

ABSTRACT

Many viral infections which are present today generates a major impact on humans. There are numerous diseases which are currently present in our society. Some are acute and some are chronic. The chronic section is further classified into two sub-categories, one is Easily diagnosed and another is laboriously diagnosed. One of such laboriously diagnosed disease that our project focusses on is Japanese Encephalitis, one of the rarest disease to be diagnosed, leading death of many children in Bihar recently. Japanese Encephalitis primarily affects children. Most adults in endemic countries have natural immunity after childhood infection, but individuals of any age may be affected.

This project deals with the data of those who are affected. The data would also consist of some resources which were obtained from the internet from various Organisation Websites. The primary step is studying the data obtained to figure out the unique and similar symptoms which are present in Japanese Encephalitis in comparison with normal Viral Fever.

The fundamental concept which is used in the project is Machine Learning. Since the project deals with the Human Body and its response to the disease, Genetic Algorithm is used. To obtain precise results along with the justification, Attribute Selection algorithm is also used.

In this advanced pharmaceutical world many rear health diseases are boomed to affect the humans via style changes, food habits and biological infections. JE is one of such viral infection disease, which is infected on brain with less symptoms and occasional inflammation of the brain occurs mostly concentrated on children and adults but it is not a chronic diseases, which takes less time to get intense. The main objective of the work is to create the preventive awareness of the disease at the initial stage. Exact the essential features of bio test from the affected person, which is taken to consideration with genetic algorithm and Attribute Selection algorithm. Genetic algorithms gives the higher quality for optimised problem and produce approximate result using Attribute Selection algorithm.

INTRODUCTION

Japanese encephalitis (JE) is the leading form of viral encephalitis in Asia. It is caused by the JE virus (JEV), which belongs to the family Flaviviridae. JEV is endemic to many parts of Asia, where periodic outbreaks take hundreds of lives.

JE, one of the leading forms of viral encephalitis worldwide, is prevalent mostly in eastern and southern Asia, covering a region with a population of more than 3 billion. The disease affects mostly children. Around 30,000–50,000 cases of JE and up to 15,000 deaths are reported annually, although these statistics may be a gross underestimation because of inadequate surveillance and reporting. About 25%–30% of JE cases are fatal, and 50% result in permanent neuropsychiatric sequelae

Signs and Symptoms:- The Japanese encephalitis virus (JEV) has an incubation period of 2 to 26 days.^[4] The vast majority of infections are asymptomatic: only 1 in 250 infections develop into encephalitis.^[5]

Severe rigours may mark the onset of this disease in humans. Fever, headache and malaise are other non-specific symptoms of this disease which may last for a period of between 1 and 6 days. Signs which develop during the acute encephalitic stage include neck rigidity, cachexia, hemiparesis, convulsions and a raised body temperature between 38–41 °C (100.4–105.8 °F). Mental retardation is usually developed.

Mortality of this disease varies but is generally higher in children. Transplacental spread has been noted. Lifelong neurological defects such as deafness, emotional lability and hemiparesis may occur in those who have had central nervous system involvement. In known cases, some effects also include nausea, headache, fever, and vomiting.

Causes:- JEV is a virus from the family Flaviviridae, part of the *Japanese encephalitis serocomplex* of 9 genetically and antigenically related viruses, some which are particularly severe in horses, and four known to infect humans including West Nile virus.^[9] The enveloped virus is closely related to the West Nile virus and the St. Louis encephalitis virus. The positive sense single-stranded RNA genome is packaged in the capsid which is formed by the capsid protein. The outer envelope is formed by envelope protein and is the protective antigen. It aids in entry of the virus into the inside of the cell. The genome also encodes several nonstructural proteins (NS1, NS2a, NS2b, NS3, N4a, NS4b, NS5). NS1 is produced as secretory form also. NS3 is a putative helicase, and NS5 is the viral polymerase.

Diagnosis:- Japanese encephalitis is diagnosed by commercially available tests detecting JE virus-specific IgM antibodies in serum and /or cerebrospinal fluid, for example by IgM capture ELISA.

JE virus IgM antibodies are usually detectable 3 to 8 days after onset of illness and persist for 30 to 90 days, but longer persistence has been documented. Therefore, positive IgM antibodies occasionally may reflect a past infection or vaccination.

