

GLOBAL SALES DATA ANALYTICS AND VISUALIZATION

Sivaguru.R¹, Abdulkalamazad.G², Babu.G³, SankarRam.R⁴, Ajay.P⁵, Madhavan.M.K⁶, Dhatchanamoorthy.U⁷

^{1,2,3}Assistant Professor/CSE & Knowledge Institute of Technology, Salem
^{4,5,6,7}CSE & Knowledge Institute of Technology, Salem

Abstract - The rise in information in daily life has resulted in a highly rapid generation of data. A vast volume of data that is challenging to analyse and utilise accumulates from numerous organisations. These data are challenging to process, analyse, and communicate. Every day, massive amounts of sales information are poured onto online purchasing platforms. For information retrieval, it is challenging to analyse and visualise this data. A system that can efficiently analyse and visualise data is therefore needed. The system that will let users apply intelligence to business, generate money, make decisions, manage business operations, and monitor task progress is the main topic of this paper's discussion. We will analyse the global sales data using IBM Cognos Analytics to find patterns, connections, and correlations. The product recommendation system can help us increase revenue quickly. To better comprehend and meet consumer wants, the customer segmentation feature is implemented. Both the pricing will be optimised and your connections with them will be better as a result. Finding solutions that can be suitably disguised for a specific problem will help to arrive at the best solution. Exploring relationships in the data and visualising the data will provide a speculation and guesstimate of occurrences.

Key Words: Data analytics, Data Visualizations, Business Intelligence, Decision making, Sales data, Business analytics.

1. INTRODUCTION

The science of analysing unstructured databases to draw inferences about the information they contain is known as data analytics. We can use it to look for patterns in the raw data and extrapolate important knowledge from them. Data analytics aid businesses in increasing market effectiveness and revenue.

Several different sorts of data analysis fall under the umbrella term of "data analytics," which is very wide. Any sort of data can be subjected to data analytics techniques to gain insight on how to make things better. Trends and metrics that would otherwise be lost in the sea of data can be revealed using data analytics approaches. Then, using this knowledge, processes can be improved to raise a business's or system's general level of effectiveness. For instance, manufacturing businesses frequently track the runtime, downtime, and work queue for various machines and then analyse the data to better plan the workloads so the machines perform closer to peak capacity.

More than just identifying production bottlenecks is possible with data analytics. Gaming businesses utilise data analytics to create player incentive schedules that keep lots of players engaged in the game. In order to keep you clicking, watching, or reorganizing content to gain another view, content providers utilise many of the same data analytics.

Data analytics is significant because it aids in performance optimisation for firms. By finding more cost-effective ways to do business and retaining a lot of data, firms can help cut expenses by incorporating it into their business strategy. Additionally, a corporation can use data analytics to improve business decisions and track consumer preferences and trends to develop fresh, improved goods and services.

1.1 PROCESS INVOLVED IN DATA ANALYTICS

Data analysis is a process that incorporates a number of processes, including:

1. The initial step is to decide on the data needs or the data grouping. Age, demography, income, and gender can all be used to separate data. Data values can either be numerical or categorical.
2. Gathering data is the second stage of data analytics. This can be done using a range of tools, including computers, the internet, cameras, environmental sources, and human employees.
3. Data must be organised so that it may be analysed after it has been gathered. This could be done using a spreadsheet or another type of programme that can handle statistical data.
4. Before analysis, the data is then cleaned up. It has been cleaned up and double-checked to make sure there are no errors or duplications and that it isn't missing anything. Before it is given to a data analyst to be analysed, this phase helps to correct any inaccuracies.

1.2 DATA ANALYTICS TECHNIQUES

To process data and extract information, data analysts might employ a variety of analytical approaches and procedures. The following is a list of some of the most used techniques.

