DrillBit

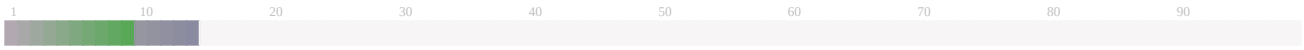
The Report is Generated by DrillBit Plagiarism Detection Software

Submission Information

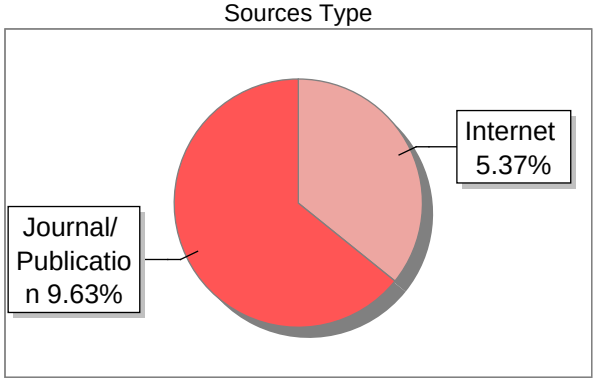
Author Name	rajesh we
Title	desease mdp
Paper/Submission ID	2202434
Submitted by	rajesheminent@gmail.com
Submission Date	2024-08-05 09:34:33
Total Pages, Total Words	11, 2290
Document type	Article

Result Information

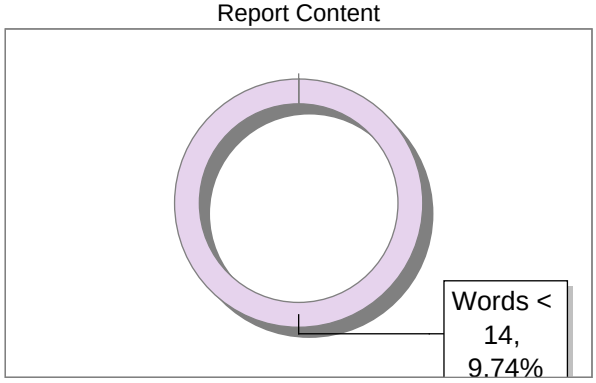
Similarity **15 %**



Sources Type



Report Content



Exclude Information

Quotes	Not Excluded
References/Bibliography	Not Excluded
Source: Excluded < 14 Words	Not Excluded
Excluded Source	**0 %**
Excluded Phrases	Not Excluded

Database Selection

Language	English
Student Papers	Yes
Journals & publishers	Yes
Internet or Web	Yes
Institution Repository	Yes

A Unique QR Code use to View/Download/Share Pdf File

DrillBit

DrillBit Similarity Report

15	22	B
SIMILARITY %	MATCHED SOURCES	GRADE

A-Satisfactory (0-10%)
B-Upgrade (11-40%)
C-Poor (41-60%)
D-Unacceptable (61-100%)

LOCATION	MATCHED DOMAIN	%	SOURCE TYPE
1	www.irjmets.com	3	Publication
2	Thesis Submitted to Shodhganga Repository	2	Publication
3	Neuronal Antibodies and Associated Syndromes by Arneth-2019	1	Publication
4	www.ncbi.nlm.nih.gov	1	Internet Data
5	sites.computer.org	1	Publication
6	springeropen.com	1	Internet Data
7	www.emerald.com	1	Internet Data
8	moam.info	1	Internet Data
9	ijircce.com	1	Publication
10	towardsdatascience.com	<1	Internet Data
11	bmjopen.bmj.com	<1	Publication
12	dl.lib.uom.lk	<1	Internet Data
13	www.ijsurgery.com	<1	Publication
14	www.journalcra.com	<1	Publication

15	Cell-derived matrices for studying cell proliferation and directional migration by Kaukonen-2017	<1	Publication
16	dovepress.com	<1	Internet Data
17	elr.tijdschriften.budh.nl	<1	Internet Data
18	journal.unair.ac.id	<1	Internet Data
19	mdpi.com	<1	Internet Data
20	www.dx.doi.org	<1	Publication
21	www.dx.doi.org	<1	Publication
22	www.linkedin.com	<1	Internet Data

Multiple Disease Predictor

Rajesh Eswarawaka, Sabari Vedhagiri, Gagana Sai M, Uoosa Mani Sai Charan

AMC Engineering College, Bengaluru

ABSTRACT

In recent times, predicting multiple diseases has become increasingly challenging in the medical field. Diseases such as heart disease, diabetes, and Parkinson's disease pose significant threats to global health. Some of these diseases progress rapidly, leading to symptoms like weakness and shivering. Our research highlights significant advances in the medical field by utilizing machine learning techniques to predict multiple diseases through a web application based on patient test datasets. This application can predict conditions like Diabetes, Heart disease and Parkinson's disease.

