Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7765892

DATA MANAGEMENT TECHNIQUES IN HEALTHCARE SYSTEM- A SURVEY

[1]Dr. Manish Bhardwaj, [2]Supriya Dubey, [3]Maneesh Pant, [4]Anurag Singh, [5]Vineet Chauhan

[1,2]Department of Computer Science and Information Technology, KIET Group of Institutions, Delhi-NCR, Ghaziabad, India

[3]Department of Computer Science, ABES Engineering College, Ghaziabad, India

[4]Department of Electrical and Electronic Engineering, G L Bajaj Institute of Technology and Management, Greater Noida, India

[5]Department of Computer Science and Engineering, COER University, Roorkee, India

Abstract

Research sectors across the board are adopting data mining as a means of extracting useful information from large amounts of data. It is one of the best ways to predict what will happen in the near future because of the current developments in the world. Massive amounts of data are now available as a result of medical advances, but the real challenge is in turning this data into practical applications. Data mining is the greatest way to overcome this obstacle. It's possible that data mining can help healthcare organizations use data more effectively and efficiently. As a result, it enhances care while also lowering expenses. Classification, grouping, association, and regression are some of the Data Mining techniques discussed in this work. In addition, it discusses healthcare data mining's uses, difficulties, and next endeavors.

Keywords: Healthcare, Data Mining, Decision Tree, Prediction Set, Classification.

Introduction

Storage of data or information was expensive in the early 1970s. However, in the last twenty-five years, we have seen a vast volume of information or data available in electronic format. Increasing database sizes are necessary to accommodate such massive data sets. In such databases, you can find a wealth of important data. If you're involved in any type of decision-making process, this data could be really helpful. Data mining or Knowledge Discovery in Databases (KDD) makes this possible (KDD). Data mining is the process of discovering previously unknown information in a huge collection of data. An abundance of correlations can be found in such a vast collection of data such as a correlation between patient data and length of stay [1]. There are five stages of the knowledge discovery process depicted in figure 1.

This stage begins and finishes with knowledge that was gleaned from the subsequent phases as depicted in figure 1 using raw data:

- Selection

This is the stage in which the data is sorted out based on a set of criteria. Subsets of data can be determined, for example, if a bicycle is owned by all those people [2].

- Pre-processing

During this stage, any information that isn't absolutely necessary is purged, such as a patient's sex, when doing a pregnancy test. Stage of data purification is another name for it.

- Transformation

Only data that is relevant to a specific research, such as demographic data, is valuable in market research at this point.

- Analysis of large amounts of data

An important part of data mining is discovering new information. In this stage, the data can be mined for patterns that have meaning.

- a process of interpreting and assessing

At this point, the system's discerned patterns are processed as new pieces of information. Making good selections will be easier if you have this information.

