VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"Jnana Sangama", Belgaum – 590 018

A PROJECT REPORT ON

"BRAIN CANCER PREDICTION USING MACHINE LEARNING"

Submitted in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING IN

INFORMATION SCIENCE AND ENGINEERING

BY

Aishwarya Reddy B (1CR15IS127)

Pranav Polepaka (1CR16IS069)

Preetika B (1CR12IS072)

Riya Pooniwala (1CR16IS082)

Under the guidance of

Dr. R Chinnaiyan

(Assistant Professor, Dept. of ISE, CMRIT)

DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING

CMR INSTITUTE OF TECHNOLOGY, BANGALORE-37

CMR INSTITUTE OF TECHNOLOGY
BANGALORE-37

Department of Information Science & Engineering

Certificate

This is to certify that the project entitled **Brain cancer prediction using Machine Learning** has been successfully completed by **Pranav Polepaka USN 1CR16IS069, Preetika B, USN 1CR16IS072, Aishwarya Reddy B, USN 1CR15IS127 and Riya Pooniwala, USN 1CR16IS082,** bonafide students of CMR Institute of Technology in partial fulfillment of the requirements for the award of degree of Bachelor of Engineering in **Information Science and Engineering** of the Visvesvaraya Technological University, Belgaum during the academic year 2019-2020. It is certified that all the corrections/suggestions indicated for Internal Assessment have been incorporated in the project report. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed for the said degree.

Name & Signature of Guide
(Dr, R Chinnaiyan)

Name & Signature of HOD
(Dr. Farida Begaum)

External Viva

Name of the Examiners | **Signature with date**

1.
2.

CMR INSTITUTE OF TECHNOLOGY
BANGALORE-37

Department of Information Science & Engineering

DECLARATION

We, **Pranav Polepaka USN 1CR16IS069, Preetika B, USN 1CR16IS072, Aishwarya Reddy B, USN 1CR15IS127 and Riya Pooniwala, USN 1CR16IS082,** bonafide students of CMR Institute of Technology, Bangalore, hereby declare that the dissertation entitled, **“”** has been carried out by us under the guidance of **Dr. R Chinnaiyan, Asst. Professor**, CMRIT, Bangalore, in partial fulfillment of the requirements for the award of the degree of Bachelor of Engineering in **Information Science Engineering**, of the Visvesvaraya Technological University, Belgaum during the academic year 2019-2020. The work done in this dissertation report is original and it has not been submitted for any other degree in any university.

Aishwarya Reddy B (USN:1CR15IS127)

Pranav Polepaka (USN:1CR16IS069)

Preetika B (USN:1CR16IS072)

Riya Pooniwala (USN:1CR16IS082)

ABSTRACT

Cell is the smallest unit of tissues, whose abnormal growth causes tumor in Brain. Machine Learning based tumor and its stages classification in brain MRI images is presented in this project. This proect is started with the enhancement of the brain MRI images which are obtained from oncology department of University of Maryland Medical Center. The integration of Temper based K-means and modified Fuzzy C-means (TKFCM) clustering algorithm used to segment the MRI images based on gray level intensity in small portion of brain image. The values of K in Temper based K-means algorithm more than the conventional K-means again, automatically updated membership of FCM eradicates the contouring problem of detecting the tumor region. Then, from the segmented images the first order statistic and region property based features are extracted. The first kind of features is used to detect and isolate tumor from normal brain MRI images with Logistic regression. There is second kind which is used to classify the tumors into benign and four malignant stages tumor with ANN.

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany a successful completion of any task would be incomplete without the mention of people who made it possible, success is the epitome of hard work and perseverance, but steadfast of all is encouraging guidance.

So, it is with gratitude that we acknowledge all those whose guidance and encouragement served as beacon of light and crowned our effort with success.

We would like to thank **Dr. Sanjay Jain,** Principal, CMRIT, Bangalore, for providing an excellent academic environment in the college and his never-ending support for the B.E program.

We would like to express our gratitude towards **Dr. Farida Begaum,** Assoc Professor and HOD, Department of Information Science and Engineering CMRIT, Bangalore, who provided guidance and gave valuable suggestions regarding the project.

We consider it a privilege and honour to express our sincere gratitude to our internal guide **Dr. R Chinnaiyan,** Assoc. Professor, Department of Information Science & Engineering and Dr. K N Sudhakar, Assoc. Professor , Department of Information Science and Engineering, CMRIT, Bangalore for their valuable guidance throughout the tenure of this project work.

We would also like to thank all the faculty members who have always been very Co-operative and generous. Conclusively, we also thank all the non-teaching staff and all others who have done immense help directly or indirectly during our project.

AISHWARYA REDDY B

PRANAV POLEPAKA

PREETIKA B

RIYA POONIWALA

`

TABLE OF CONTENTS

`

LIST OF FIGURES

LIST OF TABLES

`

PREAMBLE

1.1 Introduction:

Automated classification and detection of tumors indifferent medical images is motivated by the necessity of high accuracy when dealing with a human life. Also, the computer assistance is demanded in medical institutions due to the fact that it could improve the results of humans in such a domain where the false negative cases must be at a very low rate. It has been proven that double reading of medical images could lead to better tumor detection. Butte cost implied in double reading is very high, that's why good software to assist humans in medical institutions is of great interest nowadays. Conventional methods of monitoring and diagnosing the diseases rely on detecting the presence of particular features by a human observer. Due to large number of patients in intensive care units and the need for continuous observation of such conditions, several techniques for automated diagnostic systems have been developed in recent years to attempt to solve this problem. Such techniques work by transforming the mostly qualitative diagnostic criteria into a more objective quantitative feature classification problem In this project the automated classification of brain magnetic resonance images by using some prior knowledge like pixel intensity and some anatomical features is proposed. Currently there are no methods widely accepted therefore automatic and reliable methods for tumor detection are of great need and interest. The application of PNN in the classification of data for MR images problems are not fully utilized yet. These included the clustering and classification techniques especially for MR images problems with huge scale of data and consuming times and energy if done manually. Thus, fully understanding the recognition, classification or clustering techniques is essential to the developments of Neural Network systems particularly in medicine problems. Segmentation of brain tissues in gray matter, white matter and tumor on medical images is not only of high interest in serial treatment monitoring of "disease burden" in oncologic imaging, but also gaining popularity with the advance of image guided surgical approaches. Outlining the brain tumor contour is a major step in planning spatially localized radiotherapy (e.g., Cyber knife, iMRT) which is usually done manually on contrast enhanced T1-weighted magnetic resonance images (MRI) in current clinical practice. On T1 MR Images acquired after administration of a contrast agent (gadolinium), blood vessels and parts of the tumor, where the contrast can pass the blood brain barrier are observed as hyper intense areas. There are various attempts for brain tumor segmentation in the literature which use a single modality, combine multi modalities and use priors obtained from population atlases.

RELATED WORK

Ivana Despotovi (2013), presented a new FCM-based method for spatially coherent and noise-robust image segmentation. The contribution was 1) The spatial information of local image features is integrated into both the similarity measure and the membership function to compensate for the effect of noise and 2) An anisotropic neighborhood, based on phase congruency features, is introduced to allow more accurate segmentation without image smoothing. The segmentation results, for both synthetic and real images, demonstrate that our method efficiently preserves the homogeneity of the regions and is more robust to noise than related FCM-based methods.

Maoguo Gong (2013), presented an improved fuzzy C-means (FCM) algorithm for image segmentation by introducing a tradeoff weighted fuzzy factor and a kernel metric. The tradeoff weighted fuzzy factor depends on the space distance of all neighboring pixels and their gray-level difference simultaneously. The new algorithm adaptively determined the kernel parameter by using a fast bandwidth selection rule based on the distance variance of all data points in the collection. Furthermore, the tradeoff weighted fuzzy factor and the kernel distance measure are both parameter free. Experimental results on synthetic and real images show that the new algorithm is effective and efficient, and is relatively independent of this type of noise.
Bhagwat et al (2013) they showed that DICOM images produce better results as compared to non medical images. They found that time requirement of hierarchical clustering was least of three and that for Fuzzy C means it was highest for detection of brain tumor. K-means algorithm produces more accurate result compared to Fuzzy c-means and hierarchical clustering.[13] A.Sivaramakrishnan and Dr.M.Karnan(2013) proposed a novel and an efficient detection of the brain tumor region from cerebral image was done using Fuzzy C-means clustering and histogram. The histogram equalization was used to calculate the intensity values of the grey level images. The decomposition of images was done using principle component analysis which was used to reduce dimensionality of the wavelet co - efficient. The results of the proposed Fuzzy C- means (FCM) clustering algorithm successfully and accurately extracted the tumor region from brain MRI brain images[11] Jaskirat kaur et al (2012), described clustering algorithms for image segmentation and did a review on different tyapes of image segmentation techniques. They also proposed a methodology to classify and quantify different clustering algorithms based on their consistency in different applications. They

described the various performance parameters on which consistency will be measured. Roy et al (2012) calculated the tumor affected area for symmetrical analysis. They showed its application with several data sets with different tumor size, intensity and location. They proved that their algorithm can automatically detect and segment the brain tumor. MR images gives better result compare to other techniques like CT images and X-rays.. Image pre-processing includes conversion of RGB image into grayscale image and then passing that image to the high pass filter in order to remove noise present in image.[14] B. Sathya et al (2011), proposed four clustering algorithm; k mean, improved k mean, c mean and improved c mean algorithm. They did an experimental analysis for large database consisting of various images. They analyzed the results using various parameters Hui Zhang et al (2008), compared subjective and supervised evaluation methodology for image segmentation. Subjective evaluation and supervised evaluation, are infeasible in many vision applications, so unsupervised methods are necessary. Unsupervised evaluation enables the objective comparison of both different segmentation methods and different parameterizations of a single method.[6] Martial Heber et al (2005), presented an evaluation of two popular segmentation algorithms, the mean shift-based segmentation algorithm and a graph-based segmentation scheme.

