High	Tier 3	Supermarket Type1	343.5528	Medium	Medium	Very High	Both	No	FDO10	No
18.5	0.127469857	Snack Foods	107.7622	OUT027	1985	Medium	Tier 3	Supermarket Type3	4022.764	Medium	Medium	Very High	Both	Yes	FDP10	No
16.2	0.016687114	Frozen Foods	96.9726	OUT045	2002	Medium	Tier 2	Supermarket Type1	1076.599	Less	Less	High	Unhealthy	No	FDH17	No
19.2	0.09444959	Frozen Foods	187.8214	OUT017	2007	Medium	Tier 2	Supermarket Type1	4710.535	Less	Less	High	Unhealthy	No	FDU28	No
11.8	0	Fruits and Vegetables	45.5402	OUT049	1999	Medium	Tier 1	Supermarket Type1	1516.027	Normal	Normal	Healthy	No	No	FDY07	No
18.5	0.045463773	Dairy	144.1102	OUT046	1997	Small	Tier 1	Supermarket Type1	2187.153	Medium	Medium	Very High	Both	No	FDA03	No
15.1	0.1000135	Fruits and Vegetables	145.4786	OUT049	1999	Medium	Tier 1	Supermarket Type1	1589.265	Normal	Normal	Healthy	No	No	FDX32	No
17.6	0.047257328	Snack Foods	119.6782	OUT046	1997	Small	Tier 1	Supermarket Type1	2145.208	Medium	Medium	Very High	Both	No	FDS46	No
16.35	0.0680243	Fruits and Vegetables	196.4426	OUT013	1987	High	Tier 3	Supermarket Type1	1977.426	Normal	Normal	Healthy	No	No	FDF32	No
9	0.069088961	Breakfast	56.3614	OUT046	1997	Small	Tier 1	Supermarket Type1	1547.319	Less	Less	High	Unhealthy	No	FDP49	No
11.8	0.008596051	Health and Hygiene	115.3492	OUT018	2009	Medium	Tier 3	Supermarket Type2	1621.889	Normal	Normal	Healthy	No	No	NCB42	No
9	0.069196376	Breakfast	54.3614	OUT049	1999	Medium	Tier 1	Supermarket Type1	718.3982	Less	Less	High	Unhealthy	No	FDP49	No
18.2	0.034237682	Hard Drinks	113.2834	OUT027	1985	Medium	Tier 3	Supermarket Type3	2303.668	Less	Less	High	Unhealthy	No	DRI11	No
13.35	0.10249212	Dairy	230.5352	OUT035	2004	Small	Tier 2	Supermarket Type1	2748.422	Medium	Medium	Very High	Both	No	FDU02	No
18.85	0.138190277	Snack Foods	250.8724	OUT013	1987	High	Tier 3	Supermarket Type1	3775.086	Medium	Medium	Very High	Both	No	FDN22	No
8.02	0.035399923	Baking Goods	144.5444	OUT027	1985	Medium	Tier 3	Supermarket Type3	4064.043	Medium	Medium	Very High	Both	No	FDW12	No
14.6	0.025698134	Household	196.5084	OUT035	2004	Small	Tier 2	Supermarket Type1	1587.267	Normal	Normal	Healthy	No	No	NCB30	No
17.35	0.057556998	Baking Goods	107.6938	OUT019	1985	Small	Tier 1	Grocery Store	214.3876	Medium	Medium	Very High	Both	No	FDC37	No
13.85	0.025896485	Frozen Foods	165.021	OUT046	1997	Small	Tier 1	Supermarket Type1	4078.025	Less	Less	High	Unhealthy	No	FDR28	No

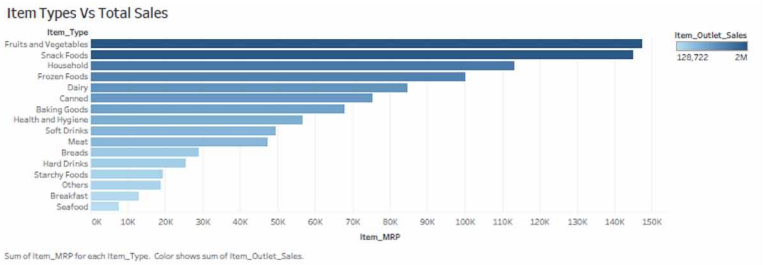
Figure 2. Data Aggregation of all the dataset attributes

