Asian Journal of Research in Computer Science

Volume 16, Issue 4, Page 336-343, 2023; Article no.AJRCOS.109402
ISSN: 2581-8260

Cost-effective Digital Prescription using Pharmaceutical Knowledge Graph

Aryan Rathore [a*], Reva Bharara [a] and Krishna Kumar Tiwari [b]

[a] *Vellore Institute of Technology, Bhopal, India.*
[b] *Head of Knowledge Graph Platform AI Center of Excellence, Jio Private Limited Bangalore, India.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJRCOS/2023/v16i4395

Open Peer Review History:

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/109402

Received: 25/09/2023
Accepted: 29/11/2023
Published: 11/12/2023

Original Research Article

ABSTRACT

In the current era of the advancing medical domain and the ever-evolving use of technology in fields of pharmaceutical research, remote monitoring, and decision support systems in healthcare, prescription management has transformed from handwritten prescriptions to digital ones. This transition however does not imply that these prescriptions are comprehensive and provide optimized treatment outcomes. These digital prescriptions still reflect the formerly used handwritten prescriptions. Thus, recipients face the daunting challenge of making cost-effective, holistic, informed, and personalized decisions without compromising the legitimacy and authenticity of the original prescription, all due to a lack of readily available alternative medicine options. This problem can be addressed by utilizing the knowledge graph that we built, which contains carefully curated medical information collected from reliable and diverse sources, ensuring the authenticity and relevance of the information. Delving into the intricate interconnections among diverse medical entities and their properties, the medical knowledge graph presents an invaluable solution, empowering the generation of smart digital prescriptions in a fast and precise manner. Specifically, this study focuses on the transformative potential of digital prescriptions, elucidating their role in streamlining healthcare processes and enhancing communication

Corresponding author: E-mail: aryanrathore13572002@gmail.com;

between healthcare providers and patients. By leveraging the insights derived from our medical knowledge graph, we aim to contribute to the advancement of digital prescription systems, fostering more effective, personalized, and technology-driven healthcare solutions.

Keywords: Knowledge graph; medical kg; pharmaceutical; medicine; symptoms.

List of Abbreviations

Abbreviation	Meaning
kg	Knowledge Graph
API	Application Programming Interface
MRP	Maximum Retail Price
JSON	JavaScript Object Notation
OCR	Optical Character Recognition
NLP	Natural Language Processing
NER	Named Entity Recognition

1 INTRODUCTION

Knowledge graphs are powerful data structures that have the capability of storing and representing information in terms of a graph where each node represents an entity (real-world object) in the information and the edges represent the relationships between them. This intricate interconnectedness of a knowledge graph paves the way for efficient data analysis and aids in the discovery of previously unknown relationships, patterns, and insights. The advent of knowledge graphs has unlocked remarkable possibilities in optimizing data management, integration, querying, and scalability, enabling a seamless and effortless experience [1-4].

Onboarding medical data to a knowledge graph offers many crucial advantages such as data integration. In this data, different rich resources like drug databases, clinical trials, patients' medical histories, and doctors' records are combined to make one centralized and unified database enabling the discovery of different patterns and relationships [5,6]. Knowledge graphs provide a representation of information visually by presenting real-life entities as nodes and relations between them as edges. In terms of the medical knowledge graph, this translates into medicines, chemicals, side effects, uses, therapeutic classes, etc represented as nodes and connected to each other through edges representing relationships such as 'medicine_has_chemical' and 'medicine_has_side_effects' [7-9]. Another advantage of knowledge graphs is semantic reasoning, a process in which a knowledge graph can infer facts and valuable information from onboarded data and relations based on ontologies and inference rules which helps and upgrades the decision-making process. This type of reasoning can be used for the identification of potentially lethal drug interactions and how they can be avoided.

