iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** VII **Month of publication:** July 2020

DOI: https://doi.org/10.22214/ijraset.2020.30385

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

Association Rules to Analyze Hospital Resources with Mortality Rates

Tej Subbaiah[1], Y R Gangadhar[2], Yashas Kumar S[3], Kullachanda Ganapathi Karan Muddappa[4], Vinay M G[5]
[1, 2, 3, 4]*Student at VVCE,* [5]*Professor at VVCE, Computer Science Department, Vidyavardhaka College Of Engineering, at VVCE, Mysuru, India*

Abstract: ***As death rates expanding continuously consistently all finished, so wellbeing administration is the most significant undertaking to diminish the death rates. It is a test issue to give restorative information and current innovation for diminishing the mortality of the populace. The thought is to show a relationship among mortality and wellbeing administration assets by utilizing Apriori calculation. Good health services is a most important job to reduce the mortality rates. Proposed system finds the hidden correlations between hospital resources such as doctors of different departments, dentists, pharmacies, nurses, technical nurses, scanning departments and mortality rates. Proposed system is an medical sector application. It aims at reducing mortality rates in hospital. System discovers the association between health resources and mortality rates by using techniques of data mining to analyze health data.***
Keywords: Association rules, Apriori, Data mining

I. INTRODUCTION

Proposed system finds the hidden correlations between hospital resources such as doctors of different departments, dentists, pharmacies, nurses, technical nurses, scanning departments and mortality rates. Apriori is the most well known and efficient data mining method. System predicts the relationship between health resources and mortality rates using data mining technique "Association Rules". Its a real world computer application with an aim to predict the mortality rate.

The rules of association are written in the form of conditions and outcomes. Association rule functions in two steps:

1) Producing item sets that cross a minimum support threshold.
2) Producing rules that cross a minimum confidence threshold.

II. LITERATURE SURVEY

Nandita Rane, Madhuri Rao proposed a system for early diagnosis of diabetes. The plan is to find the segments that reason diabetes and their relations with diabetes. Beginning advance is to increment satisfactory diabetes information, which consolidates the patients' central condition , therapeutic history, and their family members' restorative history and evaluation results. This information is acquired from Nirmay Diabetes Super Speciality Center vault. In second step, use connection calculation which can find the incessant thing sets from enormous scale educational assortment. In the resulting stage, Apriori affiliation rule mining portray thing sets that fulfill a base help criterion. FP-development affiliation mining is utilized to create rules which are valuable to distinguish general relationship in the information.
The result set is of strong connection measures is made using object and emotional constraints. Object limitations actualized are supporting, lift and sureness and certainty.[1]
Gregory Boverman and Sahika Genc made use of the vast amount of the restorative knowledge obtained in the emergency unit has been used. Such collections of information can be used as valuable tools to build and authorize predictive analytics. They concentrate on the topic of mortality expectations from respiratory failure among long-term mechanically ventilated patients using information from the publicly available MIMIC-II database, in their study.
While merely announcing values for uni-variate or multivariate regression, they establish the sparsest possible model for mortality. They developed a model that predicts ICU respiratory distress mortality with a cross-validated region under the curve (AUC) of approximately 0.74.[2]
K.M.D. Muthumali Karunarathna has proposed a framework of condensed patient data for forecasting ICU deaths. This paper focuses mainly on predicting ICU mortality rates with clustered and unclustered data where the data set is derived from MIMIC III that focuses primarily on inputs such as salinity and output such as blood , urine and patient weight. With 9059 data points, this

analysis has examined six different variables to estimate the binary outcome variable of one meaning death and 0 meaning alive. Author used K-means algorithms and methods such as Vector Machine Support, linear discriminant and logistical regression.[3]

Monali Dey and Siddharth Swarup Rautaray have proposed a system for health care decision support system using data mining algorithm. The principle point of this paper is investigation of the uniqueness of restorative information mining and diagram of human services choice emotionally supportive networks right now utilized in medication and choosing and recognizing right information mining calculation. The fundamental strategies utilized in this overview paper are choice tree calculation and neural systems for deciding heart sicknesses, gullible Baye's calculation for deciding coronary heart maladies.[4]

