



# DECISION SUPPORT SYSTEM BASED ON ECLAT ALGORITHM (DATA SCIENCE TECHNIQUE) FOR MEDICAL SECTOR

<sup>1</sup>Dr. Reshma Banu, <sup>2</sup>Hoor Afza Khanum, <sup>3</sup>Sri Raksha, <sup>4</sup>Kusumitha N, <sup>5</sup>Madhushree R

<sup>1</sup>Head of the department and Assistant Professor <sup>2,3,4,5</sup>Student

Department of Information Science and Engineering

<sup>1,2,3,4,5</sup>GSSS Institute of Engineering and Technology for Women, Karnataka, India.

**Abstract:** Online health communities continue to offer a huge variety of medical information useful for medical practitioners, system administrators, and patients alike. In this system, we collect real-time health posts from reputed websites, where patients express their views, including their experiences and side-effects on drugs used by them. We propose to perform a Summarization of user posts per drug and come out with useful conclusions for the medical fraternity as well as the patient community at a glance. Further, we propose to classify the users based on their 'emotional state of mind'. Also, we shall perform knowledge discovery from user posts, whereby useful 'patterns' about the triad drugs-symptoms-medicine is done by Association Learning.

**Index Terms - Eclat algorithm, NLP algorithm, Association learning, Unsupervised machine learning.**

## I. INTRODUCTION

Online health communities continue to offer a huge variety of medical information useful for medical practitioners, system administrators, and patients alike. In this system, we collect real-time health posts from reputed websites, where patients express their views, including their experiences and side-effects on drugs used by them.

We propose to perform a summarization of user posts per drug and come out with useful conclusions for the medical fraternity as well as the patient community at a glance. Knowledge mining of the health posts, we propose to apply different important operations like - Association Learning, Summarization, and Sentiment Analysis on data obtained from the health forum site.

Association learning is a popular and widely-known machine learning task. It is used to find out interesting relations between variables in a large database.

## A. ECLAT ALGORITHM

In the project, we use the "*eclat algorithm*" to find the relationship between medical objects such as symptoms-diseases-drugs. Eclat algorithm is one of the efficient algorithms and takes less time for data processing. This algorithm works fine for small data sets as well as large data sets.

The proposed system accepts these opinions and processes using data science algorithms and initially will identify the symptoms, disease types, and treatment details using NLP technique (lesk based algorithm), then the output of NLP is inputted to "*Eclat algorithm*" which is an unsupervised learning algorithm and processes these medical words and finds the correlation between symptoms, disease types and treatment details. This kind of pattern prediction useful to the medical sector, medical practitioners.

The main objective of the project is to

- The aim is that the system collects the posts from the users related to side effects on drugs & it summarizes all the user posts and comes out with convenient conclusions.
- It is a new online community where we collect a huge variety of medical functional convenient for medical practitioners, patients, etc.
- To discover the functional patterns based on side effects per drug.

## II. LITERATURE REVIEW

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs and limit the use of hard tabs turns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads—the template will do that for you. Finally, complete content and organizational editing before formatting. [1] “Drug disease association prediction based on neighborhood information aggregation in neural network”. In this paper the medical services and information system the base of the strategic management in health care systems. Data analysis is based on one-to-one association prediction and robust analysis by testing on different datasets. This paper Identifying the relationship between drugs and diseases, E-Health and Bioengineering Conference, 2017.[2] “An integrated model for home health care and patients cluster-based care” The model is solved with the CPLEX and the behavior of the model is concerning different scenarios. Here the advantage is the identification of patient clusters centers and corresponding assignment issues among the entities. The results indicate that the model successfully tackles the considered issues, International Conferences on Robotics and Automation Sciences, 2017. [3] “Brain-Computer Interface Based Home Automation System for Paralysed People”. In this, paper a BCI Based home controlling has been proposed. The proposed system locates the position and orientation of the individual. User attention is captured using the BCI module, which is further processed to identify the action. IEEE Recent Advances in Intelligent Computational Systems (RAICS), 2018. [4] “Mobile Based Learning Development for improving Quality of Nursing Education in Indonesia” In this paper it present’s a low-cost health sensor platform for rural health monitoring with well structured and secured interface between medical experts.Prototype of a mobile-based online learning management system (m-LMS). The study aims to determine the preferences of prospective nurses in using smartphone devices as m-learning media, 2019 IEEE Conference on Sustainable Utilization and Development in Engineering and Technologies, 2019. [5] “Reliable Machine learning Approach to Predict Patient Satisfaction for Optimal Decision Making and Quality Health Care” In this ML approach is used in determining the patient’s satisfaction in the healthcare sector. A unique approach towards the decision-making process and better quality care in healthcare applications are developed using ML. The results are validated using traditional statistical methods like binomial correlation and linear regression, 2019 International Conference on Communication and Electronics Systems (ICCES), 2019.

