iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** I **Month of publication:** January 2024

DOI: https://doi.org/10.22214/ijraset.2024.58037

www.ijraset.com

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

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue I Jan 2024- Available at www.ijraset.com

Risk Assess: A Symptom-Based Disease Predictor

Arun Kumar S[1], Niranjani K[2], Harsha M[3], Harish NG[4]
[1]*Assistant Professor, Department of Computer Science and Engineering, Presidency University, Bangalore, India*
[2, 3, 4]*Student, Department of Computer Science and Engineering, Presidency University, Bangalore, India*

Abstract: ***As of late, the medical care area has seen extraordinary changes through the joining of state of data collection. This study presents "Risk Assess", an imaginative web application fastidiously created with Flask. "Risk Assess" fills in as a complete stage, working with patient vitals for generating risk assessments. This python-based web application uses different Machine Learning models such as Classification and Regression (CART), Linear Support Vector Machine (SVM), Gaussian Naïve Bayes (NB), K-Nearest Neighbor (KNN) to analyze and build a model to predict if the given set of symptoms leads to a particular disease. Outstandingly, "Risk Assess" soothes out Cancer, Diabetes, Heart –Disease, Kidney-Disease and Liver-Disease based prediction upgrading comfort and functional productivity. Our undertaking highlights a solid login page empowering clients to make novel qualifications utilizing their email addresses. By requiring a username and secret key attached to their email, we focus on both security and client comfort. This research paper contains working engineering, plan standards, and execution complexities, featuring the pivotal job of Machine Learning. "Risk Assess" epitomizes utilization of Python-based advances and Machine Learning, exhibiting the potential for predictions based on symptoms through creative computerized arrangements.***

Keywords: Risk Assess, Python-based web application, Machine Learning, Flask, Symptoms, Disease Prediction, Risk assessment, Security.

I. INTRODUCTION

In the developing scene of medical services, the reconciliation of Machine Learning has arisen as a ground-breaking power, promising exceptional headways in diagnostics and therapy. The adequacy of Machine Learning models, especially in sickness finding, is unpredictably attached to the accessibility of broad and dependable datasets that exemplify the assorted side effects and appearances of different ailments. This examination tries to add to this development by introducing a thorough examination concerning the improvement of "Risk Assess" an internet based stage made utilizing Flask. The platform serves as a vital role for predicting the accuracy of various diseases, fostering a seamless result. "Risk Assess" adjusts progressively to the developing medical field. The web application's center advancement lies in its accurate prediction and end-to-end approach.

II. LITERATURE REVIEW

The field of Machine Learning in medical care depends vigorously on the accessibility of tremendous and solid datasets to upgrade demonstrative abilities. In this paper, we propose a strategy for joining the necessities of contemporary Machine Learning with those of conventional information combination techniques. We examine the possible results of utilizing metadata got from data consolidation processes for chipping away at the ampleness and efficiency of Machine Learning models. Towards this bearing, we look at two ordinary use cases over data storage facilities, incorporate increment and joined learning. We feature new examination valuable open doors according to the viewpoints of frameworks, portrayals, factorized learning, and united advancing by uniting Machine Learning and information reconciliation [1]. Existing information mining techniques and Machine Learning calculations need a significant amount of information to prepare, more information should be gained before they can be utilized. The multifaceted nature of the model influences the size of the document. while examining information is troublesome [2].

This work is instrumental in molding the mechanical system of the proposed web-based application, "Risk Assess", with an emphasis on lining up with the developing scene of computerized wellbeing rehearses.

A survey on Data Collection for Machine Learning, conducted on April 1 2021, by Y. Roh, G. Heo and S.E. Whang contributes essential guidance on tailoring the adaptive data collection for "Risk Assess" [3]. This paper digs into the basic convergence of Data Innovation (IT) developments, with thorough financial plan the board frameworks. The general goal is to smooth out information assortment, improve financial plan arrangement processes, and engage undertakings to form key spending arrangements. The paper investigates the combination of these innovations as well as gives execution ideas and shields to guarantee a smooth and secure progress. [4]

©IJRASET: All Rights are Reserved

This study dives into different information assortment methods and presents the implantation of Machine Learning calculations to enhance the life span of the proposed web-application. At the present time in Distant Sensor Associations, data collection expects a critical part. The principal job of data collection is to get a gigantic measure of data while to hack down data dropping in view of sensor center memory hindrances. [5].