Preprocessing and enhancement techniques are used to improve the detection of the suspicious region from Magnetic Resonance Image (MRI).This section presents the gradient-based image enhancement method for brain MR images which is based on the first derivative and local statistics. The preprocessing and enhancement method consists of two steps; first the removal of film artifacts such as labels and X-ray marks are removed from the MRI using tracking algorithm. [12]Second, the removal of high frequency components using weighted median filtering technique. It gives high resolution MRI compare than median filter, Adaptive filter and spatial filter. The performance of the proposed method is also evaluated by means of peak single-to noiseratio (PSNR), Average Signal-to-Noise Ratio (ASNR).[14]

Image segmentation is the primary step and the most critical tasks of image analysis. Its purpose is that of extracting from an image by means of image segmentation. The mechanization of medical image segmentation has established wide application in diverse areas such as verdict for patients, treatment management planning, and computer-integrated surgery. There are three broad approaches to segmentation, termed, Boundary approach (thresholding), Edgebased approach, Region-based approach.

1. Boundary Approach (Thresholding) In thresholding, pixels are allocated to categories according to the range of values in which a pixel lies. Thresholding is the simplest and most commonly used method of segmentation. Given a single threshold, t, the pixel located at lattice position (i, j), with greyscale value fij , is allocated to category 1 if fij ≤ t or else, the pixel is allocated to category 2.

2. Edge-Based Approach In edge-based segmentation, an edge filter is applied to the image, pixels are categorized as edge or non-edge depending on the filter output, and pixels which are not divided by an edge are owed to the same category. Edge-based segmentation is based on the fact that the position of an edge is given by an extreme of the first-order derivative or a zero crossing in the second-order derivative. There a pixel is classified as an object pixel judging solely on its gray value independently of the context. To improve the results, feature computation and segmentation can be repeated until the procedure converges into a stable result.

3. Region-Based Approach Region-based segmentation algorithms operate iteratively by grouping together pixels which are neighbors and have similar values and splitting groups of pixels which are dissimilar in value. Segmentation may be regarded as spatial clustering. Clustering in the sense that pixels with similar values are grouped together whereas spatial in that pixels in the same category also form a single connected component. Clustering algorithms may be agglomerative, conflict-ridden or iterative. Clustering is the group of a collected works of patterns into clusters based on similarity. [16]Patterns within a valid cluster are more analogous to each one other than they are to a pattern belonging to a dissimilar cluster. Clustering is useful in pattern-analysis, grouping, decision-making, and machine-learning situations, data mining, document recovery, image segmentation, and pattern organization. On the other hand, many such problems, there is little prior information existing about the statistics, and the decision -maker must make as few suppositions about the data as probable [4][6].

Clustering is a learning task, where one needs to identify a finite set of categories known as clusters to categorize pixels. Clustering is primarily used when module are known in progress. A resemblance criteria is defined between pixels [2] and then similar pixels are grouped together to form clusters. A good quality clustering method will produce high quality clusters with high intra-class similarity similar to one another within the same cluster low inter-class similarity and dissimilarity to the objects in further clusters. [9]The superiority of a clustering result depends on both the similarity measure used by the method and its achievement. The eminence of a clustering method is also calculated by its ability to discover. Clustering refers to the classification of objects into groups according to criteria of these objects. In the clustering techniques, an attempt is made to extract a vector from local areas in the image. A standard procedure for clustering is to assign each pixel to the nearest cluster mean. Clustering algorithms are classified as hard clustering (k- means clustering) fuzzy clustering, etc.

1.K-means algorithm is the most well-known and widely-used unsupervised clustering technique in partitioned clustering algorithms. Purpose of this algorithm is to minimize the distances of all the elements to their cluster centres. Most of the algorithms in this field are developed by inspiring or improving k-means. The algorithm upgrades the clusters iteratively and runs in a loop until it reaches to optimal solution.[14] Pseudo-code of the K-means clustering algorithm is shown. Performance of K-means algorithm depends on initial values of cluster centers. Therefore the algorithm should be tested for different outcomes with different initial cluster centers by multi-running.[15]

2.Fuzzy Clustering It is effectively used in pattern recognition and fuzzy modelling. There are various similarity measures used to identify classes depended on the data and the application. Similarity measures for example distance, connectivity, and intensity are used. Its application is in data analysis, pattern recognition and image segments. Fuzzy clustering method can be considered to be superior since they can represent the relationship between the input pattern data and clusters more naturally [14]. Fuzzy c-means is a popular softclustering method. Fuzzy cmeans is one of the most promising fuzzy clustering methods. In most cases, it is more flexible that the corresponding hard-clustering algorithm. Traditional clustering approaches generate partition, each pattern belongs to one and merely single cluster.

That‟s why; the clusters in a hard clustering technique are dislodged. Fuzzy clustering enlarges this notion to connect each pattern with every cluster by means of a membership function. The outcome of such algorithms is a clustering, although not a partition.

3.Segmentation Using ACO Ant colony optimization (ACO) is a population-based meta heuristic that can be used to find approximate solutions to difficult optimization troubles. In ACO, a set of software agents named artificial ants look for for excellent solutions to a given optimization problem. [12]To apply ACO, the optimization problem is changed into the problem of finding the best path on a weighted graph. The artificial ants incrementally build solutions by moving on the graph. The solution construction process is stochastic and is biased by a pheromone model, that is, a set of parameters related with graph components whose values are customized at runtime by the ants. 4.Segmentation Using Genetic Algorithm (GA) Thangavel and Karnan(2005) said a genetic algorithm (GA) is an optimization technique for obtaining the best possible solution in a vast solution space. Genetic algorithms operate on populations of strings, with the string coded to represent the parameter set. The intensity values of the tumor pixels are considered as initial population for the genetic algorithm. The intensity values of the suspicious regions are then converted as 8 bit binary strings and these values are then converted as population strings and intensity values are considered as fitness value for genetic algorithm. [12]Now the genetic operator„s reproduction, crossover and mutation are applied to get new population of strings 5.PSO-Based Clustering Algorithm The algorithm based on swarm intelligence has been developed by adapting the collective behavior which is shown for searching food sources. Each solution in PSO algorithm is a bird in the search space and it is called as a "particle". All particles have a fitness value evaluated by a fitness function and a velocity data that orients their fights. In the problem space, the particles move by following the existing most favourable solutions [12]. PSO algorithm starts with a group of random generated solutions (particles) and optimal solution is investigated iteratively. In each iteration, all particles are updated according to two best values. The first of these best values is that a particle found so far and is called "pbest". The other one is the best value found so far by any particles in the population. This value is the global best value for the population and called as "gbest". PSO is a numeric optimization algorithm in nature. However Omran proposed a PSO-based clustering algorithm in 2004 and he applied this method for image segmentation.

In this approach, optimal cluster centers are determined by PSO which is a population-based search technique. Thus the effects of initial conditions are reduced, compared with classic methods (k-means, fcm).

PROPOSED METHOD

The complete details of the proposed system is given below

Proposed Architecture

The architecture of the system is given below

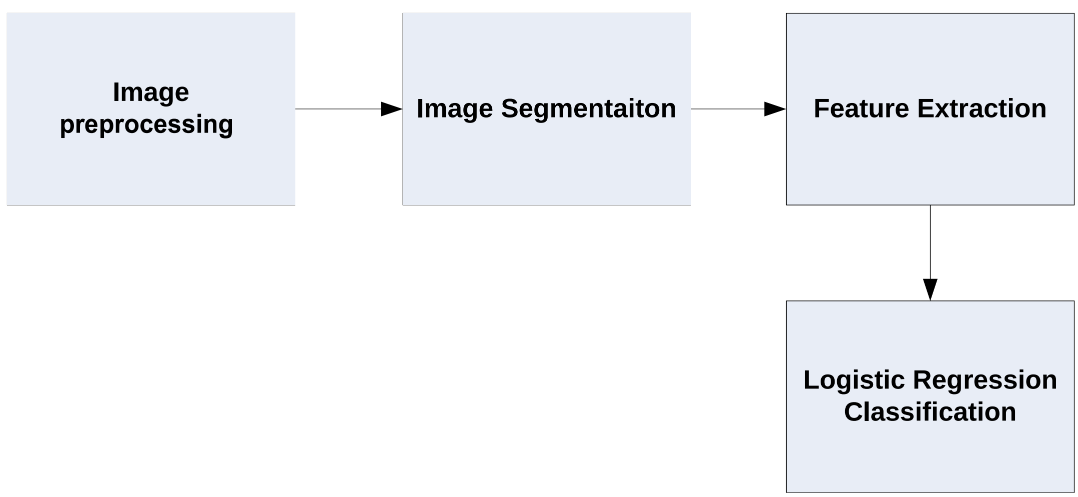


Image Preprocessing

- Smoothening and Sharpening
-    - Gabor Filter Algorithm is used
- Contrast Enhancement
-      - Histogram Equalization is used

Segmentation

- K-means algorithm is used to segment MRI images on the basis of gray level. This gray level is selected depending on the temper of the image. Then the modified Fuzzy c-means algorithm which depends on the updated membership is applied to segment the temper based K-means segmented image

Feature Extraction

- The system will extract the first and second order statistic. The first order statistic feature is used to detect exact tumor and it position in the brain MRI image and second order region based statistic feature is required for distinguishing the malignant tumor and benign tumor. Those are the parameters that will be taken as **first order; contrast, correlation, entropy, energy and Homogeneity.** The region properties of the **segmented MRI image provide area, eccentricity, and perimeter**.