By utilizing the created knowledge graph enriched with diverse medical information, we can discover alternative medicines for the medicines in the original prescription. The original prescription may sometimes have medicines prescribed in such a way that it may not be in the best interest of the patient in terms of cost-effectiveness, and completeness [10-12]. By employing several graph traversal techniques it is possible to discover alternatives for each medicine in the original prescription which could have been overlooked earlier. An alternative can be defined as a medication that has a similar dosage, chemical composition, and side effects to the original. Medicines that satisfy these requirements can easily be identified by taking advantage of the interconnected nature of the medical knowledge graph. For the purpose of building smart prescriptions, we take into account the price of these alternative medicines and compare them to the price of medicines in the original prescription in such a way that the overall cost is minimized. These smart prescriptions give patients complete, individualized, affordable, and safe treatment alternatives, enabling them to take an active role in choosing their healthcare and maximizing the effectiveness of their treatments.

2 PREREQUISITE

2.1 Data Collection

Getting the dataset and data preprocessing: A total of 38,074 rows of data were collected through web scraping, providing essential information about medicine names, manufacturers, chemical compositions, uses, side effects, habit-forming tendencies, therapeutic and chemical classes, as well as action classifications. To gather this data and enhance the knowledge graph, various websites were scraped, and APIs were utilized to access the required information. The web scraping process employed the powerful open-source web automation library called Selenium, while API calls were made using the requests library. The data thus extracted was then compiled together, preprocessed using data processing libraries like pandas and numpy, plotted for analysis using Matplotlib and Seaborn and saved as a comma-separated-values file. The preprocessing of the data involved dealing with missing values and removing duplicates to ensure consistency and avoid redundancy. It should be noted that for the sake of this research each medicine had been given a random price value which may or may not reflect its true Maximum Retail Price.

2.2 Data onboarding in Knowledge Graph

After the dataset creation process is concluded, the data is onboarded into Arnagodb, a multi-model database system that combines the features of a graph database where Documents represent nodes (entities) and Edge Documents represent edges (relationship between two documents). For the sake of uploading the data in Arangodb, we extracted all the necessary entities from the dataset such as “Medicines”, “Manufacturer” etc, and stored all the relationships between these entities in separate JSON files. The JSON files were subsequently uploaded to the database as Document collections and Edge Document collections, resulting in the construction of the comprehensive medical knowledge Graph, encompassing all interconnections.

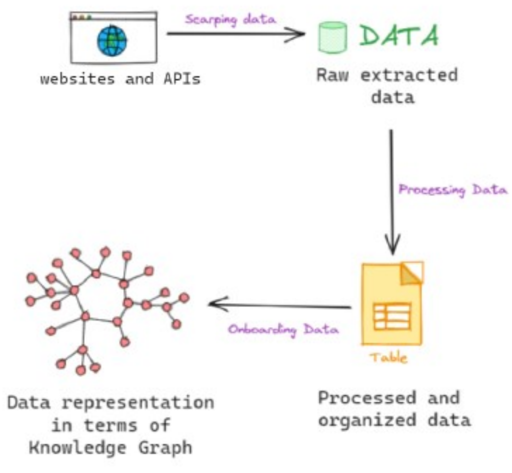

Fig. 1. Web scraping process

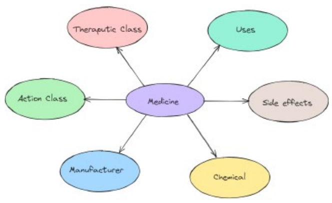

Fig. 2. Schema of the knowledge graph

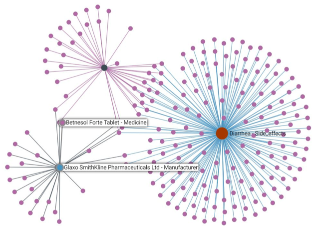

Fig. 3. General view of knowledge graph

3 METHODOLOGY

3.1 Approach

3.1.1 Prescription Entity Extraction

Once the foundational medicine database is ready, the next step is to read digital/ handwritten prescriptions for entity extraction. This can be done using optical character recognition (OCR) and natural language processing (NLP). Techniques such as Named Entity Recognition (NER) can be used to recognize and classify different entities like medication names, dosages and MRP in the original prescription. Upon processing the input data through this pipeline, the output will be a structured JSON document representing an abstract form of the original prescription. This Master JSON document will contain all the extracted entities categorized according to their respective fields, ensuring the desired format for easy analysis and further utilization.

