

Personal Healthcare Companion Using Machine Learning

Suvaranamaliya J, Sherin S, Ramya Priyadarshini N, Jemi Gold P
M.E

*Computer Science and Engineering RMK College of Engineering and Technology Pudukkottai, Thiruvallur.
Computer Science and Engineering RMK College of Engineering and Technology Pudukkottai, Thiruvallur.
Computer Science and Engineering RMK College of Engineering and Technology Pudukkottai, Thiruvallur.
Assistant Professor, Computer Science and Engineering RMK College of Engineering and Technology
Pudukkottai, Thiruvallur.*

Date of Submission: 20-11-2020

Date of Acceptance: 06-12-2020

ABSTRACT—we propose a System that would enable everyone to have access to Personal Health Care Companion. Such that it helps individuals to perform health check at regular intervals and produce dynamic health reports on the health related problems that individuals may face. Health care field has a vast amount of data, for processing those data we have used some of the Machine Learning Algorithms and detected the presence of Disease. Our System is built to predict the presence of four major diseases such as Heart Disease and Chronic Kidney Disease using Sequential Neural Network Model, Diabetes using K-Nearest Neighbors Classification Model, and skin Cancer Detection using Convolution Neural Network Model.

Index Terms—Heart Disease, Chronic Kidney Disease, Skin Cancer, Diabetes, Neural Networks, Feature Extraction, Image Processing, Classification

I. INTRODUCTION

Most doctors appear to express different views on a specific topic. Negligence of medical issues in rural areas due to lack of hospital education and insufficiency. Analysis of a issue by a Personal Health Care Partner will cost an individual a lot of money. In India alone 50, 00,000 deaths occur every year due to a misdiagnosis. Heart disease is the root cause of death in the world. This System estimates the possibilities of developing Heart Disease. The effects of this system have in percentage terms the risk of occurring heart disease. Using

Sequential Neural Network we predicted the illness. Chronic kidney disease is growing day by day and plays an important role in the life of an person. Using Sequential Neural Network

we were predicting chronic kidney disease. Diabetes is one of the world's deadliest illnesses, it is also an perpetrator of various types

of disorders such as heart failure, blindness, illnesses of the urinary organs, etc. Using K-Nearest Neighbor Classification System, we estimated a patient's diabetic risk level with greater accuracy. The most prevalent form of cancer is skin cancer, which affects the lives of millions of people each year. Early phase detection of skin cancer is a difficult and costly process. Using the Convolution Neural Network Model we predicted skin cancer is present.

II. LITERATURE SURVEY

Using Neural Networks they proposed a method for classifying skin lesions into their respective categories. This offers a way to rate complicated data with a high degree of precision. They categorized 463 images into their respective groups, with a high precision rate of 76.9% [1] They had done a lot of work in detection of Melanoma specifically. Moussaet is between different approaches. Al. detected Melanoma using its geometric features and used the k-Nearest Neighbors algorithm to distinguish it from benign lesions. The accuracy rate is 89 percent, the downside is that the dataset was small.[2] This article gathered the Data Mining and heart disease Awareness number. Information on heart failure, symptoms of heart attack, and causes of heart disease is given. The three key data mining techniques are used in this paper, namely Decision Tree, Neural Networks, and Naive Bayes Classifier. There are three techniques used to predict outcomes. The Advantages and Disadvantages of each technique can be known using this paper [3] This paper's main principle is predicting heart disease using Data Mining Techniques. The main Methodology used for prediction is KNN Algorithms, Decision Trees like ID3 Algorithms, and Naive Bayes Techniques. It has 13 medical-related features as input and Data sets it to be interpreted using certain algorithms for

that data. [4]. They had discussed about the analysis of various skills of mining to guess diabetes using Naive Bayes, Random forest, Decision Tree and J48 algorithms [5]. In this paper the authors has done a detailed research about Diabetic Retinopathy (DR) which is one of the main cause of sight inefficiency for diabetic patients. In which they reviewed the performance of a set of machine learning algorithms and applied them on particular data set [6]. They had been debating a concept of data mining called ANN (Artificial Neural Network) and Naive Bayes. For the estimation and treatment of chronic kidney disease they used these data mining classifiers. Rapid miner tool is used in this investigative work. Results obtained aArtificial Neural Network have accuracy of 72.73 percent. [7] In this paper, they had discussed about healthcare, 98% is achieved in the current status of patients healthcare. They predicted health disease as in the form of asthma and cancer. They create a cloud environment in their paper in the proposed research, "analyzing big data from healthcare with the prediction of future health condition" [8]

III. DISEASES

HEART DISEASE

Heart is important because it pumps blood around our body, delivering oxygen and nutrients to our cells and removing waste products. Based on the statistical report from WHO, one-third population worldwide died from heart disease; heart disease is found to be the leading cause of death in developing countries by 2017.