Prevention:- There is no specific treatment for Japanese encephalitis and treatment is supportive, with assistance given for feeding, breathing or seizure control as required. Raised intracranial pressure may be managed with mannitol. There is no transmission from person to person and therefore patients do not need to be isolated.

JE in India:- In India, epidemics of JE are reported from many parts of the country, and it is considered a major paediatric problem. The first recognition of JE based on serological surveys was in 1955, in Tamil Nadu, India. A total of approximately 65 cases were reported between 1955 and 1966 in Southern India. Subsequent surveys carried out by the National Institute of Virology of Pune indicated that approximately

half of the population in Southern India has neutralizing antibodies to the virus. Since 1955, many major outbreaks in different parts of the country have been reported. A major outbreak resulting in a 42.6% fatality rate was reported in the Bankura District of West Bengal in 1973. Subsequently, the disease spread to other states and caused a series of outbreaks in different parts of the country. In 1978, cases were reported from 21 states and union territories. In Uttar Pradesh, the first major JE epidemic occurred in Gorakhpur in 1978, with 1,002 cases and 297 deaths reported. Many outbreaks were reported in Gorakhpur after the 1978 JE outbreak, with varying intensity and magnitude. Since 1978 to 2005, this encephalitis has taken more than 10,000 lives in the state. The 2005 epidemic surpassed all previous reported outbreaks in the country. In that year, Uttar Pradesh faced a devastating outbreak of JE, mostly confined to Gorakhpur, with 6,061 cases and 1,500 deaths; another outbreak occurred in 2006, with 2,320 cases and 528 deaths. Similarly, JE cases in Uttar Pradesh were confined predominantly to Gorakhpur during 2007, with 3,024 cases and 645 deaths, and then onwards till 2007 there have been 103,389 reported cases in India, and 33,729 deaths. Approximately 597,542,000 people in India live in JE-endemic regions, and 1,500 to 4,000 cases are reported every year. These figures are based on total reported cases; it is possible that many cases are unreported and hence the actual magnitude of the threat of JE may be considerably higher, both in the Indian and in the global context.. The trend of JE suggests that the problem in Northern India is escalating, and larger epidemics may occur in the future.

Our Project :- The motive of our project is to find the methods to prevent JE. We tried to prevent it with the help of machine learning algorithms by applying them on our real time data. The primary step is studying the data obtained to figure out the unique and similar symptoms which are present in Japanese Encephalitis in comparison with normal Viral Fever. The main objective of the work is to create the preventive awareness of the disease at the initial stage. Extract the essential features of bio test from the affected person, which is taken into consideration with genetic algorithm and fuzzy inference system. Genetic algorithms gives the higher quality for optimised problem and produce approximate result using fuzzy inference system.

PROBLEM STATEMENT

Problem: This project intends to solve the problem of Japanese Encephalitis by detecting it using Machine Learning Algorithms.

Our Vision: The main purpose of using this project is to simplify the detection of JE which is the main cause of viral encephalitis in many countries of Asia with an estimated 68,000 clinical cases every year.

Issue Statement: The problem is important to solve because there is no cure for the disease and treatment is only focused on relieving severe clinical signs and supporting the patient to overcome the infection.

Methods: The project initially elicit the basic symptom of the disease and thereafter they are analysed. Symptoms such as chill, runny nose, headache, fever, age, blood group, RBC & WBC count, occurrence, time period, patient's current status, food choice(Veg or Non-Veg) and region(Cold or Warm) all together are analysed using Genetic algorithm.

The result obtained will be a string consisting of 1s and 0s, where 1 denotes the presence of symptom and 0 denotes the absence of the symptom.

The final attribute in the resulting string will determine the presence of JE virus. Attribute selection algorithm was also used to extract the major symptoms to regenerate the string which would elucidate the prediction.

The discovery process also includes the infected brain cell images which are compare to the uninfected brain cell images using opencv algorithm which would determine whether infection is present or not by detecting changes in the cell image i.e change in colour due to infection.