Tumor Detection

- From known Brain Tumor images , features are extracted and Logistic regression is trained.

Tumor Categorization

- Logistic regression is trained for second order statistic to classify a tumor to malign or benign

Summary

The fundamental point of this part is to see if the framework is sufficiently achievable or not. Hence various types of examination, for example, execution investigation, specialized investigation, practical examination and so forth is performed.

SYSTEM DESIGN

System design of the project is given is detail in this chapter.

4.1 System development methodology

System development method is a process through which a product will get completed or a product gets rid from any problem. Software development process is described as a number of phases, procedures and steps that gives the complete software. It follows series of steps which is used for product progress. The development method followed in this project is waterfall model.

4.1.1 Model phases

the waterfall model is a successive programming improvement process, in which advance is seen as streaming relentlessly downwards (like a waterfall) through the periods of Requirement start, Analysis, Design, Implementation, Testing and upkeep.

Prerequisite Analysis: This stage is worried about gathering of necessity of the framework. This procedure includes producing record and necessity survey.

Framework Design: Keeping the prerequisites at the top of the priority list the framework details are made an interpretation of into a product representation. In this stage the fashioner underlines on:- calculation, information structure, programming design and so on.

Coding: In this stage developer begins his coding with a specific end goal to give a full portray of item. At the end of the day framework particulars are just changed over into machine coherent register code.

Usage: The execution stage includes the genuine coding or programming of the product. The yield of this stage is regularly the library, executables, client manuals and extra programming documentation

Testing: In this stage all projects (models) are coordinated and tried to guarantee that the complete framework meets the product prerequisites. The testing is worried with check and approval.

Support: The upkeep stage is the longest stage in which the product is upgraded to satisfy the changing client need, adjust to suit change in the outside environment, right mistakes and oversights beforehand undetected in the testing stage, improve the proficiency of the product.

4.1.2 Reason for choosing waterfall model as development method

- Clear venture destinations.
- Stable undertaking necessities.
- Progress of framework is quantifiable.
- Strict close down necessities.
- Helps you to be great.
- Logic of programming improvement is plainly caught on.
- Production of a formal detail
- Better asset designation.

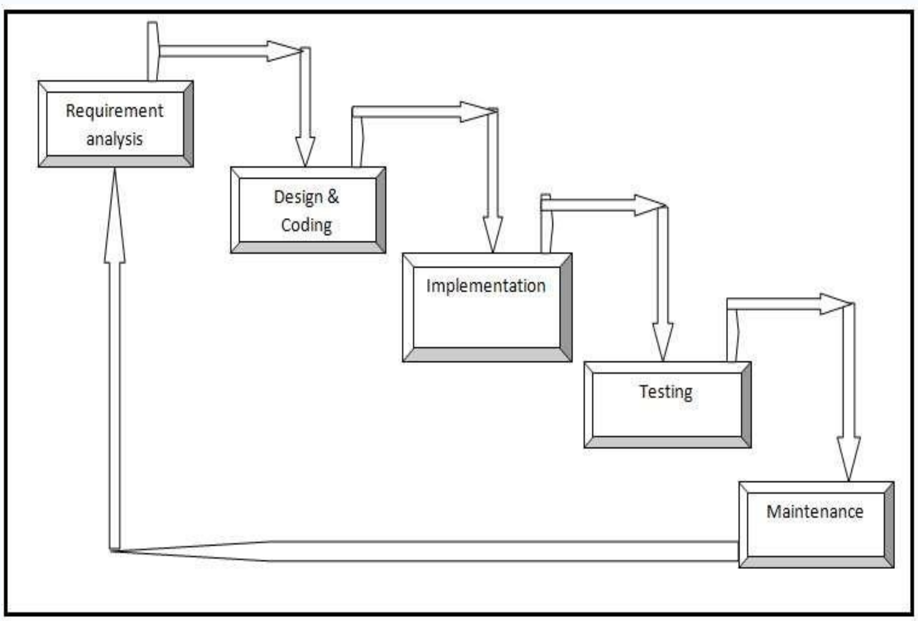
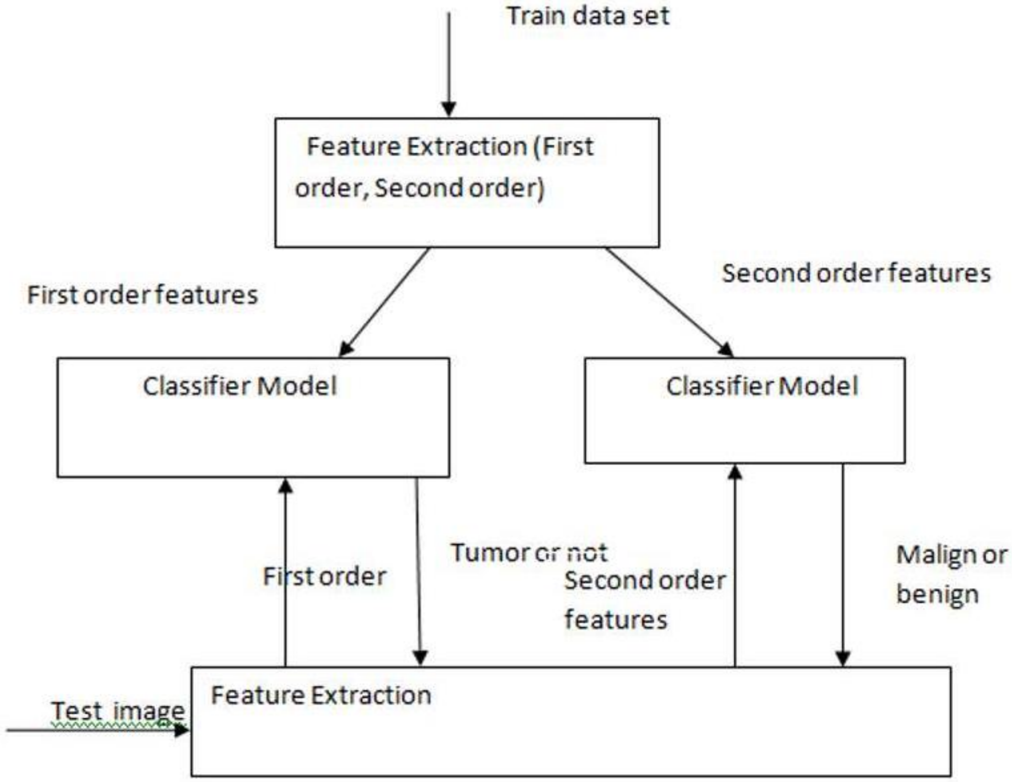


Fig :- Waterfall model

Software Architecture

The System architecture is shown below.

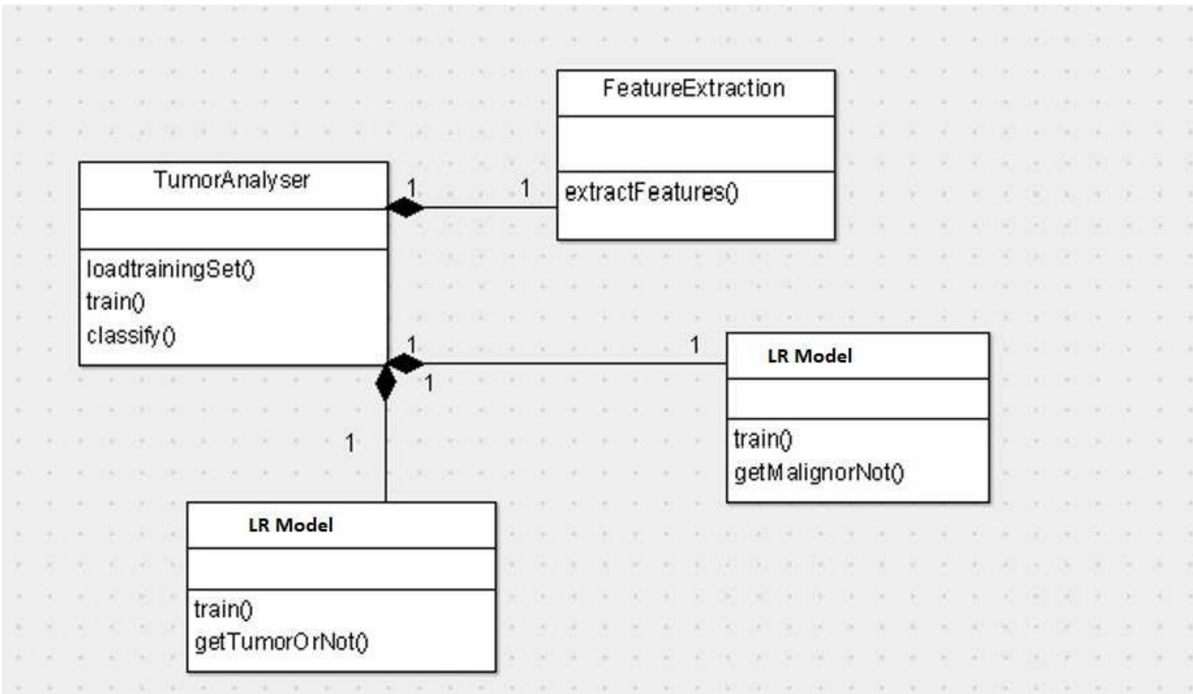


Classes Designed for the system

A class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, and the relationships between the classes.

