Project Synopsis

on

**PENTACKLES**

Submitted as a part of course curriculum for

**Bachelor of Technology**

in

**Computer Science**



**Submitted by**

Aakansha Tyagi (2000290120001)

Aakriti Gupta (2000290120002)

Ayush Jaiswal (2000290120052)

Ayush Tyagi (2000290120056)

**Under the Supervision of**

Dr. Sapna Juneja

Professor

**KIET Group of Institutions, Ghaziabad**

**Department of Computer Science**

**Dr. A.P.J. Abdul Kalam Technical University**

**2022-2023**

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S NO:** | **ESSENTIALS** | **PAGE NO:** |
| **1** | TITLE PAGE |  |
| **2** | DECLARATION |  |
| **3** | CERTIFICATE |  |
| **4** | ACKNOWLEDGEMENT |  |
| **5** | ABSTRACT |  |
| **6** | CHAPTER 1 INTRODUCTION |  |
| **7** | 1.1. Introduction |  |
| **8** | 1.2 Problem Statement |  |
| **9** | 1.2. Objective |  |
| **10** | 1.3. Scope |  |
| **11** | CHAPTER 2 LITERATURE REVIEW |  |
| **12** | CHAPTER 3 PROPOSED METHODOLOGY |  |
| **13** | 3.1 Flowchart |  |
| **14** | 3.2 Algorithm Proposed |  |
| **15** | CHAPTER 4 TECHNOLOGY USED |  |
| **16** | CHAPTER 5 CONCLUSION |  |
| **17** | REFERENCES |  |

**DECLARATION**

We hereby declare that this submission is our work and that to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

Signature of Student

Name: Aakansha Tyagi

Roll No.:2000290120001

Date:

Signature of Student

Name: Aakriti Gupta

Roll No.:2000290120002

Date:

Signature of Student

Name: Ayush Jaiswal

Roll No.:2000290120052

Date:

Signature of Student

Name: Ayush Tyagi

Roll No.:2000290120056

Date:

**CERTIFICATE**

This is to certify that Project Report entitled “**PENTACKLES**” which is submitted by **Aakansha Tyagi, Aakriti Gupta, Ayush Jaiswal, Ayush Tyagi,** in partial fulfilment of the requirement for the award of degree B. Tech. in Department of Computer Science of Dr A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

**Date: Supervisor Signature**

Dr. Sapna Juneja

(Professor)

**ACKNOWLEDGEMENT**

It gives us a great sense of pleasure to present the synopsis of the B.Tech Mini Project undertaken during B.Tech. Third Year. We owe a special debt of gratitude to **DR. SAPNA JUNEJA** (Professor), Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his/her constant support and guidance throughout the course of our work. Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his/her cognizant efforts that our endeavours have seen the light of the day.

We also take the opportunity to acknowledge the contribution of **DR. AJAY KUMAR SHRIVASTAVA** (Head of the Department of Computer Science), KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

Last but not the least, we acknowledge our friends for their contribution to the completion of the project.

Signature of Student

Name: Aakansha Tyagi

Roll No.:2000290120001

Date:

Signature of Student

Name: Aakriti Gupta

Roll No.:2000290120002

Date:

Signature of Student

Name: Ayush Jaiswal

Roll No.:2000290120052

Date:

Signature of Student

Name: Ayush Tyagi

Roll No.:2000290120056

Date:

**ABSTRACT**

Machine learning and Artificial Intelligence are playing a huge role in today’s world. From self-driving car, medical fields, we can find them everywhere. The medical industry generates a huge amount of patient data which can be processed in a lot of ways. So, with the help of machine learning, we have created a Prediction System that can detect more than one disease at a time. Many of the existing systems can predict only one disease at a time and that too with lower accuracy. Lower accuracy can seriously put a patient’s health in danger. We have considered three diseases for now that are Heart, Liver, and Diabetes and in the future, many more diseases can be added. The user has to enter various parameters of the disease and the system would display the output whether he/she has the disease or not. This project can help a lot of people as one can monitor the persons’ condition and take the necessary precautions thus increasing the life expectancy.

