Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 13, Issue 1, January 2022: 1464-1479

Predicted Infectious Disease (Corona Virus) using Data Mining and Machine Learning Algorithms Approaches

[1] Manoj Kumar Singh, [2] Dr. Kishan Pal Singh, [3] Dr.Navin Kumar Tyagi
[1]Research Scholar Mangalayatan University Aligarh, [2] AssociateProfessor Mangalayatan University Aligarh,
[3] Associate Professor MIT Engineering College Bulandshahr
[1] singhmanojmit@gmail.com,[2] drkishan.mu@gmail.com, [3]nt_74@rediffmail.com

Abstract-This Article's Abstract Corona disease prevention is a key difficulty in medicine, which focuses on diagnosis and treatment. Antimicrobial drugs, antibiotic resistance, immunization, and other reasons must all be understood in infectious disease medicine, which necessitates a detailed understanding of the prevalence and clinical presentation of infectious, viral, fungal, and parasitic illnesses. Immunomodulatory They are more equipped to deal with the factors that predispose people to infection and the core ideas of epidemiology and disseminating the disease. Most infectious disease experts are consultants to other doctors, meet patients in their office or hospital for advice, and may even longitudinally follow patients with particular disorders for therapy. The primary goal of this research is to create a rapid disease-type prediction tool for medical professionals, which will aid in clinical decision-making. E-healthcare applications using advanced data mining and machine learning methods are garnering interest from researchers in an effort to develop intelligent healthcare systems. Smart e-healthcare tools may be developed using Android's low cost and user-friendly interface. Monitoring and diagnosing patient health may also be done online. Early and accurate diagnosis may reduce mortality rates, and this e-healthcare application recommends a diet plan based on the patient's specific health profile. Electronic health care is capable of detecting and diagnosing any illness type in a matter of minutes, which will help healthcare providers enhance their services. With the use of data mining, we can establish a reliable baseline for future predictions. The application of data mining tools, on the other hand, may minimize the number of tests performed and help clinicians discover the most successful therapies and procedures. Time and performance are greatly impacted by this decreased test set. Predictive performance is based on a variety of factors that aid in the accurate diagnosis of illness, allowing patients to obtain more effective treatment at lower costs.

Keyword: infectious diseases, ML-Algorithm, Probabilistic Fuzzy Logic, MATLAB, Python.

I. INTRODUCTION

Corona viruses

Corona viruses are a widespread family of viruses that may cause anything from a simple cold to a life-threatening case of pneumonia.

Novel (new) CORONA virus

For the first time in human history, a new form of Corona virus has been discovered in humans. 2019-nCoV is the name of a new virus that has just emerged in China's People's Republic. MERS-Corona virus and SARS-Corona virus are other kinds of Corona virus that have been identified.

Wuhan City in Hubei Province, People's Republic of China, announced the first incidence of this virus on December 31, 2019. More than 6,000 instances have been documented in China as of January 28, 2020. There have also been examples outside of China in Thailand and Malaysia; Singapore; Japan; Taiwan; South Korea; Hong Kong; Macao; Vietnam; Nepal; Cambodia; Sri Lanka; Australia; France; Germany; the United States and Canada; as well as other countries. The United Arab Emirates has announced a confirmed case of Ebola today (Jan. 29).

Symptoms of the new Corona virus: - The most common symptoms include:

Fever

Cough

Shortness of Breath and Difficulty breathing

New (2019-nCoV) be transmitted from one person to another

In fact, the novel CORONA virus may be transmitted from one person to another via close contact and exposure to respiratory spray, such as in the setting of the family, workplace, or health-care facilities, for example. If the sick person's spray-contaminated surfaces aren't cleansed, they might be exposed as well.

Health care workers at risk of contracting the new Corona virus

It is true that health care professionals are more likely to transmit infections to patients than the general population since they are more likely to be in close contact with them.

Vaccine against the new CORONA virus

There is currently no vaccination that is effective against this disease. If a new illness emerges, there will be no vaccination available until one is created. A vaccination for the virus may not be available for many years. Because the novel corona virus is a member of a different viral family than the seasonal influenza vaccination, it is not protected against by the seasonal influenza vaccine this year.

Cure for the new Corona virus

The new Corona virus does not have a particular antiviral medication. Its symptoms, on the other hand, may be alleviated depending on the patient's medical state. Infected individuals may benefit greatly from supportive treatment. Drugs that have been shown useful in the treatment of other known viral infections are being utilized to treat the developing virus for medical reasons, but have not yet proven to be beneficial, and this evidence will need to be studied further in the near future.

Protect against the new Corona virus

Please take these steps to lessen the possibility of contracting the new Corona virus infection:

- If you are in an epidemic country, avoid direct contact with animals (live or dead) or animal trade marketplaces.

- Keep your distance from someone who seems to have a cold or upper respiratory infection symptoms.
- When sneezing or coughing, cover your nose and mouth with a tissue or flexed elbow.
- Use an alcohol-based hand rub or soap and water to clean your hands.
- Cook animal items thoroughly before consuming them.