Vinod L Mane , Suja S Panicker and Vidya B Patel proposed an architecture for summarization and sentiment analysis from the user post. There are 4 main phases in implementing this. The initial step is to extract the keyword from the user post. This helps in grouping up the posts , identifying the diseases and drugs. This brings us to the second phase where we use Apriori algorithm to associate the keywords. The next phase is to summarize the post about the drug family and the side effects on the patient. We use the Lesk algorithm and Wordnet dictionary in this phase . The final phase involves sentimental analysis of the patient which also helps the pharmaceutical companies wherein we classify the patient as depressed, satisfied and normal which is sent as feedback to the companies indirectly. [5]

S Leena is able to summarize the health posts of the user only because of the data available. Doctor mainly monitor the patients based on vital attributes like blood pressure, heart rate and so on. Similarly public health can be monitored by statistics. These statistics can be obtained by conducting census frequently giving the average health condition of the public. The statistics include right from the birth rate till the death rate and even the causes of the death. Most of the time the local agencies collect the data about the public health and government uses them in assessing the public need of health programs and evaluating the causes of death. These statistics are even used in health policy development and evaluations of the public health. One of the main uses of these statistics is that it helps the state and central government for budget planning , education on health and even in administrative decision making.[6]

Renata Ivancsy , Sandor Juhasz and Ferenc Kovacs present an intelligent model, Apriori calculation, which relies upon quantifiable parameters of the data dataset and on least help limit. The consistent parameters of the model can be perceived in modest number of test executions.

The presented model can be utilized not just for anticipating execution time of the Apriori calculation, yet in addition for foreseeing the reaction time of the other level_wise affiliation rule mining calculations. The model was affirmed by a couple of different datasets and the test outcomes exhibit that the general normal blunder pace of the model is under 15%.[7]

The paper proposed by Yan Yuguang, Wang Chunyan, li Min mostly talks about the Model Multi specification subject to apriori estimation in the Clinical Decision Support Program. Model multi is the multi dimensional Social Guideline study model. This is used to investigate the data from the existing database, with the goal that the center manager can handle the crisis facility by investigating information. On this model the rule applied is the Apriori-calculation of the connection rule. Get multidimensional affiliation administers by register visit predicate sets, has incredible flexibility to handle a lot of data, if the estimate is not too vast and the different characteristics are not too vast for each estimate, multidimensional display can achieve I / O dimensional predicate.[8]

III. IMPLEMENTATION

Examination of the relationship between hospital services and mortality is a critical activity in terms of data analysis for the development of policy for public health.

Good health services are a crucial role in raising the mortality rates. Program uses data analysis techniques to find associations between health services and mortality rates. Proposed program that will allow the medical departments to lower the mortality rates. The proposed program identifies secret associations between hospital services, such as physicians , dentists, pharmacy, nurses, professional nurses, scanning departments and mortality rates.

IV. APPROACH ADAPTED

The approach adopted for the development of this project is the Iterative Waterfall Model according to Software Engineering. The iterative waterfall model is a systematic approach that starts at the process of the feasibility study and progresses through research, design , coding , testing, integration and maintenance. Feedback paths are available in each step of the preceding process to allow the correction of later-phase errors committed during a step.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

V.SYSTEM DESIGN

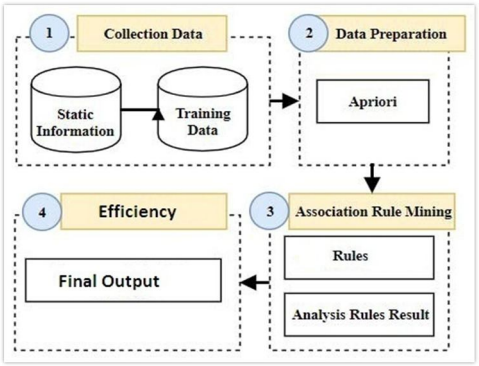


Figure 1. Flow Diagram Of System Architecture

A. Product Perspective

1) Administrator/service provider - the one who maintains the entire application.
2) Hospital(Member) - the one who receives the services.
3) Visitor - the one who visits the application.