**CHAPTER 1- INTRODUCTION**

**1.1 Introduction**

The system we have proposed is user friendly to get help and advice on health issues immediately through the online healthcare system. Now a days, with the help of the statistics and posterior distribution the problems are swiftly and easily. As the Bayesian statistics has a great success rate in the field of economic, social science and a few other fields just like that, in medical fields, people have solved various medical problems that are tiresome to be settled in classic statistics by classification and can be solved easily.The classification rules which help in solving the prediction of disease are generated by the samples trained by themselves and help in solving the problem easily.

It is approximated that greater than 70% of people in India are prone to various body diseases like viral, flu, cough, cold etc. in intervals of 2 months. As many people don’t understand that the general body diseases could be symptoms of something more harmful, 25% of this population dies or gets some serious medical problem because of ignoring the early general body symptoms and this is a very serious condition that we are facing and the problem can be proven to be a very dangerous situation for the population and can be alarming if the people will continue ignoring these diseases. Hence identifying or predicting the disease at the very basic stage is very important to avoid any unwanted problems and deaths. The systems which are available now a days are the systems that are either dedicated to a particular disease or are in development or the research for solving the algorithms related to the problem when it comes to generalized disease.

The main motive of the proposed system is the prediction of the commonly occurring diseases in the early phase as when they are not checked or examined they can turn into a disease more dangerous disease and can even cause death. The system applies data mining techniques , decision tree algorithms, Naive Bayes algorithm and Random Forest algorithm. This system will predict the most possible disease based on the given symptoms by the user and precautionary measures required to avoid the aggression of disease, it will also help doctors to analyze the patterns of diseases in the society. This project is dedicated to the Disease prediction System that will have data mining techniques for the basic stages of the dataset and the main model will be trained using the Machine Learning (ML) algorithms and will help in the prediction of general diseases.

### **Data Analysis and Data Mining**

The Data Mining is a process in which raw data is prepared and structured from the unstructured data as to take meaningful information from the data which can be used in the project. Task of making data organized and reflective about data is to way to get what this information does the data contains in it and what it does not have in it. There are so many different types of methods in which the people can make use of data analysis. It is simply very easy to use data during the analysis phase and get to some certain conclusions or some agendas. The analysis of data is a process of inspecting, cleaning, transforming, and modelling data with the objective of highlighting useful information, suggesting conclusions, and supporting decision making which are helpful to the user. Data analysis has multiple facets and approaches, encompassing diverse techniques under an array of names, in different business, science, and social science domains.

### **Machine Learning Algorithms**

KNN Algorithm

Random Forest Algorithm

XG Boost Algorithm

**1.2** **Problem Statement**

Many of the existing machine learning models for health care analysis are concentrating on one disease per analysis. For example first is for liver analysis, one for cancer analysis, one for lung diseases like that. If a user wants to predict more than one disease, he/she has to go through different sites. There is no common system where one analysis can perform more than one disease prediction. Some of the models have lower accuracy which can seriously affect patients’ health. When an organization wants to analyse their patient’s health reports, they have to deploy many models which in turn increases the cost as well as time Some of the existing systems consider very few parameters which can yield false results.

**1.3 Objective**

In multiple disease prediction, it is possible to predict more than one disease at a time. So the user doesn’t need to traverse different sites in order to predict the diseases. We are taking three diseases that are Liver, Diabetes, and Heart. As all the three diseases are correlated to each other. To implement multiple disease analyses we are going to use machine learning algorithms and Django. When the user is accessing this API, the user has to send the parameters of the disease along with the disease name. Django will invoke the corresponding model and returns the status of the patient.

**1.4 FUTURE SCOPE**

\*In the future we can add more diseases in the existing API

**\***We can try to improve the accuracy of prediction in order to decrease the mortality rate.

**CHAPTER 2- LITERATURE REVIEW**

**\***REVIEW 1

**Machine Learning Applications in Nephrology: A Bibliometric Analysis Comparing Kidney Studies to Other Medicine Subspecialities**