The research employs algorithms and Logistic regression to determine whether an individual is affected by a disease, enhancing the accuracy and efficiency of predictions. Our study predicts an individual's health status by identifying unknown values from the provided input datasets. The application reliably predicts these unknown values from user-submitted data to determine the presence of a disease. One of the primary features of our application is its ability to identify multiple diseases simultaneously within a web app, allowing for the input of extensive data for disease prediction.

Stream lit Cloud and the Stream lit library improve usability, enabling individuals to easily predict the risk of various diseases. The web app creates an accessible interface for users by offering selection options for different diseases. By selecting a specific disease within the web app, users can predict the disease based on the provided data. This study demonstrates that multiple diseases can be identified quickly, and the process is simplified through a user-friendly web app interface. High accuracy rates, ranging from 80% to 85%, can be achieved using these machine learning techniques.

Keywords

ML (Machine Learning), ICMR (Indian Council for Medical Research), Logistic Regression, SVM (Support Vector Machine)

INTRODUCTION

Chronic diseases such as Diabetes, Heart disease, and Parkinson's disease have presented substantial difficulties to global health systems in recent years, resulting in increased morbidity and death. These illnesses have a significant influence not only on individuals' quality of life but also on healthcare infrastructure and their resources.

Early detection and accurate prediction of these disorders are critical for successful disease management and intervention techniques, which can greatly improve patient outcomes while lowering healthcare expenditures. With the introduction of improved computing tools, machine learning has emerged as a strong tool in medicine, presenting efficient answers for disease prediction and diagnosis.

Machine learning algorithms can evaluate massive volumes of healthcare data, revealing intricate patterns and relationships that might otherwise go undetected. By using these skills, machine learning can provide predictive insights that allow for early diagnosis and individualized treatment strategies. This work investigates the use of machine learning approaches to forecast the start and progression of diabetes, heart disease, and Parkinson's disease. We have explored numerous machine learning models, and evaluated their performance, and determining the most effective routes to accurate prediction. Our study makes use of a large dataset that includes ICMR (Indian Council Of Medical Research), clinical, demographic, and lifestyle information, ensuring a thorough and holistic analysis of the factors that contribute to these disorders. Through this effort, we hope to demonstrate machine learning's potential to revolutionize the landscape of chronic disease management, emphasizing its role in facilitating early intervention and improving patient care.

By improving machine learning's predictive capabilities, we expect to contribute to continuing efforts to reduce the worldwide burden of chronic diseases and improve the overall well-being of those affected.

Contemporary to the medical landscape to the early detection and accurate diagnosis creates an effective treatment and management. With advancing technique of ML create an exponential growth on the health care with an significant opportunity on enhancing in predicting accuracy and efficiency.

Diabetes is characterized by certain aspects like giving number of pregnancies, glucose level, BP, skin thickness, insulin level, BMI, age which captures the progression of disease. The ML model integrates multiple factor of data and predicts the likelihood of diabetes with an greater precision rate.

Heart Disease creating an mortality world wide the millions of individuals affected, with an early detection would prevent by not letting one to lead into severe complications which furthers improves the patients outcome, its characterized by certain aspects such as resting blood pressure, serum cholesterol, fasting blood sugar, resting electrocardiographic, heart rate. The cardiovascular challenges in certain aspects like accuracy and timely diagnosis, which aims in identifying patterns and identifying heart disease, thereby leads up in facilitating prompt intervention.

Parkinson’s Disease an neurogenerative disorder which affects movement. The early detection would subtle the symptoms of the neurological conditions, using ML technique and algorithms it enhances in early detection. These are characterized by MDVP, SHIMMER, HNR.

In this successful implementation it holds an potential preventive health care by offering an non invasive, cost effective, which makes an augments decision making by empowering patients by fostering proactive management. Through this research we have aimed in evolving an AI driven health care solution paving an sophisticated accessible diagnostic technology.

LITREATURE REVIEW

This research examines multiple disease prediction using machine learning, with the goal of improving healthcare outcomes and reducing costs through early prediction and prevention. This review covers heart disease, diabetes, Parkinson's disease, kidney disease, and cancer, providing insights from both operational and technical perspectives.

Heart Disease:

The heart is a crucial organ that pumps blood to all other organs in the human body. This study aims to predict heart disease with high accuracy, determining whether a person is at risk. Data mining techniques in machine learning plays vital role in context to multiple disease prediction.

Chronic diseases using certain Machine Learning algorithms such as Naive Bayes, Logistic Regression, Support Vector Machine (SVM), and Artificial Neural Network (ANN) were used. These algorithms provided an better performance rate and accuracy in prediction.

By using ANN for heart disease diagnosis yielded an accuracy of 88% with 20 neurons in the hidden layer, indicating significant potential for heart disease prediction.

