International Journal of Research in Engineering, Science and Management
Volume-2, Issue-10, October-2019
www.ijresm.com | ISSN (Online): 2581-5792

An Analytical Study of Allergy Induced Asthma Patients Using Agent Based Data Mining Techniques

T. Rekha[1], K. Shyamala[2]
[1]*Research Scholar, Department of Computer Science, Dr. Ambedkar Government Arts College, Chennai, India*
[2]*Associate Professor, Department of Computer Science, Dr. Ambedkar Government Arts College, Chennai, India*

***Abstract*: Health promotion is very relevant today. There is a global acceptance that health and social well-being are determined by many factors outside the health system which include socio economic conditions, patterns of consumption associated with food & communication, demographic patterns, learning environments, family patterns, the cultural and social fabric of societies, socio-political and economic changes, including commercialization and trade and global environmental change. In such a situation, health issues can be effectively addressed by adopting a holistic approach by empowering individuals and communities to take action for their health, fostering leadership for public health, promoting intersect oral action to build healthy public policies in all sectors and creating sustainable health systems. Now a day many people are suffering by allergy and asthma because of their living life style, here we take some example of patient details how the allergy turns over asthma depending upon the sex and age criteria. Based on the analysis the following prediction are made.**

Data mining is used to attain this objective. Data mining techniques are used to discover models or patterns of data, and it is much helpful in the process of decision making. The concept of agent paradigm along with the clustering techniques of data mining model is used here to make an analytical study on Allergy-induced Asthma. Agent based systems involving Data mining techniques, provides a set of agents which can help us to proceed with the proposed study.

***Keywords*: data mining, allergy-induced asthma, agent mining, data mining, agent paradigm, clustering**

1. Introduction

Data mining means an important paradigm for educational assignment [1]. It is the computational process of discovering patterns in large data sets. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use [2]. The process of data mining is also called as knowledge discover in data bases abbreviated as KDD, is commonly defined with the following stages [3], [4]:

1. Developing an understanding of
   - the application domain
   - the relevant prior knowledge
   - the goals of the end-user
2. Creating a target data set: selecting a data set, or focusing on a subset of variables, or data samples, on which discovery is to be performed.

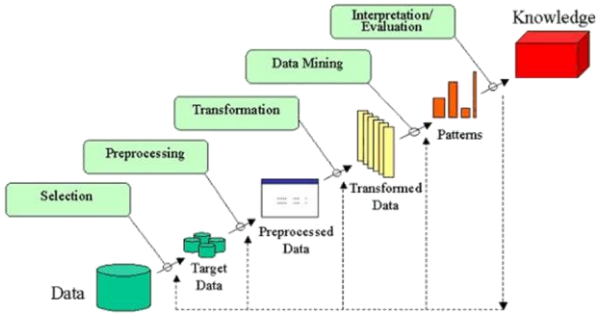

Fig. 1. An Outline of the Steps of the Knowledge Discovery in Databases Process [5]

3. Data cleaning and preprocessing.
   - Removal of noise or outliers.
   - Collecting necessary information to model or account for noise.
   - Strategies for handling missing data fields.
   - Accounting for time sequence information and known changes.
4. Data reduction and projection.
   - Finding useful features to represent the data depending on the goal of the task.
   - Using dimensionality reduction or transformation methods to reduce the effective number of variables under consideration or to find invariant representations for the data.
5. Choosing the data mining task.
   - Deciding whether the goal of the KDD process is classification, regression, clustering, etc.
6. Choosing the data mining algorithm(s).
   - Selecting method(s) to be used for searching for patterns in the data.
   - Deciding which models and parameters may be appropriate.
   - Matching a particular data mining method with the overall criteria of the KDD process.

7. Data mining.
    - Searching for patterns of interest in a particular representational form or a set of such representations as classification rules or trees, regression, clustering, and so forth [5].
    - Mining data
    - Visualization of the mined data
8. Interpreting mined patterns.
9. Consolidating discovered knowledge.

2. Intelligent agent

An agent is a software entity which functions continuously and autonomously in a particular environment… able to carry out activities in a flexible and intelligent manner that is responsive to changes in the environment… ideally, an agent that function continuously in the environment over a long period of time would be able to learn from its experience… an agent that inhabits an environment with other agents and processes to be able to communicate and co-operate with them… perhaps move from place to place in doing so. There are almost as many definitions for the term Software (Intelligent) agent as there are people employing it.