Need of the project:

Five Ws:

- **Who** the problem affects: People who live in warmer areas are more prone to this disease. Insects and mosquitos also play a major role in causing this disease, thus people who are exposed to the environment where mosquitoes breeding takes place can be a prey of this disease.
- **What** the outcome would be if the problem was not solved: The treatment is focused on relieving severe clinical signs and supporting the patient to overcome the infection. If the person is not cured, the disease can be considered as fatal.
- **Where** the problem is taking place: The major countries of Asia. Here the temperature is high and also the presence of flavivirus related to JE, dengue, yellow fever and west nile viruses which are spread by mosquitoes.
- **When** the problem needs to be fixed: As soon as the symptoms are determined in the patient, it is mandatory to perform analysis of this disease which is being discussed to prevent the death of the patient.
- **Why** is it important for the problem to be fixed: This topic for the project was selected due to the occurrence of mass death of children in the state of Bihar due to this disease. Therefore in such circumstances people must be aware of possible diseases and they must have the access to detect and treat the disease before it is too late. The major concern of this project is towards the people who do not know the treatment procedures and who are not aware of the consequences they might have to face if the disease is detected.

Aim: Considering the above mentioned points, a conclusion can be made, whether it is possible to detect JE in person based on his/her symptoms.

Literature Survey

Japanese encephalitis virus JEV is the most important cause of viral encephalitis in Asia. It is a mosquito-borne flavivirus, and belongs to the same genus as dengue, yellow fever and West Nile viruses.

The first case of Japanese encephalitis viral disease (JE) was documented in 1871 in Japan.

The annual incidence of clinical disease varies both across and within endemic countries, ranging from <1 to >10 per 100 000 population or higher during outbreaks. A literature review estimates nearly 68 000 clinical cases of JE globally each year, with approximately 13 600 to 20 400 deaths. JE primarily affects children. Most adults in endemic countries have natural immunity after childhood infection, but individuals of any age may be affected.

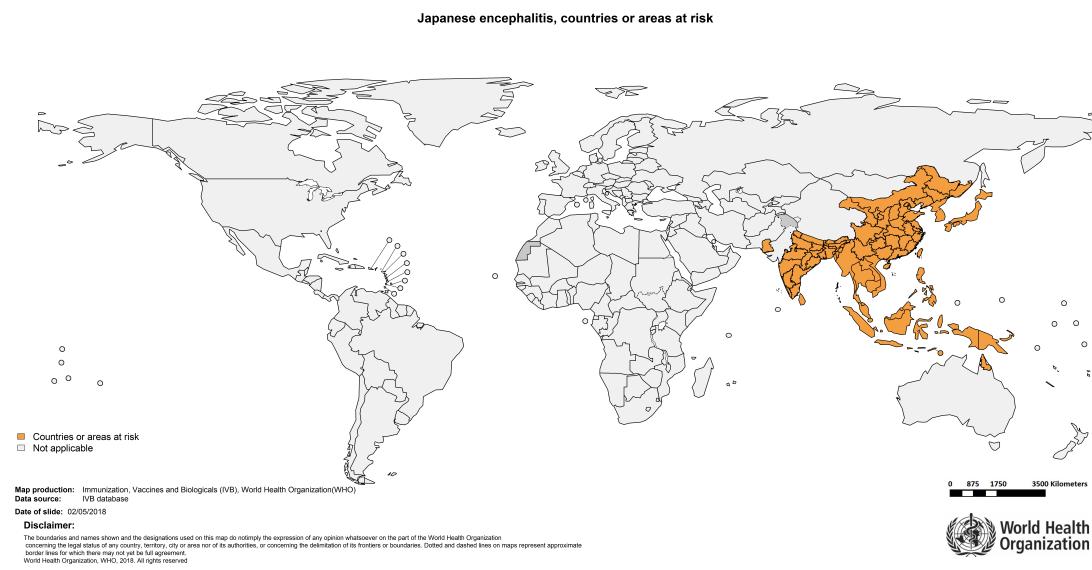


Fig1.1 Given below is the extraction from researches and thesis proposes by various analyst.

Signs and symptoms

Most JEV infections are mild (fever and headache) or without apparent symptoms, but approximately 1 in 250 infections results in severe clinical illness. The incubation period is between 4-14 days. In children, gastrointestinal pain and vomiting may be the dominant initial symptoms. Severe disease is characterised by rapid onset of high fever, headache, neck stiffness, disorientation, coma, seizures, spastic paralysis and ultimately death. The case-fatality rate can be as high as 30% among those with disease symptoms.