Table Categories Summary

Item_Type	Item_MRP	Item_Outlet_Sales	Item_Weight	Number of Records
Fruits and Vegetables	147,189	2,219,978	13,476	1,019
Snack Foods	144,949	2,194,045	12,832	988
Household	113,210	1,659,037	10,159	759
Frozen Foods	99,962	1,494,210	9,239	718
Dairy	84,526	1,244,133	7,599	566
Canned	75,053	1,155,262	6,633	539
Baking Goods	67,588	1,027,220	6,581	536
Health and Hygiene	56,384	835,171	5,651	430
Soft Drinks	49,295	735,962	4,431	374
Meat	47,160	689,281	4,319	337
Breads	28,663	437,771	2,315	204
Hard Drinks	25,262	377,067	2,086	183
Starchy Foods	19,200	301,063	1,780	130
Others	18,625	280,482	1,898	137
Breakfast	12,696	178,104	1,136	89
Seafood	7,398	128,722	640	51

Item_MRP, Item_Outlet_Sales, Item_Weight and Number of Records broken down by Item_Type.

Figure 3. Bigmart Data: Item Type vs. Total sales



computerized data. A QR code, also known as a 2D code, can be augmented using SDK and tools for AR applications. The augmented information can be displayed on the device's screen via a camera when it detects the augmented targets on the QR code image.

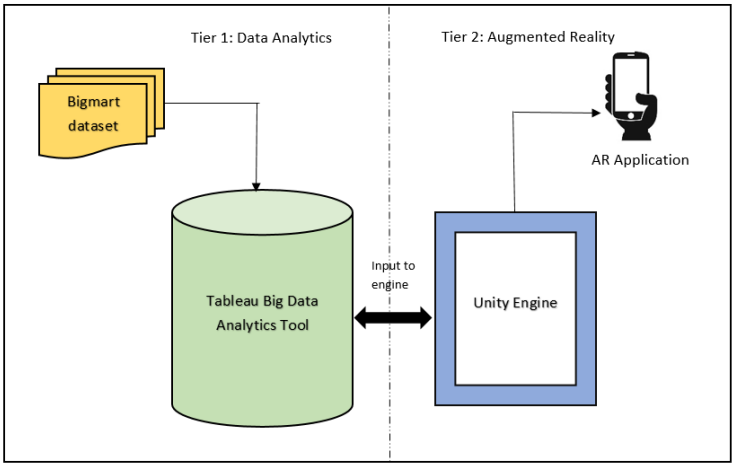
PROPOSED SYSTEM

The system implementation shown above is divided into two tiers: data analysis and AR application development. We will provide an in-depth explanation of both the tiers in this section as follows:

Table 1. Comparison of markerbased AR and marker-less AR

Comparison Aspects		Markerbased AR	Markerless AR
Methods	**Relative Position AR Software Development Kit**	**Marker dependent Commonly used**	**Depends on localization technology and gyroscope**
Position Accuracy	High/Low Influence factors	Higher Brightness	Lower Localization Technology
Stability	High/low Influence factors	Higher Markers and SDK's	Higher Localization and Gyroscope