by-Ashish Verma, Vipul C. Chitalia, Sushrut S. Waikar, and Vijaya B. Kolachalama

Machine learning is quickly becoming a crucial component of the toolkit for data analysis in a wide range of medical applications. With developments advanced machine learning in hardware and software Deep neural network frameworks are being thought as as processing a variety of biomedical data. 1, 2 Several instances of renal health and kidney illnesses include the utilisation of electronic health record data and machine learning to predict acute renal injury,3,4 Combining deep learning and digitally captured human kidney biopsies to distinguish between kidney structures 5-8 and to foretell radiological phenotypes, clinical phenotypes, and imaging data to estimate renal volume overall. Machine learning is an exciting research tool that increasingly is considered for early detection, disease diagnosis, and clinical management. We explored the use of machine learning–driven advancements in nephrology compared with other medical subspecialties. We did a bibliometric analysis employing a Web of Science database using specific search terms for organ systems and research methods. Our analysis suggested the lowest use of machine learning in nephrology compared with other medical subspecialties. Our study results highlight the importance of informing the kidney research community about this emerging data analytic tool Because all the data used in this study is available to the public and does not contain any protected health information, we did not seek institutional review board approval. The study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines for cross-sectional studies.

**\***REVIEW 2

**The role of machine learning in clinical research: transforming the future of evidence generation**

by-E. Hope Weissler, Tristan Naumann, Tomas Andersson, Rajesh Ranganath, Olivier Elemento, Yuan Luo , Daniel F. Freitag , James Benoit , Michael C. Hughes , Faisal Khan , Paul Slater, Khader Shameer , Matthew Roe, Emmette Hutchison , Scott H. Kollins , Uli Broedl, Zhaoling Meng, Jennifer L. Wong, Lesley Curtis , Erich Huang and Marzyeh Ghassemi

Machine learning has the potential to help improve the success, generalizability, patient-centeredness, and efficiency of clinical trials. Various ML approaches are available for managing large and heterogeneous sources of data, identifying intricate and occult patterns, and predicting complex outcomes. As a result, ML has value to add across the spectrum of clinical trials, from preclinical drug discovery to pre-trial planning through study execution to data management and analysis (Fig. 2). Despite the relative lack of academic and lay publications focused on ML-enabled clinical research (vìs-a-vìs the attention to ML in care delivery), the profusion of established and start-up companies devoting significant resources to the area indicates a high level of interest in, and burgeoning attempts to make use of, ML application to clinical research, and specifically clinical trials. Successful clinical trials require significant preclinical investigation and planning, during which promising candidate molecules and targets are identified and the investigational strategy to achieve regulatory approval is defined. Missteps in this phase can delay the identification of promising drugs or doom clinical trials to eventual failure. ML can help researchers leverage previous and ongoing research to decrease the inefficiencies of the preclinical process.

**\***REVIEW 3

**Big data and machine learning algorithms for health-care delivery**

by- Kee Yuan Ngiam, Ing Wei Khor

Analysis of big data by machine learning offers considerable advantages for assimilation and evaluation of large amounts of complex health-care data. However, to effectively use machine learning tools in health care, several limitations must be addressed and key issues considered, such as its clinical implementation and ethics in health-care delivery. Advantages of machine learning include flexibility and scalability compared with traditional biostatistical methods, which makes it deployable for many tasks, such as [risk stratification](https://www.sciencedirect.com/topics/medicine-and-dentistry/risk-stratification), diagnosis and classification, and survival predictions. Another advantage of machine learning algorithms is the ability to analyse diverse data types (eg, demographic data, laboratory findings, imaging data, and doctors' free-text notes) and incorporate them into predictions for disease risk, diagnosis, prognosis, and appropriate treatments. Despite these advantages, the application of machine learning in health-care delivery also presents unique challenges that require data pre-processing, model training, and refinement of the system with respect to the actual clinical problem. Also crucial are ethical considerations, which include medico-legal implications, doctors' understanding of machine learning tools, and data privacy and security. In this Review, we discuss some of the benefits and challenges of big data and machine learning in health care.