The essential examination issue in the fields of Machine Learning and information mining is lopsided information characterization. As per the perspective of practical application, unbalanced instructive assortments have a colossal degree. During the grouping, the classifier will predisposition the examples of most of classes, which will at last straightforwardly affect the arrangement brings about request to control the misfortune pace of separation. According to the investigation eventual outcomes of local and new researchers lately, most of the proposed lopsided data computations can really control the impact of data cumbersomeness on the presentation of the classifier. Thus, in light of understanding the assessment outcomes of lopsided data gathering, this paper made sense of the investigation course of computer based intelligence and proposed an estimation considering de-noising oversampling advancement. [6]

For this situation, we assembled remark information by utilizing the information assortment programming. The information was then exposed to pre-handling, word division, and opinion marking. We make a feeling classifier to consequently characterize unidentified information by changing different boundaries used to prepare the information. We show probably that using artificial intelligence methodologies to perform assessment request on thing reviews can achieve incredible results. this article adds to the expanding field of profound examination in item surveys, revealing insight into the adequacy of Machine Learning approaches. [7]

Various fields are utilizing Machine Learning to address different issues. This is being driven by the accessibility of huge datasets, minimal expense calculation assets, and the improvement of vigorous AI calculations. Some computer based intelligence applications require association of contraptions off-the-cross section for data collection and dealing with. Frameworks that can work freely during arrangement are required for these applications. The utilization of a few open-source sheets for off-network information on the basis of assortment and Machine Learning is shown in this paper [8]

By and large, the exactness of the investigation results is affected by the consistency and uprightness of the information during the course of huge information examination. Usually before the data assessment process starts, data gathering should be made on the assembled data sources to hold that resulting assessment won't make goofs there of brain due data peculiarities. Therefore, one of those vital undertakings in information grouping is to keep up with respectability [9]

Drawing insights "Amalur: Data Integration Meets Machine Learning," 2023 IEEE 39th International Conference on Data Engineering [1], and “A Survey on Data Collection for Machine Learning” [3], the platform prioritizes a user-friendly interface for patients. Data Acquisition and Processing System, inspired by a 2022study by C. Zheng, M. Zhang, Y. Wang and M Zou from Bangkok, Thailand, further enhances the holistic well-being aspect. [4]

Based on a Comprehensive study on Distinct Data Collection Techniques using Machine Learning Approaches for WSN in 2021 [5], “Risk Assess” adapts dynamically to the evolving healthcare landscape. The platform's core innovation lies in its symptom-based disease prediction accurately using machine learning models and also providing authentication.

The subject of close to home examination of item surveys is a hot area of information mining research at this moment. Whether it is in researcher or money related fields, text up close and personal examination of online business thing reviews has remarkable investigation regard. This article used simulated intelligence to lead assessment examination for clothing web business thing reviews. For this situation, we assembled remark information by utilizing the information assortment programming. The information was then exposed to pre-handling, word division, and opinion marking. We make a feeling classifier to consequently characterize unidentified information by changing different boundaries used to prepare the information. We show probably that using artificial intelligence methodologies to perform assessment request on thing reviews can achieve incredible results. this article adds to the expanding field of profound examination in item surveys, revealing insight into the adequacy of Machine Learning approaches. [7]

The excursion incorporates the cautious coordination of information assortment, pre-handling, and feeling marking, finishing in the making of an opinion classifier that stands ready to explore the intricacies of close to home articulation inside the online business space. Through observational trial and error, the review approves the reasonability of utilizing Machine Learning for opinion order, denoting a huge step in the continuous investigation of the convergence among innovation and human feeling in the computerization. [8]

In addition, our examination underscores the job of patient commitment through portable wellbeing applications. We propose the advancement of easy to understand applications, enabling patients to willfully contribute side effect information, track wellbeing measurements, and give express agree to information utilization.