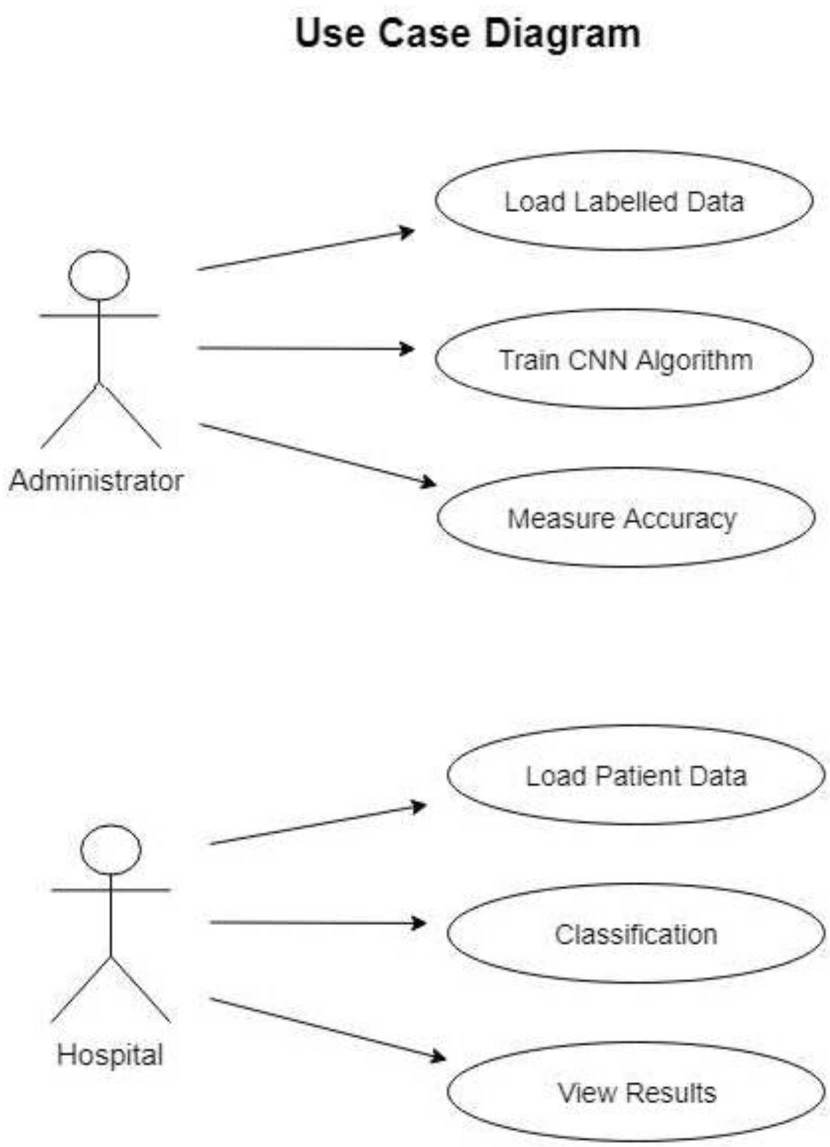
The class diagram is shown below.

The class diagram has the following classes



Use case Diagram of the system

A use case diagram is a type of behavioral diagram created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases.


Use Case Diagram
Load Labelled Data
Train CNN Algorithm
Measure Accuracy
Administrator
Load Patient Data
Classification
View Results
Hospital

There are two users

4.1.1.1 Service user

4.1.1.2 System Manager

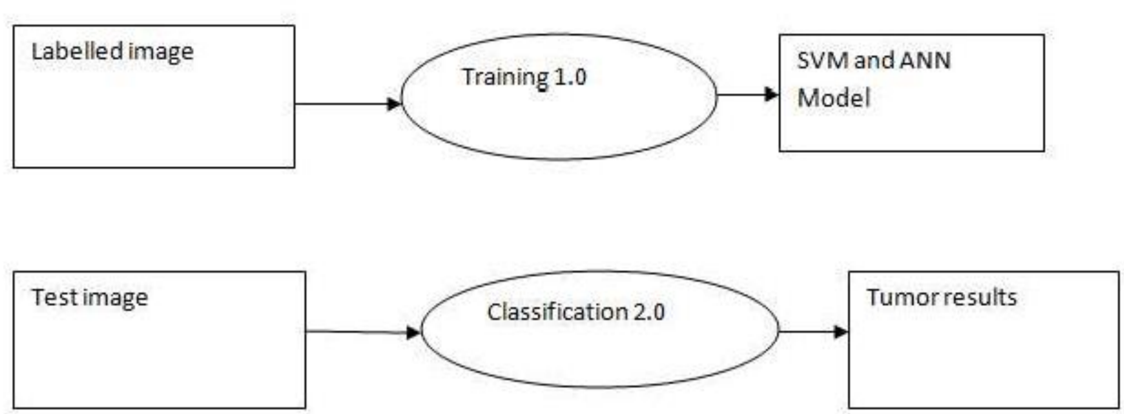
Service user can update the feedback for the service used and also can view the trust values.

System manager can store the feedback about services, and use it to calculate the trust. Calculated trust values are stored and it will provided to customers querying it.

Data Flow Diagram of the system

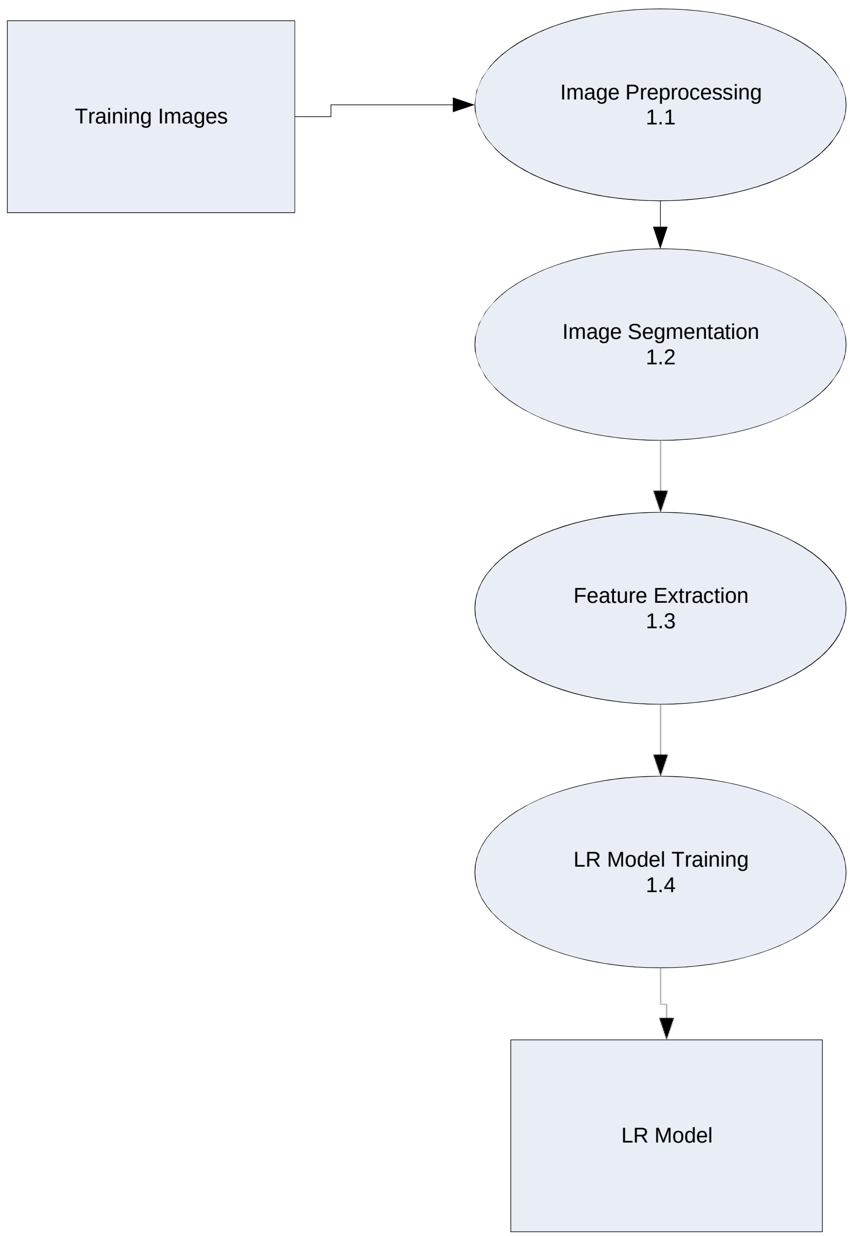
The input , output and the process flow in the system is given in this section.