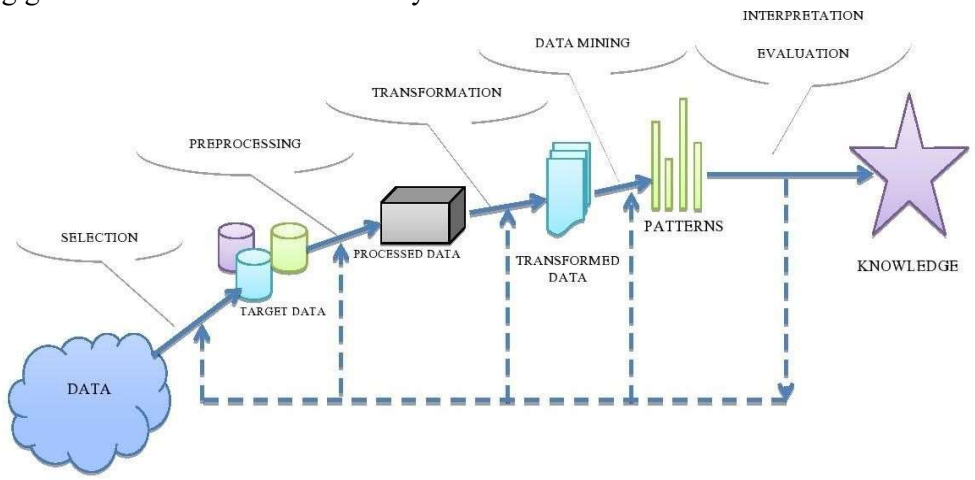

Figure 1: The Process of Discovering New Information Has Several Stages

Medical Data Mining's Importance

Healthcare data was generally kept electronically by all healthcare organizations worldwide. The majority of healthcare data consists of information about patients and those associated in the healthcare industry [3]. The amount of data that can be stored in this manner is increasing at a rapid pace. Complications arise from the ever-increasing volume of electronic healthcare data. Data in the healthcare industry is becoming increasingly difficult to analyze. It is quite difficult to extract significant information from it using typical approaches. However, thanks to advances in statistics, mathematics, and a variety of other fields, it is now possible to identify meaningful patterns. In a circumstance when a huge amount of healthcare data is available, data mining is useful [4].

Data mining mostly focuses on discovering previously unknown patterns. These patterns can then be assimilated into existing information, which can subsequently be used to make critical decisions. Data mining offers a slew of advantages [5]. Detecting fraud and abuse, providing superior medical treatments at a reasonable cost, spotting diseases in their earliest stages, and providing intelligent healthcare decision support systems are just a few of its many benefits. The healthcare industry can benefit greatly from data mining approaches [6]. Patients receive better care, and healthcare organizations benefit from their assistance in making various management decisions related to medical care. Health care data mining techniques offer a wide range of services, including: improved effective treatments, fraud insurance claims by patients as well as by providers and readmissions [7]; better treatment methods for a specific group; and successful drug recommendation systems.

Research is tremendously affected by the power of data mining for all of these reasons. Data mining is a common practice in the healthcare industry. Data mining can be done in a variety of ways [8]. Classification, grouping, regression, and the like are a few examples. Health care providers and patients alike can benefit from all medical information. Data mining, a powerful tool in the healthcare industry, is critical to its success [9]. Distributed medical environments are increasingly being used by researchers in order to give better medical services at a lower cost, better customer relationship management, better healthcare resource management, and more. It provides valuable healthcare data that may be used by management to make important decisions, such as estimating medical staff, deciding on health insurance policies, selecting therapies, and predicting illness progression, among others [10-13]. Data mining in healthcare: difficulties and challenges. Effective data mining technique is utilized to predict several diseases. Methods for bettering the healthcare system were offered, including a data mining approach, a new data mining approach, and an overall structure to guide future research.

Data mining techniques

Classification

Data mining techniques such as classification are widely employed in the healthcare industry. It's a way of categorizing the info that's collected. Each data point in the dataset is assigned to one of several classes by the classification algorithm. A risk factor can be linked to a patient's illness patterns using a categorization approach. It's a form of guided learning that relies on pre-existing notions of class. Classification can be done in two ways: binary or multilayer [14]. Patients can be classified as "High" or "Low" risk in a binary classification, however a multiclass approach considers patients in three different risk categories: "High," "Medium," and "Low."

There are separate training and testing datasets in the data set. Predicting a certain outcome based on a given input is the goal of this method [15]. A training set is a set of attributes used in the algorithm that predicts the outcome. Attribute relationships are examined in an attempt at predicting outcomes. Its outcome is its goal or projection. Prediction set is another algorithm. There are exactly the same attributes as the training set. However, in the prediction set, the prediction characteristic is still a mystery. The algorithm's accuracy is the most important factor in processing the forecast. Pawti Medical Center's training set contains all of the patient information that has been previously stored in the database [16-19]. The predictive factor here is whether or not a patient has a heart condition. Table 1 below shows examples of training sets from a database like this.

Table 1: Sets for PAWTI's Medical Data Base Training and Prediction

Training set

Age	Blood Pressure	Heart Rate	Heart Health Isuue
39	104/54	63	No
27	110/60	84	No
44	140/65	73	Yes

Prediction Set

Age	Blood Pressure	Heart Rate	Heart Health Isuue
34	141/81	87	To be calculated
44	109/57	53	To be calculated
85	140/65	82	To be calculated

Classification predicts rules for disentangling knowledge. IF-THEN rules are used to communicate predictions. The following is an example of a rule that predicts the first row of the training set:
(Age=45 AND Heart Rate>75) OR (Age>44 AND Blood Pressure>139/60) THEN Heart Problem=YES; otherwise, no.
Following is a list of healthcare classification algorithms:

K-Nearest Neighbor (K-NN)

According to K-Nearest Neighbor (K-NN) classifier is a simple classifier that uses a voting system to identify an unidentified dataset by comparing known datasets to previously recognized datasets. Consider a variety of items. The properties of one of the objects can help us forecast the characteristics of its closest neighbors, and this is because the characteristics of nearby objects are comparable. Any positive integer k, in the form of a majority vote in K-NN, can play a significant role in classifying any new occurrence (small number). It's one of the simplest data mining methods there is. Memory-based classification is commonly referred to as such since the training examples must always be in memory at run time [20]. In the case of continuous characteristics, the Euclidean distance is calculated by taking the difference between the attributes. However, when huge values drag down tiny ones, it has a very serious problem. By normalizing continuous qualities so that they have the same influence on the distance measure between distances, we may take over this important challenge.
There are several uses for K-NN, from health datasets to image analysis to cluster analysis to pattern identification to online marketing, to name a few. KNN classifiers have a number of advantages. These include: ease of use, effectiveness, intuitiveness, and competitive classification performance across a wide range of fields. It is effective and resilient to noisy training data if the training data is substantial. KNN classifiers have a major drawback in that they require a considerable amount of memory to hold the entire sample. A sequential computer will take longer to respond to a high sample size.

Decision Tree (DT)

One of the most prevalent ways to express classifiers is with DTs. We can use accessible data to build a decision tree that can solve problems in a variety of research areas. In a flowchart, every non-leaf node signifies a test on an attribute, and every branch denotes an outcome of that test, and every leaf node has a class name for that attribute. A decision tree's root node is the most important node. Medical readmission decision trees assist us determine if a patient needs to be readmitted or not, as an example. Building a solution to any problem does not necessitate subject knowledge. Operational research analysis is the most prevalent application of Decision Trees. With Decision Trees, it is possible for decision makers to select the optimal option, and traversal from root to leaf indicates a distinct class separation based on maximal

information acquisition. Many academics in the healthcare industry employ decision trees [21]. The following are a few benefits of using a decision tree:When condensed, decision trees are simple to follow and straightforward to understand. Decision trees can be used to create rules of their own. To express any discrete-value classifier, a decision tree representation is essential since it can deal with nominal and numeric input attributes of any type. Decision trees can readily deal with datasets that contain missing or incorrect values. As a result, decision trees might be regarded as nonparametric techniques. According to the above sentence, no assumptions about the distribution of space and the construction of a classifier are necessary.

Support Vector Machine (SVM)

First, Vapnik et al. introduced the concept of SVM. Of all algorithms, this one is the most accurate in terms of output quality. In spite of its origins as a tool for binary classification, it is now being used to solve challenges relating to multiclass classification. It can create single and multiple hyper planes in high-dimensional space, making it helpful for a variety of practical and efficient operations. Create a hyperplane using SVM to separate the data points is its primary goal SVM can be implemented in two different ways. Mathematical programming is used in the first method, and kernel functions are used in the second. In high dimensional space, non-linear functions can be easily mapped with the use of training datasets. Kernel functions are the only way to achieve this. Kernel functions include Gaussian, polynomial, and sigmoid functions. The hyper plane is used to classify data points in this manner. Data points must be separated as much as possible by hyper plane. The hyper plane is constructed using support vectors. SVM has a number of advantages. Here are a few examples: Firstly, it works well in high-dimensional environments. Additionally, it works best when the number of dimensions exceeds the number of samples. Third, it saves memory because it only employs a portion of the decision function's training points (known as support vectors), and second, it is flexible since multiple decision function kernel functions can be supplied.

Neural Network (NN)

For many years, it was considered the best classification algorithm before decision trees and SVMs were introduced. Many biomedical domains have turned to NN because of this, and it has since become the most extensively used classification technique in many of them. When it comes to illness diagnosis and outcome prediction, NN has proven to be a useful approach for a variety of applications, including cancer. Neurons or nodes are the basic building blocks of NN. The neurons in this network are interconnected, and they operated in concert to produce the output functions as a whole. Because they may work in parallel, they are able to generate fresh observations even when certain neurons or nodes within the network fail or go down. Each neuron in a neural network has an activation number and a weight given to it. Neural networks are mostly utilized for classification and pattern recognition applications. An NN's fundamental property is that it can reduce the error by changing its weight and structure. Only because of its adaptive nature does it reduce the inaccuracy [22]. NN can make more accurate predictions than humans. When it comes to training and classification, one of the key advantages of neural networks (NN) is that they are able to deal with noisy input.

Bayesian Methods

The Bayesian classification approach is used for probabilistic learning. We can easily get it with the help of a classification algorithm. It relies heavily on Bayes' theorem of statistics. Nave Bayes Classifier assumes that all attributes in the medical domain are independent of each

other, however in the medical domain, attributes such as patient symptoms and health state are interrelated. This is by far the biggest flaw in the Nave Bayes Classifier algorithm. As long as the attributes are uncorrelated, the Nave Bayesian classifier performs quite well in terms of accuracy [23]. They have critical functions in the healthcare industry. So they became popular amongst scientists all around the world. BBN has a number of advantages. Among its many benefits is the ease with which it facilitates the computing process. For large datasets, it performs faster and more accurately.

Regression

Data mining relies heavily on regression analysis, a key approach. Using this tool, we can quickly and simply discover the functions that are effective in demonstrating the correlation between numerous different variables It is primarily a calculating device. We may easily design it with the help of a training dataset. Consider 'P' and 'Q,' two variables. Statisticians frequently employ these two types of variables. There are two types of variables: independent and dependent. There is no limit on the number of dependent variables, but there is no limit on how many independent variables you can have. In most cases, regression is employed to examine the relationship between two variables. A variable's dependence on other variables can be solidified with ease using regression techniques. On the basis of the number of independent variables, regression can be classed as linear or non-linear. Linear regression is used to assess relationships between two types of variables, one of which is a dependent variable and the other one (or more) is an independent variable. Linear regression uses a linear function to build the linear model. This approach has a restriction because both variables are already known, and its primary goal is to trace a line that correlates between them. For categorical data, linear regression is not an appropriate tool. To use it, you must supply numerical data exclusively. Categorical data can be used with logistic regression. Logistic regression is a non-linear regression technique that uses this type of data. The chance of occurrence can be predicted using logit regression and the logit function. However, logistic regression cannot take into account linear relationships between variables. As a result of these factors, regression is frequently employed in the medical industry to make predictions about a patient's prognosis or likelihood of survival.

Clustering

Unsupervised learning happens when only independent factors are observed, but supervised learning examines both independent and dependent variables. This is what is meant by the term "clustering." Classification, on the other hand, is an unsupervised learning technique. No classes are assigned to it. As a result, clustering may be most appropriate for exploratory research involving vast amounts of data about which little is known (such as mass of data generated by microarray analysis).Clustering has a descriptive purpose, whereas classification is meant to make predictions (Veyssieres and Plant, 1998).A clustering method, also known as an unsupervised learning method, uses a similarity metric to build groups from a big database. Finding new categories and categorizing them is what clustering is all about; the new groups have their own value. Extrinsic evaluation is critical in classification tasks. Using a similarity metric, clustering was used to separate the data points. Clustering divides a dataset into subsets based on similarity, with each subset containing a different collection of occurrences. Clustering is a technique for identifying patterns in data. Each cluster's data points are more comparable to each other than data points from other clusters. Object grouping dates back to

our first attempts to describe and categorize the salient traits of people and things. Thus, it can be used to define topologies across a wide range of scientific fields, from mathematics and statistics to biology and genetics, all of which employ different terminology.

Partitional Clustering

With 'n' data points, the maximum possible number of clusters can be found in the datasets, which is 'k'. The partitioning clustering method is used to extract the 'k' clusters from the 'n' data points.

In order to employ a partitional clustering algorithm, you must provide an integer k. (which is the number of clusters).In most cases, partitional algorithms simply move objects from one k-node cluster to another. Algorithms for partitioning data are divided into three broad categories based on how they move objects, choose a cluster centroid (or representative) from inside a cluster, and assess the degree to which objects within a cluster are similar. This method necessitates the definition of the minimum number of clusters required before any datasets can be used to generate them. This approach is divided into two categories based on the similarities between items and cluster centroids. These are known as K-means and K-mediods. K-means is a well-known technique for this type of analysis. To begin, it picks k objects at random and divides them into k separate groups based on their resemblance to centroids and other objects. A cluster centriod is the average value of the items in the cluster in k-means K-mediods is the next algorithm. It employed mediods to organize the cluster. Mediod is critical since it is the most centrally positioned data point in the database. By applying k-means clustering and employing the clustering technique, Lenert et al. have been able to improve public healthcare services. Breast cancer recurrence can be detected by Belciug et al..

Hierarchical Clustering

In order to divide the data, this method can be utilized in two ways. Hierarchical data can be partitioned using either a top-down or bottom-up strategy, referred to as a tree structure. Agglomerative and Divisive are two classifications based on partitioned process. It's possible to have n data points. Among n data points, there may be a correlation between a number of data points. Using an agglomerative method, similar data points can be combined into a single unit. An iterative division into smaller groups until all data points are associated with a single cluster is the first step in the dividing strategy. Hierarchical clustering algorithms can be divided into three categories: Single-link, complete-link, and average-link are the three types. An object's group similarity is calculated using a single-link clustering technique, which selects an object's closet pair from among two groups. The complete-link algorithm compares the most remote pair of objects in two groups. Algorithms like average-link select all potential pairs of objects from two groups before averaging their distances. After assessing the degree of similarity across groups, the two with the least distance are combined. The average-link method is the most accurate of the hierarchical algorithms.

Density based clustering

These approaches play a vital role in biomedical research since they are capable of handling any arbitrary cluster shape. For the purpose of extracting significant patterns from a huge library of biomedical images, this method has been found to be effective and efficient. It is not possible to extract significant patterns from the biomedical pictures database using partitional clustering and hierarchical clustering methods because these two methods are only capable of handling circles. Density-based clustering was developed to solve the problem of patitional

clustering and hierarchical clustering approaches. DBSCAN, OPTICS, and DENCLUE are three of the most often used density-based clustering methods.

Association

KID3 is the first association rule mining algorithm was proposed by Piatetsky-Shapiro in 1991. [109] However, this method didn't get much attention because of its inefficiency and scalability difficulties. IBM Almaden Research Center's R.Agarwal presented a revolutionary association rule algorithm called Apriori that has attracted a lot of interest in relational mining. Due to Apriori's solution to KID3's difficulties, association mining may now be performed in real-world datasets, allowing for the extraction of association rules from these datasets. One of the most important techniques to data mining is to discover common patterns and intriguing correlations among a collection of data items in a data repository. Identifying connections between diseases, health status, and symptoms has a significant impact on the healthcare industry. Scientists have been using this method to study the connections between various ailments and the medications recommended for them. Companies in the healthcare insurance industry frequently employ this method in order to identify fraud and abuse. In association mining, accuracy is not an evaluation factor because every algorithm mines all possible association rules. Association mining algorithms are evaluated only on the basis of their efficiency, and this is their primary purpose.

Obstacles to Health Care Data Analytics

Although we all know, a lot of healthcare data is collected and held by many healthcare organizations. When it comes to making good decisions, however, there are a number of issues that might arise from the use of healthcare data. As a starting point, there is a disparity between how healthcare organizations manage their data. There is currently no standard format for data storage. It is possible that the absence of a common format will exacerbate pandemic conditions. For example, consider the case where the disease spreads throughout a country in its many geographic regions. For the sake of resolving the epidemic, the country's health ministry has mandated that all healthcare organizations submit their data with its central data warehouse for study. However, because the data formats are different. As a result, data analysis may be more time consuming than usual. It's possible that the situation could spiral out of control because of this. In order to improve healthcare services for patients, healthcare data can be highly beneficial in extracting relevant information from it Because we can't extract relevant information from data that lacks quality, this is a critical step. Consequently, the quality of the data is a critical issue to address. It is important to remove noise from the data and make sure that no data is missing when evaluating its quality. The quality of healthcare data must be maintained at all costs. Another key issue is coordinating the use of data. Patients and healthcare providers are both opposed to the idea of sharing their personal health information. Consequently, epidemics may get worse, effective treatment for a big population may not be possible, and fraud and misuse in healthcare insurance firms may be difficult to detect. Another problem is that it takes a long time and a lot of money to construct a data warehouse where all the healthcare institutions in a country may exchange their data.

Conclusion and future Work

The accuracy and performance of any algorithm are of the utmost significance. However, the aforementioned accuracy and performance properties of any algorithm might be considerably diminished if certain conditions exist. An algorithm of this type includes classification as well.

The classification algorithm is quite sensitive to noise. The processing power required to perform classification suffers greatly when there is any noisy data in the sample. Classification algorithm performance is hampered as a result of this. Therefore, prior to using a classification method, it is required to remove any attributes from datasets that later on are considered to be noisy attributes. For classification algorithms, feature selection approaches play a critical role in determining which traits are most beneficial.

Pattern recognition can benefit greatly from the use of clustering algorithms. When it comes to choosing an algorithm, they face a dilemma because they lack information on datasets. As soon as we know the number of clusters, we can select a partitioning algorithm. Even if the number of clusters is unknown, hierarchical clustering is still used. As the number of datasets grows, the performance of hierarchical clustering diminishes, making it less effective. Random sampling is an excellent solution to this problem.

Hierarchical clustering suffers from poor visualization if the data set is too vast to fit into a dendrogram. As an alternative, random sampling of the data could be used to build the dendrogram that can be used by users to better comprehend overall grouping/similarity of data. Cubic time complexity is the fundamental downside of hierarchical clustering techniques. Because of this, algorithms can only handle little amounts of data. Consequently, hierarchical algorithms are substantially slower than partitional clustering methods in terms of computation time. Distances between objects are also calculated using a lot of system memory.

Various data mining approaches must be combined in order to improve accuracy in disease prediction, improve survival rate in significant death-related conditions, etc.

All the essential actions must be made to establish better medical information systems that provide accurate information about patients' medical histories rather than information about their billing bills in order to obtain higher-quality medical data. It's not just individuals that benefit from high-quality medical data, but also healthcare providers and any other organizations participating in the healthcare business.

In order to obtain meaningful patterns, all steps must be taken to minimize the semantic gap across dispersed healthcare databases. These patterns can be extremely helpful in improving treatment effectiveness, detecting fraud and abuse, and improving customer relationship management around the world.

References

[1] Roy,D., Tripathy, S., Kar, S. K., Sharma, N., Verma, S. K., & Kaushal, V. (2020). Study of knowledge, attitude, anxiety & perceived mental healthcare need in Indian population during COVID-19 pandemic. *Asian Journal of Psychiatry*, 102083.

[2] Brito, J. M. C. (2016, July). Trends in wireless communications towards 5G networks—The influence of e-health and IoT applications. In *2016 International Multidisciplinary Conference on Computer and Energy Science (SpliTech)* (pp. 1-7). IEEE

[3] Hao, Y., & Helo, P. (2017). The role of wearable devices in meeting the needs of cloud manufacturing: A case study. *Robotics and Computer-Integrated Manufacturing*, *45*, 168-179.

[4] Lymberis, A., & Olsson, S. (2003). Intelligent biomedical clothing for personal health and disease management: State of the art and future vision. *Telemedicine Journal and e-health*, *9*(4), 379-386.

[5] Acampora, G., Cook, D. J., Rashidi, P., & Vasilakos, A. V. (2013). A survey on ambient intelligence in healthcare. *Proceedings of the IEEE*, *101*(12), 2470-2494.

[6] Arif, D., Ahmad, A., Bakar, M. A., Ihtisham, M. H., & Winberg, S. (2017, April). Cost effective solution for minimization of medical errors and acquisition of vitals by using autonomous nursing robot. In *Proceedings of the 2017 International Conference on Information System and Data Mining* (pp. 134-138).

[7] Yin, H., Akmandor, A. O., Mosenia, A., & Jha, N. K. (2018). Smart healthcare.

[8] Yu, H., Lee, H., & Jeon, H. (2017). What is 5G? Emerging 5G mobile services and network requirements. *Sustainability*, *9*(10), 1848.

[9] Gutierrez, J. A., Naeve, M., Callaway, E., Bourgeois, M., Mitter, V., & Heile, B. (2001). IEEE 802.15. 4: a developing standard for low-power low-cost wireless personal area networks. *IEEE network*, *15*(5), 12-19.

[10] Akpakwu, G. A., Silva, B. J., Hancke, G. P., & Abu-Mahfouz, A. M. (2017). A survey on 5G networks for the Internet of Things: Communication technologies and challenges. *IEEE access*, *6*, 3619-3647.

[11] Ahmadi, H., Arji, G., Shahmoradi, L., Safdari, R., Nilashi, M., & Alizadeh, M. (2019). The application of internet of things in healthcare: a systematic literature review and classification. *Universal Access in the Information Society*, 1-33.

[12] Dimitrov, D. V. (2016). Medical internet of things and big data in healthcare. *Healthcare informatics research*, *22*(3), 156-163.

[13]) Condoluci, M., Dohler, M., Araniti, G., Molinaro, A., & Zheng, K. (2015). Toward 5G densenets: architectural advances for effective machine-type communications over femtocells. *IEEE Communications Magazine*, *53*(1), 134-141

[14] Shen, X. (2015). Device-to-device communication in 5G cellular networks. *IEEE Network*, *29*(2), 2-3.

[15] Jain, Ishita and Bhardwaj, Dr. Manish, A Survey Analysis of COVID-19 Pandemic Using Machine Learning (July 14, 2022). Proceedings of the Advancement in Electronics & Communication Engineering 2022, Available at SSRN: https://ssrn.com/abstract=4159523 or http://dx.doi.org/10.2139/ssrn.4159523.

[16] Sharma, A., Tyagi, A., & Bhardwaj, M. (2022). Analysis of techniques and attacking pattern in cyber security approach: A survey. International Journal of Health Sciences, 6(S2), 13779–13798. https://doi.org/10.53730/ijhs.v6nS2.8625.

[17] Tyagi, A., Sharma, A., & Bhardwaj, M. (2022). Future of bioinformatics in India: A survey. International Journal of Health Sciences, 6(S2), 13767≤13778. https://doi.org/10.53730/ijhs.v6nS2.8624.

[18] Cai, S., Che, Y., Duan, L., Wang, J., Zhou, S., & Zhang, R. (2016). Green 5G heterogeneous networks through dynamic small-cell operation. *IEEE Journal on Selected Areas in Communications*, *34*(5), 1103-1115.

[19] Yeh, S. P., Talwar, S., Wu, G., Himayat, N., & Johnsson, K. (2011). Capacity and coverage enhancement in heterogeneous networks. *IEEE Wireless Communications*, *18*(3), 32-38.

[20] Philip, N. Y., & Rehman, I. U. (2016). Towards 5G health for medical video streaming over small cells. In *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016* (pp. 1093-1098). Springer, Cham.

[21] El Hassani, S., & Haidine, A. (2015, October). Roadmap towards beyond 4G: Key technologies and challenges for 5G. In *2015 International Conference on Wireless Networks and Mobile Communications (WINCOM)* (pp. 1-6). IEEE.

[22] Al-Falahy, N., & Alani, O. Y. (2017). Technologies for 5G networks: Challenges and opportunities. *IT Professional*, *19*(1), 12-20.

[23] West, D. M. (2016). How 5G technology enables the health internet of things. *Brookings Center for Technology Innovation*, *3*, 1-20.