Diabetes:

Diabetes is an metabolic disease characterized by elevated blood glucose levels, which can cause severe damage to the heart, blood vessels, eyes, nerves, and kidneys leading in providing an early result at beginning stage by using multiple disease predictor.

Patients with low health literacy rate had higher retention for diabetes. Participants in randomized group who engage in significant exercise are made an informed food choices, able to better control their diabetes.

In diabetes SVM has provided an highest accuracy rate and measuring of about 84% accuracy rate.

Parkinson's Disease:

Parkinson's disease is a disorder of the central nervous system that affects movements. It results from nerve cell damage in the brain, leading to low dopamine levels, causing uncontrolled balance, walking difficulties, shaking hands, and memory loss, which gives an analysis about the disease using the predictor.

TEHNICAL REQUIRMENTS

1. **Data Collection and Preprocessing:**

- Data Sources: Gathered data from reliable sources such as ICMR, public health databases, hospitals, or clinical studies.
- Data Cleaning: Handling missing values, outliers, and data inconsistencies.
- Feature Engineering: Selected and created relevant features for the prediction model.

2. Machine Learning Models:

- Model Selection: Chooses an appropriate models based on the type of diseases and the nature of the data. Common models includes Logistic Regression, Decision Trees, Random Forests, SVM (Support Vector Machines), and Neural Networks.
- Model Training: Trained the models using a portion of the dataset and validated their performance on a separate validation dataset.
- Model Evaluation: Using certain metrics such as accuracy, precision, recall, F1 score, and ROC-AUC to evaluate model performance.

3. Implementation:

- Programming Languages: Used languages like Python for implementing machine learning algorithms.

- Libraries and Frameworks: Utilised libraries such as scikit-learn, TensorFlow, Keras, or PyTorch for building and training models.

4. Deployment:

- Environment Setup: Built up an suitable environment for deploying our model, which could be cloud-based (AWS, Google Cloud, Azure) or on-premise.
- API Development: Developed an APIs to allow easy access to the prediction model for applications or for the end-users.

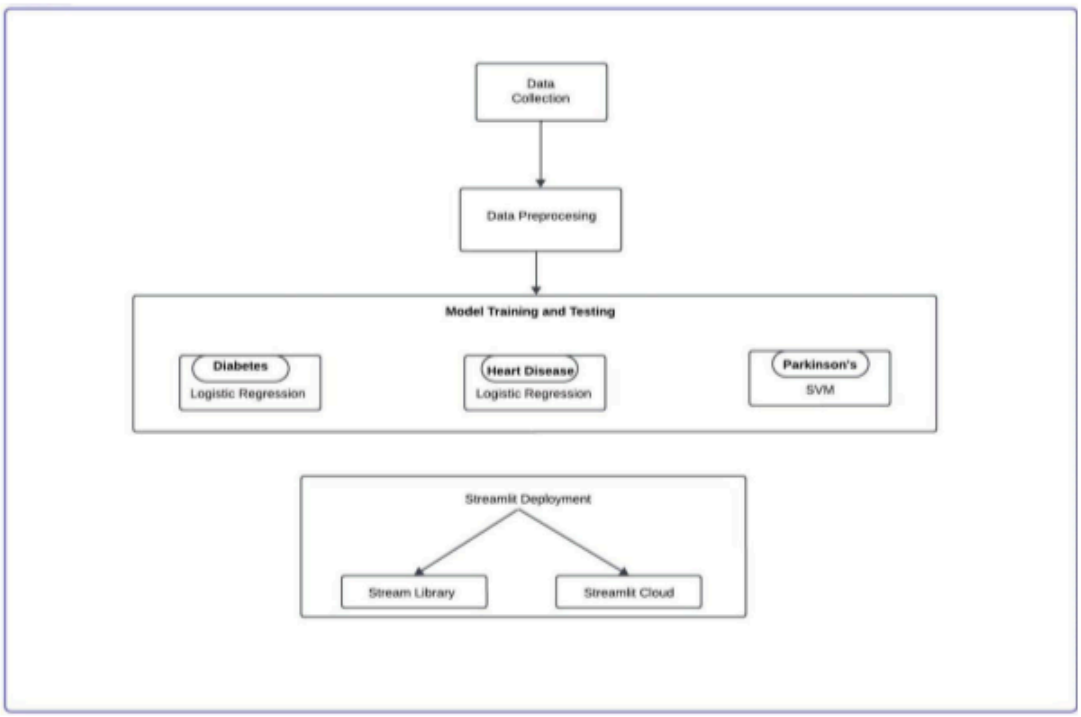
5. Documentation:

- Writing Clear Descriptions: We have documented each step of our research and methodology clearly.
- Code Repositories: We have maintained an repository for our code, ensuring it is well-documented and reproducible.