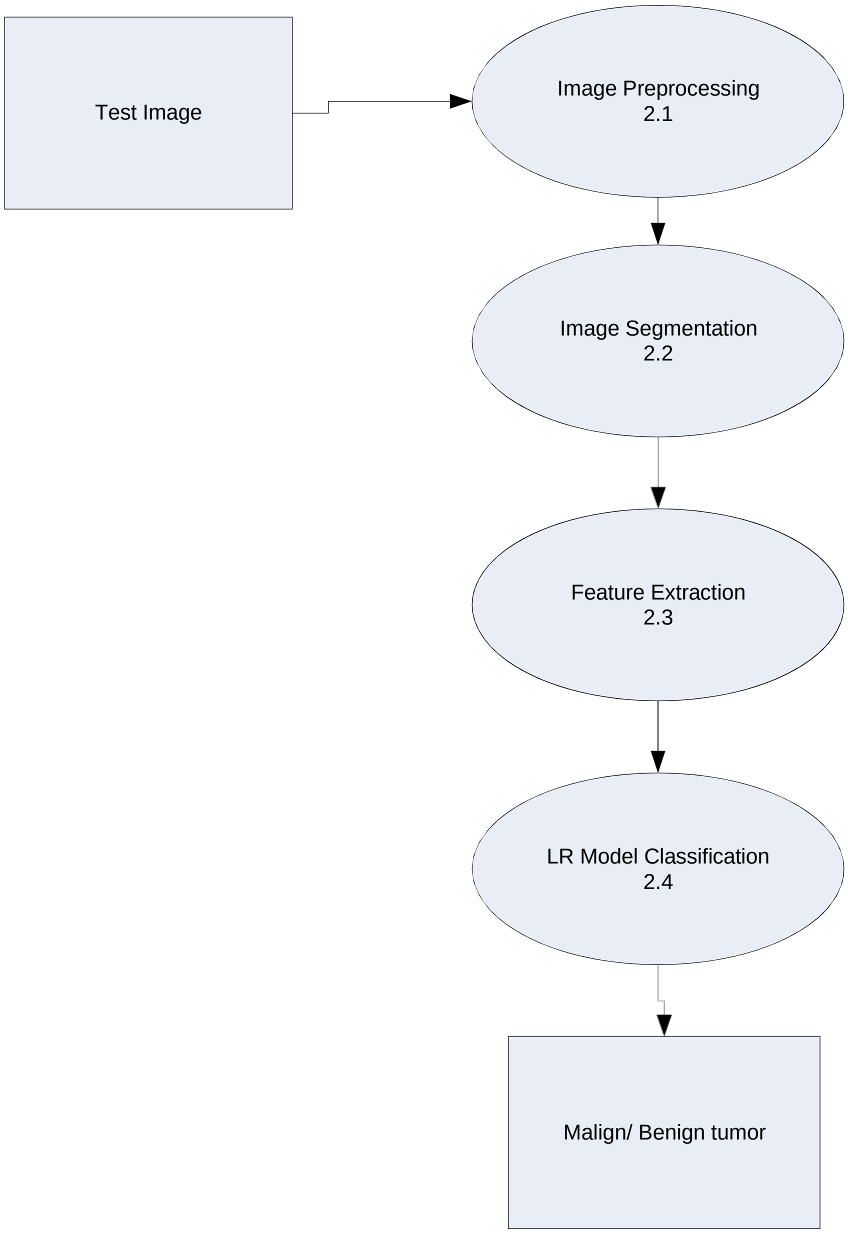
Level 0 Data flow diagram



Level 1 Data flow diagram

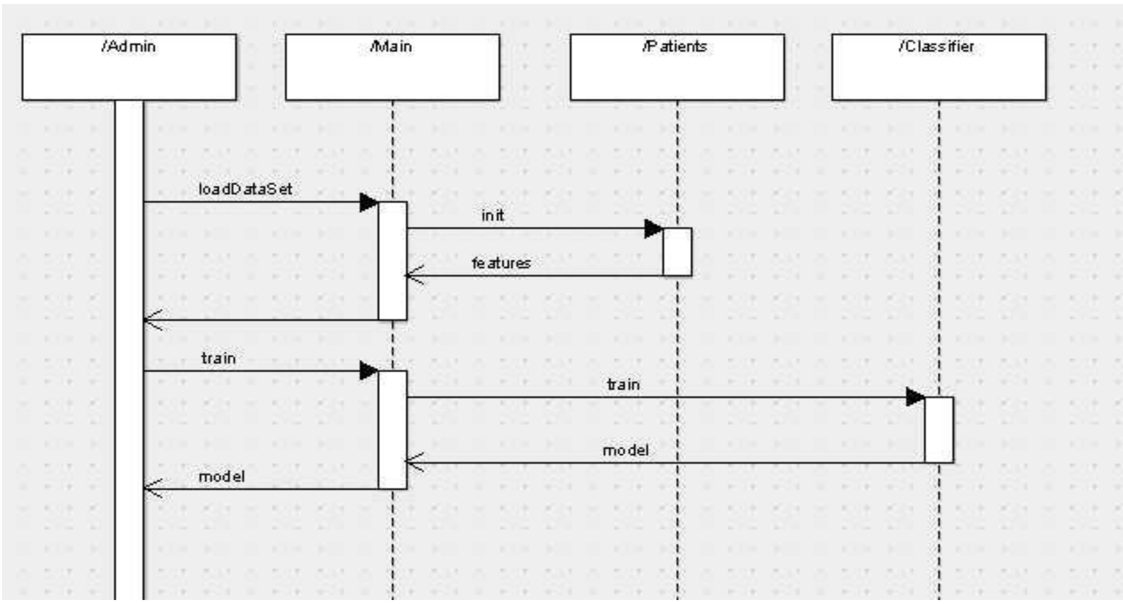
The training and classification process is split to sub process as shown below



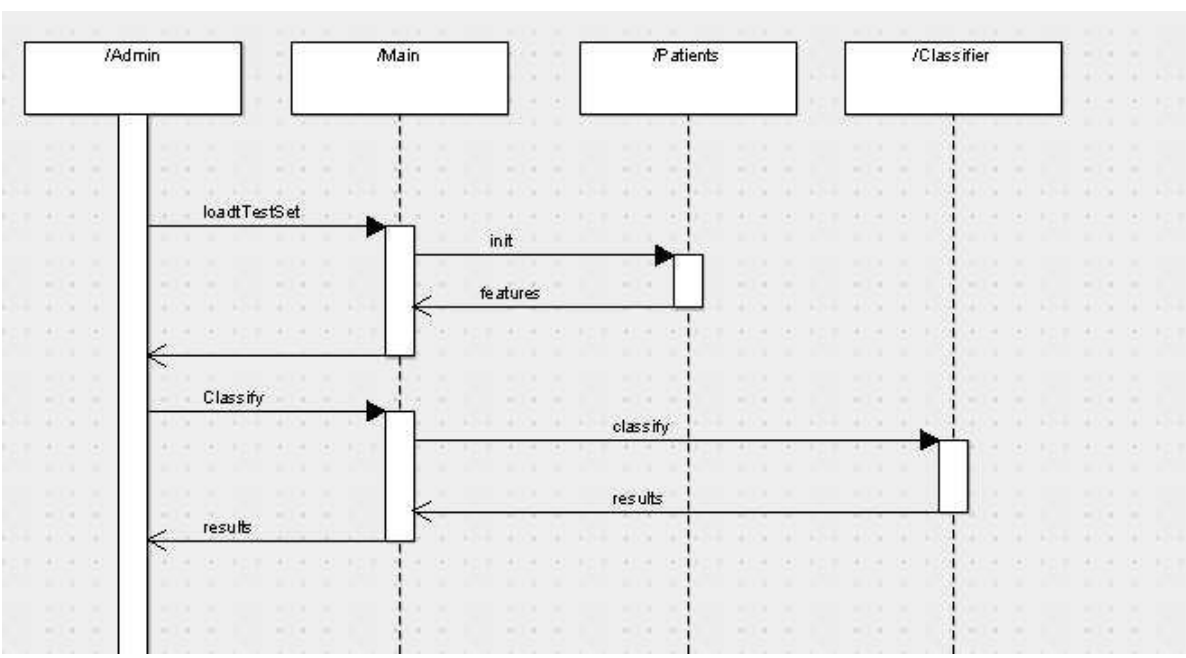

Test Image
Image Preprocessing
2.1
Image Segmentation
2.2
Feature Extraction
2.3
LR Model Classification
2.4
Malign/ Benign tumor