To additionally improve ongoing wellbeing information, including imperative signs and action levels. All through the paper, serious areas of strength are put on for executing hearty safety efforts to protect the classification and trustworthiness of the gathered information. [9]

The trial results show the way that the proposed calculation can really improve the classifier's characterization impact and has a specific enemy of commotion execution. this paper not just recognizes the criticality of tending to lopsided information characterization in Machine Learning yet additionally adds to the continuous talk by proposing a clever calculation in light of de-noising oversampling innovation. The exploration course illustrated in this paper lines up with the more extensive endeavors to refine Machine Learning philosophies, especially with regards to genuine world, imbalanced datasets. The trial approval fills in as observational proof of the calculation's viability, offering a promising road for additional investigation and reception in common sense applications. As the field keeps on developing, the proposed calculation remains as a demonstration of the imaginative arrangements arising to address the difficulties of unequal information order in contemporary AI research. [6]

As we dive into the crossing point of innovation and moral contemplations, this work attempts to impel the field of Machine Learning in medical care forward, encouraging developments that are both logically thorough and morally sound. To address privacy concerns, the project emphasizes the use of SQL technology for the use of secure stored data. By anonymously collecting patient data, we prioritize the protection of individual privacy while still extracting valuable insights that help strengthen our dataset. Furthermore, this framework recognizes the importance of transparency and credibility in the data collection process. An education campaign will be launched to inform and engage patients, ensuring they understand the purpose of the project, the safeguards in place and the potential benefits of their contribution to medical research. In conclusion, this project represents a concerted effort to address the challenges associated with collection.In conclusion, these e-health websites are supposed to predict the occurrence of diseases such as cancer, diabetes, heart disease, kidney disease and liver disease by collecting symptoms of patients in the form of data and also provide information and precautionary measure for. This study talks about a very elegant patient friendly app which can be used for self-checking. Simply put, having super accurate information helps them figure out what to do next for patients.

III. PROPOSED WORK

This platform is designed to provide symptom based predictions using machine learning, offering features like easy input of data which does not involve complexities. The undertaking's improvement draws on bits of knowledge from different insightful works, with a solid accentuation on user centered plan standards. Inspired by industry platforms like “Apollo 247” and “Orange health lab test at home”, the “Risk Assess” project adopts a distinctly human-centric approach, aiming to enhance healthcare delivery with our application, we mean to give a one-stop restorative and health care to patients. In the “Risk Assess” project, our emphasis is to give accurate predictions for cancer, diabetes, heart disease, kidney disease and liver disease by collecting patient’s data, using Machine Learning Algorithms – CART, SVM and KNN.

