**ABSTRACT**

Machine learning is a method that requires analyzing and exploring data to glean meaningful trends and patterns. In today’s period, every person on earth relies on allopathic treatments and medicines. Machine learning techniques can be applied to medical datasets that have a vast scope of opportunity for textual as well as visual data. In medical services, there are myriad obscure data that needs to be scrutinized and data mining is the key to gain useful knowledge from these data. This paper provides an application programming interface to recommend drugs to users suffering from a particular disease which would also be diagnosed by the framework through analyzing the user's symptoms by the means of machine learning algorithms. We utilize some insightful information here related to mining procedure to figure out most precise sickness that can be related with symptoms. The patient can without much of a stretch recognize the diseases. The patients can undoubtedly recognize the disease by simply ascribing their issues and the application interface produces what malady the user might be tainted with. The framework will demonstrate complaisant in critical situations where the patient can't achieve a doctor's facility or when there are situations, when professional are accessible in the territory. Predictive analysis would be performed on the disease that would result in recommending drugs to the user by taking into account various features in the database. The experimental results can also be used in further research work and for Healthcare tools.

**1.INTRODUCTION**

Since the arrival of advanced computing, the doctors’ still requires the technology in various possible ways like surgical representation process and x-ray photography, but the technology perceptually stayed behind. The method still requires the doctor’s information and experience due to alternative factors starting from medical records to weather conditions, atmosphere, blood pressure and numerous alternative factors. The huge numbers of variables are consider as entire variables that are required to understand the complete working process itself, however no model has analyzed successfully. To tackle this drawback, Medical decision support systems must be used. This system is able to assist the doctors to make the correct decision. Medical decision support system refers to both the process of attempting to determine or identify possible diseases or disorder and the opinion reached by this process. The diagnostic opinion in the sense, it indicates either degree of abnormality on a continuum or a kind of abnormality in a classification. It’s influenced by non medical factors such as power ethics and financial incentives for patient or doctor. It can be a brief summation or an extensive formulation, even taking the form of story or metaphor. It might be a means of communication such as computer code through which it triggers payment, prescription, notification, information or advice. Indication of medical diagnostic includes knowledge of what is normal and measuring of patient’s current condition. Automated decision support systems are rule based systems that are automatically providing solutions to repetitive management problems.

Medical decision could be extremely specialized and difficult job due to alternative factors or incase of rare diseases. The alternative factors include stress; tired misdiagnosis might vary from ignorance of doctors and incomplete information. Standard algorithm may go through the entire variables like prevailing conditions history of medical records, history of family records and various factors relating to the patient records, sheer magnitude of obtainable hidden factors. Differential diagnosis methods can be used to identify the presence of an entity where multiple alternatives are possible and also refers to include the candidate alternatives. This method is needs a process of elimination or obtaining information that shrinks the probability of candidate conditions to negligible levels. It contains four steps: 1) The doctor gather all information about the patients and create a symptoms list.2) The doctor should make a list of all possible causes of symptoms.3) The doctor should prioritize the list by which is the most dangerous possible cause of symptoms put in the top of the list.4) The doctor should rule out or treat the possible causes beginning with the most urgently dangerous conditions.”Rule Out” in the sense to usethe test method or other scientific method. If there will be no such diagnosis means removing the diagnosis from the list and using tests that should have distinct results, depends on which diagnosis is correct. This can be done based on the doctor’s knowledge and experience. This method is very easy to implement.

To reduce the large number of variables and find the most probable diseases by using the K-Means algorithm. This algorithm is more suitable to cluster the more number of diseases. K-Mean is one of the unsupervised learning algorithms which are used to solve the clustering problem. The main idea is to determine the k centroids, one for each cluster. Different tests performed on the patients will served as a attributes for clustering. By using this algorithm it reduce the number of iterations, boundries of clusters are well define without overlapping, to produce the accurate result for each and every diagnosis. This system uses Service oriented architecture (SOA), anyone can access with internet connections and LAMSTAR Network can be used to calculate the weight, to increase the accuracy of algorithm, overall speed test and produce the better result.