Of those who survive, 20%–30% suffer permanent intellectual, behavioural or neurological sequelae such as paralysis, recurrent seizures or the inability to speak.

Transmission

24 countries in the WHO South-East Asia and Western Pacific regions have JEV transmission risk, which includes more than 3 billion people.

JEV is transmitted to humans through bites from infected mosquitoes of the *Culex* species (mainly *Culex tritaeniorhynchus*). Humans, once infected, do not develop sufficient viraemia to infect feeding mosquitoes. The virus exists in a transmission cycle between mosquitoes, pigs and/or water birds (enzootic cycle). The disease is predominantly found in rural and periurban settings, where humans live in closer proximity to these vertebrate hosts.

In most temperate areas of Asia, JEV is transmitted mainly during the warm season, when large epidemics can occur. In the tropics and subtropics, transmission can occur year-round but often intensifies during the rainy season and pre-harvest period in rice-cultivating regions.

Diagnosis

Individuals who live in or have travelled to a JE-endemic area and experience encephalitis are considered a suspected JE case. A laboratory test is required in order to confirm JEV infection and to rule out other causes of encephalitis. WHO recommends testing for JEV-specific IgM antibody in a single sample of

cerebrospinal fluid (CSF) or serum, using an IgM-capture ELISA. Testing of CSF sample is preferred to reduce false-positivity rates from previous infection or vaccination

Surveillance of the disease is mostly syndromic for acute encephalitis syndrome. Confirmatory laboratory testing is often conducted in dedicated sentinel sites, and efforts are undertaken to expand laboratory-based surveillance. Case-based surveillance is established in countries that effectively control JE through vaccination.

Treatment

There is no antiviral treatment for patients with JE. Treatment is supportive to relieve symptoms and stabilise the patient.

Prevention and control

Safe and effective JE vaccines are available to prevent disease. WHO recommends having strong JE prevention and control activities, including JE immunisation in all regions where the disease is a recognised public health priority, along with strengthening surveillance and reporting mechanisms. Even if the number of JE-confirmed cases is low, vaccination should be considered where there is a suitable environment for JE virus transmission. There is little evidence to support a reduction in JE disease burden from interventions other than the vaccination of humans. Thus, vaccination of humans should be prioritised over vaccination of pigs and mosquito control measures.

There are 4 main types of JE vaccines currently in use: inactivated mouse brain-derived vaccines, inactivated Vero cell-derived vaccines, live attenuated vaccines, and live recombinant (chimeric) vaccines.

Over the past years, the live attenuated SA14-14-2 vaccine manufactured in China has become the most widely used vaccine in endemic countries, and it was prequalified by WHO in October 2013. Cell-culture based inactivated vaccines and the live recombinant vaccine based on the yellow fever vaccine strain have also been licensed and WHO-prequalified. In November 2013, Gavi opened a funding window to support JE vaccination campaigns in eligible countries.

To reduce the risk for JE, all travellers to Japanese encephalitis-endemic areas should take precautions to avoid mosquito bites. Personal preventive measures include the use of mosquito repellents, long-sleeved clothes, coils and vaporisers. Travellers spending extensive time in JE endemic areas are recommended to get vaccinated before travel.

Disease outbreaks

Major outbreaks of JE occur every 2-15 years. JE transmission intensifies during the rainy season, during which vector populations increase. However, there has not yet been evidence of increased JEV transmission following major floods or tsunamis. The spread of JEV in new areas has been correlated with agricultural development and intensive rice cultivation supported by irrigation programmes.

WHO responds to JE by:

- providing global recommendations for JE control, including the use of vaccines. WHO recommends JE immunisation in all regions where the disease is a recognised public health priority and supports implementation.
- providing technical support for JE surveillance, JE vaccine introduction and large-scale JE vaccination campaigns, and evaluation of JE vaccine effectiveness and programmatic impact.

Proposed System

The code containing symptoms attributes uses Genetic Algorithm and Attribute Selection or Feature Selection Algorithm. Genetic algorithms gives the higher quality for optimised problem and produce approximate result using Attribute Selection algorithm.

The discovery process also includes the infected brain cell images which are compare to the uninfected brain cell images using opencv algorithm which would determine whether infection is present or not by detecting changes in the cell image i.e change in colour due to infection.

