International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Optimizing Healthcare Data Analysis through NLP Machine Learning and Business Intelligence Solutions

Dr. Rajendra Kachhava, Sangam University, Bhilwara, India
E-mail: rajendrakachhava@gmail.com

Abstract **- All** organizations have additional data than ever do and data scientists apply advanced analytic techniques and machine learning ML models on that data to obtain value and make strategic and operational decisions. However, the collected raw data usually comes with various issues like incorrect values and duplicate records. Data preparation refers to the process of transforming such raw and data useful and clean format so that the downstream applications can consume reliably. The objective in this work is to enhance performance dashboards and investigate the use of data analysis techniques to automate data preparation pipeline operations through the use of entity matching, data enhancement, and error detection. The current demand for analytics experts vastly exceeds the supply so here analyse the solution to this problem is to increase the user friendliness of ML framework design to make them more accessible and effective automated machine learning (AutoML) mechanism is better design to solve the problem of expertise by providing fully automated off the shelf solutions for model choice and hyper parameter tuning. It clearly shows that the decisions made in medical facilities are highly data-driven. The results of the study confirm what has been analysed in the literature that medical facilities are moving towards data-based healthcare, together with its benefits.

Keywords— *BI Tools, Power BI, AutoML, ML, AI, Decomposition Tree, Prediction.*

I. INTRODUCTION

The understanding that AI-driven decision-making is essential in the fast-paced, fiercely competitive market of today has greatly increased interest in industrial machine learning (ML) applications. There is now a much more demand than there is supply for analytics expertise. Making ML frameworks more user-friendly so that even non-experts may utilize them is one way to solve this issue. The goal of automated machine learning (AutoML) is to address the skill gap by offering completely automated off-the-shelf solutions for hyperparameter tweaking and model selection. A significant factor in enabling this new approach to decision-making is business analytics [1].

Doctrine based on medical data is an essential part of the therapeutic process and a way to show the effectiveness of treatments. Keeping correct records helps patients, doctors, and other healthcare professionals communicate and understand one another. Sharing paper-based medical records with individuals in need must take a long time. Furthermore, the paper-based medical record storage system is not scalable and lacks backups [2].

Furthermore, many workers are reluctant to embrace AI technology out of fear that they may be supplanted by it as a result of advancements in automation and artificial intelligence. An expanding amount of data-driven decision-making based on artificial intelligence (AI) and machine learning (ML) is necessary in the fast-paced, globally competitive digital economy. All things considered, machine learning (ML) as a general-purpose technology for data-driven prediction will continue to permeate corporate analytics and influence the information management industry [3].

II. RELATED WORK

In this section contains the literature review. The literature review starts with an overview of the transformation of the healthcare sector before explaining big data. After giving a general introduction to AI, machine learning and its applications in both general and healthcare settings are discussed. After that, some of the promises and advantages of simulated intelligence for healthcare are discussed, along with some of its disadvantages. The ethical, legal, and regulatory issues that AI confronts are many and contribute to the limitations and challenges that have already been discussed [4].

When it comes to completely adopting modern technology and all that it has to offer, the private sector is expressively ahead of the health care and other government organizations. With the ongoing advancements in computers, new techniques for data collecting and diagnostics have emerged. The connectivity and analysis of this massive amount of data, however, are lacking. The healthcare system is ready for the seismic changes that come with the unavoidably quickening speed of technological advancement. The goal of the new healthcare system is to promote wellbeing from the inside out, while the majority of existing healthcare focuses on treating symptoms [5].

III. METHODOLOGY

In this research work, the design of an automatic machine learning system using the Pycaret library can make it easier for users by reducing some of the operations of machine learning in general.

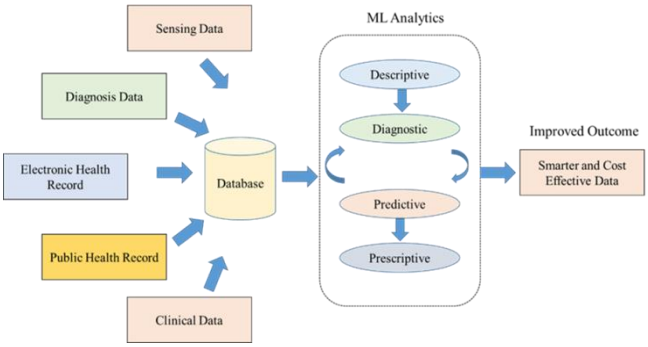

Figure 1. Data Workflow of ML Analytics

Here figure 1 Illustrates how big data in healthcare may come from a variety of sources. It is not necessary for the information to come from reputable sources like scientific studies, electronic health records (EHR), or medical supplies. Additionally, information gleaned from social media and search engines may be very helpful in making connections. This section outlines the methods for gathering and analyzing data as well as the overall quality of the research. Over the last 20 years, business intelligence has seen significant transformation [6].