* 1. **Problem statement:**

Since the arrival of advanced computing, the doctors’ still requires the technology in various possible ways like surgical representation process and x-ray photography, but the technology perceptually stayed behind. The method still requires the doctor’s information and experience due to alternative factors starting from medical records to weather conditions, atmosphere, blood pressure and numerous alternative factors. The huge numbers of variables are consider as entire variables that are required to understand the complete working process itself, however no model has analyzed successfully. To tackle this drawback, Medical decision support systems must be used. This system is able to assist the doctors to make the correct decision. Medical decision support system refers to both the process of attempting to determine or identify possible diseases or disorder and the opinion reached by this process.

## Motivation:

Medical decision could be extremely specialized and difficult job due to alternative factors or incase of rare diseases. The alternative factors include stress; tired misdiagnosis might vary from ignorance of doctors and incomplete information. Standard algorithm may go through the entire variables like prevailing conditions history of medical records, history of family records and various factors relating to the patient records, sheer magnitude of obtainable hidden factors. Differential diagnosis methods can be used to identify the presence of an entity where multiple alternatives are possible and also refers to include the candidate alternatives. This method is needs a process of elimination or obtaining information that shrinks the probability of candidate conditions to negligible levels.

**LITERATURE SURVEY**

# MobDBTest: A machine learning based system for predicting diabetes risk using mobile devices

**Abstract:** Diabetes mellitus (DM) is reaching possibly epidemic proportions in India. The degree of disease and destruction due to diabetes and its potential complications are enormous, and originated a significant health care burden on both households and society. The concerning factor is that diabetes is now being proven to be linked with a number of complications and to be occurring at a comparatively younger age in the country. In India, the migration of people from rural to urban areas and corresponding modification in lifestyle are all moving the degree of diabetes. Deficiency of knowledge about diabetes causes untimely death among the population at large. Therefore, acquiring a proficiency that should spread awareness about diabetes may affect the people in India. In this work, a mobile/android application based solution to overcome the deficiency of awareness about diabetes has been shown. The application uses novel machine learning techniques to predict diabetes levels for the users. At the same time, the system also provides knowledge about diabetes and some suggestions on the disease. A comparative analysis of four machine learning (ML) algorithms were performed. The Decision Tree (DT) classifier outperforms amongst the 4 ML algorithms. Hence, DT classifier is used to design the machinery for the mobile application for diabetes prediction using real world dataset collected from a reputed hospital in the Chhattisgarh state of India.

# Heart Disease Prediction System using Associative Classification and Genetic Algorithm

**Abstract:** Associative classification is a recent and rewarding technique which integrates association rule mining and classification to a model for prediction and achieves maximum accuracy. Associative classifiers are especially fit to applications where maximum accuracy is desired to a model for prediction. There are many domains such as medical where the maximum accuracy of the model is desired. Heart disease is a single largest cause of death in developed countries and one of the main contributors to disease burden in developing countries. Mortality data from the registrar general of India shows that heart disease are a major cause of death in India, and in Andhra Pradesh coronary heart disease cause about 30%of deaths in rural areas. Hence there is a need to develop a decision support system for predicting heart disease of a patient. In this paper we propose efficient associative classification algorithm using genetic approach for heart disease prediction. The main motivation for using genetic algorithm in the discovery of high level prediction rules is that the discovered rules are highly comprehensible, having high predictive accuracy and of high interestingness values. Experimental Results show that most of the classifier rules help in the best prediction of heart disease which even helps doctors in their diagnosis decisions.

# Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques

**Abstract:** Cancer is the most important cause of death for both men and women. The early detection of cancer can be helpful in curing the disease completely. So the requirement of techniques to detect the occurrence of cancer nodule in early stage is increasing. A disease that is commonly misdiagnosed is lung cancer. Earlier diagnosis of Lung Cancer saves enormous lives, failing which may lead to other severe problems causing sudden fatal end. Its cure rate and prediction depends mainly on the early detection and diagnosis of the disease. One of the most common forms of medical malpractices globally is an error in diagnosis. Knowledge discovery and data mining have found numerous applications in business and scientific domain. Valuable knowledge can be discovered from application of data mining techniques in healthcare system. In this study, we briefly examine the potential use of classification based data mining techniques such as Rule based, Decision tree, Naïve Bayes and Artificial Neural Network to massive volume of healthcare data. The healthcare industry collects huge amounts of healthcare data which, unfortunately, are not “mined” to discover hidden information. For data preprocessing and effective decision making One Dependency Augmented Naïve Bayes classifier (ODANB) and naive creedal classifier 2 (NCC2) are used. This is an extension of naïve Bayes to imprecise probabilities that aims at delivering robust classifications also when dealing with small or incomplete data sets. Discovery of hidden patterns and relationships often goes unexploited. Diagnosis of Lung Cancer Disease can answer complex “what if” queries which traditional decision support systems cannot. Using generic lung cancer symptoms such as age, sex, Wheezing, Shortness of breath, Pain in shoulder, chest, arm, it can predict the likelihood of patients getting a lung cancer disease. Aim of the paper is to propose a model for early detection and correct diagnosis of the disease which will help the doctor in saving the life of the patient.

# Data Mining Techniques to Predict Chronic Kidney Disease

**Abstract:** Chronic Kidney Disease incorporates the state where the kidneys fail to function and reduce the potential to keep a person suffering from the disease healthy. When the condition of the kidneys gets worse, the wastes in the blood are formed in high level. Data mining has been a present pattern for accomplishing analytic outcomes. Colossal measure of un-mined data is gathered by the human services industry so as to find concealed data for powerful analysis and basic leadership. Data mining is the way towards extricating concealed data from gigantic datasets. The goal of our paper is to anticipate CKD utilizing the classification strategy Naïve Bayes. The phases of CKD are anticipated in the light of Glomerular Filtration Rate (GFR). Chronic Kidney Disease (CKD) is one of the most widespread illnesses in the United States. Recent statistics show that twenty-six million adults in the United States have CKD and million others are at increased risk. Clinical diagnosis of CKD is based on blood and urine tests as well as removing a sample of kidney tissue for testing. Early diagnosis and detection of kidney disease is important to help stop the progression to kidney failure. Data mining and analytics techniques can be used for predicting CKD by utilizing historical patient’s data and diagnosis records. In this research, predictive analytics techniques such as Decision Trees, Logistic Regression, Naive Bayes, and Artificial Neural Networks are used for predicting CKD. Pre-processing of the data is performed to impute any missing data and identify the variables that should be considered in the prediction models. The different predictive analytics models are assessed and compared based on accuracy of prediction. The study provides a decision support tool that can help in the diagnosis of CKD.

**SYSTEM ANALYSIS**

**EXISTING SYSTEM:**

* Medical decision could be extremely specialized and difficult job due to alternative factors or incase of rare diseases. The alternative factors include stress; tired misdiagnosis might vary from ignorance of doctors and incomplete information. Standard algorithm may go through the entire variables like prevailing conditions history of medical records, history of family records and various factors relating to the patient records, sheer magnitude of obtainable hidden factors. Differential diagnosis methods can be used to identify the presence of an entity where multiple alternatives are possible and also refers to include the candidate alternatives.

**DISADVANTAGES OF EXISTING SYSTEM:**

This method is needs a process of elimination or obtaining information that shrinks the probability of candidate conditions to negligible levels.

It contains four steps:

1) The doctor gather all information about the patients and create a symptoms list.

2) The doctor should make a list of all possible causes of symptoms.

3) The doctor should prioritize the list by which is the most dangerous possible cause of symptoms put in the top of the list.