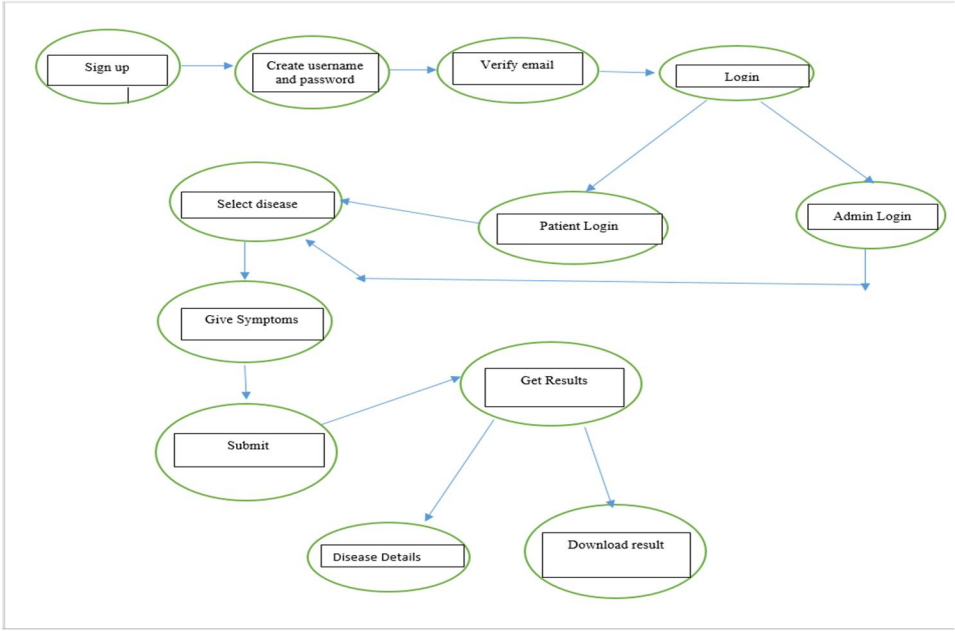


Fig 1. Flow Diagram

Users of "Risk Assess" application can easily sign up through our website using their email id. They can create their own username and password for smooth login experiences in future requirements. After which their email id will be verified for authentication purposes to enhance security. Once the user has successfully signed up and created an account, they can login easily with their username and password created by them. Login can be done by patient as well as by the admin. An admin logs in on behalf of patient, if the patient faces any difficulties during the sign up process and also uploads patient's symptoms vitals. After which he passes on the result to the patient by downloading it. Similarly patient can also login, give their symptom details and view their accurately predicted results. They can even download it for further requirements. We use various kinds of Machine Learning Models such as CART, which is decision tree predictive model used to predict the target variable's value based on the factors provided, SVM, which is used to handle both classification and regression, NB, which is used for text classification and KNN, which is used to obtain highly accurate predictions.The architecture of this application involves various modules. These modules include:

- Admin Module
- Patient Module
- Symptoms Module
- Result Module

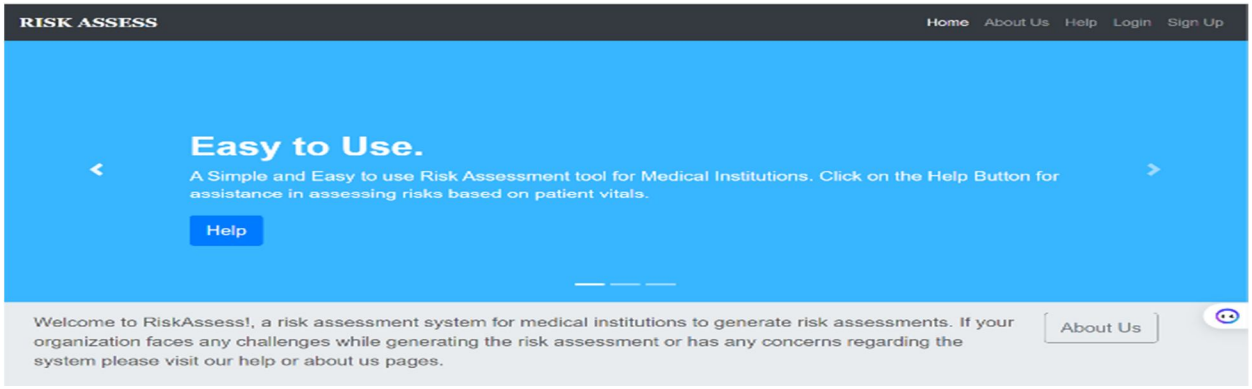


Fig 2. Home Page

A. Admin Module

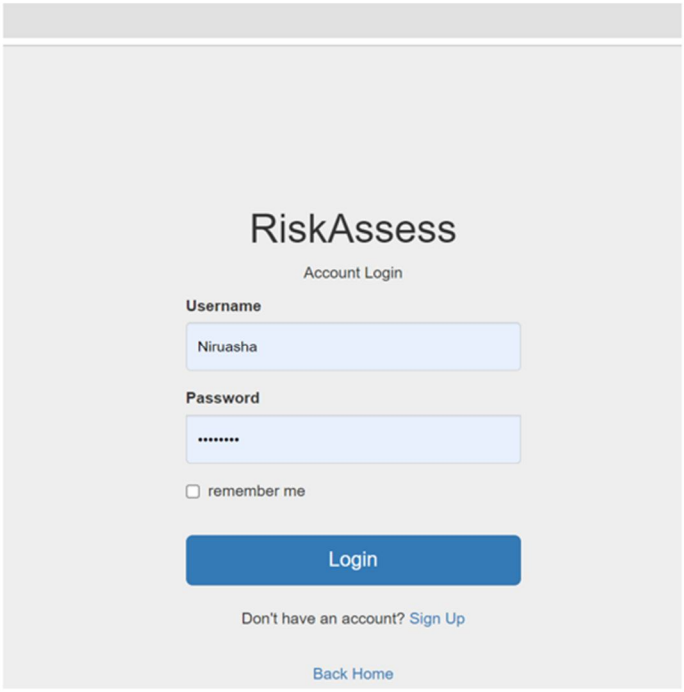


Fig 3. Admin Login Module

Used by admins to login on behalf of patients when they are unable to login due to any technical issues. They collect patient's symptom details and feed them into our application. Later on they provide the downloaded copy of the accurate predictions obtained from the application to the patient.