1. To identify how a change in one dependent variable may affect the change in another, regression analysis involves examining the relationship between the dependent variables.
2. To do a factor analysis, a huge data set must be reduced to a smaller data set. By using this technique, it is hoped to uncover tendencies that might have been harder to spot otherwise.
3. The practise of segmenting a data collection into groups of related data, frequently segmented by a customer demography, is known as cohort analysis. This enables data analysts and other data analytics users to go deeper into the statistics pertaining to a certain subset of data.
4. Monte Carlo simulations simulate the likelihood that various events will occur. These simulations, which frequently include many values and variables and

frequently have better predicting abilities than other data analytics techniques, are frequently utilised for risk mitigation and loss prevention.

5. Time series analysis tracks data through time and establishes a connection between a data point's value and its occurrence. This method of data analysis is frequently employed to identify cyclical patterns or to forecast financial outcomes.

1.3 IMPORTANCE OF DATA ANALYTICS

Data analytics is significant because it aids in performance optimisation for firms. Companies can assist cut costs by locating more effective ways to do business by incorporating it into their business strategy. Additionally, a corporation can use data analytics to improve business decisions and track consumer preferences and trends to develop fresh, improved goods and services.

Many industries, including the airline and hospitality sector, whose turnaround times are often rapid, have embraced data analytics. This sector can gather client information and identify any problems and their causes. Another industry that uses both structured and unstructured data in large quantities is healthcare, where data analytics can aid in speedy decision-making. Similar to this, the retail sector makes extensive use of data to satisfy customers' shifting needs.

Data analytics assists people and organisations in ensuring the accuracy of their data in a world that is relying more and more on information and statistics collection. A set of raw numbers can be turned into instructive, educative insights that guide decision-making and considerate management using a range of tools and methodologies.

5. LITERATURE REVIEW

In today's technologically advanced age, every company wants to equip its sales force with a sustainable sales force automation system to improve sales performance and customer relationship capabilities. This study examines the impact of big data analytics on the sales performance of the organizations. Current advances in information technology and the development trend of social networks have had changed the way salespeople perform their daily activities. Data related to customer buying behavior is being generated at an unprecedented rate due to the technological revolution and the advent of sources such as social networks. Sales performance helps to efficiently and effectively achieve sales process goals by looking at opportunities and improving close rates. The data analytics established in this study as a technology or system provides useful insights into customer behavior by uncovering hidden patterns in BD to aid in the development of effective strategies for sales. In the era of the big data revolution, the method of strategy formulation in sales has changed, and organizations need to use data analytics systems to meet their needs. Individual characteristics is said to be the individual perception of big data analytics. In order to improve the objectivity of the comparison results, companies can add other models to participate in the comparison, so as to obtain accurate data analysis results. Data analytics is of great significance in this era of data overflow, and can provide unforeseen insights and benefits to decision makers in various areas. If properly exploited and applied, big data analytics has the potential to provide a basis for advancement. By applying such analytics to the data, valuable information can be extracted and exploited to enhance decision making and support informed decisions.

The technique of studying raw data to conclude a specific piece of information is known as data analytics. It is employed to assist people and organizations in making sense of data. They are applied to the analysis of raw data to discover trends and insights. We can infer conclusions about the information they contain by looking at select datasets and identifying trends. Data analytics is carried out using specialized hardware and software. These tools and methods are frequently employed in a variety of commercial sectors to empower businesses to take wise business decisions. Additionally, the analytics give companies the ability to react quickly to changing market trends and acquire an advantage over rival companies. Various efforts can benefit from some of the components of this analytics process. A good data analytics initiative will give you a clear picture of where you are, where you have been, and where you should go by merging these elements. To improve corporate performance, however, is data analytics' ultimate objective. Depending on the specific application, the data that is evaluated may be made up of new data that has been processed for real-time analytics or historical records. For the most effective data manipulation, data analytics uses a variety of software tools, including spreadsheets, data visualization, and reporting tools, data mining software, or open-source programming languages. Inside the data analytics process, the data analytics applications involve more than just analyzing data, especially on advanced analytics projects. After the data are analyzed, it will produce charts and other info graphics that can be designed to make findings easier to understand.