Sequence Diagram of the system

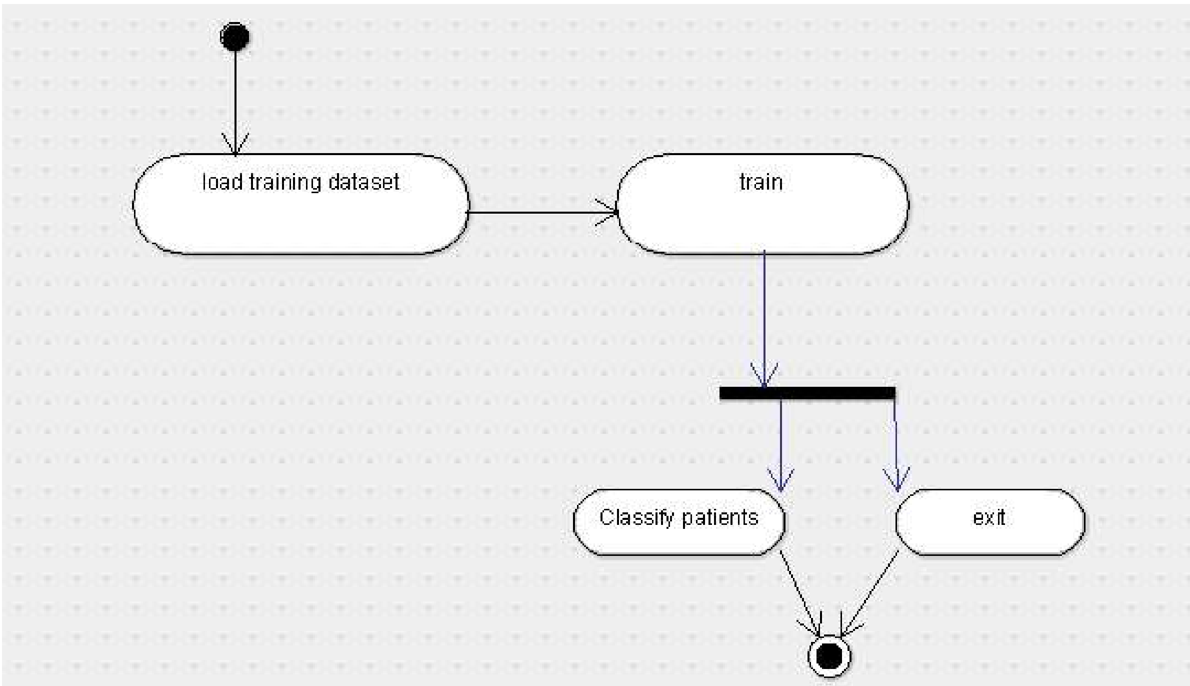
Training Flow



Classification Flow



Activity Diagram



Summary

This chapter mainly concentrates on system architecture, class diagram, sequence diagram, use-case diagram, data flow diagram etc.

IMPLEMENTATION

The implementation details of the project are given in detail in this section.

5.1 Pre-processing

The code snippet for pre processing is given

```
function im=applyHE(img_gray)

[r c]=size(img_gray);

b=zeros(r,c);

hp_fil=[-1 2 -1;0 0 0;1 -2 1];



im=histeq(img_gray);



b=imfilter(img_gray,hp_fil);

c=b+img_gray+25;

medfilt2(c);

im=c;
```

5.2 Segmentation

The code snippet for segmentation is given

```
T = graythresh(c);

bw = im2bw(c,T+0.3);

 bw5=watershed(bw);
```

5.3 Feature Extraction

The code snippet for feature extraction is given below

```
function [noofparts,avgsol,Ea,Ed] = getTumorFeature(filename)



inp = imread(filename);

inp=rgb2gray(inp);

his=applyHE(inp);

[seg,bw]=applySegment(his);

seg=imclearborder(seg,4);

measurements = regionprops(seg, 'BoundingBox', 'Area');
```

```
numt=0;

totsol=0;

for k = 1 : length(measurements)

thisBB = measurements(k).BoundingBox;

 ar = measurements(k).Area

 if ar>100

numt=numt+1;

      sol=measurements(k).Area;

      totsol=totsol+sol;

  end

end

noofparts=numt;

if numt==0

   avgsol=0;

else

   avgsol=totsol*1.0/numt;

end
```

```
[C,L] = wavedec(inp,4,'db5');

[Ea,Ed] = wenergy(C,L);
```

5.4 Tumor Classfication

```
np=rgb2gray(inp);

his=applyHE(inp);

axes(handles.axes2);

imshow(his,[]);

[seg,bw]=applySegment(his);

axes(handles.axes3);

imshow(bw);

axes(handles.axes4);

imagesc(seg);

seg=imclearborder(seg,4);

measurements = regionprops(seg, 'BoundingBox', 'Area');
```

```
numt=0;

for k = 1 : length(measurements)

  thisBB = measurements(k).BoundingBox;

  ar = measurements(k).Area;

  if ar>70
      rectangle('Position', [thisBB(1),thisBB(2),thisBB(3),thisBB(4)],...
      'EdgeColor','r','LineWidth',2 )
      numt=numt+1;
  end

end

%fprintf(1,'The value of numt=%d \n',numt);

if numt>=2

    set(handles.mc,'String','Tumor Detected');
    [noofparts,avgsol,Ea,Ed]=getTumorFeature(vfilename);

    [r c]=size(tumor);
    ldis=99999;
    ldiswho=-1;

    for i=1:r
```

```
dis= abs(noofparts-tumor(i,1)) + abs(avgsol-tumor(i,2)) + abs(Ea-tumor(i,3)) + abs(Ed(1)-tumor(i,4)) + abs(Ed(2)-tumor(i,5)) + abs(Ed(3)-tumor(i,6)) + abs(Ed(4)-tumor(i,7));

      if dis<ldis
          ldis=dis
          ldiswho=i;
      end
    end
    le=tumor(ldiswho,8);



    if le==1
    set(handles.seve,'String','Malign');
    else
      set(handles.seve,'String','Benign');
    end

    set(handles.numparts,'String',num2str(noofparts));

else
    set(handles.mc,'String','No Tumor Detected');
     set(handles.seve,'String','None');
end
```

Summary

This chapter mainly concentrates on implementation of proposed solution.

6. TESTING

This section details the testing carried out on the project code

.

6.1 Testing

Since we have followed OOPS, each class is a unit, so unit test cases for written for each class and tested.

Classes integrated	Tests done	Remarks
Class: **TumorAnalyzer**	Class tested for all user interface working	Success
Class: **FeatureExtraction**	Class tested to check whether features are extracted	Success
Class: **ANNModel**	Class tested to check whether congnitive neural network is working	Success
Class: **SVMModel**	Class tested to check whether SVM model is working	Success

6.2 Validation Testing

At the culmination of integration testing, software is completed and assembled as a package. Interfacing errors are uncovered and corrected. Validation testing can be defined in many ways. Here the testing validates the software function in a manner that is reasonably expected by the customer.

Functionality to be tested	**Input**	**Tests done**	**Remarks**
Working of Front-End	User interaction with help of a mouse and keyboard	Appropriate forms open when buttons are clicked	Success
Working of training	User has to upload images in training folder and press train button	Training is done and model is created	Success
Working of classificaiton	User has to browse and choose the image and press classify	Tumor detected and results displayed	Success

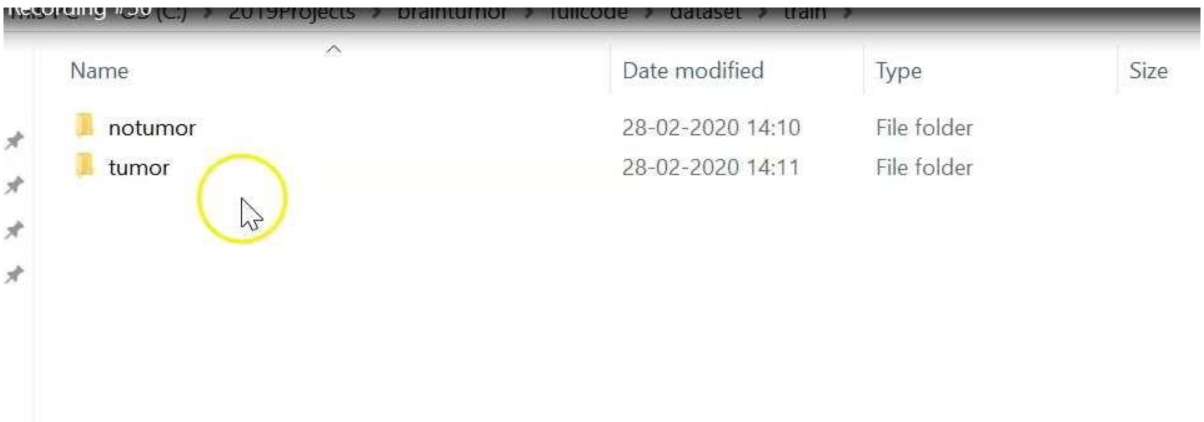
Summary

This code is well tested and all defected are removed from the code.

INTERPRETATION OF RESULT

The following snapshots define the results or outputs that we will get after step by step execution of all the modules of the system.