B. Patient Module

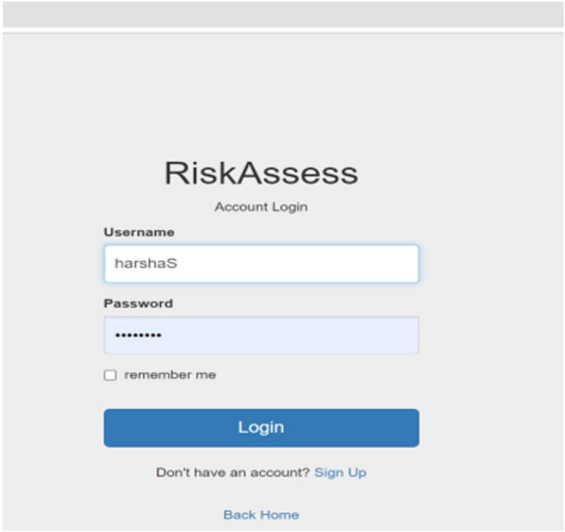


Fig 4. Patient Login Module

Once the patient register, they can easily login through the username and password created by them in the authentication phase. This enables smooth login each time. After logging in the patient can choose a particular disease from caner, diabetes, heart disease, liver and kidney disease.

C. Symptoms Module

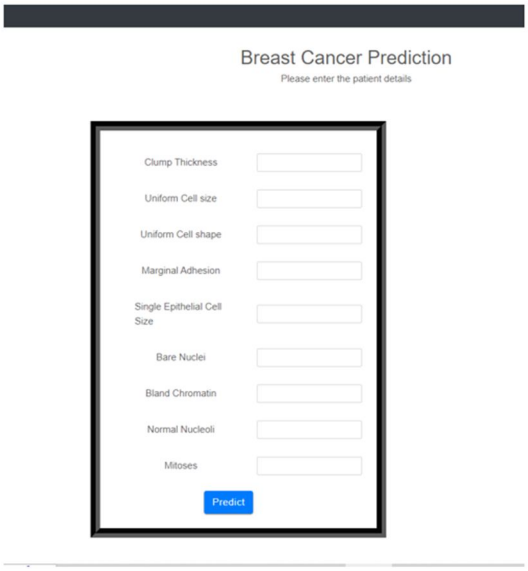


Fig 5. Symptoms Collection

In this module, patient's numeric data based on symptoms is collected for various diseases such as breast cancer, diabetes, heart disease, kidney disease and liver disease. They must enter each value manually and later click on predict button to get accurate predictions based on their symptoms.

D. Result Module

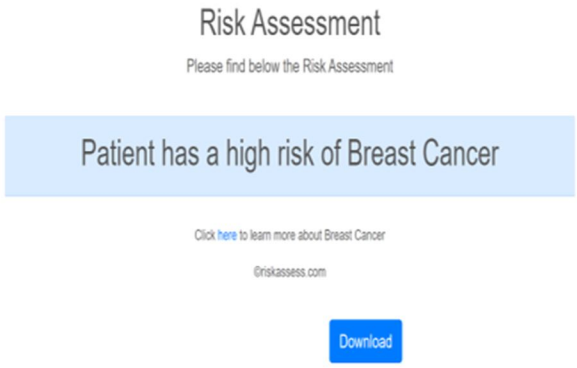


Fig 6. Result

In this module, patients can view their result based on the symptoms given by them. Through which they can even download and print their results. They can also learn more about what causes this disease and how they can prevent it further.