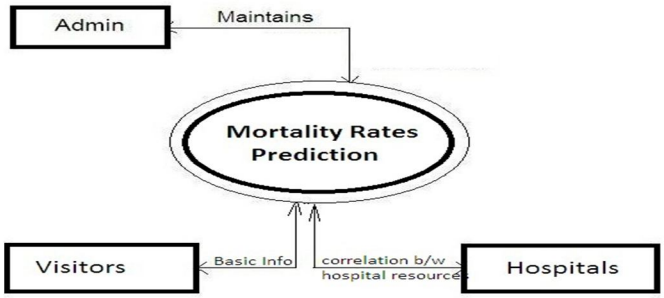


Figure 2. Context Flow Diagram

VI. INPUT AND OUTPUT

1) Input – Unique hospital services, such as surgeons, dentists, pharmacy, nurses, professional staff, scanning departments and previous year mortality rates.
2) Output - Uncovers the rules of association between medical care and mortality.
3) Platform : DOT NET -- VISUAL STUDIO 2008/2010

A. Hardware Requirements

1) Minimum of 2GB RAM
2) Pentium IV or higher
3) Minimum of 40GB HARD DISK
4) Standard PC configuration to carryout challenging computing.

B. Software Requirements

1) Operating System : Windows OS XP version, 7 and Higher
2) Design Tool: Visual Studio.
3) Language: C#.NET
4) Back End: SQL Server

C. Communications Interfaces

The HTTP protocol is used to promote client-server communications.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

VII. DATA PREPARATION

Different hospital resources such as doctors of different departments, dentists, pharmacies, nurses, technical nurses, cardiologists, neurosurgeons, etc. and previous years' mortality rates are chosen from the database and stored in excel sheet format, which is then imported by the hospital to determine the rules of association between medical resources and mortality rates.

	InchargerId	Date	Neurologist	Cardiologist	Gynecologist	Orthopedics	Surgeon	Physician	Beds	ICU	Nurses	MortalityRate
2	Hospital Ayush	30-01-2016 00:00	14	8	5	14	5	7	29	13	29	5
3	Hospital Ayush	30-01-2016 00:00	7	13	15	12	15	7	40	20	37	10
4	Hospital Ayush	30-01-2016 00:00	14	12	13	13	11	9	36	21	35	7
5	Hospital Ayush	30-01-2016 00:00	13	10	10	9	11	5	30	16	34	7
6	Hospital Ayush	30-01-2016 00:00	13	15	10	11	13	12	23	14	38	10
7	Hospital Ayush	30-01-2016 00:00	14	14	5	9	7	12	33	23	45	9
8	Hospital Ayush	30-01-2016 00:00	9	12	11	11	13	10	33	22	35	9
9	Hospital Ayush	30-01-2016 00:00	15	11	7	9	5	12	20	17	38	6

Figure 3 .Data Preparation table

VIII. METHODOLOGY

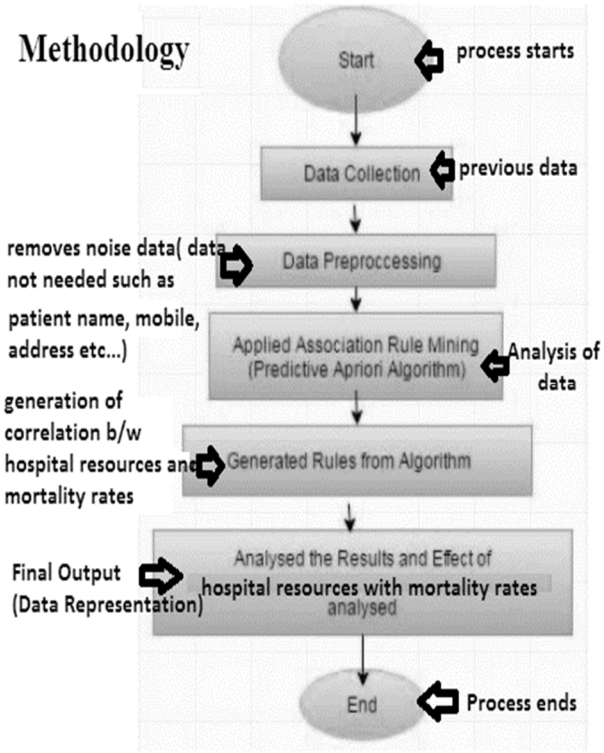


Figure 4. Methodology

A. Apriori Algorithm

It is one of the methods most common in data mining. This is used to locate two or more sets of data relationships for a trend that happens regularly within a broad data set and is used to evaluate the relationship or forecast, typically a transactional database. Apriori is an algorithm in the algorithm relationship, which decides frequent item sets in a given data. Apriori uses minimum support as a parameter, mining the data to produce the rules of association. Then developed rules are filtered using vilification techniques like "goodness metrics," trust and lifting combined with knowledge of the domain. The goals of this algorithm are to find fascinating patterns, inspiration and information from the data hidden in it and to attempt to explain some business sense. The rules of association are written as the form of conditions and outcomes, e.g. A---> B where conditions are A, and outcomes are B. The formula tells that B arises when A appears. Rules of Association are important if the values of support and confidence are greater than the thresholds set and can be obtained as follows.