The following are some examples

- "An agent is anything that can be viewed as perceiving its environment through sensors and acting on that environment through effectors" [12].
- "Autonomous agents are computational systems that inhabit some complex dynamic environment, sense and act autonomously in this environment and by doing so realize a set of goals or tasks for which they are designed" [13].
- "An intelligent agent is software that assists people and acts on the behalf. Intelligent agents work by allowing people to delegate work that they could have done to the agent software. Agents can, just as assistants can, automate repetitive task, remember things you forgot, intelligently summarize complex data, learn from you and even make recommendations to you" [14].
- An Agent is a piece of software that performs a given task using information gleaned from its environment to act in a suitable manner so as to complete the task successfully. The software should be able to adapt itself based on changes occurring in its environment, so that a change in circumstances will still yield the intended result" [15].
- Data mining is the process of analysing data from different perspectives and summarizing it into useful information. Data mining is a knowledge based systems, and information can be retrieved. Data mining is used to find correlation or patterns [11].
- Data mining and intelligent agents can make people to help in the decision making process, to elaborate decisional models and take good decisions in real time [11].

3. Objectives of the study

The study presented under this survey aims to analyse the following:

- To analyse the statistical value for asthma patients from the allergy patient details. Here we collect more than 200 patient values from the allergy and asthma health care centre.
    - Value of the FEV1 test used to divide the patients into two categories of allergy and allergy-induced asthma.
    - From the predicted value of asthma patients, we divide them into two categories depending upon the sex.
    - From the predicted value of asthma patients, we divide them into three categories depending upon the age for showing which category of patients affect more.
    - From the predicted value of asthma patients, we divide them into three categories depending upon the season for showing which season affect the patients more.

4. Literature review

The study presented here, mainly deals with the concept of Allergy-induced Asthma, as we all know, what is allergy and how it affects us.

- Allergic Asthma is caused by breathing in an allergen, which is a substance that can cause an allergic reaction. Airborne, year-round allergens, such as pet dander or dust mites are commonly called allergic triggers [6]. Allergic asthma is the most common type of asthma.
- There are approximately 26.5 million people in the US with asthma. About 60% (6 out of 10) of people with asthma have allergic asthma. For school-age children, this number is closer to 80%. There are 300 million asthmatics worldwide with 1/10th of those living in India. A recent review analysis of 15 epidemiological studies showed that the mean prevalence of asthma among children was 7.24%. The prevalence of childhood asthma has continued to increase in last 10 years on the Indian subcontinent [7].
- According to India's largest community of verified doctors Curofy in a poll of 1040 doctors, 82% said that there has been an increased asthma incidence in children due to increased environmental pollution [8].

5. Spirometry

- Spirometry helps diagnose all types of asthma including allergy induced, which is shown in figure 2. The test involves taking a deep breath and measuring the amount of air you can inhale and exhale as well as how fast you breathe in and out. Spirometry also helps your doctor determine the severity of your asthma and determine if

treatment is effective. FEV1 is one of the most common indices used to assess airway obstruction. It is automatically calculated during spirometry or pulmonary function testing. It is calculated using a spirometer.

6. How FEV1 is used in asthma treatment

Most commonly Forced Expiratory Volume will be ordered by your doctor as part of complete pulmonary function tests. Your doctor may do this to assess your symptoms before an asthma diagnosis has been made or monitor your asthma control as part of your asthma action plan. Symptoms such as the following may trigger your doctor to order these tests:

- Wheezing
- Chest tightness
- Cough
- Shortness of breath

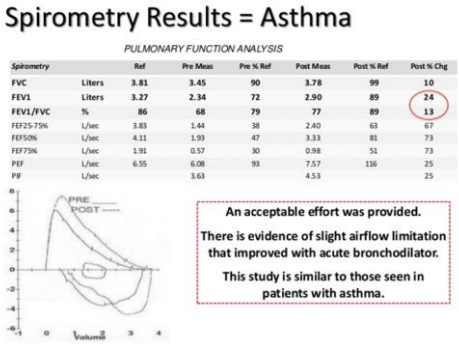

Fig. 2. Spirometry test values

7. Decreasing allergens

Preventing allergy-induced asthma flare-ups is one of the most vital aspects of treatment. Prevention involves identifying and reducing allergens. Depending on what allergen is triggering an asthma attack, there are several ways to decrease exposure [9]. Association rules are if-then statements that help to show the probability of relationships between data items within large data sets in various types of databases. Association rule mining has a number of applications and is widely used to help discover sales correlations in transactional data or in medical data sets. How association rules work