**\***REVIEW 4

**Predicting self-intercepted medication ordering errors using machine learning**

by- Christopher Ryan KingId , Joanna Abraham, Bradley A. Fritz , Zhicheng Cui , William Galanter, Yixin Chen , Thomas Kannampallil

Current approaches to understanding medication ordering errors rely on relatively small manually captured error samples. These approaches are resource-intensive, do not scale for computerized provider order entry (CPOE) systems, and are likely to miss important risk factors associated with medication ordering errors. Previously, we described a dataset of CPOE-based medication voiding accompanied by univariable and multivariable regression analyses. However, these traditional techniques require expert guidance and may perform poorly compared to newer approaches. In this paper, we update that analysis using machine learning (ML) models to predict erroneous medication orders and identify its contributing factors. We retrieved patient demographics (race/ethnicity, sex, age), clinician characteristics, type of medication order (inpatient, prescription, home medication by history), and order content. We compared logistic regression, random forest, boosted decision trees, and artificial neural network models. Model performance was evaluated using area under the receiver operating characteristic curve (AUROC) and the area under the precision-recall curve (AUPRC). The dataset included 5,804,192 medication orders, of which 28,695 (0.5%) were voided. ML correctly classified voids at reasonable accuracy; with a positive predictive value of 10%, ~20% of errors were included. Gradient boosted decision trees achieved the highest AUROC (0.7968) and AUPRC (0.0647) among all models. Logistic regression had the poorest performance. Models identified predictive factors with high face validity (e.g., student orders), and a decision tree revealed interacting contexts with high rates of errors not identified by previous regression models. Prediction models using order-entry information offers promise for error surveillance, patient safety improvements, and targeted clinical review. The improved performance of models with complex interactions points to the importance of contextual medication ordering information for understanding contributors to medication errors.

**\***REVIEW 5

**Significance of machine learning in healthcare: Features, pillars and applications**

by- Mohd Javaid , Abid Haleem , Ravi Pratap Singh , Rajiv Suman , Shanay Rab

Machine Learning (ML) applications are making a considerable impact on healthcare. ML is a subtype of Artificial Intelligence (AI) technology that aims to improve the speed and accuracy of physicians’ work. Countries are currently dealing with an overburdened healthcare system with a shortage of skilled physicians, where AI provides a big hope. The healthcare data can be used gainfully to identify the optimal trial sample, collect more data points, assess ongoing data from trial participants, and eliminate data-based errors. ML-based techniques assist in detecting early indicators of an epidemic or pandemic. This algorithm examines satellite data, news and social media reports, and even video sources to determine whether the sickness will become out of control. Using ML for healthcare can open up a world of possibilities in this field. It frees up healthcare providers’ time to focus on patient care rather than searching or entering information. This paper studies ML and its need in healthcare, and then it discusses the associated features and appropriate pillars of ML for healthcare structure. Finally, it identified and discussed the significant applications of ML for healthcare. The applications of this technology in healthcare operations can be tremendously advantageous to the organisation. ML-based tools are used to provide various treatment alternatives and individualised treatments and improve the overall efficiency of hospitals and healthcare systems while lowering the cost of care. Shortly, ML will impact both physicians and hospitals. It will be crucial in developing clinical decision support, illness detection, and personalised treatment approaches to provide the best potential outcomes.

\*REVIEW 6

**Machine learning in health care and laboratory medicine: General overview of supervised learning and Auto-ML**

by- Hooman H. Rashidi , Nam Tran , Samer Albahra , Luke T. Dang

Artificial Intelligence (AI) and machine learning (ML) have now spawned a new field within health care and health science research. These new predictive analytics tools are starting to change various facets of our clinical care domains including the practice of laboratory medicine. Many of these ML tools and studies are also starting to populate our literature landscape as we know it but unfamiliarity of the average reader to the basic knowledge and critical concepts within AI/ML is now demanding a need to better prepare our audience to such relatively unfamiliar concepts. A fundamental knowledge of such platforms will inevitably enhance cross-disciplinary literacy and ultimately lead to enhanced integration and understanding of such tools within our discipline. In this review, we provide a general outline of AI/ML along with an overview of the fundamental concepts of ML categories, specifically supervised, unsupervised, and reinforcement learning. Additionally, since the vast majority of our current approaches within ML in laboratory medicine and health care involve supervised algorithms, we will predominantly concentrate on such platforms. Finally, the need for making such tools more accessible to the average investigator is becoming a major driving force for the need of automation within these ML platforms. This has now given rise to the automated ML (Auto-ML) world which will undoubtedly help shape the future of ML within health care. Hence, an overview of Auto-ML is also covered within this manuscript which will hopefully enrich the reader's understanding, appreciation, and the need for embracing such tools.