Heart Disease risk factors include:

- High blood pressure
- High Cholesterol
- Consuming too much alcohol
- Smoking
- Diabetics
- Family history of coronary illness
- Being overweight or obese

Symptoms of Heart Disease:

- Shortness of breath
- Pain and discomfort in chest
- Fatigue
- Cold sweat and unsteadiness
- Rapid or irregular heart beat
- Heart burn or abnormal pain

CHRONIC KIDNEY DISEASE

Chronic Kidney Disease is a kind of situation in which kidneys get damaged and toxics cannot easily filtered out from our body. An

individual realized the chronic kidney disease when kidneys function lower down 25% of the normal kidney disease is considered as one of the major global issue concerned with the person's health.

Chronic Kidney Disease risk factors include:

- Cardiovascular Disease
- Abnormal Kidney Structure
- High blood pressure
- Obesity
- Family history of kidney disease
- Smoking
- Diabetes

Symptoms of Chronic Kidney Disease:

- Nausea and Vomiting
- Too much urine or not enough urine
- Swelling in your feet and ankles
- Muscle cramps
- Not feeling hungry
- Trouble sleeping

SKIN CANCER

It is the abnormal growth of skin cells .This type of cancer can also occur in many areas of the skin that are not usually exposed to sunlight. Three primary causes include skin cancer, basal cell carcinoma, squamous cell carcinoma and melanoma.

Checking for unusual changes on the skin will help you to diagnose skin cancer in its earliest stages. Early detection of skin cancer gives you the best chance of successfully curing skin cancer. Skin Cancer risk factors include:

- Fair skin
- A history of sunburns
- Excessive sun exposure
- Moles
- A family history of skin cancer
- A weekend immune system

Symptoms of Skin Cancer:

- A large brownish spot with darker speckles.
- A mole that changes in color, size or feel or that bleeds.
- A small lesion with an irregular border and portions that appear red, pink, white, blue or blue-black.
- A painful lesion that itches or burns.

DIABETES

Diabetes is a condition that impairs the body's ability to process blood glucose, otherwise known as blood sugar. Different forms of Diabetes can occur, and managing the condition depends on the type. Not all forms of diabetes stem from person being overweight or leading an inactive

lifestyle. In fact, some are present from childhood. Three major diabetes types can develop: Type 1, Type 2, and gestational diabetes.

Diabetes risk factors include

- Being overweight
- A family history of diabetes
- A history of high blood pressure
- Having gestational diabetes or giving birth to a child
- Having a sedentary lifestyle
- A history of PCOS

Symptoms of Diabetes:

- Increased thirst
- Frequent urination
- Hunger
- Fatigue

IV. EXISTING SYSTEM

Predictions of various diseases are done already. But they have failed to meet the accuracy level. And this occurs due to the choice of Machine Learning models. The Dataset used were also very limited. Even if the models were constructed, they were not available with user interface, and were not user friendly.

A normal person felt difficulty in accessing it. Some of the existing system can show only the symptoms of diseases, they don't have the capability to analyze and generate reports regarding health issues in an individual. Only a limited number of health issues are taken into consideration, due to which proper awareness was not created among people.

Health related checkups are available but they are more expensive which makes the rural people less affordable.

V. PROPOSED SYSTEM

Personal HealthCare Companion is built to predict the presence of four major diseases.

- Heart Disease
- Chronic Kidney Disease
- Diabetes
- Skin Cancer

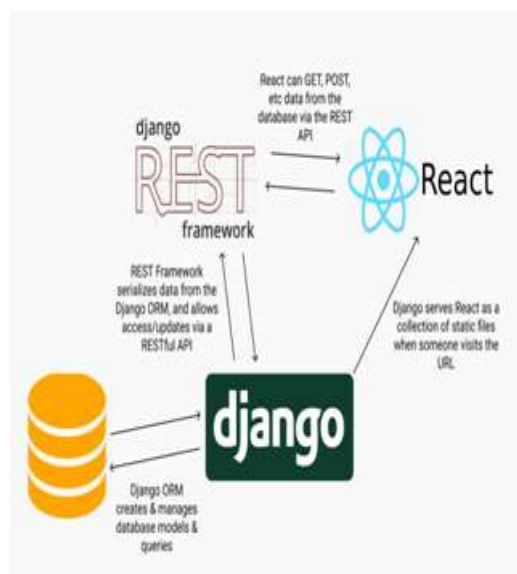
We have proposed a system that will integrate all these disease prediction. It has a single user interface, such that people who have the symptoms of any one of these diseases can easily access their Personal HealthCare Companion through internet and can get to know whether they have the disease or not through predictions.

