Analysing of Data using automated Data Analytics

Mr. Vinod Rajput[1]
Research Scholar, SAM Global
University, Bhopal, MS

Dr. Subhodini Gupta[2]
Research Guide, SAM Global
University, Bhopal, MS

Ms. Kajal Jaisinghani[3]
Research Scholar, University of
Mumbai, MS

Abstract:- **Data analysis, therefore, plays a key role in distilling this information into a more accurate and relevant form, making it easier for researchers to do to their job. Data analysis also provides researchers with a vast selection of different tools, such as descriptive statistics, inferential analysis, and quantitative analysis. The process of data analysis, or alternately, data analysis steps, involves gathering all the information, processing it, exploring the data, and using it to find patterns and other insights. Automated data analytics is the practice of using computer systems and processes to perform analytical tasks with little or no human intervention. Many enterprises can benefit from automating their data analytics processes. For example, a reporting pipeline that requires analysts to manually generate reports could instead automatically update an interactive dashboard.**

Keyword: - Data Analysis, automated, exploring, pattern

INTRODUCTION

Automation in data analytics is particularly useful when you're dealing with **big data**, and it can be used for a variety of tasks, such as data discovery, data preparation, data replication, and data warehouse maintenance. Automated analytics mechanisms vary in complexity. They range from simple scripts that fit records to a pre-established data model, to full-service tools that perform exploratory data analysis, feature discovery, model selection, and statistical significance tests. Automated data analytics can make decisions on behalf of enterprise stakeholders and create useful feedback mechanisms, such as an analytics system that regularly runs a study on data, then uses the results to automatically improve business processes while adjusting study inputs or parameters in real time. Automation in data analytics can provide insights that might be otherwise unavailable to an enterprise. A cybersecurity firm might use a classification algorithm to categorize large swathes of web activity, then deliver information about these categories in an interactive dashboard for their clients, who are hoping to protect their own customers. Feedback and customer input to this dashboard can be automatically reverted into the classification model, improving it in real time without intervention from the team that first implemented it.

What Is the Data Analysis Process?

☐ Data Requirement Gathering: Ask yourself why you're doing this analysis, what type of data you want to use, and what data you plan to analyze.

☐ Data Collection: Guided by your identified requirements, it's time to collect the data from your sources. Sources include case studies, surveys, interviews, questionnaires, direct observation, and focus groups. Make sure to organize the collected data for analysis.

☐ Data Cleaning: Not all of the data you collect will be useful, so it's time to clean it up. This process is where you remove white spaces, duplicate records, and basic errors. Data cleaning is mandatory before sending the information on for analysis.

☐ Data Analysis: Here is where you use data analysis software and other tools to help you interpret and understand the data and arrive at conclusions. Data analysis tools include Excel, Python, R, Looker, Rapid Miner, Chartio, Metabase, Redash, and Microsoft Power BI.

☐ Data Interpretation: Now that you have your results, you need to interpret them and come up with the best courses of action based on your findings.

☐ Data Visualization: Data visualization is a fancy way of saying, "graphically show your information in a way that people can read and understand it." You can use charts, graphs, maps, bullet points, or a host of other methods. Visualization helps you derive valuable insights by helping you compare datasets and observe relationships.

What Is the Importance of Data Analysis in Research?

A huge part of a researcher's job is to sift through data. That is literally the definition of "research." However, today's Information Age routinely produces a tidal wave of data, enough to overwhelm even the most dedicated researcher. Data analysis, therefore, plays a key role in distilling this information into a more accurate and relevant form, making it easier for researchers to do to their job. Data analysis also provides researchers with a vast selection of different tools, such as descriptive statistics, inferential analysis, and quantitative analysis. So, to sum it up, data analysis offers researchers better data and better ways to analyse and study said data.

When to automate data analytics

Automation can enhance data analytics, but how do you know when and where to use automation? As a general rule, it's most appropriate for tasks that are rules-based, performed often, and part of a stable business process. Automating a specific one-time study makes little sense. But automating data discovery processes in an organization that employs many data scientists, each working with varied data sources, would be more effective. Many analytical tasks are good candidates for automation. Creating dashboards, and reporting in general, are ideal candidates for automation. Automated analytics systems can stream, process, and aggregate data for publishing to interactive plots and live data summaries. Automation simplifies data maintenance tasks such as modifying and tuning a **data warehouse**. An enterprise should take advantage of the many tools that facilitate automatically integrating new data sources or migrating data from legacy systems. For example, Stitch parent Talend's suite of data integration applications allows customers to create compartmentalized data migration jobs that users can schedule and automate.

How to automate data analytics

- **Delineate your objectives**:- Data analytics are often cross-functional, so many teams may need to be involved in the planning process, including marketing, operations, and human resources. Set clear goals and expectations for the automation process in advance to facilitate cooperation and understanding between teams as the process moves forward.

- **Determine metrics: -** for measuring the performance and utility of the automated processes. This codifies the chosen objectives and helps ensure that they're met. Metrics also provide a reference for future projects or when extending the initial automated system.

- **Select reliable, well-supported automation tools: -** such as R or Python's NumPy, Pandas, and SciPy packages. Development focus for these programming languages is geared toward making studies shareable among academics and analytics practitioners (as exemplified by the **Jupyter project**). This focus makes it easier to move code and processes between humans and improves collaboration. Many data analytics tasks can be automated with these packages in combination with **other tools**. The cloud platforms that host **organizations' data warehouses** may provide tools for automated analytics. For example, Google Analytics includes a built-in **Analytics Intelligence tool** that uses machine learning to flag anomalies in time series data at the click of a button. Not all data tools lend themselves to automation. **Hadoop**, for instance, is great for a variety of big data tasks, but tools in the Hadoop ecosystem require extensive human involvement and can be difficult to automate.

- **Develop, test, iterate**: - Once you've prototyped an automated process, test it extensively. The automation should reduce repetitive work. An automated analytics system prone to failing or propagating errors can end up costing more time and taking more resources than a manual system. Implement the automated process and monitor its performance. Most automated data analytics systems have logging and reporting built in, so they can function with minimal oversight until failures occur or adjustments are required.

CONCLUSION

Companies that deal with big data may benefit from automating a portion of their data analytics infrastructure. Data lakes are stuffed with unstructured data that machines can analyse faster than humans. In addition, today's data warehouses are characterized by strict requirements for data modelling and processing, which can easily be streamlined by automated data analytics.

REFERENCES

[1] H. P. Luhn, A Business Intelligence Systems. IBM Journal of Research and Development, 1958, pp.314-319.
[2] M. Berthold, D. Hand, Intelligent Data Analysis. Springer, Berlin, 2009.
[3] P. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining. 2005. ISBN 0-321-32136-7
[4] G. Shmueli, N. R. Patel, P. C. Bruce, Data Mining for Business Intelligence. 2006. ISBN 0- 470-08485-5
[5] D. Pokorná, Business Data Analyses Possibilities. Diploma thesis. Faculty of Applied Informatics,Tomas Bata University in Zlín. 2010.
[6] D. Power, Dssresources.com [online]. 2007 [cit. 2010-06-07]. A Brief History of Decision SupportSystems