Genetic Algorithm:

In computer science and operations research, a genetic algorithm (GA) is a metaheuristic inspired by the process of natural selection that belongs to the larger class of evolutionary algorithms (EA). Genetic algorithms are commonly used to generate high-quality solutions to optimization and search problems by relying on bio-inspired operators such as mutation, crossover and selection. John Holland introduced genetic algorithms in 1960 based on the concept of Darwin's theory of evolution; afterwards, his student David E. Goldberg extended GA in 1989.

Fig 1.2 Genetic Algorithm Concept

In a genetic algorithm, a population of candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem is evolved toward better solutions. Each candidate solution has a set of properties (its chromosomes or genotype) which can be mutated and altered; traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible.

The evolution usually starts from a population of randomly generated individuals, and is an iterative process, with the population in each iteration called a generation. In each generation, the fitness of every individual in the population is evaluated; the fitness is usually the value of the objective function in the optimization problem being solved. The more fit individuals are stochastically selected from the current population, and each individual's genome is modified (recombined and possibly randomly mutated) to form a new generation. The new generation of candidate solutions is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population.

A typical genetic algorithm requires:

- a genetic representation of the solution domain,
- a fitness function to evaluate the solution domain.

Feature Selection Algorithm:

Feature selection is also called variable selection or attribute selection.

It is the automatic selection of attributes in our data (such as columns in tabular data) that are most relevant to the predictive modeling problem we are working on.

Feature selection is the process of selecting a subset of relevant features for use in model construction.

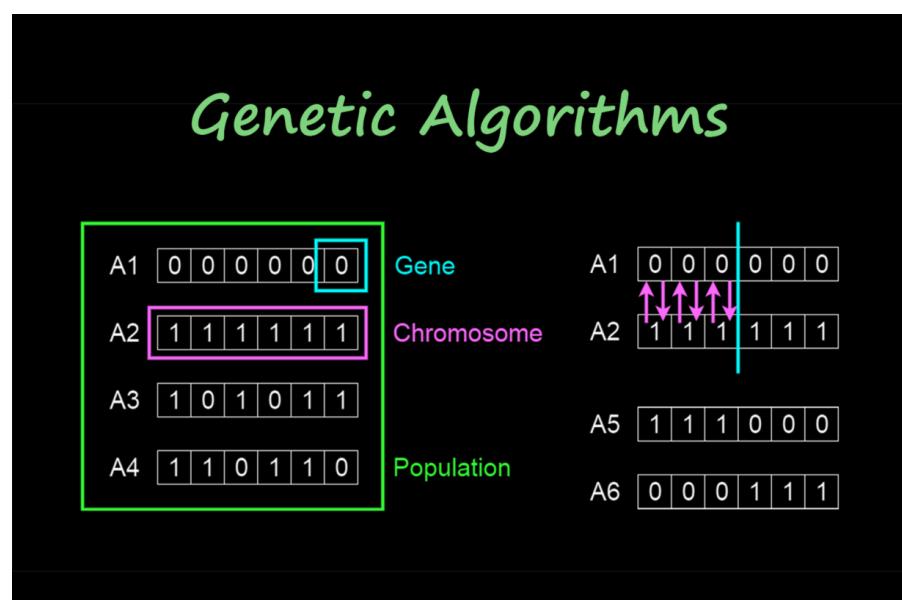
Feature selection is different from dimensionality reduction. Both methods seek to reduce the number of attributes in the dataset, but a dimensionality reduction method

do so by creating new combinations of attributes, whereas feature selection methods include and exclude attributes present in the data without changing them.

Feature selection is itself useful, but it mostly acts as a filter, muting out features that aren't useful in addition to our existing features.