3.1.2 Finding alternative medicine

The Master JSON document, which encompasses all the medicines and their corresponding attributes from the original prescription, serves as the input for the encoded API. This API facilitates access and traversal of the knowledge graph to identify alternative medicines based on specific criteria. An alternative medicine can be defined as a medication that shares similarities in dosage, chemical composition, and side effects with the original medicine. Consequently, any medicines within the knowledge graph that satisfy these criteria are selected as alternative medicines. The API provides a JSON response containing all the alternative medicines for each medicine listed in the Master JSON, regardless of their price.

3.1.3 Cost-effective prescription

The JSON document received from the Internal API is subsequently processed in order to identify medicines that possess an MRP lower than the medicines specified in the original prescription. Once such medicines are found, the alternative medicine with the lowest MRP is selected as a replacement for the original medicine in the smart prescription. Additionally, multiple medicines with the lowest price can be selected as alternative options, thereby providing the patient with a variety of choices to consider. These alternative medicines are then combined in various combinations to form the smart prescription, offering the patient a range of diverse options. By presenting such options, patients are empowered to make an informed decision based on their specific requirements and preferences.

3.2 Mathematical Formulation

Original prescription (OP):

Medicine 1 (m_1) $\rightarrow$ Price 1 (m_1^p)
Medicine 2 (m_2) $\rightarrow$ Price 2 (m_2^p)
. .
. .
Medicine n (m_n) $\rightarrow$ Price n (m_n^p)

Where m_i is the medicine and each medicine has a non-negative price attribute represented by m_i^p.

Smart prescription (SP):

Alternative Medicine 1 (a_1) $\rightarrow$ Alternative Price 1 (a_1^p)
Alternative Medicine 2 (a_2) $\rightarrow$ Alternative Price 2 (a_2^p)
. .
. .
Alternative Medicine n (a_n) $\rightarrow$ Alternative Price n (a_n^p)

Where a_i is the medicine and each medicine has a non-negative price attribute represented by a_i^p.

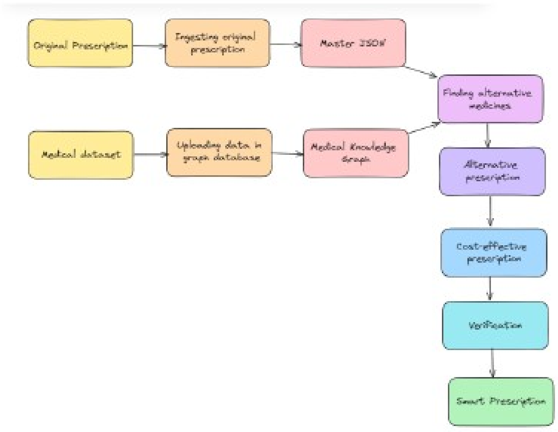

Fig. 4. Approach for generating smart prescription

3.3 Proof

Total cost of the original prescription (OP): $m_1^p + m_2^p + m_3^p + \ldots\ldots\ldots\ldots + m_n^p = \sum_{i=0}^{n} m_i^p$

Total cost of the smart prescription (SP):

$$a_1^p + a_2^p + a_3^p + \ldots\ldots\ldots\ldots + a_n^p = \sum_{i=0}^{n} a_i^p$$

Let,

$m_1^p = a_1^p + k_1$

$m_2^p = a_2^p + k_2$

.

.