The symptoms of infectious illnesses may be used to predict the level of public concern about them. Monitoring technologies to track and control diseases and advise policymakers on changing prevention and reduction initiatives have traditionally been the focus of public health. Delays in reporting and reviewing data, on the other hand, seem to impede traditional monitoring technology. New risk modelling approaches are being developed that will allow for more accurate monitoring of public concerns and early warning of prospective health impacts, as well as the combination of the two. All climate risk evaluations have strengths and weaknesses, and this work aims to explain how they might drive public health efforts to prevent, identify, and reduce the effects of climate change on infectious illnesses. Data may now be gathered and interpreted from previously inaccessible locations thanks to advances in machine learning analytics. Smart meters, for example, may transmit a wide range of data to external medical networks, allowing for better diagnosis and feedback. Table Analyzer automatically analyses medical data and then publishes the results in real time via national patient groups.The Epic Star and the Wellness Chart combine health care data and monitoring expertise. Methods for preventing the spread of infectious diseases like Zika and H1N1. Networks, machines or artefacts with sensors that allow for direct data transmission between the devices without the need for a larger web. Using the information listed below, it is possible to fulfil the regional scientific criteria for predictive modelling and non-proliferation implementation for real-time illness monitoring.

People's lives have been ruined and many have died as a result of the new corona virus illness (COVID-19). As of mid-July 2020, the illness has spread to virtually every nation, killing over 580,000 people out of the over 13,379,000 confirmed cases, according to WHO figures [11]. In response to the COVID-19 pandemic, governments in a number of nations have recommended a variety of interventions. In this unique and tumultuous period, science and technology have played a vital role in the execution of these program. Drones and robots may be used for a variety of purposes, including as cleaning streets and public places, delivering food and medication to patients with the corona virus, and more. There are a number of medical experts scrambling to examine treatments and medicines to treat sick people while others are working on vaccinations to avoid the illness. X-ray pictures and computed tomography (CT) scans, on the other hand, have been used by computer science researchers to discover infected patients at an early stage. Artificial intelligence (AI) includes approaches like these, and it has been effectively used in a variety of disciplines. Focusing on the application of Machine Learning in the fight against COVID-19, this study examines the role of these technologies. Several long-standing industries, such as computer vision, natural language processing (NLP), voice recognition, and video games have seen significant development thanks to machine learning, particularly deep learning. Deep learning has a huge edge over typical machine learning approaches since it is able to manage and make sense of a wide range of data kinds, including large and unstructured data, such as text, picture, video, and audio. With the use of deep learning and artificial intelligence (AI), a wide range of sectors such as electrical devices and automobiles as well as food processing and agriculture have seen greater results and advantages.

Artificial intelligence (AI) tools, such as those discussed in the next section, are therefore likely to be useful in the battle against the COVID-19 virus pandemic.

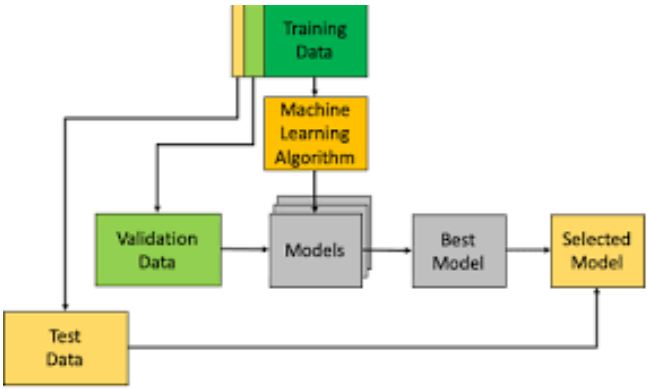


Figure- Machine Learning Model validation framework

II. EMERGENCE OF COVID-19 DISEASE

At some point in our lives, the majority of us will come into contact with a corona virus. People who have only heard about SARS-CoV-2, the cause of COVID-19, may be alarmed to learn that there are other corona viruses.

SARS-CoV-2 is only one of several corona viruses. There are literally hundreds of different Corona viruses. It's common for them to infect animals such as birds and mammals. There have been a few instances when viruses that normally only affect one species have mutated such that they may infect another. Cross-species transmission (sometimes known as "spillover") is a term used to describe this phenomenon.

Early in the 1930s, a corona virus was initially identified in chickens. Until the 1960s, no one knew that human corona viruses existed. Seven coronaviruses have been identified so far as capable of causing illness in people. It's common knowledge that there are four endemic diseases, which are present in a certain population or geographic region. However, three of these diseases may be very dangerous and even deadly.

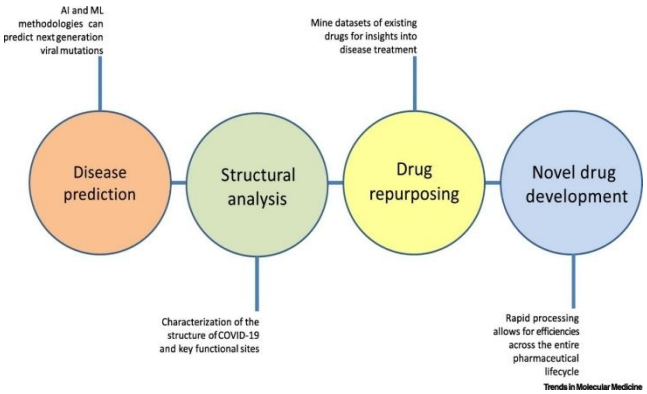


Figure- Emergence of Covid-19 Disease

List of COVID Infection Risk Factors (Fever >=37.50 C, Respiratory Symptoms, Prior Contact with COVID Patients, Family History of COVID Infection)

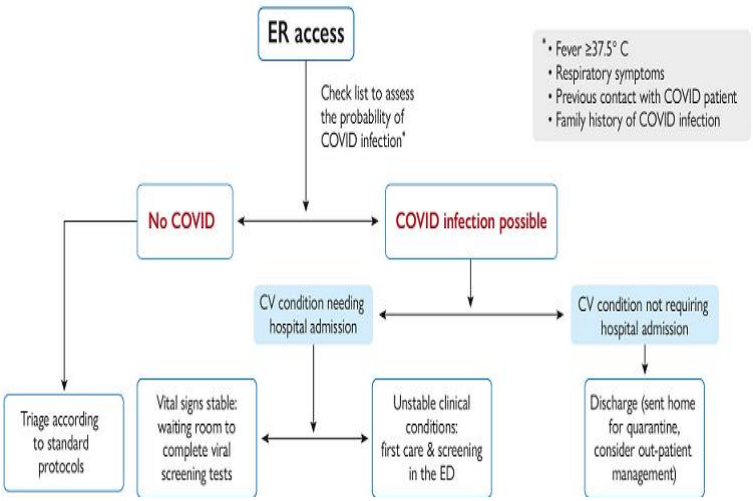


Figure - Algorithm for Triaging Patients Admitted to the Emergency Room for a suspect acute corona virus disease

III. DATA MINING APPROACHES

For the majority of companies, data and information are becoming essential assets. A crucial part of medical database knowledge discovery is data mining. As the name suggests, databases are structured collections of data. We refer to the software used to create and manage this data as DBMS. Database knowledge discovery is the whole process of extracting information from data. Models and patterns may be extracted from vast data sources using data mining methods.As stated by Frawley et al. (1991), the foundation of KDD is data mining that extracts intriguing patterns from data that are simple to understand, analyses and alter. It is a branch of mathematics that deals with analyzing large amounts of data in order to discover meaningful patterns. There are a few phases in the KDD process before new knowledge may be generated.

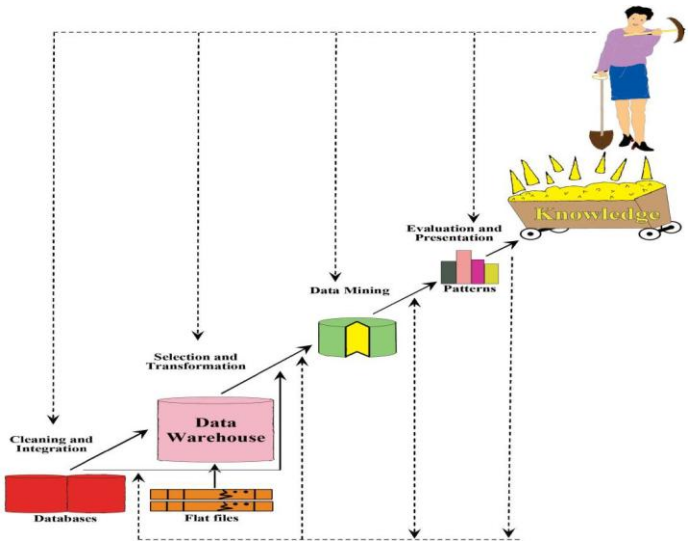


Figure: - Knowledge discovery as a process

IV. Subbalakshmi et al. (2011) Knowledge discovery is described as an iterative sequence of cleaning, integration, data selection, data mining pattern recognition, and presentation of the findings. An important aspect of data mining is to find out the linkages and global patterns that are buried in massive databases of data. Before any data mining methods can be employed, a target data set must be prepared. In order to examine multivariate data sets, data marts and warehouses are typical sources of

data, and preprocessing is a need. Validating the patterns found by data mining algorithms across all of the data sets is the last stage in discovering new knowledge from data. The discovered knowledge may include rules that explain the data's attributes, patterns that occur often, and items that are clustered in the database, among other things.

IV. DATA MINING TASKS

Machine learning often involves the classification of data. In the mid-1950s, artificial intelligence was initially recognised as a field. The ability to learn is a prerequisite for all forms of intelligent conduct. Researchers currently believe that learning is the only way to get intelligence. It's no secret that machine learning has been at the heart of AI research since its inception. The field of computer science known as "machine learning" focuses on the creation of algorithms that enable computers to learn on their own. It may be used to create systems that boost the overall system efficiency and effectiveness.