\*REVIEW 7

**Interpretable Machine Learning in Healthcare**

by- Muhammad Aurangzeb Ahmad, Carly Eckert, Ankur Teredesai, and Greg McKelvey

The drive towards greater penetration of machine learning in healthcare is being accompanied by increased calls for machine learning and AI based systems to be regulated and held accountable in healthcare. Interpretable machine learning models can be instrumental in holding machine learning systems accountable. Healthcare offers unique challenges for machine learning where the demands for explainability, model fidelity and performance in general are much higher as compared to most other domains. In this paper we review the notion of interpretability within the context of healthcare, the various nuances associated with it, challenges related to interpretability which are unique to healthcare and the future of interpretability in healthcare.

\*REVIEW 8

**Secure and Robust Machine Learning for Healthcare: A Survey**

by-Adnan Qayyum , Junaid Qadir , Muhammad Bilal , and Ala Al-Fuqaha

Recent years have witnessed widespread adoption of machine learning (ML)/deep learning (DL) techniques due to their superior performance for a variety of healthcare applications ranging from the prediction of cardiac arrest from one-dimensional heart signals to computer-aided diagnosis (CAD) using multi-dimensional medical images. Notwithstanding the impressive performance of ML/DL, there are still lingering doubts regarding the robustness of ML/DL in healthcare settings (which is traditionally considered quite challenging due to the myriad security and privacy issues involved), especially in light of recent results that have shown that ML/DL are vulnerable to adversarial attacks. In this paper, we present an overview of various application areas in healthcare that leverage such techniques from security and privacy point of view and present associated challenges. In addition, we present potential methods to ensure secure and privacy-preserving ML for healthcare applications. Finally, we provide insight into the current research challenges and promising directions for future research.

\*REVIEW 9

**Causal machine learning for healthcare and precision medicine**

**by-**Pedro Sanchez, Jeremy P. Voisey, Tian Xia1 , Hannah I. Watson , Alison Q. O’Neil1,2 and Sotirios A. T saftaris

## Causal machine learning (CML) has experienced increasing popularity in healthcare. Beyond the inherent capabilities of adding domain knowledge into learning systems, CML provides a complete toolset for investigating how a system would react to an intervention (e.g. outcome given a treatment). Quantifying effects of interventions allows actionable decisions to be made while maintaining robustness in the presence of confounders. Here, we explore how causal inference can be incorporated into different aspects of clinical decision support systems by using recent advances in machine learning. Throughout this paper, we use Alzheimer’s disease to create examples for illustrating how CML can be advantageous in clinical scenarios. Furthermore, we discuss important challenges present in healthcare applications such as processing high-dimensional and unstructured data, generalization to out-of-distribution samples and temporal relationships, that despite the great effort from the research community remain to be solved. Finally, we review lines of research within causal representation learning, causal discovery and causal reasoning which offer the potential towards addressing the aforementioned challenges.

\*REVIEW 10

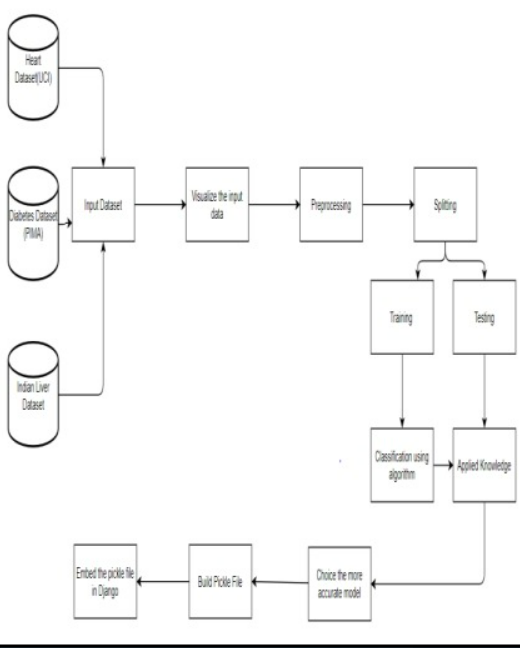
**A MACHINE LEARNING APPROACH TO LOW-VALUE HEALTH CARE: WASTED TESTS, MISSED HEART ATTACKS AND MIS-PREDICTIONS**