Support of A = a / T

When the no. of transactions that A appears, T is the no. of transactions and confidence is the proportion of the transactions that contains A which also contains B.

Confidence (A->B) = Support (A ∪ B) / Support of A.

Steps required to generate the result are as follows

1) Check the collection of data and identify support(s) for each item.
2) Generate L1, Frequent item set 1.
3) Utilize Lk-1, join Lk-1 to produce the group of candidate k - itemset.
4) Check the candidate k item and generate support for every k-itemset of candidates.
5) Insert into frequent itemset, until C = Nullset.
6) For every element in the frequent itemset, all non-empty subsets are to be generated.
7) Find the confidence for every non-empty subset. If the confidence is >= to a stated confidence, attach to Strong Association rule.

IX. RESULT

The result shows the predictions by utilizing the association rule mining method in order to show which all hospital resources affect the mortality rates for a given hospital. The result will be as shown below:

Rule X	->	Rule Y	Confidence
Beds_Low	->	MortalityRate_High	52.38%
Beds_Medium,Cardiologist_Medium,Gynecologist_Medium	->	MortalityRate_High	51.11%
Beds_Medium,Gynecologist_Medium,Neurologist_Medium	->	MortalityRate_High	55.00%
Beds_Medium,Gynecologist_Medium,Orthopedics_High	->	MortalityRate_High	54.76%
Beds_Medium,Gynecologist_Medium,Physician_Medium	->	MortalityRate_High	51.02%
Beds_Medium,Neurologist_Medium,Orthopedics_High	->	MortalityRate_High	55.26%
Cardiologist_High,Orthopedics_High	->	MortalityRate_High	52.27%
Cardiologist_High,Surgeon_High	->	MortalityRate_High	51.02%
Cardiologist_Medium,Gynecologist_Medium	->	MortalityRate_High	50.91%
Cardiologist_Medium,Physician_High	->	MortalityRate_High	62.86%
Cardiologist_Medium,Surgeon_Medium	->	MortalityRate_High	52.17%
Gynecologist_Medium,Neurologist_Medium	->	MortalityRate_High	52.08%
Gynecologist_Medium,Nurses_High	->	MortalityRate_High	57.50%
Gynecologist_Medium,Orthopedics_High	->	MortalityRate_High	54.72%