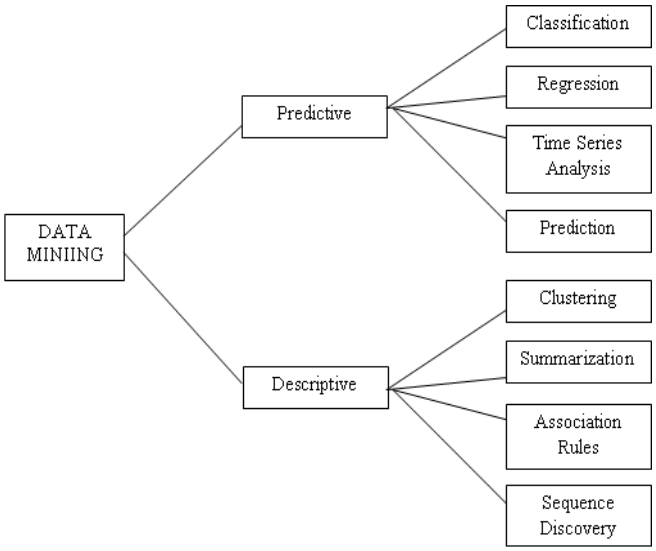


Figure: - Data mining tasks

V.DATA MINING CLASSIFICATION TECHNIQUES

Data warehousing's early adopters considered data mining as a component of the warehousing process. Data mining is now regarded as a separate industry, and a warehouse may be a suitable source of data. There are a large variety of data mining methods and techniques that may be used to solve a wide variety of problems. Large datasets naturally lend themselves to data mining. Learning and classifying data are both included in the categorization process. Training data is examined using classification algorithms, while classification rules are tested using test data to determine their correctness.Classification in data mining is used to predict data instances' membership in groups, according to Lashari et al. (2013). Using advanced data analysis methods, data mining aims to uncover hidden relationships within massive datasets. It is common practise in data mining to apply decision tree-based categorization algorithms for decision support purposes. Medical data categorization might greatly benefit from using data mining methods.

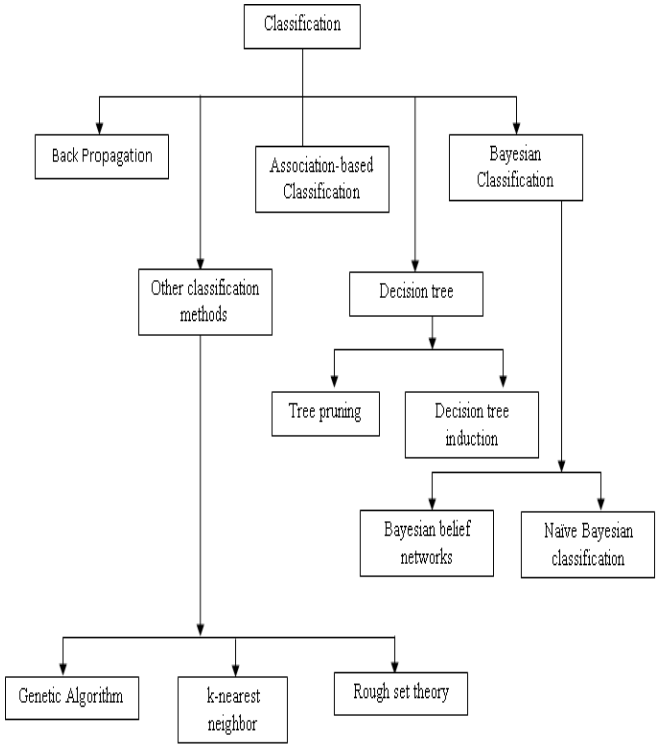


Figure: - Data mining classification methods

VI.DATA MINING IN HEALTH INFORMATICS

Statistics, computer learning, neural networks, and pattern recognition all come together in data mining. It deals with the extraction of hidden knowledge structures expressed in models and patterns from big data sources.Healthcare is a data-heavy industry. Every second, fresh data is generated by a slew of operations. As a result, it is also the most heavily funded by taxpayers. Health care has experienced a growth in computers and new algorithms, and can no longer disregard these growing instruments. As a consequence, health informatics has been formed, which brings together healthcare and computers. Typically, they use a combination of medical data and clinical skills to do their job.

Among the medicinal uses of data mining are the following:

- Health care cost projections.
- To determine the best therapy for a condition.
- Diagnosis and prognosis of a wide range of illnesses,

When it comes to health information technology, health informatics is described as a growing scientific subject that focuses on the collection, storage, retrieval... Research in this discipline focuses on improving patient care and population health via clinical care, public health and biological research.