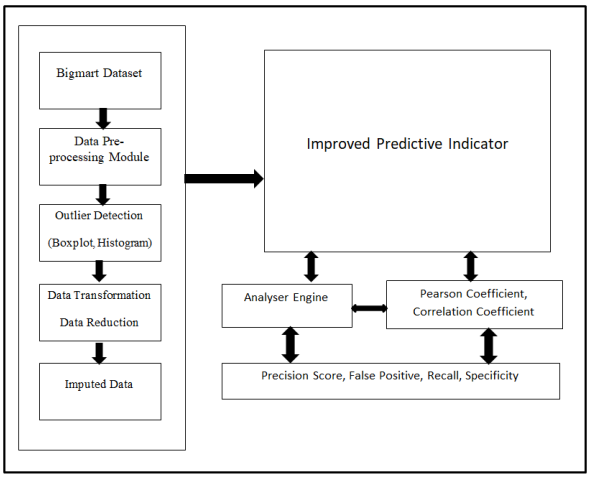
Figure 4. System Architecture



- Data Analytics Tier 1 - The Bigmart data is processed and given input to the tableau big data analytics tool. The business intelligence tool processes the data with several analytical computations by finding a statistical pattern to share with the consumers. After obtaining the visualizations, the diagrams are saved and embedded in the AR mobile application developed in the Unity engine and aided by the Vuforia plugin.
- AR Application Development - The augmented reality application developed in the Unity engine is given the input from tier 1. The graphs are embedded in the application for displaying on the users' screens. The virtual reality plugin required with unity is Vuforia and can be installed as a plugin in the unity engine application.

The proposed model consists of a predictive enhancer algorithm that takes the pre-processed bigmart dataset as the input and outputs the predictions in the form of health, unhealthy, and both

Figure 5. Proposed Model for prediction of health status for bigmart dataset



target variable categories. The bigmart dataset consisted of missing values, and therefore, the noise was cleaned using outlier detection practices such as boxplot and histogram visualization. The dataset was transformed using data transformation techniques such as data generalization, aggregation and data smoothing at the final step. The class labels are assigned manually after computing the values of all 13 attributes of the dataset to make it favourable for supervised learning. Our machine learning classifier forwarded the final cleaned and imputed dataset for predictive analysis. Algorithm to construct improved predictive indicator for bigmart dataset

Input: Bigmart dataset.
Output: Visualization of Predictions.
Step 1: Input the bigmart dataset into data pre-processing module.
Step 2: Analyze the dataset for noise and clean the noise using outlier detection techniques. Replace missing values with the median value.
Step 3: Forward the dataset to the data transformation module to perform data generalization, aggregation, and smoothing operations. Calculate the class label values to make them useful for supervised learning.
Step 4: Send the final imputed dataset to the improved predictive indicator module to obtain coefficient values, model accuracy, and predictions for the dataset.
Step 5: Evaluate the predictions and accuracy values.
Step 6: Perform parameter tuning and repeat steps 4 and 5 if the accuracy is not up to the mark.

RESULTS

A. Results of Pearson correlation coefficient

Bigmart dataset collected from the public grocery store has 8526 records and 14 attributes inclusive of the class label. The target variable determines the health status of the food products, such as healthy, unhealthy and both. The significant attributes are shown in table 2, along with their correlation coefficient values against the target output, that is, health status. Other high influencing factors are proteins, carbohydrates and item types.

Dataset 2 from bigmart has 4320 records along with 13 attributes. Dataset was originated from three different bigmart outlet types, and six attributes were selected after applying Pearson correlation for target output, i.e. health status. These selected attributes are shown in table 3, along with correlation coefficient values. The attribute with the highest correlation with the health status target attribute is outlet type.

On comparison of these results with the selected algorithms in the below table which shows the comparison between Random Forest, decision trees (J48, C5.0) and SVM. It is clearly observed that

Table 2. .Selected attributes and their Pearson correlation coefficient value for Dataset

Attribute name	Pearson correlation coefficient value
Item Weight	0.66
Item Visibility	0.62
Item Type	0.54
Item_MRP	0.11
Outlet Identifier	0.40
Outlet_Establishment_Year	0.13
Outlet Size	0.33
Outlet_Location_Type	0.29

Table 3. Attributes and their correlation values for Dataset

Attribute name	Correlation value
Outlet Type	0.5272
Item_Outlet_Sales	0.3080
Proteins	0.4353
Carbohydrates	0.3758
Fats Content	0.3890

decision tree C5.0 outclasses other algorithms by achieving an accuracy of 98% on the validation set consisting of 131 instances.