IV. METHODOLOGY

With the improvement of the general public, our lives have gotten more occupied and subsequently monitoring wellbeing, development and arrangement dates has turned into an undertaking. Notwithstanding this keeping up with the records is an issue, consequently we have concocted an application that takes care of the relative multitude of above issues, All the client needs to do is to make a profile and give their numeric symptom data, assuming that they have the past arrangement subtleties they can enter that too. The application will give you ideal tokens of the following arrangement date accordingly making the whole experience bother free. Notwithstanding the above includes the application giving symptom-based prediction using Machine Learning. The goal of our proposed strategy is to foster an exhaustive gamble evaluation application that gives exact forecasts to different sicknesses including malignant growth, diabetes, coronary illness, kidney infection, and liver infection. This application intends to offer a one-stop answer for patients looking for clinical and medical care administrations.

A. Data Collection

1) *Patient Information Assortment:* We will gather many patient information including clinical history, side effects, analytic experimental outcomes, way of life factors, and hereditary data.
2) *Data Sources:* Information will be obtained from patiens.

B. Machine Learning Algorithm

1) *Calculation Determination:* We will use a mix of regulated AI calculations like Order and Relapse Tress(CART), Liner support Vector Machines(SVM), KNN.
2) *Model Preparing:* The chose calculations will be prepared utilizing the gathered patient information to foster sickness expectation models intended for malignant growth, diabetes, coronary illness, kidney infection, and liver infection.
3) *Model Assessment:* The exhibition of the created models will be assessed utilizing measurements like exactness, accuracy, review, and F1 score to guarantee their adequacy in foreseeing sickness gambles.

C. Application Development

1) *User Connection Point Plan:* We will plan a natural and easy to understand interface for the gamble evaluation application to work with simple contribution of patient information and show of prescient outcomes.
2) *Integration of Models:* The created AI models will be incorporated into the application backend to empower constant forecasts in view of info information

D. Ethical Considerations

1) *Data Protection:* Severe measures will be carried out to guarantee the security and classification of patient information in consistence with medical services guidelines like HIPAA.
2) *Bias Moderation:* Steps will be taken to relieve predispositions in the gathered information and AI models to guarantee fair and unprejudiced forecasts.

V. RESULTS AND DISCUSSIONS

From the above screen captures, it is apparent that the normal result for this medical services application has been gotten. The normal results were Will be permitted to add a patient and specialist profile. The application will set up space for obtaining the patients symptoms values. It will feed these data into various machine learning models for accurate predictions. Which the patients can even download for future reference and can also learn about what caused them this disease in the 1st place. This multitude of results have been effectively carried out.

VI. CONCLUSION

In this innovative healthcare project, data is collected through a user-friendly website, and fed into a machine learning model. The project emphasizes user engagement, ethical data practices, and transparent communication. By fostering collaboration with healthcare providers, contributing to the development of a robust machine learning model. The scalable system design ensures adaptability to growing datasets and sustained user participation. Overall, this project represents a transformative approach to healthcare research, leveraging technology to empower individuals, advance disease diagnosis, and contribute to the broader healthcare ecosystem.