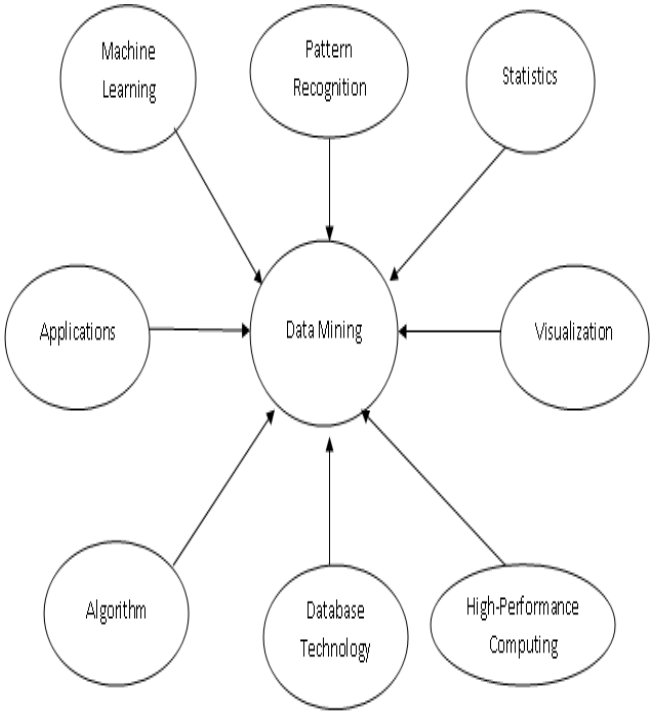


Figure: - Origin of Data mining

VII.MACHINE LEARNING ALGORITHM USE IN CORONA VIRUS DISEASES

Analysis of data using machine learning automates the creation of analytical models. In this field of AI, computers are taught to learn from data, spot patterns, and make choices on their own with the least amount of human input possible.An important aspect of machine learning is the ability of computers to complete tasks without being explicitly taught how to do so. In order to do specific jobs, computers must learn from the data that is presented to them. When computers are tasked with basic tasks, algorithms may be programmed that instruct the machine on how to carry out all steps necessary to solve the issue at hand; the computer itself does not need to learn anything. Humans may have difficulty creating complex algorithms on their own for increasingly complex jobs. Rather of relying on human programmers to describe every step, it may prove more efficient to let the machine figure out its own method.Machine learning is a branch of computer science that uses a variety of techniques to educate computers to do jobs for which there is no totally suitable solution. It's common in circumstances when there are a lot of possible solutions to identify some of them as genuine. Data such as this may be sent into the computer to help it fine-tune its algorithms for determining the proper replies. Digital character recognition systems often train on the MNIST dataset, which contains handwritten digits.A broad range of applications, including email filtering and computer vision, require machine learning algorithms because traditional methods are either ineffective or impossible to create.Statistical learning is closely connected to a subset of machine learning, however not all machine learning is based on statistics. In the discipline of machine learning, the study of mathematical optimization provides tools, theory, and domains for application. Unsupervised learning and exploratory data analysis are two related fields of research. Machine learning is also known as predictive analytics when it is used to solve business challenges.

VIII. MACHINE LEARNING APPROACHES

According on the nature of the "signal" or "feedback," machine learning algorithms are generally split into three basic categories:

Supervised learning: As an example of input and output, a "teacher" gives examples to the computer in order to teach it a general rule for mapping inputs to output.

Unsupervised learning: Rather, the learning algorithm is left to its own devices in search of patterns in the data it receives. For example, unsupervised learning may be an end in and of its own (finding patterns in data), or it can be used as a tool to achieve a specific objective (feature learning).
Reinforcement learning: Computer program interact with a constantly changing environment to accomplish a certain task (such as driving a vehicle or playing a game against an opponent). The software receives input equivalent to incentives as it makes its way across the issue space, and it strives to optimize that feedback.

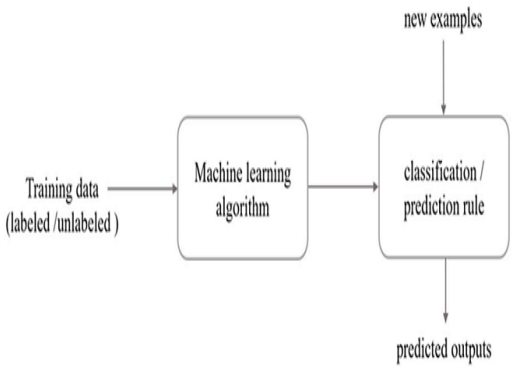


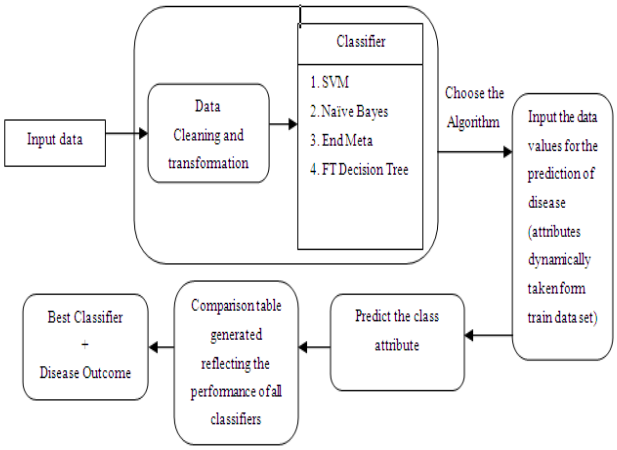
Figure- Using Machine Learning Algorithm predicted Outputs

IX. CURRENT TREND IN PREDICTION MODEL

To predict future occurrences, predictive analytics relies on historical data. A model based on previous data is used to predict the current trend. The predictive model is then used to test or actual data in order to predict the future. Researchers are becoming more interested in predictive analytics as a result of new supporting methods, such as machine learning and data mining. Here are some of the current methods for forecasting the future:
Mathematical and Statistical Approach: Large amounts of data are handled by databases in the fields of commerce, education, politics, and health research at the moment. The vast majority of the data is statistical. As with search engines, weather forecasts are based on statistical data. Complex computing is required to extract any data from these massive databases. We need a statistically-based predictive model to make these kinds of data predictions. Mathematical functions are used to extract information from statistical data or anticipate the future.

Regression: In order to discover the connection between two variables, regression analysis is an effective tool. Regression is another method that helps in forecasting. As a result, independent attributes are used as predictors (also known as changing variables) while dependent variables are used as responses (also known as response variables). Regression methods include linear regression, multiple regression, logistic regression, and non-linear regression.
Machine Learning: The application of machine learning methods for categorization and prediction has been increasing in recent years High-quality classification models may be built using a wide range of classifiers, including support vector machine (SVM), Genetic Algorithm, Naive Bayes, Decision Tree, and Neural Network.