After obtaining the predictions from the unknown data that the user enters after training the algorithms, the visualization of the predictions is embedded in the augmented reality system developed in the Unity engine.

As shown in Table 5 above, the researchers [20], in their experimentation, proposed a DTI (Decision Tree Induction) Model in which they obtained an accuracy of 77.5% with the J48 algorithm and 79% with multi-class SVM on multiple medical datasets such as heart disease, breast cancer and diabetes. On the other hand, the authors [21], in their latest research of 2021, illustrated a comparative study of supervised learning algorithms for evaluating student performance on students datasets. The survey involved three different datasets and two outclassing algorithms such as J48 with 85% accuracy and multi-class SVM with 77%.

The results of this experimentation are displayed in this section in detail. Figure 6 describes the QR code required for augmenting the virtual information on top of the primary physical layer. Vuforia is an augmented reality SDK that primarily focuses on mobile applications. The SDK combines computer vision and machine learning which allows advanced artificial intelligence. Vuforia supports 2D and 3D targets, including markerless image targets. The image provided in the Vuforia database is augmented by the unity engine itself through the target manager online.

The QR code augmented in the vuforia engine is aligned with the 3D image imprinted on the smartphone device, as shown in figure 7 above. The augmented information is a surface view of the products with low, medium and high fats, proteins, and carbohydrates. The surface view better visualises healthy and unhealthy food products.

Table 4. Performance comparison between Random Forest, PART, J48, Bayes Net and C5.0 for Dataset 2

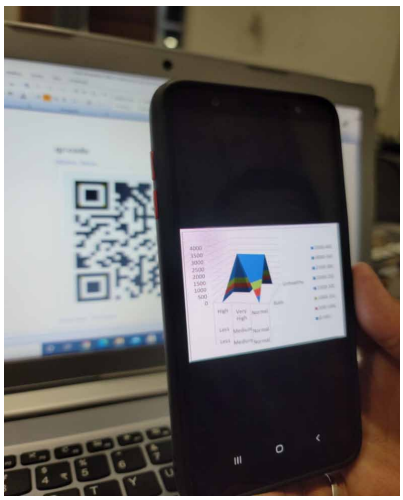
Algorithm	Number of instances	Correctly classified instances	Accuracy (in %)
J48	300	219	73
Multi-class SVM	300	196	65.33

Table 5. Comparative Study of Research works

Machine learning algorithms	Decision tree Induction Model- Accuracy	Performance prediction Model- Accuracy	Improved Predictive Indicator - Accuracy
J48	77.5%	85%	90.10%
Multi-class SVM	79%	77%	83.25%

Figure 6. Augmentable Image with Targets in Vuforia

Figure 7. 3D Augmented marker-based image displayed on smart phone device



CONCLUSION

A systematic approach to data analytics is implemented with a public business intelligence tool. This research module focuses on developing an Augmented Reality application responsible for displaying the virtual representation of statistics obtained from data analytics. The already existing IDE and correspondingly available plugins such as Unity engine and Vuforia AR are used for comparing the accuracy. The applications of this research work could range from silent meeting rooms in schools, offices and colleges to e-learning in laboratories, and many others added in the future. The scope and methodology of this research are expanded by considering Virtual Reality and mixed reality in combination with Augmented Reality to remove the inclusion of virtual reality devices such as Oculus VR and Sony VR.

ACKNOWLEDGMENT

The authors of this publication declare there is no conflict of interest. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Broucke, S. V., & Deligiannis, N. (2019, October). Visualization of real-time heterogeneous smart city data using virtual reality. In *2019 IEEE International Smart Cities Conference (ISC2)* (pp. 685-690). IEEE.

Buyya, R., Calheiros, R. N., & Dastjerdi, A. V. (2016). *Big data: Principles and paradigms*. Morgan Kaufmann.

Chandra, Ramaseri, Jamiy, & Reza. (2019). Augmented reality for big data visualization: A review. *International Conference on Computational Science and Computational Intelligence (CSCI)*.