Machine learning (ML) is a component of AI. It is used to comprehend data and identify trends. These patterns may be used to comprehend many medical scenarios, such as making a diagnosis based on the provided information. With machine learning (ML), programs may train and learn from data without the need for conventional programming by using mathematical models based on probability and statistics. Learning algorithms are used to facilitate learning, and machine learning (ML) characterizes the capacity to learn by identifying patterns in big datasets [7].

The learning algorithm/program may then be given unlabelled data to generate predictions after the labelled data has served as a basis. There are several approaches, but ultimately, it comes down to feeding the software a large amount of data so it can learn how to obtain the statistically most correct result. For instance, gathering information from several text-based sources, such clinical notes and medical records. Since these are often disorganized and challenging for computer algorithms to comprehend, natural language processing (NLP) may be useful in obtaining the relevant data to support the story [8].

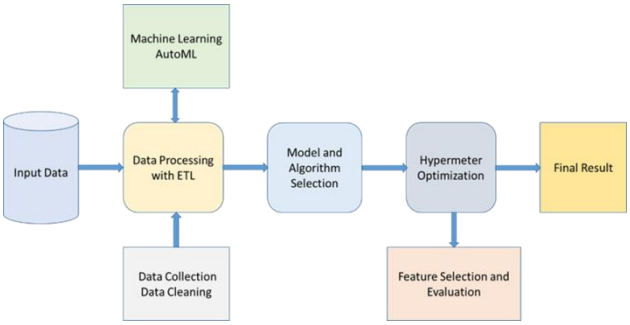

Figure 2. AutoML Design Model

Here in figure 2 show AutoML design model which define the way input data processing with model evaluation and feature selection with performance analysis and prediction based on the dataset which improve accuracy and decision making with use of AutoML pre trained model [9].

IV. RESULTS AND DISCUSSION

The updated dashboard appears as a summary page that serves as a master page for various offices, providers, payments, and insurance tabs. Additionally, date and office slicers were employed on this summary page to filter the data in accordance with user requests. View important revenue data, including revenue by department, revenue loss, and revenue year over year, time to goal revenue, and coverage. Additionally, the team will have the option to filter by department, facility, year, and month. This page summary displays all actual data that is updated in accordance with the datasets.

This summary page includes tables and a funnel with important influencer AI visuals for effectively displaying data. The updated dashboard appears as a monthly costs page that makes it simple to update and share with other users by displaying various offices, charges, and monthly visuals. Additionally, date and office slicers were employed on this charges page to filter the data in accordance with user requests. This page displays all scheduled and refreshed data that is updated in accordance with the datasets. With the help of our custom Power BI healthcare dashboards, BI healthcare dashboards will help healthcare organizations of all sizes gain insights into one or more facilities, patients, departments, clinical records, revenue, etc. so they can respond quickly and offer quality patient support systems while also assisting in the organization's growth.

These dashboards are helpful, for instance, in providing an overview of the locations of certain specialists throughout different areas in connection to the patients who need their particular treatment. Analyzing individual employee performance, clinician communication, nursing units, prescriptions, revenues, expenditures, etc. at a more daily level is made possible by this dashboard.

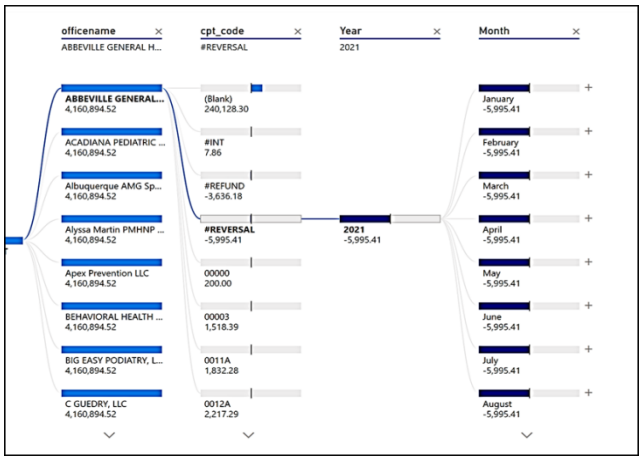

Figure 3: Healthcare Category Information Dashboard

The Patient Visit Summary gives valuable information on the number of patients, changes from year to year, and the number of patients by age group, case type, race, and gender. Obtaining data insights regarding facility performance, revenue summary, departmental information, and more is essential for healthcare businesses with many locations to work toward developing an effective patient support system.

The latest current AI visualization from the Power BI team is called the Decomposition Tree. You may deconstruct (divide down) a measure across several dimensions by multiple characteristics with this dynamic graphic aid. You may use this graphic for impromptu exploratory research to get a deeper understanding of your data. In addition, root-cause analysis may be performed with its assistance thanks to the AI capabilities included in. For example, while drilling into the tree, you may choose "High value" or "Low value." When you choose one of these, the graphic does a "AI split," determining which field needs to be drilled into next and displays the greatest or lowest number, depending on the situation. Now that the Decomposition Tree is in public beta, more enhancements should be forthcoming in the next months. Wherever necessary, elaborate on the tables and figures without repeating their contents. Interpret the findings in view of the results obtained in this and in past studies on this topic. State the conclusions in a few sentences at the end of the paper. However, valid colored photographs can also be published [10].