Disease Prediction Model: -



.

Figure-Overall Architecture of Proposed System

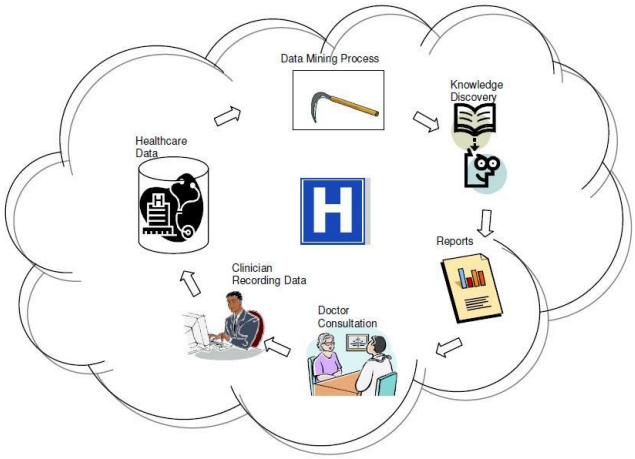


Figure- Data mining for medicine and health care architecture

Intelligent techniques for extracting patterns in data are used in Data Mining, which offers a significant potential to help doctors cope with the massive amounts of data they must deal with. In a specialized medical field, medical data mining has been used to accurately classify and rapidly forecast prognosis and diagnosis of patients. Using a variety of classifiers on the data sets is one of the guidelines for classifying medical data mining. Medical databases are expanding at an astronomical rate. This fast increase is the key motivator for academics to harvest relevant information from these medical databases. Data mining methods serve a critical role in discovering patterns and extracting information to improve patient care and diagnostic skills as the amount of recorded data grows. Data mining is a broad term that encompasses a wide range of specialized methods and tools. It is possible to use statistical approaches, visualization, and machine learning to uncover hidden patterns in data, giving healthcare workers new information from which to draw when making choices.

X. Research Background

Albahri et al. (2020),Interest in creating AI solutions to solve medical issues is on the rise. More than 350,000 people have died as a result of the worldwide epidemic, which has infected most of the globe. Expert mining and data mining approaches were used to assess the CoV prediction tools in this research. Data mining and machine learning methods are becoming more important in the medical area, according to academics.

Alimadadi et al. (2020),Coronavirus 2 is the cause of SARS-CoV-2, a severe and disabling illness. President Trump has asked AI experts from across the globe to help in COVID-19 research. Regular updates to COVID will be carried out in a joint effort with premier institutions.

Kumar et al. (2020),Efforts to contain the corona virus (COVID-19) pandemic are putting a lot of strain on healthcare systems. Any pressure on the healthcare system may be relieved if this infection was discovered early. X-rays of the chest have proven quite useful in the diagnosis of a variety of illnesses, including pneumonia.

Lalmuanawma et al. (2020),SARS-CoV2 and its related outbreak will be studied in this research in terms of AI and machine learning's involvement in predicting, forecasting, contacting, and producing drugs. This study demonstrates the advances made in the fight against the last pandemic through machine learning and artificial intelligence.

Kukar et al. (2020),Machine learning was applied to forecast COVID-19 based on previous data. Blood samples from 5,333 people were used to create the models. Patients with a severe cold virus infection have blood variables that resemble those of a bacterial illness rather than a viral infection. RT-PCR and chest CT studies claimed diagnostic accuracy is at the very least equivalent and may even be complimentary to it.

Brinati et al. (2020),In terms of influenza testing, real-time RT-PCR is the most accurate method available right now, but it also comes with a plethora of drawbacks. Two machine learning algorithms have been created to determine whether a patient has the SARS coronavirus. This research emphasized the need of rRT-PCR supplies and specialized labs in impoverished nations across the globe.

Yao et al. (2020),The recent coronavirus disease-2019 (COVID-19) epidemic in China and the rest of the globe posed major challenges to human civilization. The illness detection model was built using machine learning methods. The model had an accuracy of 0.8148 after being trained with 28 features.

Zoabi et al. (2021),Since its first epidemic, the SARS-CoV-2 coronavirus has spread to more than 200 countries, causing about 200,000 fatalities. There are no known strains, viruses, or the most up-to-date scientific procedures for confirming their existence. In this study, the criteria for faster, cheaper, and more equitable research methodologies are identified.

Zoabi& Shomron(2020),SARS-CoV-2 testing provides for a more rapid and accurate diagnosis. For the purpose of estimating the likelihood of infection, many models have been created. To diagnose COVID-19, the Israeli Ministry of Health developed a system that asks patients many straightforward questions about their health and family history. Only eight characteristics in the suggested model are needed to accurately estimate an Ebola patient's infection level.