## METHODOLOGY

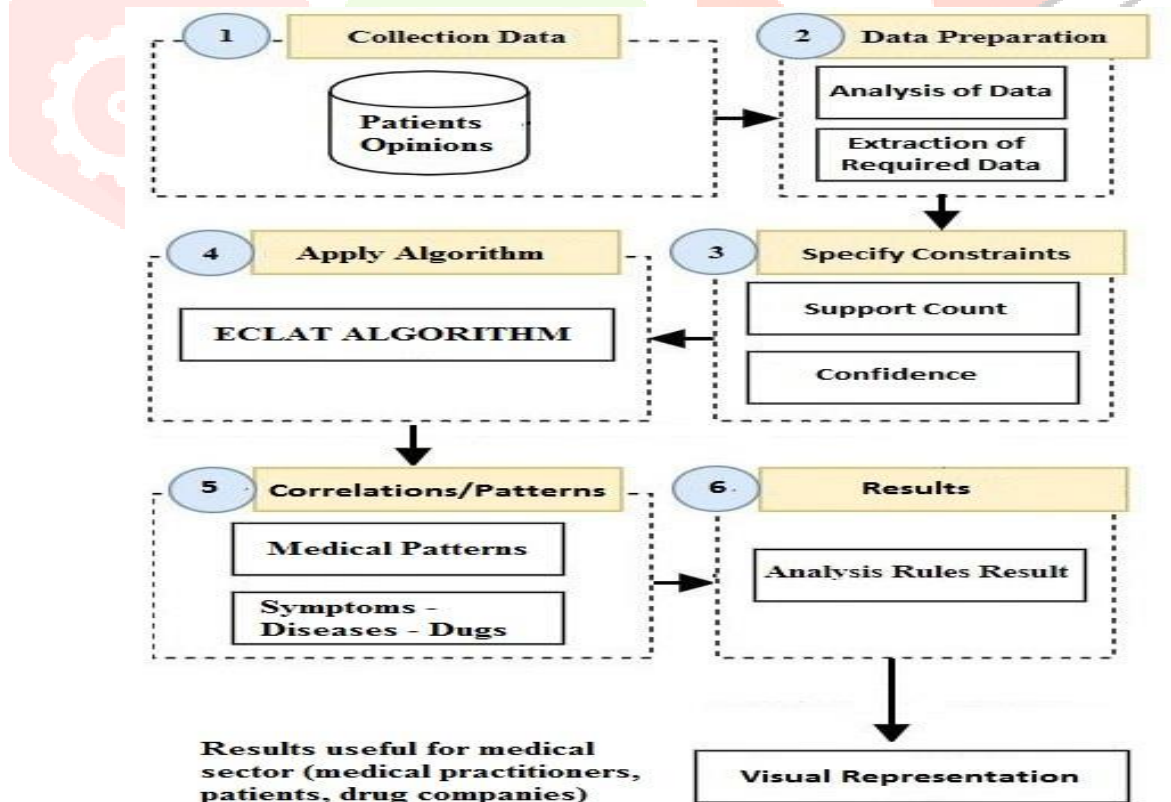


Fig 1: system architecture

**Machine Learning:** Machine learning is the process of studying a system based on data. Machine learning is a part of data science where we use machine-learning algorithms to process data.

**Unsupervised Learning:** A Descriptive model is worn for tasks that would benefit from the insight gained from summarizing data in new and absorbing ways. There are no predefined labels in the unsupervised learning technique. The goal is to explore the data and find some structure within it. Unsupervised learning works well on transactional data.