4) The doctor should rule out or treat the possible causes beginning with the most urgently dangerous conditions.”Rule Out” in the sense to use the test method or other scientific method. If there will be no such diagnosis means removing the diagnosis from the list and using tests that should have distinct results, depends on which diagnosis is correct. This can be done based on the doctor’s knowledge and experience. This method is very easy to implement.

# 3.1Proposed System:

Employing the XGBoost, AdaBoost, MLP classifier, catboost, SVM, and KNN algorithms to decrease the enormous number of variables and identify the diseases that are most likely to occur. These algorithms are more suited to grouping more disorders. One unsupervised learning approach used to address the clustering issue is SVM. Different tests carried out on the patients will be used as clustering attributes.

# Advantages of proposed system:

1. These techniques decrease the amount of iterations and provide well defined, non-overlapping cluster boundaries.
2. The ability to accurately diagnose each and every case.

**SYSTEM REQUIREMENTS:**

**HARDWARE REQUIREMENTS:**

* System : Pentium IV 2.4 GHz.
* Hard Disk : 40 GB.
* Floppy Drive : 1.44 Mb.
* Monitor : 15 VGA Colour.
* Mouse : Logitech.
* Ram : 512 Mb.

**SOFTWARE REQUIREMENTS:**

* **Operating System:** Windows
* **Coding Language**: Python 3.7

**SYSTEM STUDY**

### **FUNCTIONAL REQUIREMENTS**

1. Data Collection

2. Data Preprocessing

3. Training And Testing

4. Modiling

5. Predicting

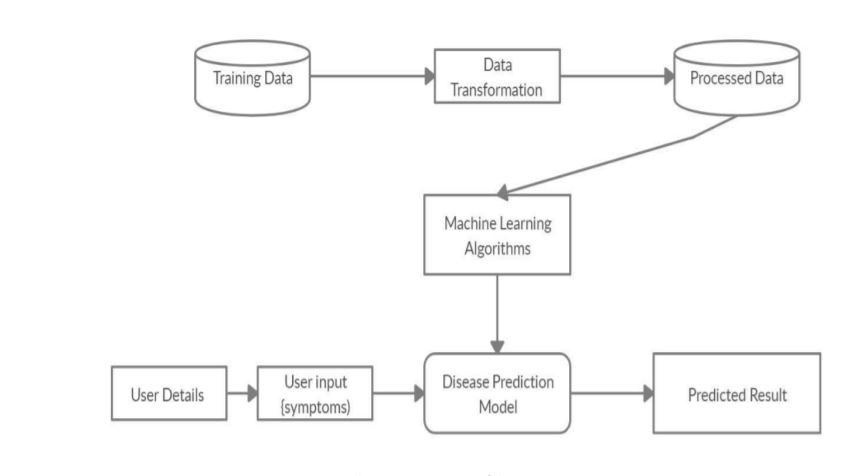
### **NON FUNCTIONAL REQUIREMENTS**

NON-FUNCTIONAL REQUIREMENT (NFR) specifies the quality attribute of a software system. They judge the software system based on Responsiveness, Usability, Security, Portability and other non-functional standards that are critical to the success of the software system. Example of nonfunctional requirement, *“how fast does the website load?”* Failing to meet non-functional requirements can result in systems that fail to satisfy user needs. Non- functional Requirements allows you to impose constraints or restrictions on the design of the system across the various agile backlogs. Example, the site should load in 3 seconds when the number of simultaneous users are > 10000. Description of non-functional requirements is just as critical as a functional requirement.