A customer in the insurance industry needed to monitor his operations while they were carried out at several sites. It shouldn't be at all difficult to grasp the report. They may take the necessary steps to increase their business by using this report, which helps them identify which branches, areas, and policies are doing better and which are performing worse. The client works in the health insurance sector and is responsible for manually gathering, storing, and analyzing enormous amounts of data. The company wants to automate its data reporting on a daily, weekly, and monthly basis. Additionally, the company wants to use Power BI to create a reporting dashboard so that it can get daily transaction data and create a business plan [11].

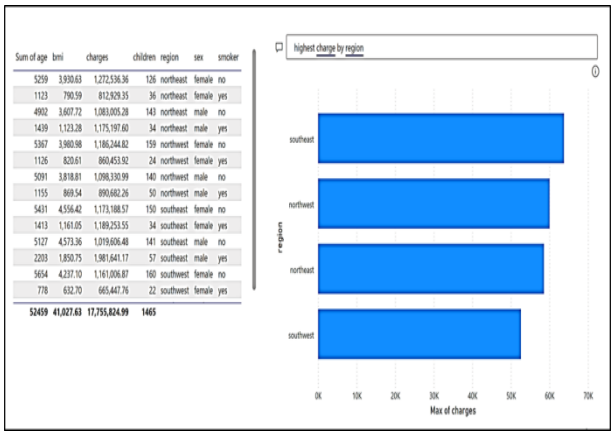

Figure 4: Use of NLP for Insight.

V. CONCLUSION AND FUTURE SCOPE

In this study, the technical, legal, and ethical aspects were prioritized. Potential study topics include pricing and other factors. Additionally, since laws and regulations may differ from one country to the next, there is a chance to research the subject abroad. Furthermore, research in the future could concentrate on only one of these three aspects and use an approach that focuses more on that specific element [9].

Our understanding of how AI interacts with ethical issues, legal and regulatory frameworks, and the healthcare industry is expanded by this study. Although previous study has been conducted on each of these three criteria, this thesis is the first to look at them all at once.

Furthermore, our analysis goes one step further and closes a data gap by dissecting the crisis division's assent to these restrictions. This is also accepted in Sweden, where it isn't taken care of before work [11].

REFERENCES

[1] M. Nassar et al., "The Role of Business Intelligence in Health Care: An Empirical Study," Journal of King Saud University - Computer and Information Sciences, vol. 30, no. 3, pp. 311-320, 2018. DOI: 10.1016/j.jksuci.2017.06.006.

[2] J. Waring et al., "Automated machine learning: Review of the state-of-the-art and opportunities for healthcare," Artificial Intelligence in Medicine, vol. 104, pp. 101822, 2020. DOI: 10.1016/j.artmed.2020.101822.

[3] A. Akshara et al., "Precipitation Climate Data Analysis using Power BI," 2024 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2024, pp. 1-4, doi: 10.1109/IC3IoT60841.2024.10550277.

[4] J. Virani et al., "Mental Healthcare Analysis using Power BI & Machine Learning," 2023 4th International Conference on Signal Processing and Communication (ICSPC), Coimbatore, India, 2023, pp. 73-76, doi: 10.1109/ICSPC57692.2023.10125680.

[5] J. K. Mathew et al., "Live Streaming Data Analysis using Distributed Stochastic Bi-LSTM Model," 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2022, pp. 1-4, doi: 10.1109/IC3IOT53935.2022.9767934.

[6] M. Jaipurkar et al., "IPL Data Analysis and Visualization Using Microsoft Power BI Tool," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-6, doi: 10.1109/ICCCNT56998.2023.10306783.

[7] A. Bafna et al., "A Novel approach to Data Visualization by supporting Ad-hoc Query and Predictive analysis : (An Intelligent Data Analyzer and visualizer)," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 113-119, doi: 10.1109/ICCS45141.2019.9065380.

[8] S. Mukherjee et al., "Using Microsoft Power BI Tool for Data Visualization in Healthcare," 2020 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE). IEEE, pp. 257-262. DOI: 10.1109/ICCIKE48582.2020.9279514.

[9] M.K. Hasan et al., "A Framework for Real-Time Dashboard Reporting using Power BI Tool," 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2). IEEE, pp. 1-4. DOI: 10.1109/IC4ME247184.2019.9036648.

[10] M.A. Alanezi et al., "The Application of Business Intelligence in Healthcare: Case Study in Saudi Arabia," Journal of Healthcare Engineering, vol. 2018, Article ID 7243483. DOI: 10.1155/2018/7243483.

[11] V. Koutkias et al., "Health Information Systems Integration in Healthcare Institutions: A Systematic Review of Methodologies," Journal of the American Medical Informatics Association, vol. 23, no. 6, pp. 1361-1369, 2016. DOI: 10.1093/jamia/ocw080.