$m_n^p = a_n^p + k_n$

Where k is a positive integer value

Adding all equations:

$m_1^p + m_2^p + m_3^p + \ldots\ldots\ldots\ldots + m_n^p = (a_1^p + a_2^p + a_3^p + \ldots\ldots\ldots\ldots + a_n^p) + (k_1 + k_2 + k_3 + \ldots\ldots\ldots\ldots + k_n)$

$\sum_{i=0}^{n} m_i^p = \sum_{i=0}^{n} a_i^p + \sum_{i=0}^{n} k_i$

$\sum_{i=0}^{n} m_i^p \geq \sum_{i=0}^{n} a_i^p$

Therefore, the total cost of the smart prescription is always less than or equal to the original prescription.

4 RESULT

The approach to addressing the real-life problem of finding alternative medicines to those specified in a given prescription, with the objective of generating a cost-effective prescription that empowers patients to make informed and personalized decisions, began with data collection through web scraping and APIs. A comprehensive dataset comprising 38,074 rows was gathered, yielding valuable insights and inferred facts. Fig. 5. shows the analysis of the data which revealed that approximately one in every 40 medicines possessed addictive properties, with a significant portion belonging to the therapeutic class "NEURO CNS." Further investigation uncovered that Neuro/CNS drugs can induce addiction by interacting with reward pathways in the brain, leading to dependence and cravings, thus validating the insight.

The analysis also revealed valuable insights on the distribution of medicines based on action class, therapeutic class, chemical class, and manufacturer, as depicted in Fig. 6. Notably, Sun Pharmaceutical Industries Ltd emerged as the market leader, offering an impressive range of 1350 distinct types of medicines. Furthermore, the analysis highlighted the primary focus of each manufacturer, with Glaxo SmithKline Pharmaceuticals Ltd specializing in the production of anti-infectives, while Alembic Pharmaceuticals Ltd had a significant presence in the manufacturing of cardiac-related medicines as shown in Fig. 7. These findings provide valuable knowledge for various applications and contribute to the enrichment of the knowledge graph.

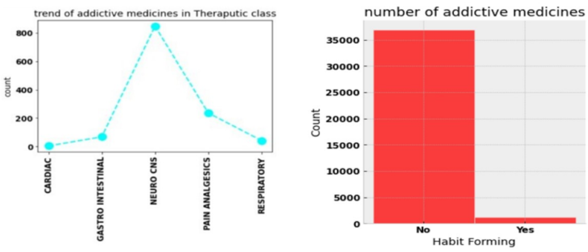

Fig. 5. Trend of addictiveness of medicines

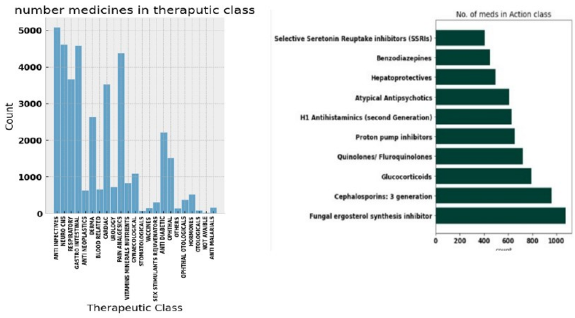

Fig. 6. Distribution of medicines based on Therapeutic and Action classes

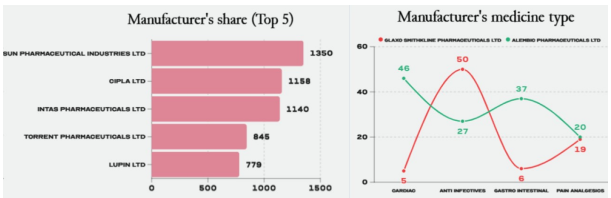

Fig. 7. Insights on medicine manufacturers

Subsequently, the enriched knowledge graph was extensively explored and traversed to identify alternative medicines based on composition and other relevant criteria. The primary objective remained the generation of smart prescriptions that were both cost-effective compared to the original prescription and valid. To achieve this, entities in the original prescription were extracted using NLP, OCR, and named entity recognition techniques. These entities served as inputs for querying the knowledge graph, enabling the identification of more affordable alternatives. The amalgamation of these economical alternatives resulted in an alternative prescription capable of legitimately and accurately substituting the original prescription.