Place the training images in the training folder



Run the extract to extract the features from the training images

[redacted]

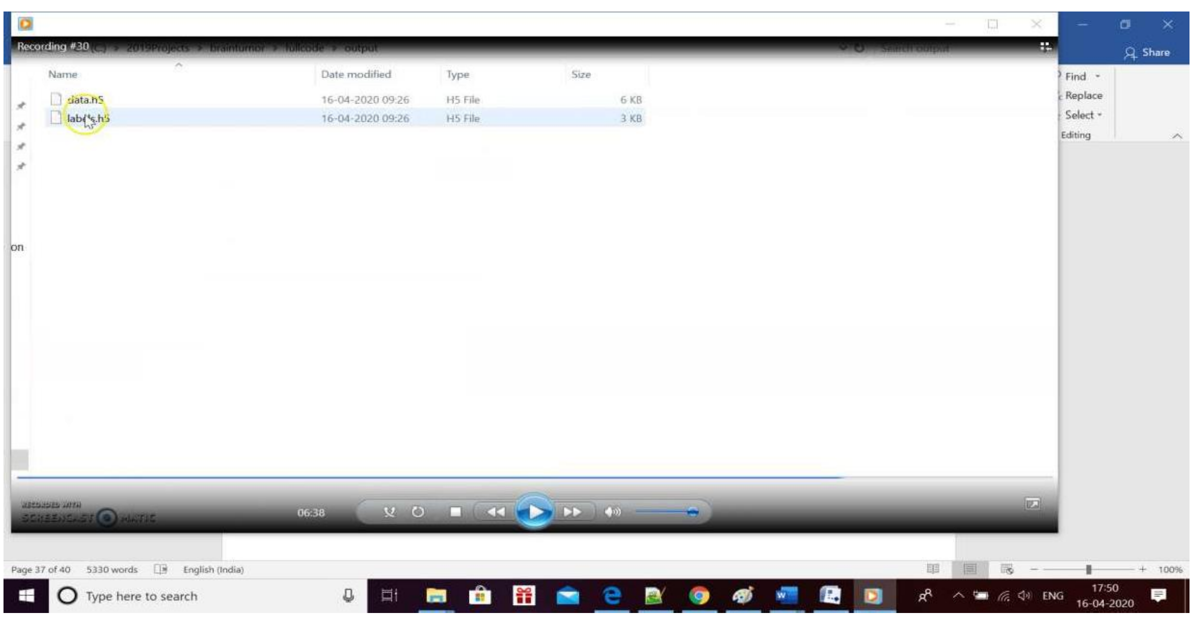
File Edit Format n Options Window Help

```
    # comput      unModule      FS      at use vector                                    a
    gray=ima   Run.:Customied SW%'
    haralick                           ick(gray).mean(axis=0)

    l   xetu£n    Check Module     Alt+X
    Cet.t iW.     Python Shell

# feature-descriptor-3: Color Histogram
def fd_histogram(image, mask=N‹:roe):
    # convert the image to HSV color—space
    image - cv2.cvtColor(image, cv2.COLOR_BGR2HSV)
    # compute the color histogram
    hist  = cv2.calcHist([image), [0, 1, 2], N:n-, {bins, bins, bins], [0, 256,
    # normalize the histogram
    cv2.normalize(hist, hist)
    # return the histogram
    return hist.flatten()

# get the training labels
train_labels = os.listdir(trainpath)
print(train_labels)

# sort the training labels
train_labels.sort()
print(train_labels)

# empty lists to hold feature vectors and labels
global_features = []
labels = []
```

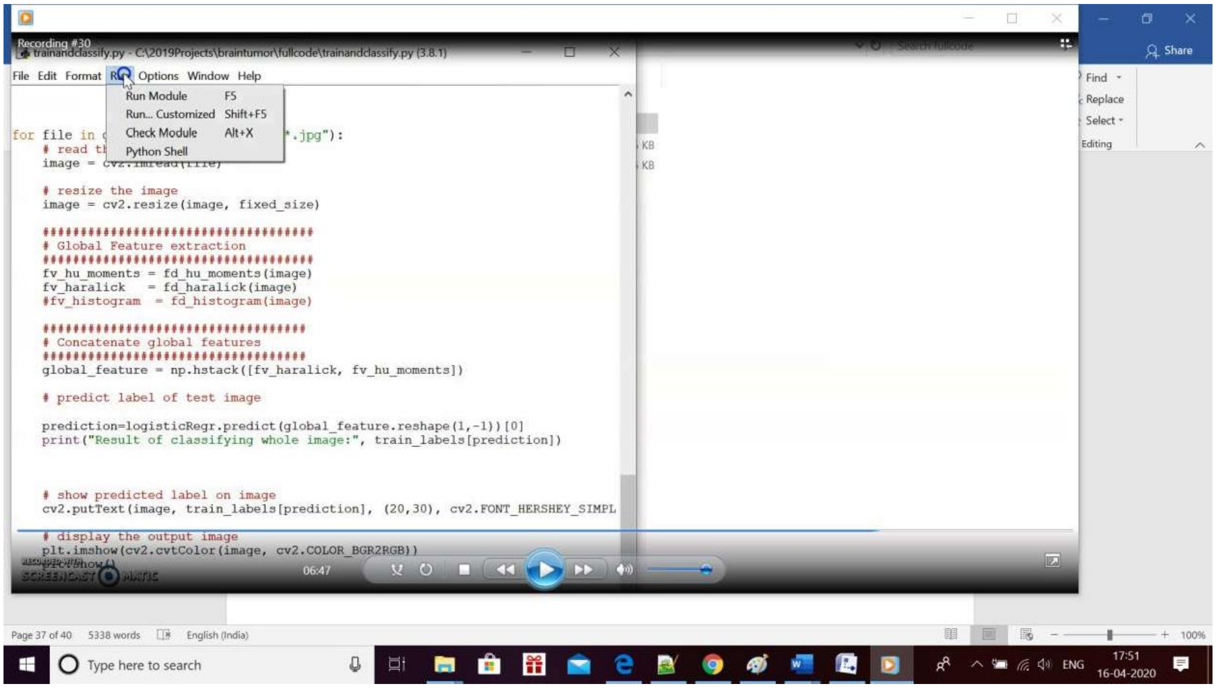
Extracting features from the training images

```
Type "help", "copyright", "credits" or "license()" for more information.
>>>
=========== RESTART: C:\2019Projects\braintumor\fullcode\extractf.py ===========
['notumor', 'tumor']
['notumor', 'tumor']
dataset/train\notumor/1.jpg
dataset/train\notumor/2.jpg
dataset/train\notumor/3.jpg
dataset/train\notumor/4.jpg
dataset/train\notumor/5.jpg
dataset/train\notumor/6.jpg
dataset/train\notumor/7.jpg
dataset/train\notumor/8.jpg
dataset/train\notumor/9.jpg
dataset/train\notumor/10.jpg
[STATUS] processed folder: notumor
dataset/train\tumor/1.jpg
dataset/train\tumor/2.jpg
dataset/train\tumor/3.jpg
dataset/train\tumor/4.jpg
dataset/train\tumor/5.jpg
dataset/train\tumor/6.jpg
dataset/train\tumor/7.jpg
dataset/train\tumor/8.jpg
dataset/train\tumor/9.jpg
dataset/train\tumor/10.jpg
[STATUS] processed folder: tumor
[STATUS] completed Global Feature Extraction...
[STATUS] feature vector size (20, 20)
[STATUS] training Labels (20,)
[STATUS] training labels encoded...
```

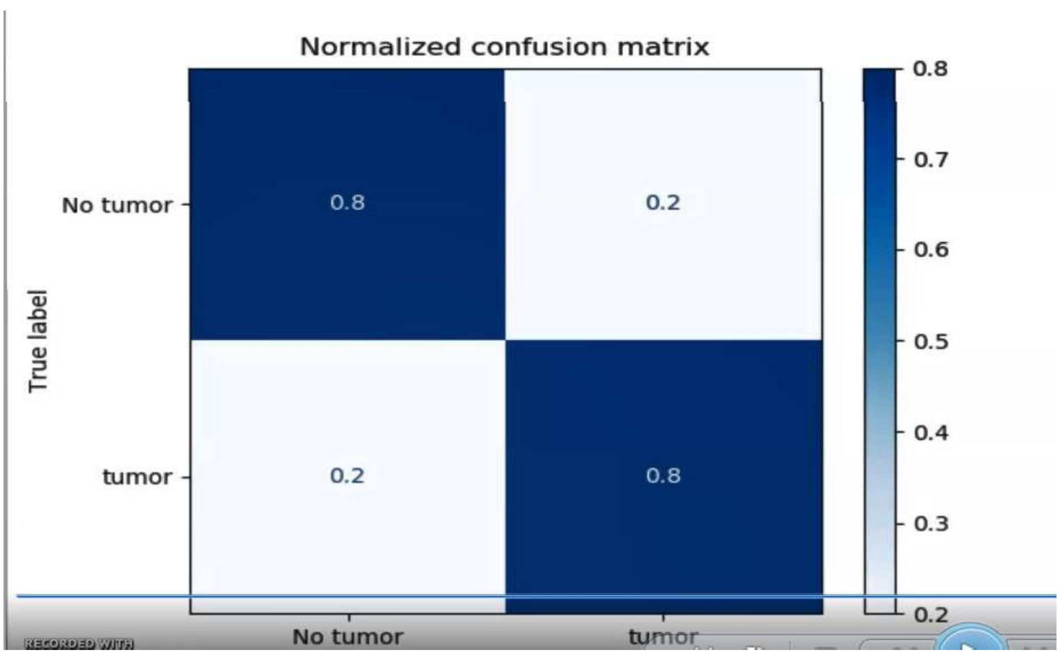
Training features and labels written to a file