* Usability requirement
* Serviceability requirement
* Manageability requirement
* Recoverability requirement
* Security requirement
* Data Integrity requirement
* Capacity requirement
* Availability requirement
* Scalability requirement
* Interoperability requirement
* Reliability requirement
* Maintainability requirement
* Regulatory requirement
* Environmental requirement

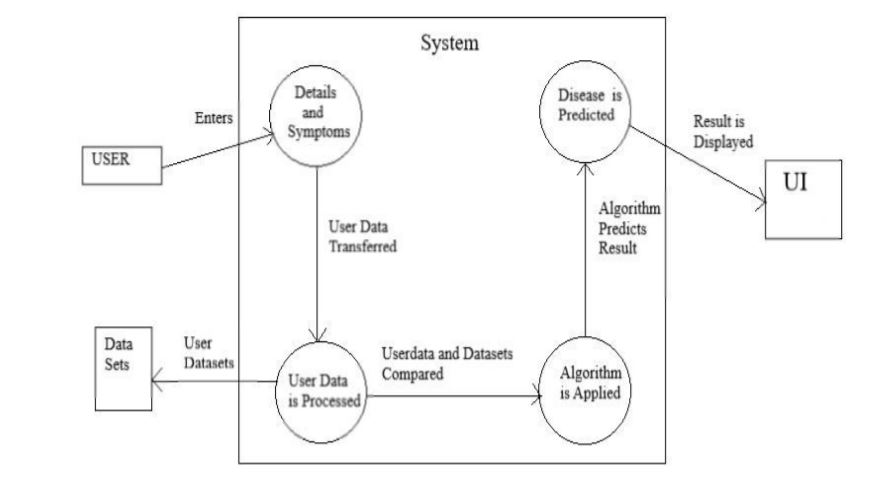
**SYSTEM DESIGN**

**SYSTEM ARCHITECTURE:**

****

**Flow Chart:**

A flowchart is a type of diagram that represents a workflow or process. A flowchart can also be defined as a diagrammatic representation of an algorithm, a step-by-step approach to solving a task. The flowchart shows the steps as boxes of various kinds, and their order by connecting the boxes with arrows.

****

**DATA FLOW DIAGRAM:**

1. The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.
2. The data flow diagram (DFD) is one of the most important modeling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.
3. DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.
4. DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

**User**

**Check**

**Unauthorized user**

**Yes NO**

**Select Symptoms**

**Build Xgboost**

Build Adaboost

Build MLP classifier

Build Catboost

Build SVM

Build KNN

Drug Prediction

**End process**

**4.3 UML DIAGRAMS**

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modeling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems.

The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

**Use case diagram:**

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.



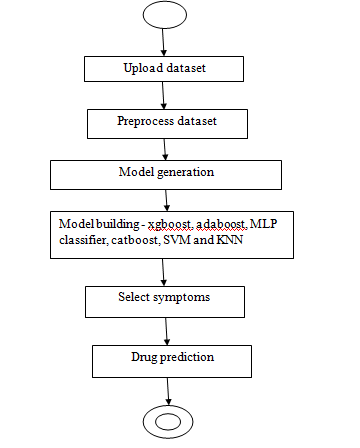
**Class diagram:**

The class diagram is used to refine the use case diagram and define a detailed design of the system. The class diagram classifies the actors defined in the use case diagram into a set of interrelated classes. The relationship or association between the classes can be either an "is-a" or "has-a" relationship. Each class in the class diagram may be capable of providing certain functionalities. These functionalities provided by the class are termed "methods" of the class. Apart from this, each class may have certain "attributes" that uniquely identify the class.

****

**Activity diagram:**

The process flows in the system are captured in the activity diagram. Similar to a state diagram, an activity diagram also consists of activities, actions, transitions, initial and final states, and guard conditions.

****

**Sequence diagram:**

A sequence diagram represents the interaction between different objects in the system. The important aspect of a sequence diagram is that it is time-ordered. This means that the exact sequence of the interactions between the objects is represented step by step. Different objects in the sequence diagram interact with each other by passing "messages".



**Collaboration diagram:**

A collaboration diagram groups together the interactions between different objects. The interactions are listed as numbered interactions that help to trace the sequence of the interactions. The collaboration diagram helps to identify all the possible interactions that each object has with other objects.