ARCHITECTURE DIAGRAM



TECHNOLOGY USED

In our proposed work, web service REST API using Django Framework, User Interface using REACT JS, Construction of ML models using Python Libraries like SCIKIT-LEARN, KERAS are used.



Django Rest framework:

Django Rest framework is powerful and flexible toolkit for building Web APIs. Main advantages are Simplicity, Flexibility, quality, and

test coverage of source code. Powerful serialization engine compatible with ORM and non-ORM data sources. Pluggable and easy to customize emitters, parsers, validators and authenticators.

React JS

React JS is an open source JavaScript library, which is used primarily for single-applications to create user interfaces. It's used for web and mobile devices screen layer handling. React also lets us create reusable UI components React lets developers build massive web applications that can alter data without reloading the page

Python Library

Keras is a Python-written Neural Network Library that is high in design-making it incredibly easy and intuitive to use. It works as a wrapper to low-level libraries like tensor flow written in python.

Scikit learning is potentially the most powerful machine learning library in Python. This collection contains many effective machine learning and statistical modeling resources including classification, regression, clustering, and reduction in dimensionality

DATA COLLECTION

Heart disease

This dataset includes a variety of variables, along with a target condition of having heart disease or not. There are 13 features are available in this dataset We will try to use this data to create a Sequential Neural Network model which tries predict if a patient has this disease or not. The "goal" field refers to the patient's existence of heart disease.

Some of the features:

- age: The person's age in years
- sex: The person's sex
- cp: The chest pain experienced
- trestbps: person's resting blood pressure
- chol: The person's cholesterol
- fbs: The person's fasting blood sugar

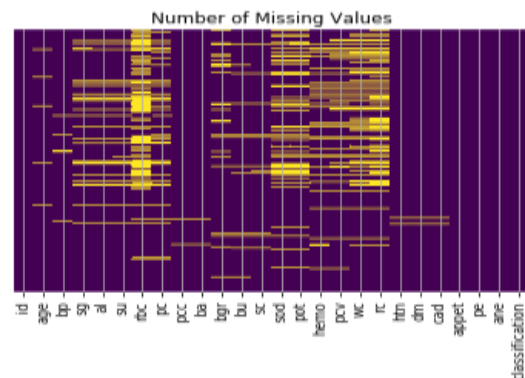


Chronic Kidney disease

The data is taken over a 2-month period in India with 25 features. The target is the 'classification', which is either 'ckd' or 'notckd' - ckd=chronic kidney disease. We will try to detect the Kidney disease using the **Sequential Neural Network Model**.

Some of the features:

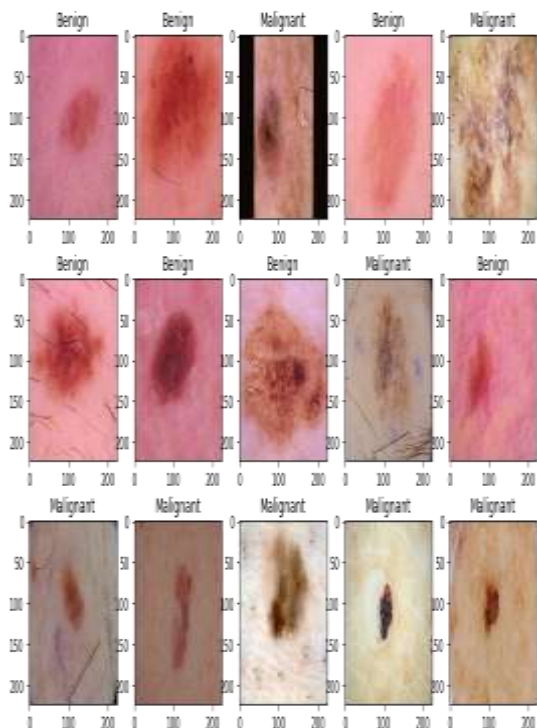
- red blood cell count
- white blood cell count
- Blood pressure



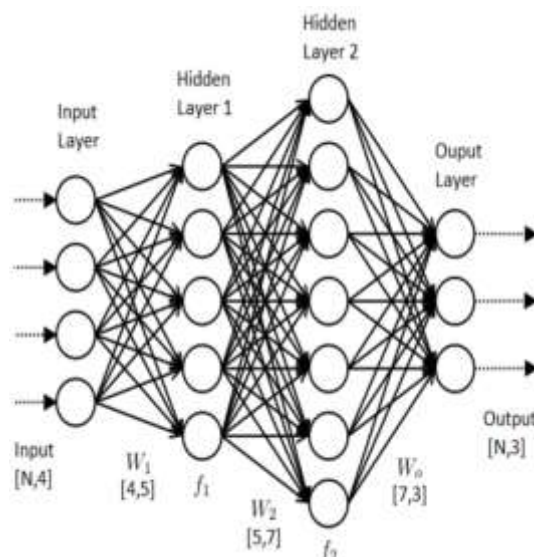
Skin Cancer

The dataset is taken from the ISIC (International Skin Image Collaboration) Archive. It consists of 1800 pictures of benign moles and 1497 pictures of malignant classified moles.

It has 2 different classes of skin cancer: Benign, Malignant We will try to detect using Convolution Neural Network Model with keras tensorflow in backend



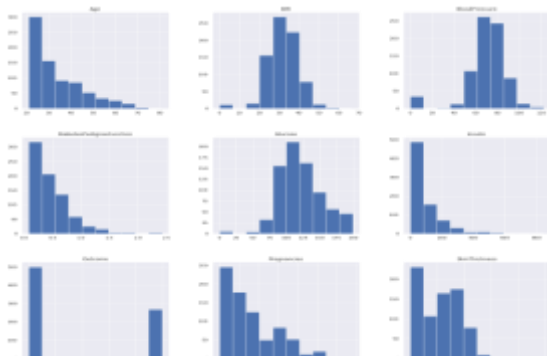
systems, such as the brain. Artificial neural network uses the supervised learning which classifies the input information into the desired product. The artificial neurons consist of weighted interconnections which regulate the effect of the respective neurons



Diabetes

The Dataset for diabetes consist of 768 entries. It has eight different features available in dataset. Three major types of diabetes are Type 1, Type2, and gestational diabetes Some of the features:

- Pregnancies
- Glucose
- BloodPressure
- SkinThickness
- Insulin
- BMI



MACHINE LEARNING ALGORITHMS

ARTIFICIAL NEURAL NETWORK

An Artificial Neural Network is an information processing model inspired by the way the information is processed by biological nervous

Firstly, the neural network collects and defines the data as an input to the network when classifying disease.

The network is trained with a given training dataset and choose the training algorithm. ANN is evaluated after the training phase to acquire the network reaction which states whether the disease is magnificently classified or not. We used ANN for predicting heart disease and chronic kidney Disease.

ML model for heart disease

```

classifier = Sequential()

# Adding the input layer and the first hidden layer
classifier.add(Dense(output_dim = 11, init = 'uniform', activation = 'relu', input_dim = 11))

# Adding the second hidden layer
classifier.add(Dense(output_dim = 11, init = 'uniform', activation = 'relu'))

# Adding the output layer
classifier.add(Dense(output_dim = 1, init = 'uniform', activation = 'sigmoid'))

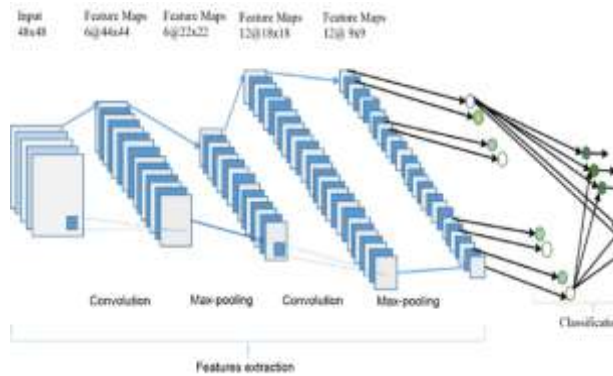
# Compiling the ANN
classifier.compile(optimizer = 'adam', loss = 'binary_crossentropy', metrics = ['accuracy'])
    
```

ML model for chronic kidney disease

```
model = Sequential()
model.add(Dense(100, input_dim=X.shape[1], activation='relu'))
model.add(Dense(50, activation='relu'))
model.add(Dense(25, activation='relu'))
model.add(Dense(1, activation='sigmoid'))
model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
```

CONVOLUTIONAL NEURAL NETWORK

A convolutional neural network in deep learning is a form of profound neural networks, most widely used to explore visual imagery. This has one or more convolutional layers and is primarily used for image recognition, classification, segmentation, and other auto-related data as well.