Figure 5. Result table

X. CONCLUSION

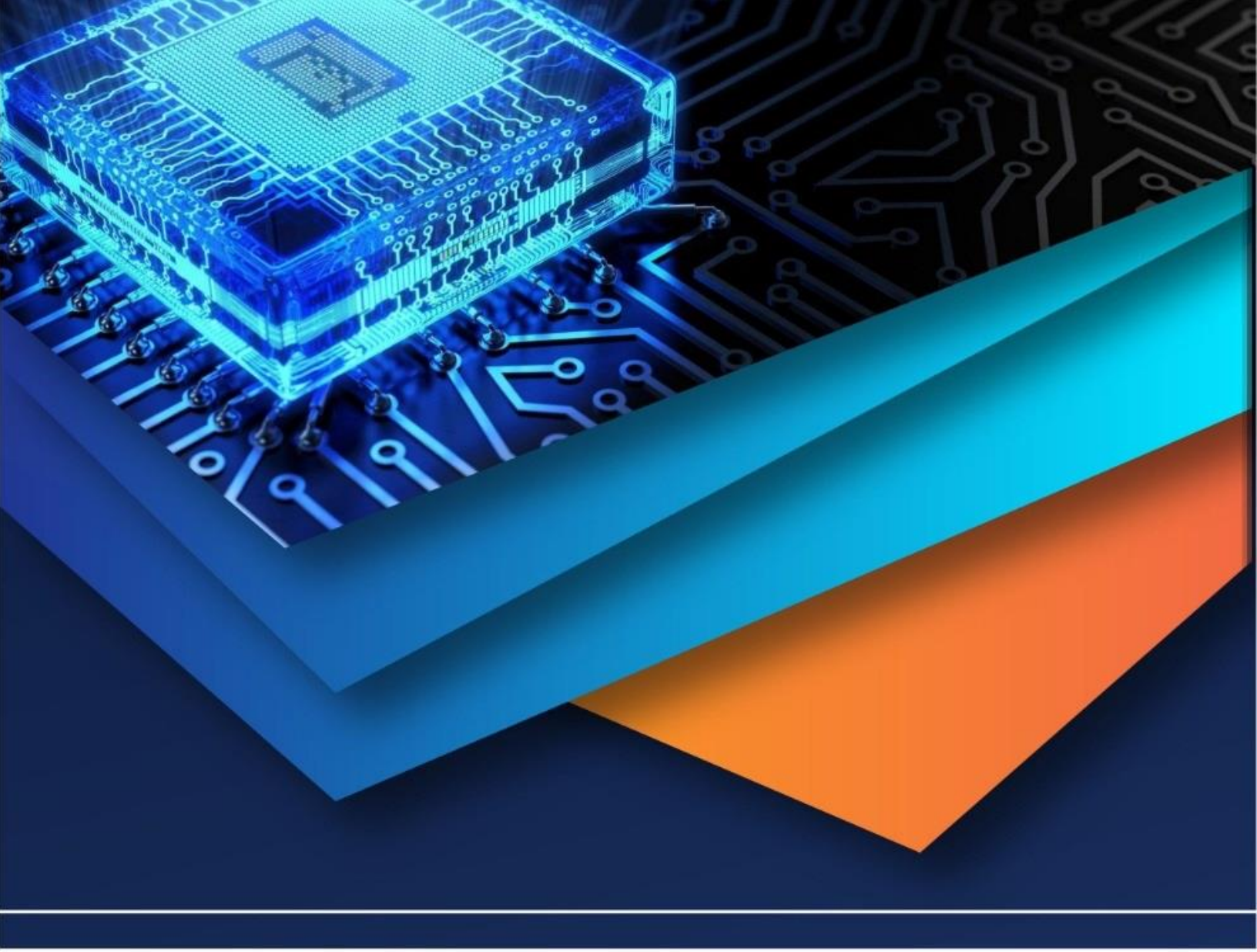
Reducing mortality rates is necessary in order to increase the economy of the country in a way . As data in a hospital is only for data storage purposes, the data present the databases are not analyzed, this has to be resolved. In order to resolve these problems, we can build systems which predict the mortality rates in the data given. These systems are built using various data mining techniques like apriori algorithm to find the correlation between hospital resources and mortality rates . In this paper we have discussed about data mining methods and how it is implemented in the systems and how it is used in predicting the mortality rates.

XI. ACKNOWLEDGEMENT

The authors are thankful to Dr. Ravi Kumar V, Head of the Department CS&E and Prof. Vinay M G, CS&E, Vidyavardhaka College of Engineering Mysuru, for providing required resources and infrastructure to conduct this project work and for their encouragement.

REFERENCES

[1] Nandita Rane and Madhuri Rao , “Association Rule Mining on Type 2 Diabetes using FP-growth association rule “.(2013)
[2] Gregory Boverman and Sahika Genc, “Prediction of Mortality from Respiratory Distress Among Long-Term Mechanically Ventilated Patients”.
[3] K.M.D Muthumali Karunarathna, “Predicting ICU Death With Summarized Patient Data”.
[4] Monali Dey and Siddharth Swarup Rautaray, “Study and Analysis of Data mining Algorithms for Healthcare Decision Support System”,(2014).
[5] Vinod L Mane , Suja S Panicker and Vidya B Patel “Summarization and sentiment analysis from user health posts”.(2015)
[6] S Leena,“ Role of statistics in public health”.
[7] Renata Ivancsy , Sandor Juhasz and Ferenc Kovacs , “Performance Prediction for Association Rule Mining Algorithms”.
[8] Yan Yuguang, Wang Chunyan, li Min “Application of the Model_Multi based on Apriori algorithm in Supporting System of Medical Decision”.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY