iJRASET
International Journal For Research in Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** XI **Month of publication:** November 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.56897

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 11 Issue XI Nov 2023- Available at www.ijraset.com*

# Different Disease Prediction

Aayush Mankar[1], Atharv Pawar[2], Akash Pore[3], Amar Waghmare[4]
*Department of Information Technology, Dhole Patil College of Engineering, Pune*

***Abstract:*** ***One of the most significant subjects of society is human healthcare. It is looking for the best one and robust disease diagnosis to get the care they need as soon as possible.The task of following new approaches is challenging these disciplines, moving beyond the conventional ones.The actual number of new techniques makes it possible to provide a broad overview that avoids particular aspects. To this end, we suggest a systematic analysis of human diseases related to machine learning. This research concentrates on existing techniques related to machine learning growth applied to the diagnosis of human illnesses in the medical field to discover exciting trends, make unimportant predictions, and help decision-making.This paper analyzes unique machine learning algorithms used for healthcare applications to create adequate decision support.This paper intends to reduce the research gap in creating a realistic decision support system for medical applications. .Datasets are used in this project for predictions. Random Forest algorithm is to create multiple decision trees during training and then combine their predictions to obtain a more accurate and robust model. Ensemble models based on deep learning have made significant contributions to the medical field, particularly in the area of disease prediction. Breast cancer is a highly aggressive disease with a high mortality rate. Timely and effective prediction of breast cancer can reduce the risk of it progressing to later stages and the need for unnecessary medications. This study proposes a novel machine learning algorithm to predict the AD progression utilising a multi-task ensemble learning approach. Specifically, we present a novel tensor multi-task learning (MTL) algorithm based on similarity measurement of spatio-temporal variability of brain biomarkers to model AD progression. In this model, the prediction of each patient sample in the tensor is set as one task, where all tasks share a set of latent factors obtained through tensor decomposition. Furthermore, as subjects have continuous records of brain biomarker testing, the model is extended to ensemble the subjects' temporally continuous prediction results utilising a gradient boosting kernel to find more accurate predictions***
***Keywords: Human disease, Healthcare, Machine learning, Deep learning, Convolutional Neural Networks.***

## I. INTRODUCTION

In an era characterized by unprecedented advancements in technology and an ever-increasing volume of healthcare data, the field of disease prediction has witnessed a remarkable transformation. Predicting diseases before they manifest clinically is a pivotal aspect of modern healthcare that holds the promise of early intervention, improved patient outcomes, and more efficient resource allocation within the healthcare system. The conventional methods of disease diagnosis and prediction, heavily reliant on the expertise of healthcare professionals and manual processes, often fall short in terms of accuracy, speed, and scalability. This is where the convergence of healthcare and machine learning emerges as a game-changer.

Machine learning, a subset of artificial intelligence, has revolutionized various industries, and healthcare is no exception. Its capacity to harness the power of big data, extract meaningful patterns, and make data-driven predictions has opened up new horizons for disease prevention and management. This report delves into the exciting intersection of healthcare and machine learning, where predictive algorithms are employed to anticipate disease onset, progression, and outcomes with increasing accuracy.

Nowadays, Healthcare services are so expensive so everybody can't afford it So there is need to develop cost effective system to detect the diseases.The main advantage of this project is that we can get the test result immediately at our home with a just few clicks.Many of existing machine learning models are concentrating on one disease per analysis.For example,one for diabetes,one for heart disease ,one for lung disease and many such.If user wants to predict more than one disease ,he/she has to go through different sites.There is no common site where one analysis can perform more disease prediction.Some models have inconsistent accuracy which can seriously affect the patients

## II. PROCESSED SYSTEM

1) *User Registration:* Users start by registering on the platform, providing essential details such as name, email,aadhar number password and mobile number.
2) *Security measures:* After registration the user will get the conformation mail on their mail id which consist of login information.

©IJRASET: All Rights are Reserved 

3) *Choose Action:* Upon registration user needs to enter the login information to access the website where they will check their disease.
4) *Select Disease:* User needs to select the disease which he/she wants to check among the available diseases.
5) *Enter details:* The user will require to enter his details which are being asked by the system.
6) *System Details:* User will be required to enter his name,gender, age ,values or upload the required files.
7) *User Interface:* After completion of all the steps the user will get a pop up screen which will consist of the result with all the details of User which he/she has entered.