By-Sendhil Mullainathan Ziad Obermeyer

We use machine learning to better characterize low-value health care and the decisions that produce it. We focus on costly tests, specifically for heart attack (acute coronary syndromes). A test is only useful if it yields new information, so efficient testing is grounded in accurate prediction of test outcomes. Physician testing decisions can therefore be benchmarked against tailored algorithmic predictions, which provide a more precise way to study low-value care than the usual approach—looking at average test yield. Implemented in a large national sample, this procedure reveals significant over-testing: 52.6%of high-cost tests for heart attack are wasted. At the same time, it also reveals significant under-testing: many patients with predictably high risk go untested, then experience frequent adverse cardiac events including death in the next 30 days. At standard clinical thresholds, these event rates suggest that testing these patients would indeed have been highly cost-effective. Of the potential welfare gains from more efficient testing, 42.8% would come from addressing under-use. Existing policy levers, however, appear too blunt a tool to address both over- and under-use inefficiencies. We find that they cut testing across the board, for low-risk (reducing over-use) and high-risk patients (exaggerating under- use). Finally, we uncover two behavioural mechanisms for physician testing errors: (i) bounded rationality, in which physicians use an overly narrow set of variables, but make effective use of that set; and (ii) representativeness, in which they over-weight how \representative" heart attack is for a patient, above and beyond the conditional probability. Together, these results suggest the need for models of low-value care that incorporate mis-prediction so as to account for both over- and undertesting.

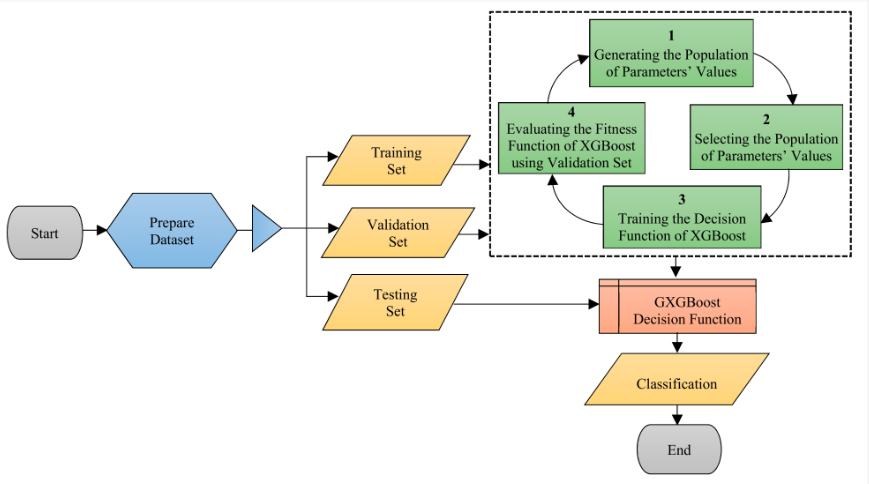
**CHAPTER 3-PROPOSED METHODOLOGY**

**3.1Flowchart**

****

The first step is to the dataset for heart disease, diabetes disease and liver

disease we have imported the UCI dataset, PIMA dataset and Indian liver dataset respectively. Once we have imported the dataset then visualization of each imputed data takes place. After visualization pre-processing of data takes place where we check for outliers, missing values and also scale the dataset then on the updated dataset we split the data into training and testing. Next is on the training dataset we had applied K-NN and random forest algorithm and applied knowledge on the classified algorithm using testing dataset. Then we build a pickle file for all the disease and then integrated the pickle file with the Django framework for the output of the model on the webpage.

****

**3.2Algorithm Proposed**

\*KNN Algorithm

\*Random Forest Algorithm

\*XG Boost Algorithm

**CHAPTER 4- TECHNOLOGY USED**

## **Django Framework**

The Django is a high-level Python framework that helps in the rapid development and clean, pragmatic design, django makes it easier to build applications more quickly, efficiently and with less code. Django is used for creating the User Interface (UI) for the application. The UI created by Django is easy to use so that the person which are from the non-technical field can also use the application for the prediction of disease without going anywhere any saving time and money

**Heart disease**

Urban areas record between 400 or 500 cases in every 100,000 people, while rural populations record 100 cases per 100,000 people.