Association rule mining, at a basic level, involves the use of machine learning models to analyse data for patterns, or co-occurrence, in a database. It identifies frequent if-then associations, which are called association rules. An association rule has two parts: an antecedent (if) and a consequent (then). An antecedent is an item found within the data. A consequent is an item found in combination with the antecedent.

Association rules are created by searching data for frequent if-then patterns and using the criteria support and confidence to identify the most important relationships. Support is an indication of how frequently the items appear in the data. Confidence indicates the number of times the if-then statements are found true. A third metric, called lift, can be used to compare confidence with expected confidence. Association rules are calculated from item sets, which are made up of two or more items. If rules are built from analysing all the possible item sets, there could be so many rules that the rules hold little meaning. With that, association rules are typically created from rules well-represented in data.

8. The analytical study

The proposed study is designed to analyse the healthcare of patience history. The data presented in the study were collected to determine the challenges encountered and the extent to which the project goals were achieved. This study examined how the allergy turn into allergy – induced asthma. The carefully crafted mix of research method included data collected and observed test value. The findings of predicted value study attempts to reveal that purpose of how the people suffered by the asthma nowadays.

Implementing this in the field of computer science is the first of its kind. The concept of data mining is used to perform this analysis. WEKA3.8, one of the prominent tools of data mining is used for this study.it provides a collection of machine learning algorithms for data mining tasks. WEKA contains tools for data preprocessing, classification, association and clustering rules and visualization [10], [11]. To perform the analysis, the test data was collected from allergy and asthma centre. Nearly 200 patients test values are taken for this analysis. Out of this patients 40% male patients and 60%are female patients the analysis made among them are as follows.

9. Algorithms

Step 1. *Get test value of fev1*
Step 2. *Classify the data set into two of allergy patients and asthma patients using given predicted values (fev1)*
Step 3. *Separate the allergy patients and asthma patients using if statement*
Step 4. ***If*** *(fev1 chd>= 12)* ***then***
Step 5. *Add into asthma patient group* ***else*** *Add into allergy patient group*
Step 6. *Calculate how many patients affect by allergy and asthma*
Step 7. *Classify and calculate how many children and adult predict by allergy and asthma by using Age rule.*
Step 8. ***If*** *(age<14)* ***then*** *add into children group*
Step 9. ***If*** *(age> 14 && <40)* ***then*** *add into adult group*
Step 10. ***If*** *(age > 40)* ***then*** *add into middle age group*
Step 11. *Classify and calculate which gender patients predict by allergy and asthma using Sex rule.*
Step 12. If *(sex== male)* ***then*** *add into male group* ***else***
Step 13. *Add into female group*
Step 14. *Classify and calculate which season affects patients more than other season using Season rule.*
Step 15. *Make the result using data mining techniques like graph chart and clusters.*

10. Result and Discussion

To perform the analysis, the test data was collected from the Allergy and Asthma health care centre. Nearly 200 patient data are taken for this analysis. It contains 18 attributes of various data type. Some of the attributes are as follows: PATIENT ID, ETHNIC GROUP, SMOKNIG HISTORY, HEIGHT, WEIGHT, AGE, SEX, ADDRESS, OCCUPATION, AND SOME TEST VALUES etc.