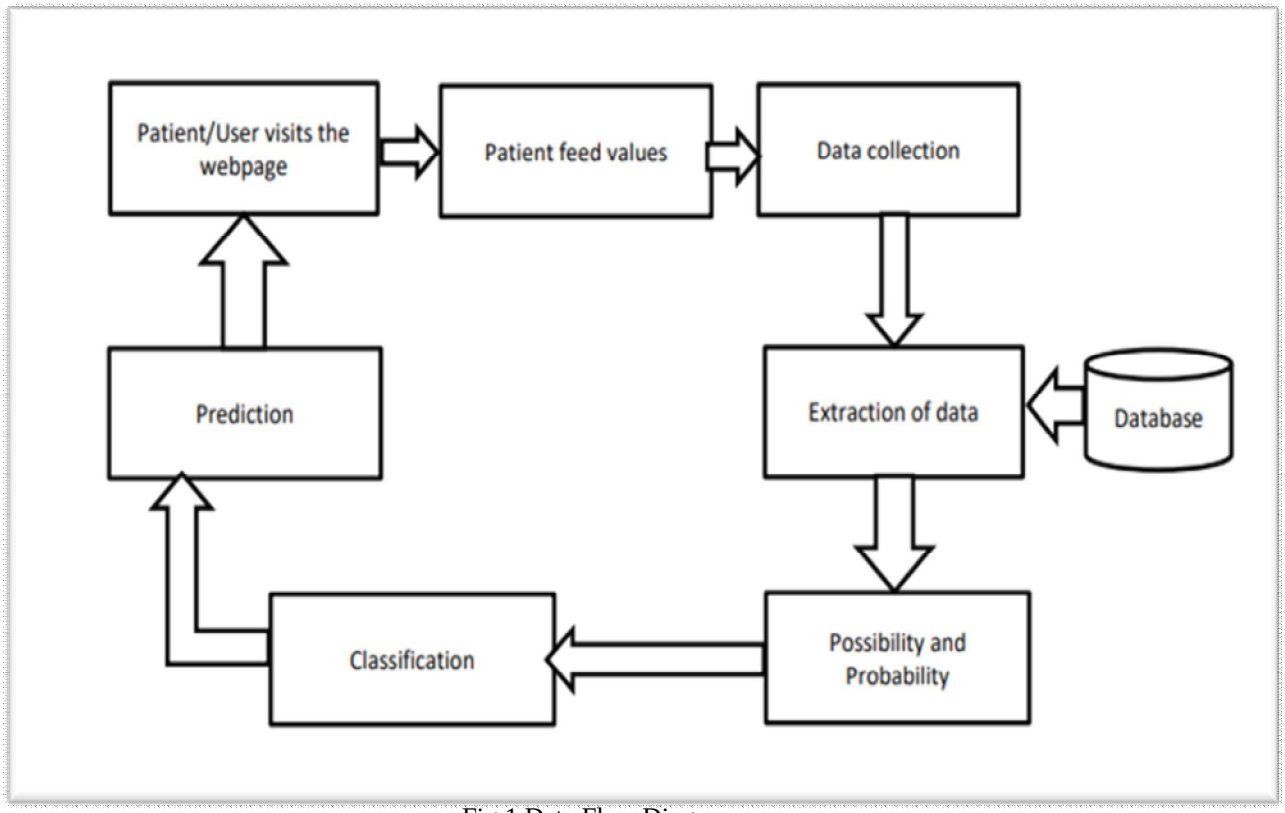



Fig.1 Data Flow Diagram

## III. LITERATURE REVIEW

Table 1 Literature Survey

| Sr. No. | Paper Name | Author, Year of publishing journals | Work |
|---|---|---|---|
| 1 | Deep Learning-Based Multi-Modal Ensemble Classification Approach for Human Breast Cancer Prognosis | Ehtisham Khan Jadoon, Fiaz Gul Khan, Sajid Shah, Ahmad Khan, And Muhammed Elaffendi Year:2023 | Breast cancer prognosis with selective features |
| 2 | Explainable Tensor Multi-Task Ensemble Learning Based on Brain Structure Variation for Alzheimer's Disease Dynamic Prediction | Yu Zhang, Tong Liu, Vitaveska Lanfranchi And Po Yang Year: 2022 | Numerous preceding brain science studies have focused on the differences in brain structure variation |

| | | | |
|---|---|---|---|
| 3 | Prediction of Diabetes Empowered With Fused Machine Learning | Usama Ahmed, Shabib Aftab, Muhammad Farhan Khan<br>Year: 2022 | Early detection of diabetic patients using machine learning fusion. |
| 4 | Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques | Senthilkumar Mohan, Chandrasegar Thirumalai, And Gautam Srivastava<br>Year: 2019 | Prediction model for heart disease with the hybrid random forest with a linear model |
| 5 | Similar Disease Prediction With Heterogeneous Disease Information Networks | Jianliang Gao, Ling Tian, Jianxin Wang, Yibo Chen<br>Year: 2020 | The goal of predicting similar diseases is to obtain the diseases with the best similarity score from networks of multiple diseases. |

## IV. METHODOLOGY

### A. *Algorithms:*

*1)* Machine learning: Machine learning is the most sought-after approach, nowadays, which is based on several statistical algorithms that make the machine learn about the data and the relationships between features of the data, in which, primary task is to gain the knowledge, predict or classify the data, all the process is based on the data, known as 'training data'. Machine learning has variety of applications in almost all the fields of the businesses and research.