****

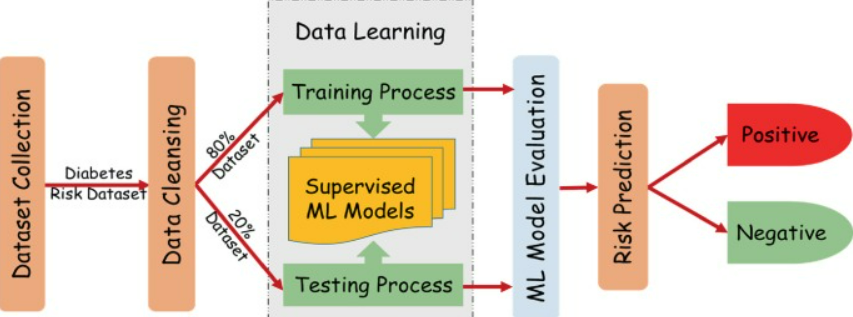
**Liver Disease**

Liver harm is the one of the best deadliest ailment on the planet. The fundamental driver of liver harm are Fatty liver, Liver Fibrosis, Cirrhosis, hepatitis and diseases

****

**Diabetes**

Currently, 25.2 million adults are estimated to have IGT, which is estimated to increase to 35.7 million in the year 2045. India ranks second after China in the global diabetes epidemic with 77 million people with diabetes.

****

**CHAPTER 6- CONCLUSION**

The main objective of this project was to create a system that would predict more than one disease and do so with high accuracy. Because of this project the user doesn’t need to traverse different websites which saves time as well. Diseases if predicted early can increase your life expectancy as well as save you from financial troubles. For this purpose, we have used various machine learning algorithms like Random Forest, XG Boost, and K nearest neighbour (KNN) to achieve maximum accuracy.

**REFERENCES**

• Esteva, A. Robicquet, B. Ramsundar, et al

• F. Wang, L.P. Casalino, D. Khullar • N. Tomašev, X. Glorot, J.W. Rae, et al

• J.L. Koyner, K.A. Carey, D.P. Edelson, M.M. Churpek • Ginley, B. Lutnick, K.Y. Jen, et al.

• S. Chaudhuri, D. Oudejans, H. Thompson and G. Demiris, "Real-World Accuracy and Use of a Wearable Fall Detection Device by Older Adults", Journal of the American Geriatrics Society, vol. 63, no. 11, pp. 2415-2416, 2015.

• T. Dietterich, "Editorial Exploratory research in machine learning", Machine Learning, vol. 5, no. 1, 5-9,

• K. Liao, T. Cai, G. Savova, S. Murphy, E. Karlson, A. Ananthakrishnan, V. Gainer, S. Shaw, Z. Xia, P. Szolovits, S. Churchill and I. Kohane, "Development of phenotype algorithms using electronic medical records and incorporating natural language processing", BMJ, vol. 350, no. 2411, pp. h1885-h1885, 2015.

• G. Meyfroidt, F. Güiza, J. Ramon and M. Bruynooghe, "Machine learning techniques to examine large patient databases", Best Practice & Research Clinical Anaesthesiology, vol. 23, no. 1, pp. 127-143, 2009.

• R. Pirracchio, M. Petersen and M. van der Laan, "Improving Propensity Score Estimators' Robustness to Model Misspecification Using Super Learner", American Journal of Epidemiology, vol. 181, no. 2, pp. 108-119, 2014

• Samuel AL. Some studies in machine learning using the game of checkers. IBM J Res Dev 1959; 3: 210–29.

• LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015; 521: 436–44.

• Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. Health Aff (Millwood) 2014; 33: 1123–31.

• Faust O, Hagiwara Y, Hong TJ, Lih OS, Acharya UR. Deep learning for healthcare applications based on physiological signals: a review. Comput Methods Programs Biomed 2018; 161: 1–13.

• Jha S, Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists. JAMA 2016; 316: 2353–54.

• Ammenwerth E, Schnell-Inderst P, Machan C, Siebert U. The effect of electronic prescribing on medication errors and adverse drug events: a systematic review. JAMIA. 2008;15(5):585–600. pmid:18579832

• Bates DW, Leape LL, Cullen DJ, Laird N, Petersen LA, Teich JM, et al. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. Jama. 1998;280(15):1311–6. Epub 1998/10/30. pmid:9794308.