Table 1
Separation of predicted values

Chg Analysis	Value	Result	Rule
1	6	Asthma	{chg>12, Female, age < 15}
2	43		{chg>12, Female, 14<age>40}
3	13		{chg>12, Female, age>39}
4	6		{chg>12, Male, age < 15}
5	27		{chg>12, Male, 14<age>40}
6	22		{chg>12, Male, age>39}
7	1	Allergic	{chg<12, Female, age < 15}
8	36		{chg<12, Female, 14<age>40}
9	21		{chg<12, Female, age>39}
10	5		{chg<12, Male, age < 15}
11	8		{chg<12, Male, 14<age>40}
12	12		{chg<12, Male, age>39}

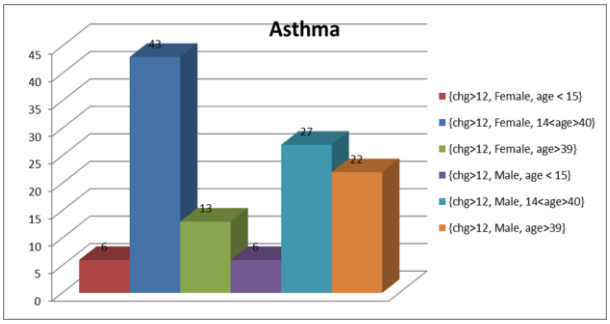

Fig. 3. Graph of Asthma patients

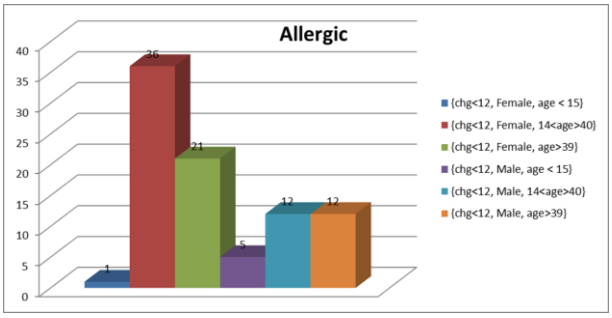

Fig. 4. Graph of Asthma patients

A. Based on sex rule

The major aim of this work is to predict the statistical value for asthma patients from the allergy patient details. Here we collect more than 200 patient values from the hospital and the value of the FEV1 test used to divide the patients into two categories of allergy and allergy-induced asthma. From the predicted value of asthma patients, we divide them into two categories depending upon the sex.

Table 2
Sex Distribution

Sex Distribution	Value	Rule
1	80	Female
2	120	Male

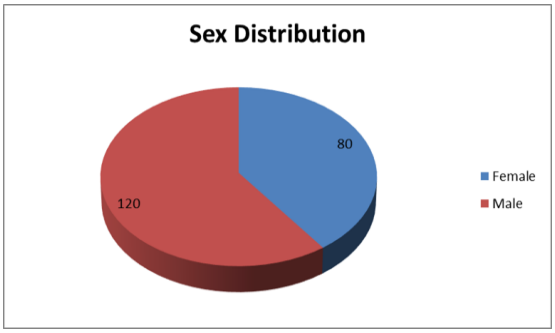

Fig. 5. Graph of Sex Distribution

B. Based on age rule

The major aim of this work is to predict the statistical value for asthma patients from the allergy patient details. Here we collect more than 200 patient values from the hospital and the value of the FEV1 test used to divide the patients into two categories of allergy and allergy-induced asthma. From the predicted value of asthma patients, we divide them into three categories depending upon the age for showing which category of patients affect more.

Table 3
Age Distribution

Age Distribution	Value	Rule
1	18	age < 15
2	114	14<age>40
3	68	age>39

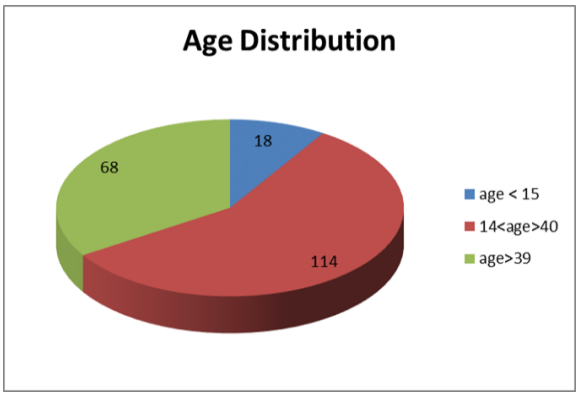

Fig. 6. Graph of Age Distribution

C. Based on season rule

The major aim of this work is to predict the statistical value for asthma patients from the allergy patient details. Here we collect more than 200 patient values from the hospital and the value of the FEV1 test is used to divide the patients into two categories of allergy and allergy-induced asthma. From the predicted value of asthma patients, we divide them into three categories depending upon the season for showing which

season affect the patients more.