Comparing the total cost of the original prescription to that of the alternative prescription yielded significant savings. For instance, in a test prescription comprising seven medicines with a total cost of Rs. 2577, the alternative prescription amounted to just Rs. 114. A visual representation highlighting the cost comparison between each medicine in the original prescription and its corresponding alternative in the smart prescription revealed substantial deviations as depicted in Fig. 10.

```
medicine_names = ["Augmentin 625 Duo Tablet", "Bignac SP Tablet", "Doxypen 100 Capsule",
                  "Azithral 500 Tablet", "ZX 200 DT Tablet", "Xomet 10mg Tablet", "Vesbal 24 Tablet"]
```

Fig. 8. Input (original prescription)

```
Augmentin 625 Duo Tablet, (452) Rs./- has the alternate medicine: Womclav 500 mg/125 mg Tablet, price: (12) Rs./-
Bignac SP Tablet, (415) Rs./- has the alternate medicine: Zolron SP 100mg/325mg/15mg Tablet, price: (12) Rs./-
Doxypen 100 Capsule, (469) Rs./- has the alternate medicine: Doxy 100mg Tablet, price: (14) Rs./-
Azithral 500 Tablet, (105) Rs./- has the alternate medicine: Welzith 500mg Tablet, price: (10) Rs./-
ZX 200 DT Tablet, (237) Rs./- has the alternate medicine: Johncef 200mg Tablet, price: (15) Rs./-
Xomet 10mg Tablet, (459) Rs./- has the alternate medicine: Voxitin 10 Tablet, price: (49) Rs./-
Vesbal 24 Tablet, (440) Rs./- has the alternate medicine: Vertin MDS Orally Disintegrating Strip, price: (12) Rs./-
```

Fig. 9. Output (smart prescription)

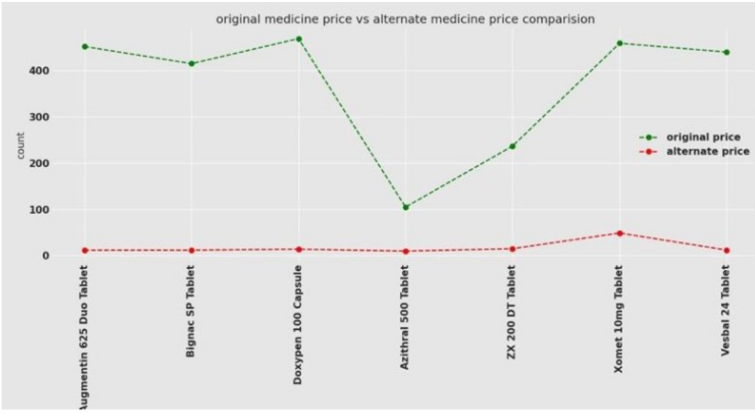

Fig. 9. Comparison between cost of medicines in original prescription and smart prescription

Therefore, the developed approach successfully tackled the challenge of providing cost-effective alternative medicines in smart prescriptions

5 CONCLUSION

In this research paper, we accomplished the task of generating a pharmaceutical knowledge graph using with generation of cost-effective medical prescription was possible.Through the collection and analysis of a comprehensive dataset, important insights were gained regarding the prevalence of addictive medicines and the distribution of medications across various therapeutic classes and manufacturers. the utilization of a medical knowledge graph has proven to be a valuable approach in addressing the challenge of finding cost-effective alternative medicines for patients. Through the collection and analysis of a comprehensive dataset, important insights were gained regarding the prevalence of addictive medicines and the distribution of medications across various therapeutic classes and manufacturers. The knowledge graph, enriched with diverse medical information, enabled the discovery of alternative medicines that share similar composition and properties to those in the original prescription. Overall, the research highlights the significance of knowledge graphs in the medical field, offering a promising solution for generating cost-effective and personalized prescriptions. As technology continues to advance and medical data expands, the application of knowledge graphs holds great potential for enhancing healthcare decision-making and improving patient outcomes.