Author (Year)	**Methodology Used**	**Result**	**Disease**
Albahri et al. **(2020**	The MERS-CoV dataset was subjected to three different machine learning algorithms in order to find the most accurate model for binary and multiclass classification.	CoV prediction systems based on data mining and machine learning were examined in this research.	Novel Corona virus (COVID-19)

Alimadadi etal. (2020)	Personalized protective strategies, novel diagnostic approaches using machine learning algorithms.	Aim to speed up diagnosis, create new and more efficient treatments, and maybe identify the most vulnerable individuals based on their genetic and physiological traits	Corona virus disease 2019 (COVID-19)
Kumar et al.(2020)	ResNet152 was used to classify the derived deep feature recovered from COVID-19 and Pneumonia patients on chest X-ray pictures using machine learning techniques.	Random Forest predictors are accurate 0.973 and XGBoost predictors are accurate 0.977.	corona virus (COVID-19)
Lalmuanawma et al. (2020)	Machine Learning and Artificial Intelligent for tackling Covid-19 pandemic.	Human involvement has been reduced in the Covid-19 epidemic because to advances in AI and machine learning.	Covid-19
Kukar et al. (2020)	machine learning predictive model for COVID-19	COVID-19 diagnosis is attainable using ML on data from routine	corona virus disease (COVID)

	diagnosis	blood tests	
Brinati et al. **(2020)**	Two machine learning classification models using hematochemical values from routine blood exams	Using blood tests analysis and machine learning instead of rRT-PCR to detect COVID-19 positive patients is feasible and clinically sound.	COVID-19

XI. RESEARCH DESIGN AND METHODOLOGY

To practise and double-check the data collection, the illness data might be split 70-30. Many characteristics, such as age, ethnicity, and other medical criteria, are found in studies used to make a medical diagnosis, such as

a. Modelling of data

The inputs of the logical method are reflected in this step. UCI machine learning is used to collect data.

b. Treatment of Missing Values

All characteristics from the vector to the field of the device have been adjusted using noise cancellation and data normalisation as a priori model.

c. Data Analysis

A small number of machine learning approaches are being used to alter data collecting.

d. Construct the model for Real Time

In the end, there is a requirement to model for real-time testing. The suggested real-time model should be used to identify and predict coronary disease in the custom framework.

Phase 1:first and first, be certain that the datasets are in fact relevant. It is decided to do statistical analysis on the characteristic in our dataset that has the lowest and biggest values.

Phase 2: Tests the normalcy of the data using mathematical patterns of data.

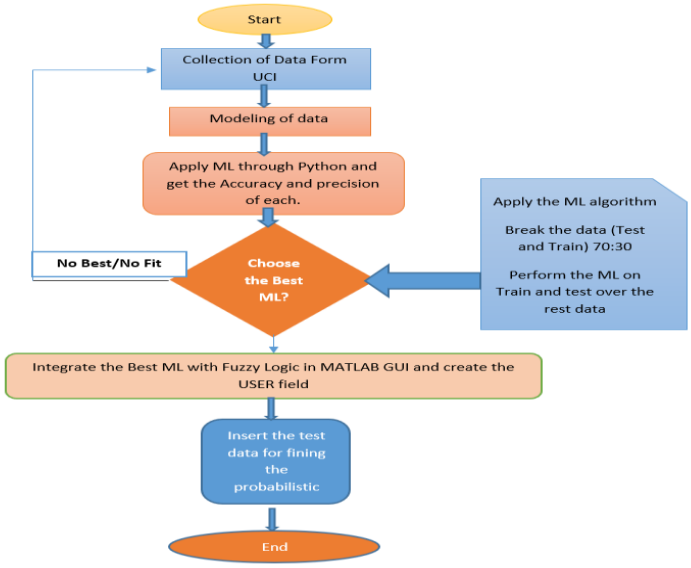
Phase 3: The "Evaluate column mean" should be used in the "missing values" area.

Phase 4: Recommendation: Fill up missing numbers using median and mean values.

Phase 5: Use a 70% to 30% split to train the data using ML algorithms and then do additional testing.

Phase 6: Use the train data sets to run the machine learning algorithm.

Phase 7: The outcomes of the tests should be compared to the data sets used in the tests.



Infectious illness simulation software is discovered to be used for a wide range of purposes by users, and the tools available are as diverse. A few instruments addressed the inclusion of compatibility trials or implementation plans in the research, although the instrument design was inconsistently stated. Many studies have shown problems with data sharing, security and consistency. There are a wide range of features and functions supplied by contemporary tools, although they are often used for a single application. The method's general appeal has been hampered by a lack of organizational resources, access issues, and misconceptions about how to use it.

XII. **CONCLUSION**

The healthcare industry's primary focus is on disease identification and treatment. Using data mining and machine learning in medical diagnostics and forecasting is critical to the success of these processes. There is a growing demand for better data mining approaches that may enhance the quality of medical services as technology progresses. In this research, we developed a disease prediction system based on an integrated application-based data mining approach. A better grasp of the history of virus, fungus, and parasite issues in humans as well as the knowledge of medications and immunizations may be gained through studying infectious illnesses, Experts in environmental, occupational, and host conditions have a thorough grasp of how these circumstances impact susceptibility. In the healthcare industry, quality management is applied. Specialists or a single clinic are responsible for the most critical procedures. In addition to seeing patients in their office or hospital, some doctors also act as mentors to others and aid patients in obtaining treatment for their diseases Essential clinical services are often the duty of doctors in these fields. In addition to giving lectures and conducting experiments, professors at universities also serve as mentors and advisors to their students and other professionals. For the sake of public health, we held a number of research sessions to acquire further data. Visualizations have been shown to be difficult to understand and apply. These methods were shown to be successful in the construction of a prediction model utilizing different classification strategies to predict coronary artery disease and their performance in predicting. Confidence was found to be an important element in the use of these ways It also indicates that data mining may be used to forecast or categories data in medical databases with a good degree of accuracy.