****

**Component diagram:**

The component diagram represents the high-level parts that make up the system. This diagram depicts, at a high level, what components form part of the system and how they are interrelated. A component diagram depicts the components culled after the system has undergone the development or construction phase.



**Deployment diagram:**

The deployment diagram captures the configuration of the runtime elements of the application. This diagram is by far most useful when a system is built and ready to be deployed.

****

**4.3 IMPLEMENTATION:**

**MODULES:**

* Select symptoms: Using this module enter the Name of patient and enter the symptoms of the patient to prediction the disease.
* Xgboost: Using this module given symptoms it predicted with algorithm
* Adaboost: Using this module given symptoms it predicted with algorithm
* MLP classifier: Using this module given symptoms it predicted with algorithm
* Catboost: Using this module given symptoms it predicted with algorithm
* SVM: Using this module given symptoms it predicted with algorithm
* KNN: Using this module given symptoms it predicted with algorithm
* Drug Prediction: Using this module models was predicted the Drug for the disease.

**Algorithm:**

* Xgboost: The XGBoost (eXtreme Gradient Boosting) is a popular and efficient open-source implementation of the gradient boosted trees algorithm. Gradient boosting is a supervised learning algorithm that attempts to accurately predict a target variable by combining an ensemble of estimates from a set of simpler and weaker models.
* Adaboost: AdaBoost algorithm, short for Adaptive Boosting, is a Boosting technique used as an Ensemble Method in Machine Learning. It is called Adaptive Boosting as the weights are re-assigned to each instance, with higher weights assigned to incorrectly classified instances.
* MLP classifier: MLPClassifier stands for Multi-layer Perceptron classifier which in the name itself connects to a Neural Network. Unlike other classification algorithms such as Support Vectors or Naive Bayes Classifier, MLPClassifier relies on an underlying Neural Network to perform the task of classification.
* Catboost: CatBoost is an algorithm for gradient boosting on decision trees. It is developed by Yandex researchers and engineers, and is used for search, recommendation systems, personal assistant, self-driving cars, weather prediction and many other tasks at Yandex and in other companies, including CERN, Cloudflare, Careem taxi.
* SVM: Support vector machines (SVMs) are a set of supervised learning methods used for classification, regression and outliers detection. The advantages of support vector machines are: Effective in high dimensional spaces. Still effective in cases where number of dimensions is greater than the number of samples.
* KNN: K Nearest Neighbor algorithm falls under the Supervised Learning category and is used for classification (most commonly) and regression. It is a versatile algorithm also used for imputing missing values and resampling datasets.

**7.SCREENSHOTS**

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application, Teams

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

**CONCLUSION**

This paper gave a diagram of utilization of information machine learning procedures in regulatory, clinical, inquire about, furthermore, instructive parts of Clinical Predictions. This paper set up that while the current down to earth utilization of information machine learning in wellbeing related issues is constrained, there exists an extraordinary potential for information mining systems to enhance different parts of Clinical Predictions. Besides, the inescapable ascent of clinical information will build the potential for information mining systems that enhances the quality and reduces cost of social insurance.

**Future work:**

This system has large scope as it has the following features which are: • Automation of Disease Diagnosis. • Paper free work helping the environment. • To increase the efficiency, accuracy for the patients to help them in future. • Managing the information related to diseases.

**9. REFERENCES**

1. Mário W. L. Moreira, Joel J. P. C. Rodrigues, Antonio M. B. Oliveira, Ronaldo F. Ramos, Kashif Saleem, "A preeclampsia diagnosis approach using Bayesian networks", Communications (ICC) 2016 IEEE International Conference on, pp. 1-5, 2016, ISSN 1938-1883.