*2)* Supervised Machine Learning: The labeled output indicates usage of supervised method of learning. These algorithms learn from previously inputted data, known as training data, execute analyses, and utilize the results to predict future occurrences of any new data within the existing categories. To accurately forecast test data, you'll need a lot of data with better data pattern analysis. Having the comparative analysis of predicted vs real or ground truth values and thereby refining the methods to maximize the accuracy based on the previous mistakes the ML algorithms further can be developed and trained.

*3)* COMPUTER VISION: In the field of machine learning known as computer vision, pictures and videos are analysed and understood. It is used to educate computers how to "see" and how to use visual data to do tasks that humans can do with visuals.In order to translate visual data, computer vision models need features and contextual data that are discovered during training. In order to perform prediction or decision-making tasks, models can now understand photos and videos.

*4)* RANDOM FOREST: Popular machine learning algorithm Random Forest is a part of the supervised learning methodology. It can be applied to ML issues involving both classification and regression. It is built on the idea of ensemble learning, which is a method of integrating various classifiers to address difficult issues and enhance model performance. According to what its name implies, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead than depending on a single decision tree, the random forest uses forecasts from each tree and predicts the result based on the votes of the majority of predictions.

*5)* XGBOOST: XGBoost is a distributed gradient boosting library that has been optimised for quick and scalable machine learning model training. A number of weak models' predictions are combined using this ensemble learning technique to get a stronger prediction. Extreme Gradient Boosting, or XGBoost, is one of the most popular machine learning algorithms because it can handle big datasets and perform at the cutting edge in many machine learning tasks like regression and classification.Its effective handling of missing values, which enables it to handle real-world data with missing values without requiring a lot of pre-processing, is one of the key characteristics of XGBoost. Additionally, XGBoost includes integrated parallel processing capability, allowing you to train models on huge datasets quickly.

## V. CONCLUSION

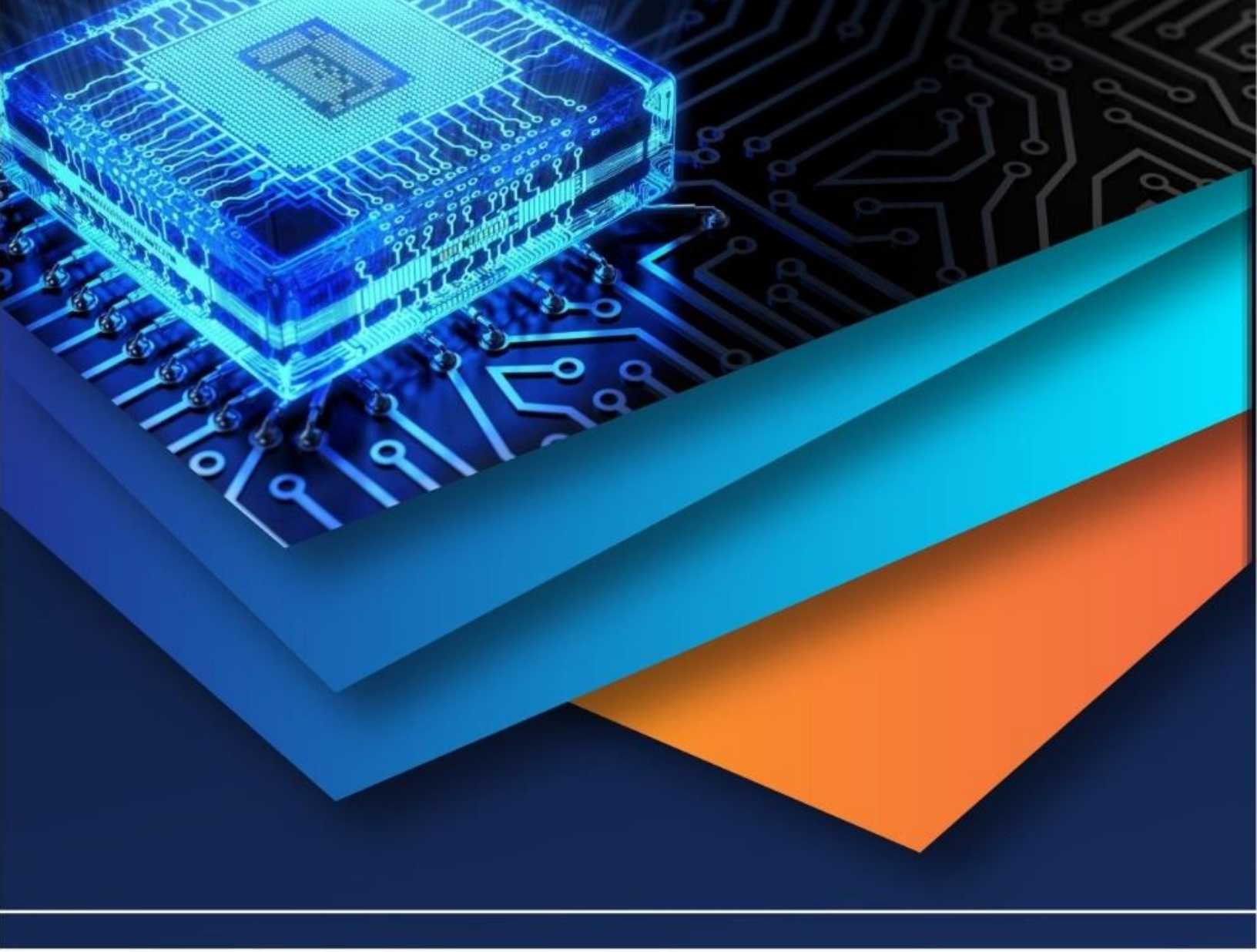
Multiple diseases prediction projects have the potential to revolutionize healthcare by shifting the focus from treatment to prevention.While there are challenges and complexities to address, the benefits of early disease detection, personalized healthcare, and improved health outcomes make these projects a valuable pursuit in the field of medicine and data science. By carefully navigating the associated issues, we can work toward a future where healthcare is more proactive,individualized and effective.