**Descriptive Model:** Descriptive model developed using clustering techniques and association learning techniques. We have many efficient algorithms such as eclat algorithm AIT algorithm SFIT algorithm STEM Algorithm FP Growth algorithm K Means algorithm Fuzzy C Means algorithm etc

In the project, we use the “**eclat algorithm**” to find the relationship between medical objects such as symptoms-diseases-drugs. Eclat algorithm is either efficient algorithm and takes less time for data processing. This algorithm works fine for small data sets too large data sets.

**Proposed System**

- Machine learning is a process of studying a system based on data. Machine learning is a part of data science where we use machine learning algorithms to process data.
- A Descriptive model is worn for tasks that would benefit from the insight gained from summarizing data in new and absorbing ways. There are no predefined labels in the unsupervised learning technique.
- The goal is to explore the data and find some structure within it. Unsupervised learning works well on transactional data.

In the project, we use the “*eclat algorithm*” to find the relationship between medical objects such as symptoms-diseases-drugs. Eclat algorithm is either efficient algorithm and takes less time for data processing. This algorithm works fine for small data sets too large data sets.

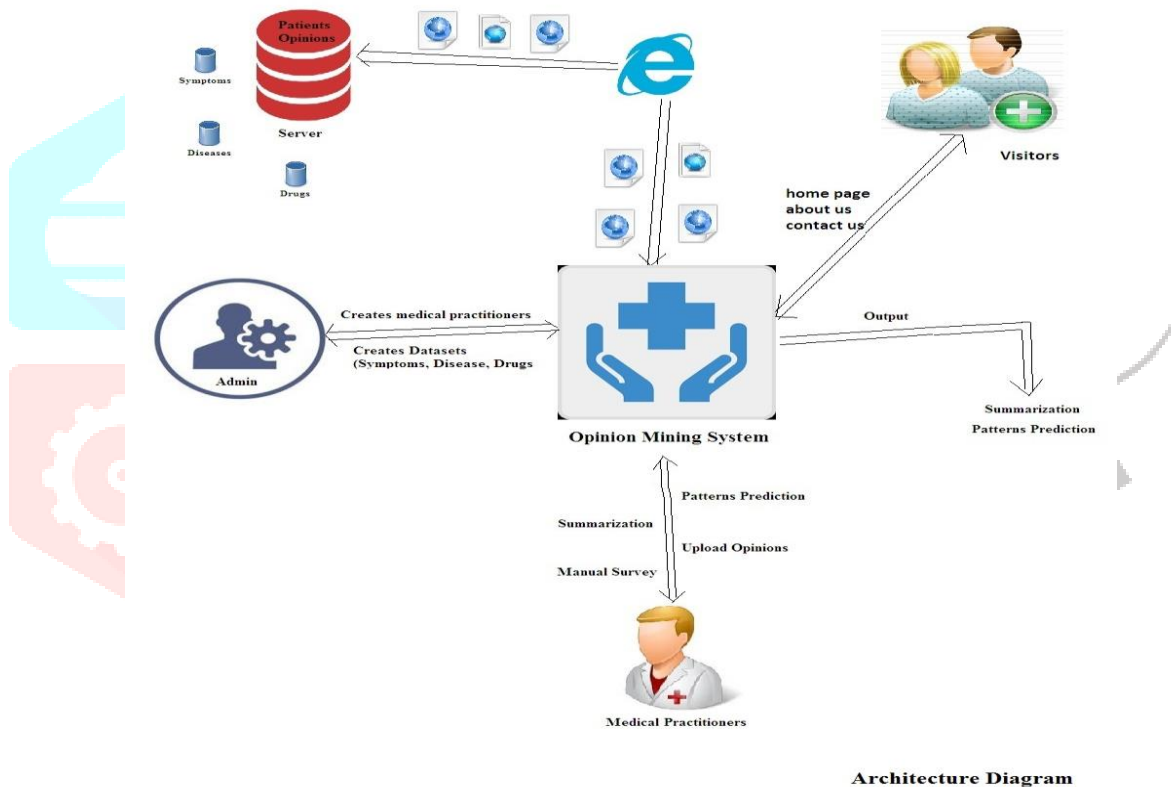


Fig 2: System Architecture for representation of the relationship between drugs, disease, and symptoms.