6. Ethical Considerations:

- Data Privacy: We Ensured that the patient data is anonymized and complies with data protection regulations like HIPAA or GDPR.
- Bias and Fairness: We have address certain potential biases in our data and model to ensure fair and unbiased predictions.

ARCHITECTURE



The data is collected from ICMR or from public health data base which are in form of demographic, medical history and lab results. The data integration merges up the data from various sources and creates an unified data by using certain tools, and the data cleaning process handles up the missing value and remove the duplicate corrected inconsistencies by using python and certain libraries, further the data is transformed by numerical standard range and converts the categorical data into an numerical format and creates up an feature existing an improved performance model.

Data preprocessing splits up the data as training set of about 80% and 20% as testing and validates the set of the training data, it balances up the data by using certain technique and feature is selected using certain technique like recursive feature.

The model is trained and tested with diabetes, where the model is trained with logistic regression algorithm by using Scikit learn in python by processing the trained model using the datasets by using validation set the hyper parameter

optimizes using certain techniques like random search, the model is evaluated with accuracy, precision, recall and F1-score, AUC-ROC.

Heart disease model is trained with logistic regression and uses certain library as of diabetes as Scikit-learn in python and hyperparameter gets optimized accordingly.

Parkinson's disease uses sum function approach like algorithms such as decision tree, SVM (Support Vector Machine) by using library like Scikit-learn and pickle which would further train the datasets

Its deployed with certain frameworks for creating interactive web application, the process builds and its developed by user friendly interface using stream lit with the prediction model and its hosted by applications on cloud platform. Stream lit is integrated with trained model into stream lit application, and the features are uploaded with patient data for prediction and displays results and interpretability insights, it visualizes data and model performance metrics.

Its deployed in cloud environment by setting up an virtual machine with an serverless environment further its installed with certain libraries and dependencies, it implements continuous integration and continuous

deployment pipeline for automatic updates and maintenance. And the scalability ensures that it handles multiple users of large datasets.

The model is trained and tested with diabetes, where the model is trained with logistic regression algorithm by using Scikit learn in python by processing the trained model using the datasets by using validation set the hyper parameter optimizes using certain techniques like random search, the model is evaluated with accuracy, precision, recall and F1-score, AUC-ROC.

Heart disease model is trained with logistic regression and uses certain library as of diabetes as Scikit-learn in python and hyperparameter gets optimized accordingly.

Parkinson's disease uses sum function approach like algorithms such as decision tree, SVM (Support Vector Machine) by using library like Scikit-learn and pickle which would further train the datasets

Its deployed with certain frameworks for creating interactive web application, the process builds and its developed by user friendly interface using stream lit with the prediction model and its hosted by applications on cloud platform. Stream lit is integrated with trained model into stream lit application, and the features are uploaded with patient data for prediction and displays results and interpretability insights, it visualizes data and model performance metrics.

Its deployed in cloud environment by setting up an virtual machine with an serverless environment further its installed with certain libraries and dependencies, it implements continuous integration and continuous deployment pipeline for automatic updates and maintenance. And the scalability ensures that it handles multiple users of large datasets.

RESULTS:

The structure approch in ensuring research process with clear findings, making it to undestand easier and replicate.

Here below is the accuracy rates of the disease using diffrent algorithms:

Disease	Logistic Regression	SVM	KNN	Accuracy
Diabetes	77.60	75.64	75.52	77.60
Heart Disease	82.50	78.87	83.84	83.84
Parkinson’s	86.42	84.63	87.17	87.17

CONCLUSION

Diabetes is a sizable persistent circumstance that impacts many people international, often by way of disrupting the frame's capacity to modify blood sugar ranges. With out right control, it could result in critical headaches like cardiovascular sickness, nerve harm, and kidney problems. No matter development in treatments and technology, diabetes remains a tremendous fitness situation. Encouraging healthy existence, making sure early prognosis, and advancing medical remedies are critical in handling the disease. Persisted research and public education are crucial to improving the lives of these with diabetes and lowering its worldwide impact.

Heart disease is a essential worldwide fitness difficulty, accounting for a substantial variety of deaths every year. This huge category consists of conditions like coronary artery ailment, coronary heart failure, and arrhythmias, often influenced by means of way of life factors, genetics, and different fitness situations. Notwithstanding advancements in medical generation and treatment, heart disorder remains a major project. Efforts to promote coronary heart-wholesome behaviour, enhance remedy alternatives, and lift consciousness are essential. Ongoing research and public fitness tasks play a important position in preventing heart ailment and improving the fine of life for the ones affected.

Parkinson’s disease, in this research we have demonstrated an potential ML model, particularly SVM (Support Vector Machine) in predicting the disease with high accuracy. By its significant feature of symptoms the model promises for early diagnosis and an improved management, it should focus towards the future on broader validation by enhancing model interpretability and by integrating the tool for clinical and widespread practices.