• Bates DW, Teich JM, Lee J, Seger D, Kuperman GJ, Ma’Luf N, et al. The impact of computerized physician order entry on medication error prevention. JAMIA. 1999;6(4):313–21. pmid:10428004

• Kuperman GJ, Gibson RF. Computer physician order entry: benefits, costs, and issues. Annals of internal medicine. 2003;139(1):31. pmid:12834316

• van Rosse F, Maat B, Rademaker CM, van Vught AJ, Egberts AC, Bollen CW. The effect of computerized physician order entry on medication prescription errors and clinical outcome in pediatric and intensive care: a systematic review. Pediatrics. 2009;123(4):1184–90. pmid:19336379

• A. Abdelaziz, M. Elhoseny, A.S. Salama, A.M. Riad

• D.S. Char, M.D. Abràmoff, C. Feudtner

• M.A. Ahmad, C. Eckert, A. Teredesai

• P. Kaur, M. Sharma, M. Mittal

• M.A. Sarwar, N. Kamal, W. Hamid, M.A. Shah

• Shalf J. The future of computing beyond Moore's Law. Philos Trans A Math Phys Eng Sci. 2020; 378(2166): 20190061. • Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial intelligence and machine learning in pathology: the present landscape of supervised methods. Acad Pathol. 2019; 3(6).

• van Otterlo M, Wiering M (2012) Reinforcement Learning and Markov Decision Processes. In: M Wiering, M van Otterlo, eds. Reinforcement Learning. Adaptation, Learning, and Optimization, vol 12. Heidelberg: Springer, Berlin.

• Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018; 2(10): 719- 731.

• Kratz A, Lee SH, Zini G, Riedl JA, Hur M, Machin S. International Council for Standardization in Haematology. Digital morphology analyzers in hematology: ICSH review and recommendations. Int J Lab Hematol. 2019; 41(4): 437- 447.

• Ahmad, Muhammad Aurangzeb, Zoheb Borbora, Jaideep Srivastava, and Noshir Contractor. "Link prediction across multiple social networks." In Data Mining Workshops (ICDMW), 2010 IEEE International Conference on, pp. 911--918. IEEE, 2010.

• l-Shedivat, Maruan, Avinava Dubey, and Eric P. Xing. "Contextual Explanation Networks." arXiv preprint arXiv:1705.10301 (2017).

• l-Shedivat, Maruan, Avinava Dubey, and Eric P. Xing. "The Intriguing Properties of Model Explanations

• S. Latif, J. Qadir, S. Farooq, and M. Imran, “How 5G wireless (and concomitant technologies) will revolutionize healthcare?” Future Internet, vol. 9, no. 4, p. 93, 2017.

• Z. Yan et al., “Multi-instance deep learning: Discover discriminative local anatomies for bodypart recognition,” IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1332–1343, May 2016.

• M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, “Lung pattern classification for interstitial lung diseases using a deep convolutional neural network,” IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1207–1216, May 2016.

• W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, “Multi-scale convolutional neural networks for lung nodule classification,” in Proc. Int. Conf. Inf. Process. Med. Imag., 2015, pp. 588–599.

• Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. 2017A survey on deep learning in medical image analysis. Med. Image Anal. 42, 60-88.

• Bica I, Alaa AM, Lambert C, Schaar M. 2021From real-world patient data to individualized treatment effects using machine learning: current and future methods to address underlying challenges. Clin. Pharmacol. Ther. 109, 87-100.

• Holland PW. 1986Statistics and causal inference. J. Am. Stat. Assoc. 81, 945-960.

• Abaluck, Jason et al. (2016). “The Determinants of Productivity in Medical Testing: Intensity and Allocation of Care”. In: American Economic Review 106.12, pp. 3730– 3764.

• Abelson, Reed and Julie Creswell (2012). “Hospital chain inquiry cited unnecessary cardiac work”. In: New York Times 6.

• Al-Lamee, Rasha et al. (2018). “Percutaneous coronary intervention in stable angina (ORBITA): a double-blind, randomised controlled trial”. In: The Lancet 391.10115, pp. 31–40