2. Ameera M. Almasoud, Hend S. Al-Khalifa, Abdulmalik Al-Salman, "Recent developments in data mining applications and techniques", Digital Information Management (ICDIM) 2015 Tenth International Conference on, pp. 36-42, 2015

3. K Sowjanya, Ayush Singhal, Chaitali Choudhary, "MobDBTest: A machine learning based system for predicting diabetes risk using mobile devices", Advance Computing Conference (IACC) 2015 IEEE International, pp. 397-402, 2015.

4. Bharathan Venkatesh, Danasingh Asir Antony Gnana Singh, Epiphany Jebamalar Leavline, Advance in Intelligent Systems and Computing, vol. 517, pp. 633, 2017, ISSN 2194-5357, ISBN 978- 981-10-3173-1.

5. Felix Gräßer, Surya Kallumadi, Hagen Malberg, and Sebastian Zaunseder. 2018. Aspect-Based Sentiment Analysis of Drug Reviews Applying Cross-Domain and Cross-Data Learning. In Proceedings of the 2018 International Conference on Digital Health (DH '18). ACM, New York, NY, USA, 121-125. 6. Effieient Algorithms to find Frequent Itemset Using Data MiningSagar Bhise1, Prof. Sweta Kale International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056.

7. Text Data Mining of Care Life Log by the Level of Care Required Using KeyGraph Muneo Kushima, Kenji Araki, Tomoyoshi Yamazaki, Sanae Araki, Taisuke Ogawa, Noboru Sonehara Proceedings of the International MultiConference of Engineers and Computer Scientists 2017 Vol I, IMECS 2017, March 15 - 17, 2017, Hong Kong.

8. Chen, W., Guo, H., Zhang, F., Pu, X., and Liu, X. (2012). Mining Schema Matching Between Heterogeneous Databases. In Consumer Electronics, Communications and Networks (CECNet), 2012 2nd International Conference on, pages 1128–1131. IEEE.

9. Han, J. and Gao, J. (2009). Research Challenges for Data Mining in Science and Engineering. Next Generation of Data Mining, pages 1– 18.

10. H. Abe, S. Hirano and S. Tsumoto, “Evaluation Temporal Models Based on Association Mining From Medical Documents,” Japan Association for Medical Informatics, vol. 27, no. 1, pp. 33-38, 2007.

11. Sun, Y., Han, J., Yan, X., and Yu, P. S. (2012). Mining Knowledge from Interconnected Data: A Heterogeneous Information Network Analysis Approach. Proceedings of the VLDB Endowment.

12. M. Akhil jabbar & Dr. Priti Chandrab “Heart Disease Prediction System using Associative Classification and Genetic Algorithm” International Conference on Emerging Trends in Electrical, Electronics and Communication Technologies-ICECIT, 2012.

13. Nikhil N. Salvithal“ Appraisal Management System using Data mining “International Journal of Computer Applications (0975 – 8887) Volume 135 – No.12, February 2016.

14. Tanvi Sharma, Anand Sharma & Vibhakar Mansotra “Performance Analysis of Data Mining Classification Techniques on Public Health Care Data” International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 6, June 2016.

15. V.Krishnaiah , Dr.G.Narsimha, Dr.N.Subhash Chandra “Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques”(IJCSIT)International Journal of Computer Science and Information Technologies, Vol. 4 (1) , 2013, 39 - 45.

16. Jaimini Majali, Rishikesh & Niranjan, Vinamra Phatak “Data Mining Techniques For Diagnosis And Prognosis Of Cancer” International Journal of Advanced Research in Computer and CommunicationEngineering Vol. 4, Issue 3, March 2015.

17. Data Mining Techniques to Predict Diabetes Influenced Kidney Disease Swaroopa Shastri, Surekha, SaritaInternational Journal of Scientific Research in Computer Science, Engineering and Information Technology © 2017 IJSRCSEIT | Volume 2 | Issue 4 | ISSN : 2456-3307.

18. DeFronzo Ralph. From the triumvirate to the ominous octet: a new paradigm for the treatment of type 2 diabetes mellitus. Diabetes. 2009;58:773-95.