Data analytics enables organizations to analyze all of their data to identify patterns and generate insights to inform and, in some cases, automate decisions by relating Smart and actionable. Today's best solutions support end-to-end analytics, from accessing, preparing and analyzing data to operating analysis and monitoring results. When analyzing data, the main task is defined objects to analyze and separate data time period analyzed, to ensure the eccentricity of data analysis results. Data is useless if it cannot be analyzed, understood and applied in context. With this information, businesses can make meaningful changes to their future plans to maximize profits and success. Most raw data, especially large-scale databases, are worthless in their unprocessed state. We can extract valuable insights from this bit store using Power BI tools. The main goal here is to read and analyze the available data sets to generate business insights and overviews. The only part of the organization that earns revenue and money and delivers profits. The importance of selling is as follows: Sales data is a broad word that includes many types of metrics, but in general if you can measure something based on sales process is the actionable sales data. Through visualization, data analysis helps students understand concepts. Much technology is available to perform business data analysis, but Power BI visualization technique is the most popular techniques to learn the basics of data analysis. With the help of visualization techniques, data interpretation and data representation can be done quickly and easily. This strategy is useful for a more solid conceptual design.

Data analytics strategies can screen developments and metrics data might in any other case be misplaced with inside the mass of facts. These facts can then be used to optimize procedures to growth the overall performance of a commercial enterprise or system. Data analytics is the technology of studying uncooked records to make conclusions approximately that facts. Many of the strategies and procedures of records analytics has been automatic into mechanical procedures and algorithms that paintings over uncooked records for human

consumption. Data analytics is the technology of reading uncooked statistics to make conclusions approximately data information. The strategies and approaches of statistics analytics were computerized into mechanical approaches and algorithms data paintings over uncooked statistics for human consumption. Data analytics assist a business optimize its performance. Companies everywhere in the international try and get the advantages from get entry to the statistics to improve their overall performance and boom their revenue, however processing heterogeneous varieties of information to extract the precious information is a massive hassle that many businesses try and solve. One of the most essential developments is “Big Data Analytics”, a generation for Storing, Processing, and analyzing the information; groups are managing information to apply it in new ranges and direct decision-makers. Companies can use the insights they advantage from statistics analytics to tell their decisions, main to higher outcomes. Data analytics removes lots of the guesswork from planning marketing campaigns, selecting what content material to create, growing products, and more. Data analytics additionally offers you beneficial insights into how your campaigns are appearing so that you can fine-track them for topos-of-the-line outcomes. Data analytics offers you extra insights into your customers, permitting you to tailor customer support to their needs, offer extra personalization, and construct more potent relationships with them.

The fields of analytics, data visualisation, and business intelligence have all been completely transformed by Power BI. With the help of the web tool Power BI, users can find data, manipulate it, visualise it, and share the reports and dashboards they produce with other users in the same or different departments/organizations as well as the general public. As of February 2017, more than 200,000 companies across 205 nations employed Power BI. The unique Quick Insights feature of Power BI is based on a growing library of cutting-edge analytical algorithms. After uploading a data set to Power BI, this feature may be engaged with a single click, and it automatically generates a variety of reports based on the analysis of the data without requiring user input. Additionally, it helps to lessen human error in statistical calculations, which might result in unreliable research. Power BI accepts even Excel files as input and is suitable for implementation as a platform for research data analysis. The emergence of mobile-friendly analytics and business intelligence tools has been accelerated by the coronavirus pandemic.

6. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM

To identify skin cancer in an image, the current approach employs transfer learning and deep neural networks. In order to forecast skin cancer, the system used the pre-trained model Mobile Net. The Mobile Net model's lightweight DNN architecture makes it the perfect choice for embedded and mobile vision applications. The 2014 Image Net Challenge produced 12,80,000 photos with 1,000 different object classes, which were used to train the Mobile Net model. The HAM1000 database was used to collect the dataset. The dataset consists of 10015 dermoscopy images in total, including 115 Dermatofibroma images, 142 Vascular images, 6705 Melanocytic nevi images, 1113 Melanoma images, 1099 Benign keratosis images, 514 Basal cell carcinoma images, 327 Actinic keratosis images, and 600X450 resolution images of benign keratoses.

There is an uneven distribution of photos among the seven classes in the HAM10000 dataset. With the use of data augmentation, it is possible to rebalance the dataset's classes in a way that benefits other minority classes. Data augmentation is a useful technique for increasing the size of training data by arbitrarily changing various image-related parameters, such as rotation range, zoom range, horizontal and vertical flip, fill mode, etc. In order to create around 6000 images for each of the minority classes in the dataset—melanoma, benign keratosis, basal cell carcinoma, actinic keratosis, vascular lesion, and dermatofibroma—we performed data augmentation. This gave the training set a total of 38,569 images.

The final model will contain the Mobile Net model. The output of the model will be given to a convolutional neural network and a set of Batch 12 normalization and Dropout are applied on the model. The output will be finally obtained from the output layer which is a soft max function. The trained model was then deployed as a web application.

3.2 PROBLEM IDENTIFICATION

1. There are lots of data available, but interpreting that data correctly is vital. We will be at an extreme disadvantage if we misinterpret facts and make decisions based on that.
2. Market research involves a detail process of collecting and analysing data, which is time-consuming.
3. We can't fully trust the data. According to a study, 70% of marketers admitted they have poor quality and inconsistent data.

3.3 PROPOSED SYSTEM

This system aims to reduce the manual work in getting information from the data. It uses dataset to gain insights from the data. The workflow of the system is shown below.

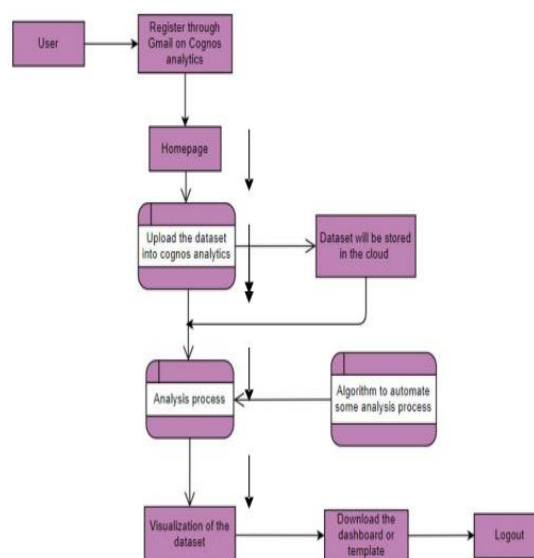


Fig-1 THE WORKFLOW OF THE SYSTEM

4. SYSTEM ARCHITECTURE

The proposed system is based on the dataset that is used. The dataset is from kaggle (Global sales data of super store). The data are visualized to get an idea about their attributes. Finally, the data are made into patches and visualized and the workflow contains the following modules:

1. Dataset Exploration
2. Preparing the Dataset
3. Data Visualizations
4. Deployment (Webpage)

Data collection is the process of measuring and gathering the data, or any variables in a systemized and well-established manner. It enables the collector to answer or test hypothesis and evaluate the outcomes of the collection.

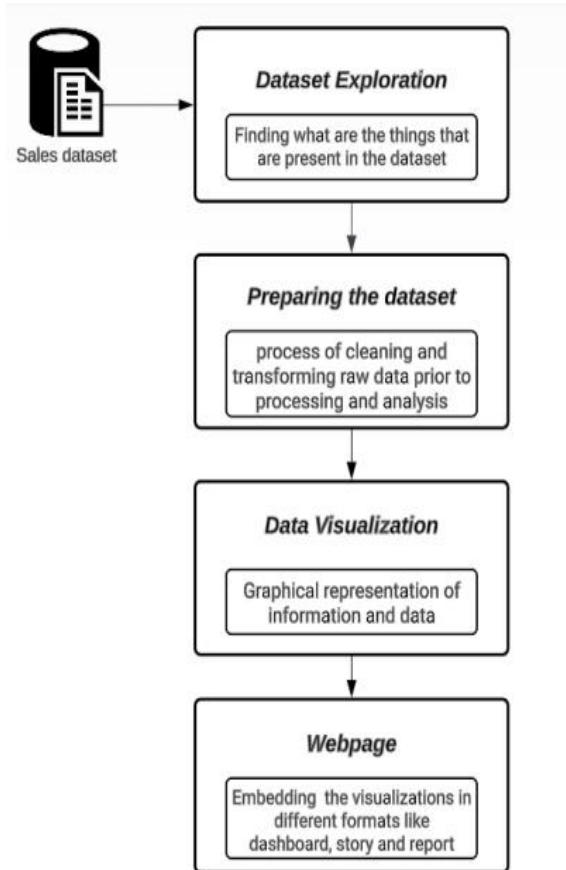


Fig-2 SYSTEM ARCHITECTURE

5. SYSTEM IMPLEMENTATION

The crucial project stage of system implementation occurs when the theoretical design is adjusted for use in actual systems. System implementation needs to happen after appropriate testing and validation. All the processes involved in converting an existing system to a new one are referred to as system implementation. The new system might be unique. A significant change to an existing system could be the replacement of a manual or automated system.

The implementation of our proposed solution begins with launching of Jupyter notebook leads to importing of certain necessary packages such as pandas, numpy, keras, Tensorflow etc.

The pretrained model will increase the accuracy more since it was already trained on a large amount of image dataset. The six-execution stage involves data collection, data analysis, preprocessing, model creation, training, testing. Using CNN max-pooling and also by using ResNet15V2 pretrained model improves training and test speed. To be efficient in an implementation phase a lot of activities have to be performed across various departments.

After importing all the packages, various machine learning is implemented for identifying an algorithm with high accuracy. The algorithm which is found to be more accurate is embedded with GUI (Graphical User Interface) backend for database connectivity.



Fig-3 DASH BOARD

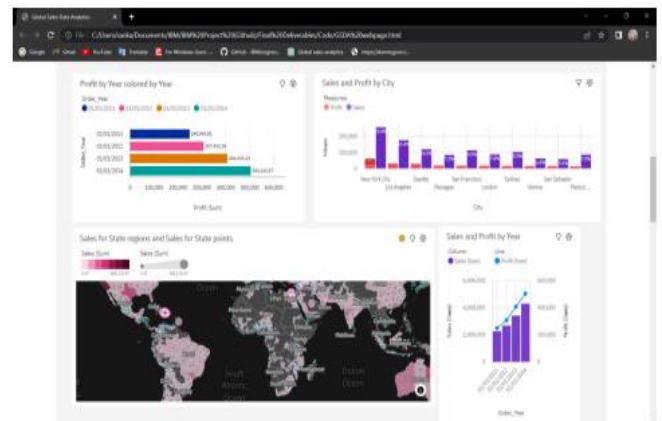


Fig-4 SALES DATA

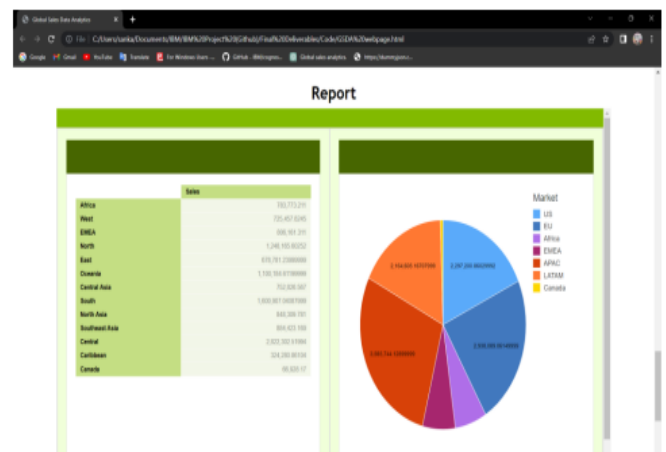


Fig-5 REPORT

6. CONCLUSIONS

In this project, automated augmentation of skin images dependent on deep neural networks were carried out. The data is trained with various architectures and their efficiency were analyzed. The new two-track design (which can model both the local information and the global environment) is used for high efficiency and the modeling of local mark dependencies through the stacking of two CNNs and ResNet15V2 model.

Learning is focused on a two-step process which, when the label distribution becomes unbalanced, helps one to train CNNs efficiently. The resulting segmentation method is very quick and 85 percent effective.

7. FUTURE ENHANCEMENT

In future research it is important to better model prior information by more suitable geometric modeling and closeness which codes classification errors in those areas more accurately.

REFERENCES

1. https://www.researchgate.net/publication/354780227_Survey_on_Growth_of_Business_using_Data_Analytics_for_Business_Intelligence_in_Real-Time_world.
2. <https://developer.ibm.com/tutorials/data-visualization-a-powerful-tool-for-visualizing-insight-data-overview/>
3. <https://www.ijraset.com/research-paper/data-analysis-and-visualization-of-sales-dataset-using-power-bi>.
4. https://www.researchgate.net/publication/359692412_Data_mining_with_its_role_in_marketing_sales_support_and_customer_identification_data_analysis.
5. <https://www.ibm.com/docs/en/cognos-analytics/11.1.0?topic=startedgetting-cognos-analytics>.
6. Eggert A, Serdaroglu M. Exploring the Impact of Sales Technology on Salesperson Performance: A Task-Based Approach. *J Mark Theory Pract*. 2011;19: 169–186.
7. Block J., Golterman J., Wecksell J., Scherburger K. and WC. Building blocks for technology-enabled selling. *Gart Gr Res Rep R-100-104 Stamford, CT Gart Gr*. 1996;1: 1–12.
8. Fan S, Lau RYK, Zhao JL. Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix. *Big Data Research*. 2015. pp. 28–32.
9. Agnihotri R, Dingus R, Hu MY, Krush MT. Social media: Influencing customer satisfaction in B2B sales. *Ind Mark Manag*. 2016;53: 172–180
10. Xu X, Sheng QZ, Zhang L-J, Fan Y, Dustdar S. From Big Data to Big Service. *Computer (Long Beach Calif)*. 2015;48: 80–83.
11. Shin DH. Demystifying big data: Anatomy of big data developmental process. *Telecomm Policy*. 2016;40: 837–854.
12. Shahbaz M, Gao C, Zhai LL, Shahzad F, Khan I. Environmental air pollution management system: Predicting user adoption behavior of big data analytics. *Technol Soc*. 2021;64.
13. Shahbaz M, Gao C, Zhai L, Shahzad F, Arshad MR. Moderating Effects of Gender and Resistance to Change on the Adoption of Big Data Analytics in Healthcare. *Complexity*. 2020;2020: 1–13.
14. Jeffrey DA. Testing the Technology Acceptance Model 3 (TAM 3) with the Inclusion of Change Fatigue and Overload, in the Context of Faculty from Seventh-day Adventist Universities: A Revised Model. *Dissertations*. 2015;3: 166. Available: <http://digitalcommons.andrews.edu/dissertations>
15. Weerakkody V, Kapoor K, Balta ME, Irani Z, Dwivedi YK. Factors influencing user acceptance of public sector big open data. *Prod Plan Control*. 2017;28: 891–905.

BIOGRAPHIES



Mr. R. Sivaguru Working As an Assistant Professor in Computer Science and Engineering Department at Knowledge Institute of Technology, Salem.