Defeyter, M. A., Russo, R., & McPartlin, P. L. (2009). The picture superiority effect in recognition memory: A developmental study using the response signal procedure. *Cognitive Development*, *24*(3), 265–273. doi:10.1016/j.cogdev.2009.05.002

Donalek, C., Djorgovski, S. G., Cioc, A., Wang, A., Zhang, J., Lawler, E., Yeh, S., Mahabal, A., Graham, M., Drake, A., Davidoff, S., Norris, J. S., & Long, G. (2014). Immersive and collaborative data visualization using virtual reality platforms. *IEEE International Conference on Big Data (Big Data)*, 609–614.

Dubel,S., Rohlig, M., Schumann, H., & Trapp, M. (2014, November). 2D and 3D presentation of spatial data: A systematic review. In *2014 IEEE VIS international workshop on 3DVis (3DVis)* (pp. 11-18). IEEE.

El Jamiy, F., Daif, A., Azouazi, M., & Marzak, A. (2015). An optimized file system for high performance computing (hpc). *International Journal of Application or Innovation in Engineering & Management*, *3*(4), 3–5.

El Jamiy, F., & Marsh, R. (2019). Survey on depth perception in head mounted displays: Distance estimation in virtual reality, augmented reality, and mixed reality. *IET Image Processing*, *13*(5), 707–712.

Franke, I. S., M¨uller, M., Grunder, T., & Groh, R. (2014). *FlexiWall: Interaction in-between 2D and 3D Interfaces in HCI International 2014 - Posters' Extended Abstracts*. Springer International Publishing.

Li, X., Lv, Z., Zhang, B., Wang, W., Feng, S., & Hu, J. (2015). *Webvrgis based city bigdata 3d visualization and analysis*. arXiv preprint arXiv:1504.01051.

Makela, A. (2018). *Data visualization in virtual reality* [Undergraduate Thesis]. JAMK University of Applied Sciences.

Millais, P., Jones, S., & Kelly, R. (2018). Exploring data in virtual reality: Comparisons with 2D data visualizations. In *CHI 2018 – Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.

Moran, A., Gadepally, V., Hubbell, M., & Kepner, J. (2015). Improving big data visual analytics with interactive virtual reality. *IEEE High Performance Extreme Computing Conference (HPEC)*.

Olshannikova, E., Ometov, A., Koucheryavy, Y., & Olsson, T. (2015). Visualizing big data with augmented and virtual reality: Challenges and research agend. *Journal of Big Data*, *2*(1), 1–2. doi:10.1186/s40537-015-0031-2

Sanji, S., Murata, M., Hartono, P., & Hashimoto, S. (2001). A virtual reality system for high-dimentional data visualization. *IEEE International Conference on Multimedia and Expo,* 1049–1052.

Wang, L., Wang, G., & Alexander, C. A. (2015). Big data and visualization: Methods, challenges and technology progress. *Digital Technologies*, *1*(1), 33–38.

Wang, L., Wang, G., & Alexander, C. A. (2015). Big data and visualization: Methods, challenges and technology progress. *Digital Technologies.*, *1*(1), 33–38.

Sumit Arun Hirve received the Bachelor of Computer Engineering (BE) From Pune University and Master of Technology in Computer Engineering (M-Tech) degree from Bharati Vidyapeeth Deemed University Pune. He is currently working as Assistant Professor in the College of Engineering Pune, Pune University. He has a total of 14 years of experience in both teaching and research His research interest includes big data Analytics, Cyber Security, Cloud Computing. He is doing research in his desired fields by combining his work experiences and research interests.

Pradeep Reddy C. H. is currently working as Professor in VIT-AP University, Amaravati, India. He has a total of 13 years of experience in both teaching and research. He received his B.Tech in Computer Science and Engineering from PBR VITS, JNTU Andhra Pradesh, India and M.Tech in Computer Science and Engineering from VIT University, Vellore, India. He did his PhD in Computer Science and Engineering from VIT University, Vellore. He has published various journals and served as reviewer and editor for reputed journals. His research interests include Wireless Networks, IoT, Security systems and Cloud computing.