**System Modules:**

- Administrator  
An Administrator is a person who maintains the entire application. The Administrator has full authority.
- Medical Practitioner (doctors, receptionists, surgeon)  
In the medical profession, a general practitioner (GP) is a medical doctor who treats acute and chronic illnesses and provides preventive care and health education to the victims. A Medical practitioner is a person who receives the services from the application.
- Visitor  
The Visitor is a new user, the one who visits the application.

**CONCLUSION AND FUTURE SCOPE**

We are proposing this system to identifying the list of Symptoms, disease, and drug from each patient opinion and we make a summarization. Once the summarization is done we predict medical patterns (which shows the relationship between symptoms, disease, and drugs). Thus our work shall equally benefit the medical fraternity, patient community, and pharmaceutical companies.

- This project system pattern prediction was useful to the medical sector, medical practitioners, and aid in improving the medical system in the country.
- The projects help to collect the posts from the users related to side effects of the drug & it summarizes all the user posts and come out with useful conclusions.
- It helps to collect a huge variety of medical information useful for medical practitioners, patients, etc...to discover the useful patterns based on side effects per drug.

**REFERENCES**

- [1] Jayashree R, Srikanta Murthy K, Basavaraj .S.Anami, "Categorized Text Document Summarization in the Kannada Language by Sentence Ranking" 12th International Conference on Intelligent Systems Design and Applications (ISDA), pp 776-781, 2012.
- [2] Alok Ranjan Pal, DigantaSaha, "An Approach to Automatic Text Summarization using WordNet" IEEE International Advance Computing Conference (IACC), 2014.
- [3] Jesmin Nahar, Tasadduq Imam, Kevin S. Tickle, Yi-Ping Phoebe Chen, "Association rule mining to detect factors which contribute to heart disease in males and females" J. Nahar et al. / Expert Systems with Applications 40 (2013) 1086–1093, Elsevier, 2012.
- [4] Lakshmi K.S, G. Santhosh Kumar, "Association Rule Extraction from Medical Transcripts of Diabetic Patients" IEEE,2014.
- [5] WalaaMedhat, Ahmed Hassan, HodaKorashy, "Sentiment analysis algorithms and applications: A survey" In press, Elsevier, 2014.
- [6] Rafael Ferreira, FredericoFreitas, Luciano de Souza Cabral, Rafael DueireLins, Rinaldo Lima, Gabriel Franca, Steven J. Simske, and Luciano Favaro, "A Context Based Text Summarization System",11th IAPR International Workshop on Document Analysis Systems,pp 66-70, 2014.
- [7] C. Lakshmi Devasenal and M. Hemalatha, "Automatic Text Categorization and Summarization using Rule Reduction" IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012), pp 594-598, 2012.
- [8] Sara Keretna, CheePeng Lim, Doug Creighton, "A Hybrid Model for Named Entity Recognition Using Unstructured Medical Text" Proc.Of the 2014 9th International Conference on System of Systems Engineering (SOSE), Adelaide, Australia- June 9-13, pp 85-90, 2014.
- [9] SaeedMohajeri, AfsanehEsteki, Osmar R. Zaiane, and DavoodRafiei, "Innovative Navigation of Health Discussion Forums based on Relationship Extraction and Medical Ontologies" IEEE International Conference on Bioinformatics and Biomedicine, pp 13-14, 2013.
- [10] Yi Chen, Yunzhong Liu, "Connecting the Dots: Knowledge Discovery in Online Healthcare Forums" ICEC'14 August 05 - 06 2014, ACM.
- [11] Subhabrata Mukherjee, Gerhard Weikum, CristianDanescu-Niculescu-Mizil, "People on Drugs: Credibility of User Statements in Health Communities" KDD '14, August 24 - 27 2014, New York, ACM, 2014.
- [12] Khairullah Khan, BaharumBaharudin, Aurangzeb Khan, Ashraf Ullah, "Mining opinion components from unstructured reviews: A review" Journal of King Saud University – Computer and Information Sciences (2014), Elsevier, 2014.
- [13] Miller G, Beckwith R, Fellbaum C, Gross D, Miller K., "WordNet: an on-line lexical database", OxfordUniversity Press; 1990.