Run the trianandclassify file



Classifier is trained and normalized confusion matrix is printed



Test image is classified and result is shown

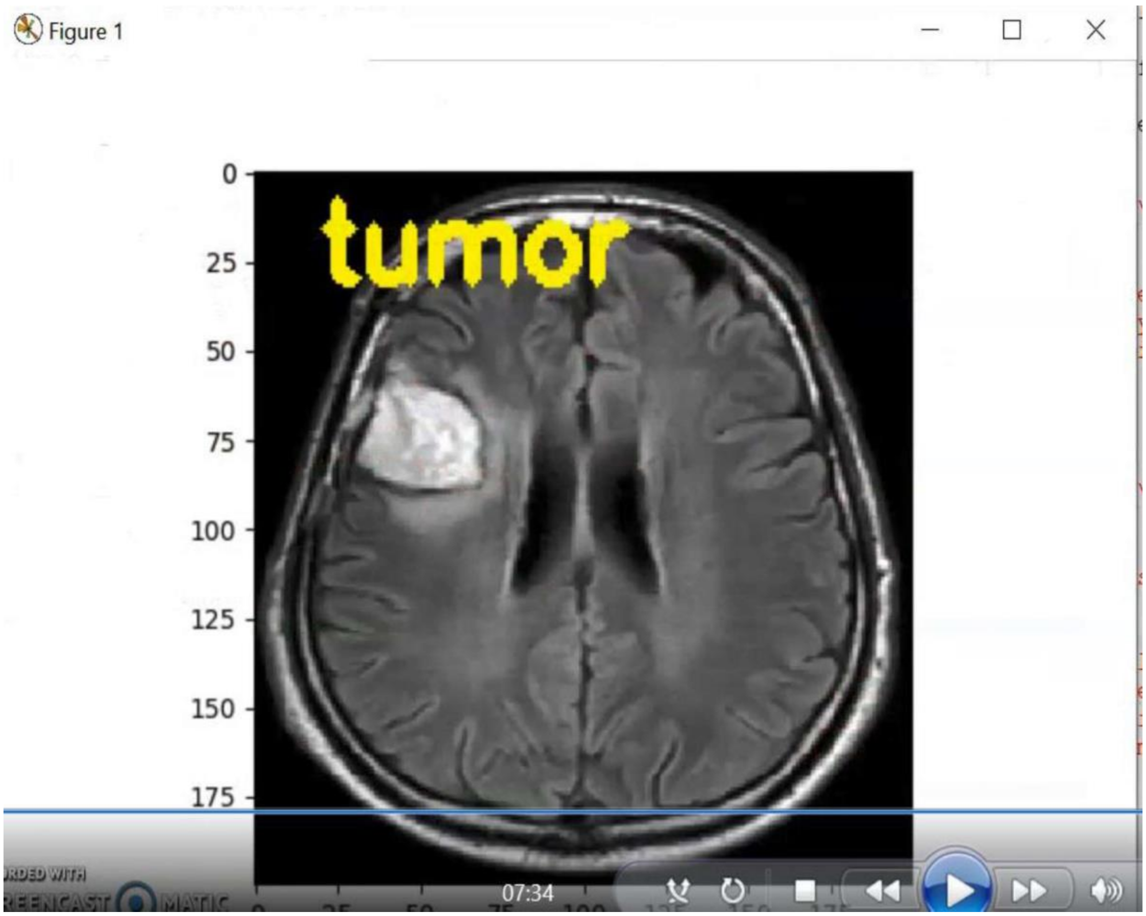


Figure 1
tumor
0
25
50
75
100
125
150
175
07:34

CONCLUSION

The abnormality of the tissues is needed to be identified or classified for the betterment of the human body. In this paper, the incorporated SVM and ANN based classification technique is proposed. The processed brain MRI images are firstly segmented using TKFCM algorithm. In which the k value and updated membership are different from conventional process. There are two kind of features have been extracted from segmented images for the purpose of isolating and classifying tumor. The first kind of statistical features are used for the classifying normal and abnormal brain MRI image using Logistic Regression. The tumor categories and malignant tumor stages are classified through Logistic regression.

References

[1] Marshland, S. (2009) Machine Learning an Algorithmic Perspective. CRC Press, New Zealand, 6-7. 4

[2] Sharma, P. and Kaur, M. (2013) Classification in Pattern Recognition: A Review. International Journal of Advanced Research in Computer Science and Software E ngineering , 3, 298.

[3] Rambhajani, M., Deepanker, W. and Pathak, N. (2015) A Survey on Implementation of Machine Learning Techniques for Dermatology Diseases Classification. I nternational Journal of Advances in Engineering & Technology , 8, 194-195.

[4] Kononenko, I. (2001) Machine Learning for Medical Diagnosis: History, State of the Art and Perspective. Journal of Artificial Intelligence in Medicine , 1, 89-109.

[5] Otoom, A.F., Abdallah, E.E., Kilani, Y., Kefaye, A. and Ashour, M. (2015) Effective Diagnosis and Monitoring of Heart Disease. International Journal of Software En gineering and Its Applications . 9, 143- 156.

[6] Vembandasamy, K., Sasipriya, R. and Deepa, E. (2015) Heart Diseases Detection Using Naive Bayes Algorithm. IJISET - International Journal of Innovative Science , Engineering & Technology , 2, 441-444.

[7] Chaurasia, V. and Pal, S. (2013) Data Mining Approach to Detect Heart Disease. International Journal of Advanced Computer Science and Information Technology (IJACSIT), 2, 56-66.

[8] Parthiban, G. and Srivatsa, S.K. (2012) Applying Machine Learning Methods in Diagnosing Heart Disease for Diabetic Patients. International Journal of Applied I nformation Systems (IJAIS), 3, 25-30.

[9] Tan, K.C., Teoh, E.J., Yu, Q. and Goh, K.C. (2009) A Hybrid Evolutionary Algorithm for Attribute Selection in Data Mining. Journal of Expert System with Appl ications , 36, 8616-8630. https://doi.org/10.1016/j.eswa.2008.10.013

[10] Karamizadeh, S., Abdullah, S.M., Halimi, M., Shayan, J. and Rajabi, M.J. (2014) Advantage and Drawback of Support Vector Machine Functionality. 2014 IEEE I nternational Conference on Computer , Communication and Control Technology (I 4 CT), Langkawi, 2-4 September 2014, 64-65. https://doi.org/10.1109/i4ct.2014.6914146

[11] Iyer, A., Jeyalatha, S. and Sumbaly, R. (2015) Diagnosis of Diabetes Using Classification Mining Techniques. International Journal of Data Mining & Knowledge Ma nagement Process (IJDKP), 5, 1-14. https://doi.org/10.5121/ijdkp.2015.5101

[12] Sen, S.K. and Dash, S. (2014) Application of Meta Learning Algorithms for the Prediction of Diabetes Disease. International Journal of Advance Research in Compu ter Science and Management Studies , 2, 396-401.

[13] Kumari, V.A. and Chitra, R. (2013) Classification of Diabetes Disease Using Support Vector Machine. International Journal of Engineering Research and Applic ations (IJERA), 3, 1797-1801.

[14] Sarwar, A. and Sharma, V. (2012) Intelligent Naïve Bayes Approach to Diagnose Diabetes Type-2. Special Issue of International Journal of Computer Applications (0975-8887) on Issues and Challenges in Networking , Intelligence and Computing Technologies - ICNICT 2012, 3, 14-16.

[15] Ephzibah, E.P. (2011) Cost Effective Approach on Feature Selection using Genetic Algorithms and Fuzzy Logic for Diabetes Diagnosis. International Journa l on Soft Computing (IJSC), 2, 1-10. https://doi.org/10.5121/ijsc.2011.2101

[16] Archana, S. and DR Elangovan, K. (2014) Survey of Classification Techniques in Data Mining. International Journal of Computer Science and Mobile Applications , 2, 65-71

[17] Vijayarani, S. and Dhayanand, S. (2015) Liver Disease Prediction using SVM and Naïve Bayes Algorithms. International Journal of Science , Engineering and Tec hnology Research (IJSETR), 4, 816-820.

[18] Gulia, A., Vohra, R. and Rani, P. (2014) Liver Patient Classification Using Intelligent Techniques. (IJCSIT) International Journal of Computer Science and Inform ation Technologies , 5, 5110-5115.

[19] Rajeswari, P. and Reena,G.S. (2010) Analysis of Liver Disorder Using Data Mining Algorithm. Global Journal of Computer Science and Technology , 10, 48-52.

[20] Tarmizi, N.D.A., Jamaluddin, F., Abu Bakar, A., Othman, Z.A., Zainudin, S. and Hamdan, A.R. (2013) Malaysia Dengue Outbreak Detection Using Data Mining Models. Journal of Next Generation Information Technology (JNIT), 4, 96-107.