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

Intelligent data processing is a social necessity for identifying, as soon as possible, of useful and robust disease detections to provide patients with appropriate care within the shortest possible time. This detection has been carried out in recent decades by detecting exciting patterns in databases. Smart data processing is emerging as a requirement for effective and robust diseases to be found by society. Detection of patients providing the necessary treatment as soon as possible within the shortest possible period. This identification has been achieved in recent decades through the method of identifying exciting patterns in databases. A comprehensive overview of intelligent data analysis tools in the medical sector is given in this paper. Some examples of some algorithms used in these medical field areas are also presented, examining potential patterns based on the target searched, the methodology used, and the application field. Given the pace at which new works emerge in this emerging field, a systematic analysis such as the one we have just presented may become obsolete in a short period.

## REFERENCES

[1] G. Chugh, S. Kumar, and N. Singh, ‘‘Survey on machine learning and deep learning applications in breast cancer diagnosis,’’ Cognit. Comput., vol. 13, no. 6, pp. 1451–1470, Nov. 2021.
[2] S. Kumar and S. Sharma, ‘‘Sub-classification of invasive and non-invasive cancer from magnification independent histopathological images using hybrid neural networks,’’ Evol. Intell., vol. 15, no. 3, pp. 1531–1543, Sep. 2022.
[3] N. Arya and S. Saha, ‘‘Multi-modal classification for human breast cancer prognosis prediction: Proposal of deep-learning based stacked ensemble model,’’ IEEE/ACM Trans. Comput. Biol. Bioinf., vol. 19, no. 2, pp. 1032–1041, Mar. 2022.
[4] Z. S. Khachaturian, ‘‘Diagnosis of Alzheimer’s disease,’’ Arch. Neurol., vol. 42, no. 11, pp. 1097–1105, 1985.
[5] S. Tabarestani et al., ‘‘A distributed multitask multimodal approach for the prediction of Alzheimer’s disease in a longitudinal study,’’ NeuroImage, vol. 206, pp. 1–15, Feb. 2020.
[6] X. Wang, X. Zhen, Q. Li, D. Shen, and H. Huang, ‘‘Cognitive assessment prediction in Alzheimer’s disease by multi-layer multi-target regression,’’ Neuroinformatics, vol. 16, nos. 3–4, pp. 285–294, Oct. 2018.
[7] R. S. Doody, V. Pavlik, P. Massman, S. Rountree, E. Darby, and W. Chan, ‘‘Predicting progression of Alzheimer’s disease,’’ Nature BMC, vol. 2, no. 1, pp. 1–9, 2010.
[8] F. Islam, R. Ferdousi, S. Rahman, and H. Y. Bushra, Computer Vision and Machine Intelligence in Medical Image Analysis. London, U.K.: Springer, 2019.
[9] WorldHealth Organization(WHO).(2020).WHORevealsLeading Causes of Death and Disability Worldwide: 2000–2019. Accessed: Oct. 22, 2021. [Online]. Available: https://www.who.int/news/item/09-12-2020-whoreveals-leading-causes-of-death-and-disability-worldwide-2000-2019
[10] A. Frank and A. Asuncion. (2010). UCI Machine Learning Repository. Accessed: Oct. 22, 2021. [Online]. Available: http://archive.ics.uci.edu/ml



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)