XIII. REFERENCES

1. Albahri, A. S., Hamid, R. A., Alwan, J. K., Al-Qays, Z. T., Zaidan, A. A., Zaidan, B. B., &Madhloom, H. T. (2020). Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review. *Journal of medical systems*, *44*, 1-11.

2. Alimadadi, A., Aryal, S., Manandhar, I., Munroe, P. B., Joe, B., & Cheng, X. (2020). Artificial intelligence and machine learning to fight COVID-19.
3. Kumar, R., Arora, R., Bansal, V., Sahayasheela, V. J., Buckchash, H., Imran, J., ... & Raman, B. (2020). Accurate prediction of COVID-19 using chest x-ray images through deep feature learning model with smote and machine learning classifiers. *MedRxiv*.
4. Lalmuanawma, S., Hussain, J., &Chhakchhuak, L. (2020). Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. *Chaos, Solitons & Fractals*, 110059.
5. Kukar, M., Gunčar, G., Vovko, T., Podnar, S., Černelč, P., Brvar, M., ... &Notar, M. (2020). COVID-19 diagnosis by routine blood tests using machine learning. *arXiv preprint arXiv:2006.03476*.
6. Brinati, D., Campagner, A., Ferrari, D., Locatelli, M., Banfi, G., &Cabitza, F. (2020). Detection of COVID-19 infection from routine blood exams with machine learning: a feasibility study. *Journal of medical systems*, *44*(8), 1-12.
7. Yao, H., Zhang, N., Zhang, R., Duan, M., Xie, T., Pan, J., ... & Wang, G. (2020). Severity detection for the coronavirus disease 2019 (covid-19) patients using a machine learning model based on the blood and urine tests. *Frontiers in cell and developmental biology*, *8*, 683.
8. Zoabi, Y., Deri-Rozov, S., & Shomron, N. (2021). Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *npj Digital Medicine*, *4*(1), 1-5.
9. Zoabi, Y., & Shomron, N. (2020). COVID-19 diagnosis prediction by symptoms of tested individuals: a machine learning approach. *MedRxiv*.
10. Umarani, V., &Subathra, M. (2020). Data Mining and Machine Learning Techniques in Prediction of Covid-19 Outbreaks-A Recent Review.Tierärztliche Praxis, 40, 1437-1447
11. World Health Organization (2020). WHO coronavirus disease (COVID-19) dashboard. https://covid19.who.int/. Accessed on 17 July 2020.
12. U LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradientbased learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324
13. 13. Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (pp. 1097-1105).
14. 14.Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D.,and Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-9).
15. Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
16. He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770-778).
17. 17. Bai, H. X., Hsieh, B., Xiong, Z., Halsey, K., Choi, J. W., Tran, T. M. L.,... and Jiang, X. L. (2020). Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. Radiology, 200823.
18. 18. Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B. ... and Cao, K. (2020). Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology, 200905.
19. Dympna O'Sullivan, William Elazmeh, Szymon Wilk, Ken Farion, Stan Matwin, Wojtek Michalowski, and Morvarid Sehatkar, "Using Secondary Knowledge to Support Decision Tree Classification of Retrospective Clinical Data", MCD 2007, LNAI 4944, pp. 238–251, 2008.
20. Raghupathi V, Raghupathi W, "Benchmarking Hospital Performance using Health Analytics", in Journl of Health & Medical Informatics, Vol.6 Issue.2, 2015.
21. Hian Chye Koh and Gerald Tan,"Data Mining Applications in Healthcare", journal of Healthcare Information Management – Vol 19, No 2, 2015.

22. Megha Rathi, Vikas Pareek, "Spam Mail Detection through Data Mining – A Comparative Performance Analysis", IJMECS, vol.5, no.12, pp.31-39, 2013.DOI: 10.5815/ijmecs.2013.12.05
23. Vijayarani, S & Sudha, S 2013, ' An Effective Classification Rule Technique for Heart Disease Prediction', *International Journal of Engineering Associates*, vol. 1, no.4, pp. 81-85
24. Vikas Chaurasia & Saurabh Pal 2013, 'Data Mining Approach to Detect Heart Diseases', *International Journal of Advanced Computer Science and Information Technology*, vol. 2, no.4, pp. 56-66
25. Witten, IH & Frank, E 2011, '*Data Mining: Practical machine learning tools and Technique'* , 3rd edition, Morgan Kaufmann, San Francisco, USA
26. Varun Kumar & Ezaz Ahmed, MD 2011, ' An Empirical Study of the Applications of Web Mining Techniques in Health Care', *International Journal of Advanced Computer Science and Applications,* vol. 2, no.10, pp. 91-94
27. Vijaya Lakshmi, K & Padmavathamma, M 2013, ' Modeling an Expert System for Diagnosis of Gestational Diabetes Mellitus Based On Risk Factors', *International Organization of Scientific Research Journal of Computer Engineering,* vol. 8, no.3, pp. 29-32
28. Tina Patil, R & Sherekar, SS 2013, 'Performance Analysis of Naive bayes and J48 Classification Algorithm for Data Classification', *International Journal of Computer Science and Applications,* vol. 6, no.2, pp. 256-261