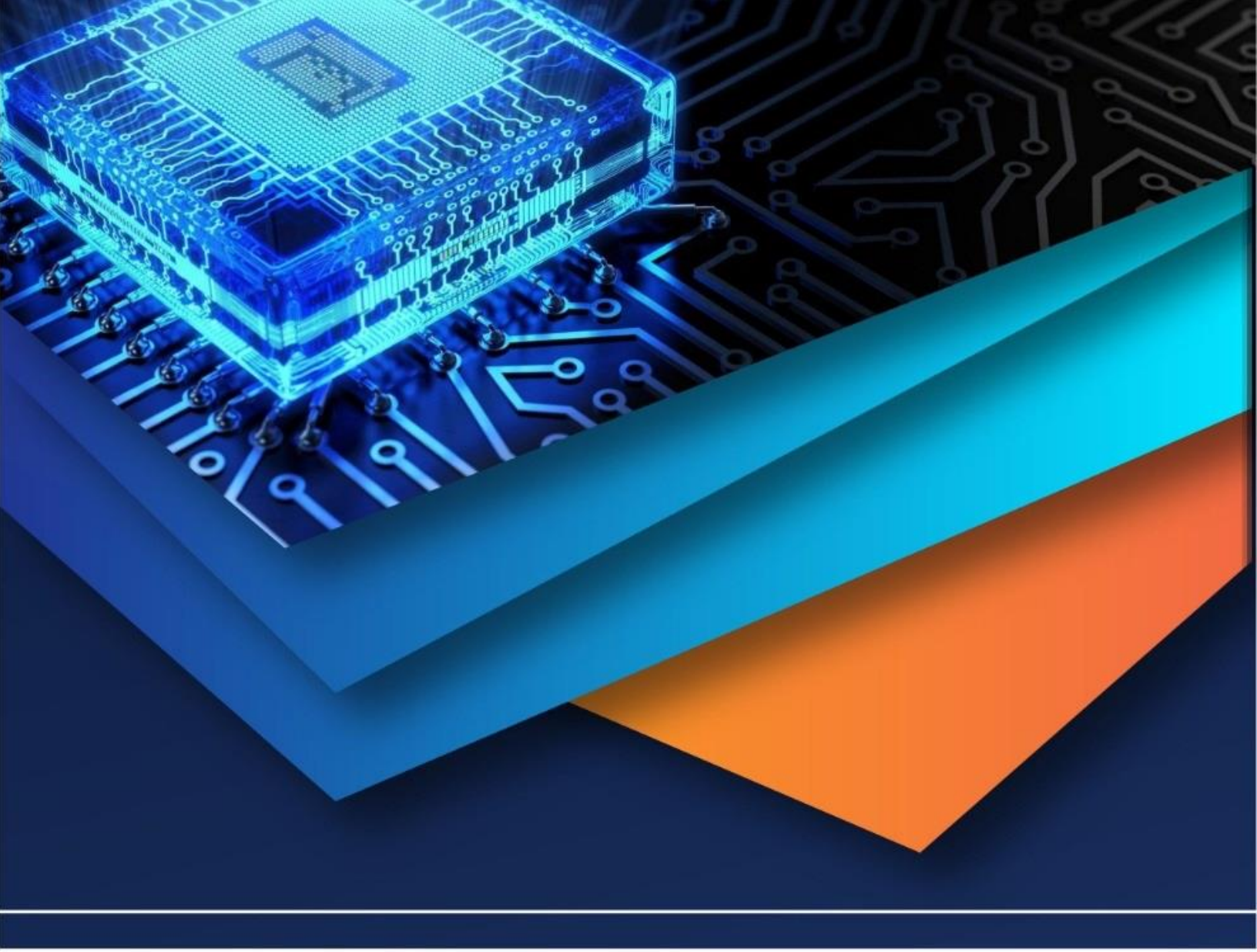
VII. FUTURE SCOPE

There is a ton of extension to foster this task further. The application can restrict with multispecialty emergency clinics on a membership base and permit its primary care physicians to utilize the application which will help in better administration and legitimate association of the information. During the conversation, it was likewise recommended to incorporate a stage that interfaces Patients with emergency clinic Day Cares. By adding these highlights, the application will turn into a one-stop stage for all patient's prerequisites

REFERENCES

[1] R. Hai et al., "Amalur: Data Integration Meets Machine Learning,"
[2] S. K. Jena, P. Sahu and S. Mishra, "Dynamic Data Mining for Multidimensional Data Based On Machine Learning Algorithms," 2021
[3] Y. Roh, G. Heo and S. E. Whang, "A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective, April 2021.
[4] C. Zheng, M. Zhang, Y. Wang and M. Zou, "Application of PRA and Machine Learning Algorithm in Budget Data Acquisition and Processing System," 2022.
[5] G. Sulakshana and G. R. Kamatam, "A Comprehensive study on Distinct Data Collection Techniques using Machine Learning Approaches for WSN," 2021
[6] Y. Zhou, W. Yang and X. Deng, "Research Status of Unbalanced Data Classification Based on Machine Learning," 2022
[7] Y. Hong and X. Shao, "Emotional Analysis of Clothing Product Reviews Based on Machine Learning," 2021
[8] G. Kiarie, J. Kabi, L. Mugambi and C. W. Maina, "The use of Open-Source Boards for Data Collection and Machine Learning in Remote Deployments,"
[9] Y. Liu and X. Gou, "A multi-model Imputation Approach for Missing Data Based on Machine Learning," 2022

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)