We used CNN to predict Skin Cancer. It recognizes and defines skin lesions as benign or malignant based on photographs taken from general cameras. The photographs are segmented, features extracted using the ABCD rule and a Neural Network is trained to identify lesions to a high degree of accuracy.

ML model for skin Cancer

```
model = Sequential()
model.add(Conv2D(64, kernel_size=(3, 3), padding = 'Same', input_shape=input_shape,
activation=activ, kernel_initializer = 'glorot_uniform'))
model.add(MaxPool2D(pool_size = (2, 2)))
model.add(Dropout(0.25))

model.add(Conv2D(64, kernel_size=(3, 3), padding = 'Same',
activation =activ, kernel_initializer = 'glorot_uniform'))
model.add(MaxPool2D(pool_size = (2, 2)))
model.add(Dropout(0.25))

model.add(Flatten())
model.add(Dense(128, activation='relu', kernel_initializer=init))
model.add(Dense(num_classes, activation='softmax'))
model.summary()
```

K-NEAREST NEIGHBOR

K- nearest neighbors is a simple algorithm which stores all available cases and classifies new cases based on a

measure of similarity (e.g., distance functions). KNN has been used for statistical estimation and pattern recognition.

In KNN it's not easy to find the significance of k. A small value of k means that noise will have a greater effect on the result and a large value will cause it to expect computationally.

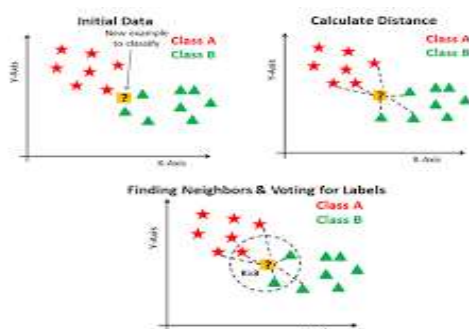
2. Another simple approach is to select k is set $k = \sqrt{n}$.

Distance functions

Euclidean $\sqrt{\sum_{i=1}^k (x_i - y_i)^2}$

Manhattan $\sum_{i=1}^k |x_i - y_i|$

Minkowski $\left(\sum_{i=1}^k (|x_i - y_i|^q) \right)^{1/q}$



VI. CONCLUSION

As technology keeps improving, our standards keep improving as well.

Predictions would be of great help when it comes to the health of a person.

Since we have a combination of prediction algorithms all in a capsule, the benefit is that we get is massive.

Machine Learning will be a part of every individual's life in future. This would be a better way to keep track of your health.

REFERENCES

- [1]. Pratik Dubal, Sankirtan Bhatt, Chaitanya Joglekar, Dr. Sonali Patil "Skin Cancer Detection and Classification"
- [2]. R. Moussa, F. Gerges, C. Salem, R. Akiki, O. Falou and D. Azar, "Computer-aided detection of Melanoma using geometric features", 2016 3rd Middle East Conference on Biomedical Engineering (MECBME), Beirut, 2016, pp. 125-128.
- [3]. Rairikar, A., Kulkarni, V., Sabale, V., Kale, H., & Lamgunde, A. "Heart disease prediction using data mining techniques." In 2017 International Conference on Intelligent Computing and Control (I2C2)
- [4]. Methaila, Aditya, Prince Kansal, Himanshu Arya, and Pankaj Kumar. "Early heart disease prediction using data mining techniques." Computer Science & Information Technology Journal (2014): 53-59.
- [5]. Dr. M. Renuka Devi and J. Maria Shyla, "Analysis of Various Data Mining Techniques to Predict Diabetes Mellitus", International Journal of Applied Engineering Research
- [6]. Ridam Pal, Dr. Jayanta Poray, and Mainak Sen, "Application of Machine Learning Algorithms on Diabetic Retinopathy", 2017 2nd IEEE International Conference On Recent Trends In Electronics Information & Communication Technology, May 19-20,

2017, India

- [7]. A.Sai Sabitha, Abhay Bansal, Khushboo Chandel, Veenita Kunwar, "Chronic Kidney Disease Analysis Using Data Mining Classification Techniques," 6th rence on Cloud System and Big Data Engineering, 2016.
- [8]. Prasan Kumar Sahoo, Suwendu Kumar Mohapatra, Shih-Lin Wu, "Analyzing Healthcare Big Data with Prediction for Future Health Condition" Journal Publications, IEEE, 2016.