Table 4
Season Distribution

Season Distribution	Value	Rule
1	109	Winter
2	59	Summer
3	32	Monsoon

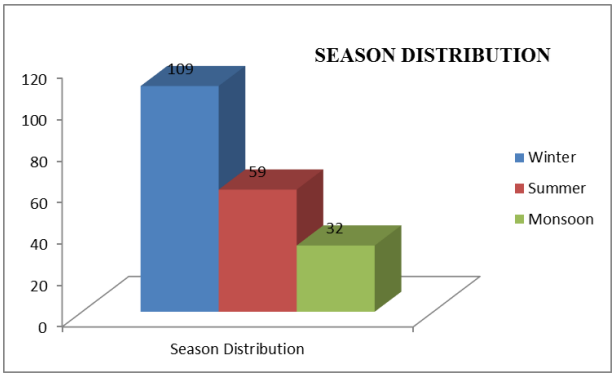

Fig. 7. Graph of Season Distribution

11. Conclusion

Data mining is the procedure of mining knowledge from data which can be widely used in the process of decision making. An agent is a computer system located in some environment, and is capable of autonomous action in this environment to meet its design objective. Multi agent systems are systems made up of multiple interacting agents. In this research work, the mining process is carried out by different intelligent agent with the help of the data collected. The results are analysed, evaluated and visualized by forming and detecting the clusters using used interface, data mining and visualization agents.

The major aim of this research work is to predict how many people are suffered from asthma and we classified them in to sex categories, age category and to find which season affect them lot. Therefore, we used 200 data of allergy affected patients, which were all collected from Allergy asthma healthcare centre. From these data we need to predict whether the patients are affected by asthma or only affected by allergy for down immune or climate changing etc. As per our survey we analysed that the issue of Allergy-induced asthma. The classification of Allergy and Asthma patients are done with using some rules like the patient age, sex and season. From this analysis, many people are affected on winter seasons, from sex rule many female patients are predicted with asthma and using age rule the between age 14 to 40 years old are predicted with asthma. From total number of 200 patients 117 people are predicted for asthma and remaining 83 people are suffered by only allergy.

12. Future enhancements

The study is further extend by building a data mining model using agent paradigm. Its scope for future work can also be extended in the part of gathering data by increasing the number of instances from different health care centre. Comparative study can also be conducted between different ethnic group patient history.

References

[1] Steven L. Tanimoto," Improving the prospects for Educational Data Mining". University of Washington.
[2] http://en.wikipedia.org/wiki/Data_mining.
[3] Atkotiya Kishore, and Chandra Hansrajbhai "Analtical study and computational modelling of statistical methods for data mining", Ph.D. thesis.
[4] Usama M. Fayyad, Gregory Piatetsky-Shapiro and Padhraic Smyth, "From data mining to knowledge discovery: an overview," Advances in knowledge discovery and data mining, American Association for Artificial Intelligence Menlo Park, CA, USA, 1996, pp. 1-34.
[5] http://www2.cs.uregina.ca/~dbd/cs831/notes/kdd/1_kdd.html
[6] ips-to-fight-breathlessness/articleshow/52088776.cms.
[7] https://www.news18.com/news/india/india-has-10-of-worlds-asthma-patients-survey-1238186.html
[8] https://www.thehealthsite.com/news/world-asthma-day-3rd-may-4-out-of-5-doctors-say-environmental-pollution-is-the-biggest-cause-of-asthma-ag0516-396386/
[9] https://www.pubfacts.com/detail/29520161/Knowledge-on-self-management-and-levels-of- asthma-control-among-adult-patients-in-Ho-Chi-Minh-City-V
[10] R. Jayabrabhu, V. Saravanan, and K. Vivekanandan, "A frame work: cluster detection and multidimensional visualization of automated data mining using intelligent agent", International Journal of Artificial Intelligence & Applications, vol. 3, January 2012.
[11] weka-x64.sharewarejunction.com
[12] J. Russell and P. Norvig, "Artificial intelligence: A modern Approach," upper saddle River, N. J. Prentice Hall, 1995.
[13] P. Maes, "Artificial Intelligence Meets entertainment: Life-Like Autonomous Agents," Communications of the ACM, November 1995.
[14] D. Gilbert, "Intelligent Agents: The Right Information at the Right Time," IBM white paper, May 1997.