6 FUTURE WORK

In future upgrades, our research project aims to enhance the performance and efficiency of the knowledge graph by incorporating the true Maximum Retail Prices (MRPs) of medicines. This improvement will significantly contribute to improving cost-effectiveness. Furthermore, we plan to expand the knowledge graph by integrating diverse medical and pharmaceutical information beyond the scope of medicines and their details. We will source this information from verified government databases, ensuring its reliability, safety, and affordability. Enriching the knowledge graph with such comprehensive data will encompass various aspects, including drug interactions, evaluation of potentially toxic medications, and safety advice for patients with specific health conditions. These enhancements will enable the knowledge graph to provide even more personalized and accurate medical prescriptions, empowering advanced healthcare decision-making.

COMPETING INTERESTS

Authors have declared that they have no known competing financial interests OR non-financial interests OR personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

[1] Nahin RL, Straus SE. Research into complementary and alternative medicine: problems and potential. BMJ. 2001;322(7279):161-4.
DOI: 10.1136/bmj.322.7279.161
PMID: 11159578; PMCID: PMC1119420

[2] Kern SE, Jaron D. Complementary and alternative medicine in the technology age. in IEEE Engineering in Medicine and Biology Magazine. 2005;24(2):28-29.
DOI: 10.1109/MEMB.2005.1411344

[3] Shailaja K, Seetharamulu B, Jabbar MA. Machine learning in healthcare: A review. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India. 2018;910-914.
DOI: 10.1109/ICECA.2018.8474918

[4] Martins A, Maia E, Praça I. Herb–Drug interactions: A holistic decision support system in healthcare. 2022 IEEE International Conference on E-health Networking, Application & Services (HealthCom), Genoa, Italy. 2022:1-6.
DOI: 10.1109/HealthCom54947.2022.9982729

[5] Zaki N, Tennakoon C, Ashwal HA. Knowledge graph construction and search for biological databases. Proc. Int. Conf. Res. Innov. Inf. Syst. (ICRIIS). 2017;1-6.

[6] Abu-Salih B. Domain-specific knowledge graphs: A survey. Journal of Network and Computer Applications, 2021;185:103076.
ISSN 1084-8045
DOI: https://doi.org/10.1016/j.jnca.2021.103076
bibitemb3 Glez-Peña D, Lourenco A, López-Fernández H, Reboiro-Jato M, Fdez-Riverola F. Web scraping technologies in an API world. Briefings in Bioinformatics. 2013;15.
DOI: https://doi.org/10.1093/bib/bbt026

[7] Khder M. Web scraping or web crawling: State of art, techniques, approaches and application. International Journal of Advances in Soft Computing and its Applications. 2021;13:145-168.
DOI: 10.15849/IJASCA.211128.11

[8] Yu S, Yuan Z, Xia J, Luo S, Ying H, Zeng S, Ren J, Yuan H, Zhao Z, Lin Y, Lu K, Wang J, Xie Y, Shum HY. BIOS: An Algorithmically Generated Biomedical Knowledge Graph.
Retrieved from arXiv:2203.09975

[9] Ji S, Pan S, Li S, Zhu X, Chang E. A survey on knowledge graphs: Representation, acquisition, and applications. IEEE Transactions on Neural Networks and Learning Systems. 2021;33(2):494-514.

[10] Asada M, Gunasekaran N, Miwa M, Sasaki Y. Representing a heterogeneous pharmaceutical knowledge-graph with textual information. Front Res Metr Anal. 2021;1;6:670206.
DOI: 10.3389/frma.2021.670206
PMID: 34278204; PMCID: PMC8281808

[11] Akter, Salima, Hasan, Mohammad, Begum, Rokeya, Akhter, Hajara, Gazi, Md, Sabrin, Farah, Kim, Sung. Alternative Medicine: A Recent Overview; 2021.
10.5772/intechopen.97039

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle5.com/review-history/109402