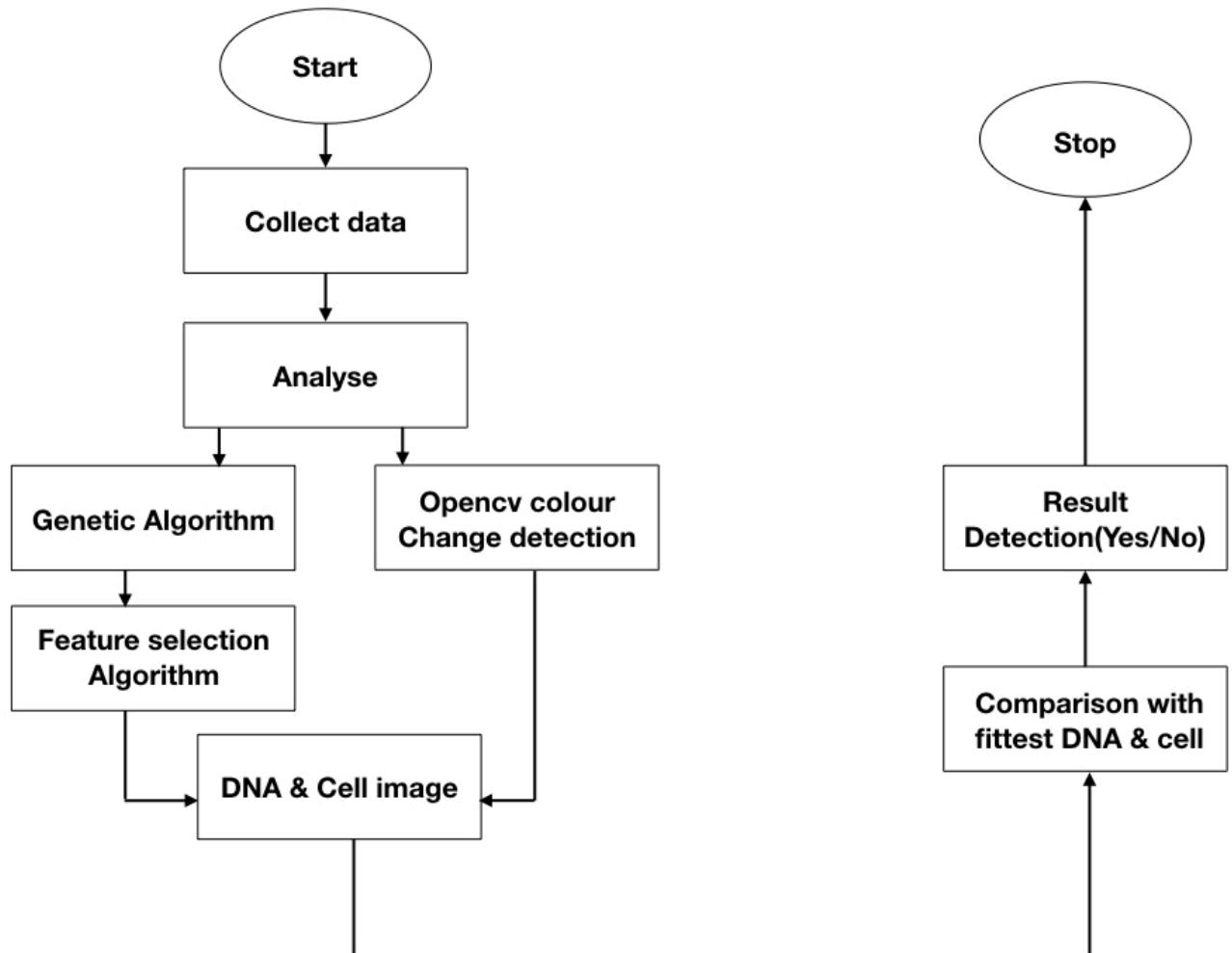
Opencv colour change detection in infected cell:

Opencv colour change detection is used to detect the change in the colour of the Brain cell. This concept detects whether the cell is parasitised or uninfected.



Architecture

Flow chart representation of Project



Algorithms

Genetic Algorithm:

The genetic algorithm is a method for solving both constrained and unconstrained optimisation problems that is based on natural selection, the process that drives biological evolution. The genetic algorithm repeatedly modifies a population of individual solutions. At each step, the genetic algorithm selects individuals at random from the current population to be parents and uses them to produce the children for the next generation. Over successive generations, the population "evolves" toward an optimal solution.

Genetic algorithm can be applied to solve a variety of optimisation problems that are not well suited for standard optimisation algorithms, including problems in which the objective function is discontinuous, non differentiable, stochastic, or highly nonlinear. The genetic algorithm can address problems of mixed integer programming, where some components are restricted to be integer-valued.

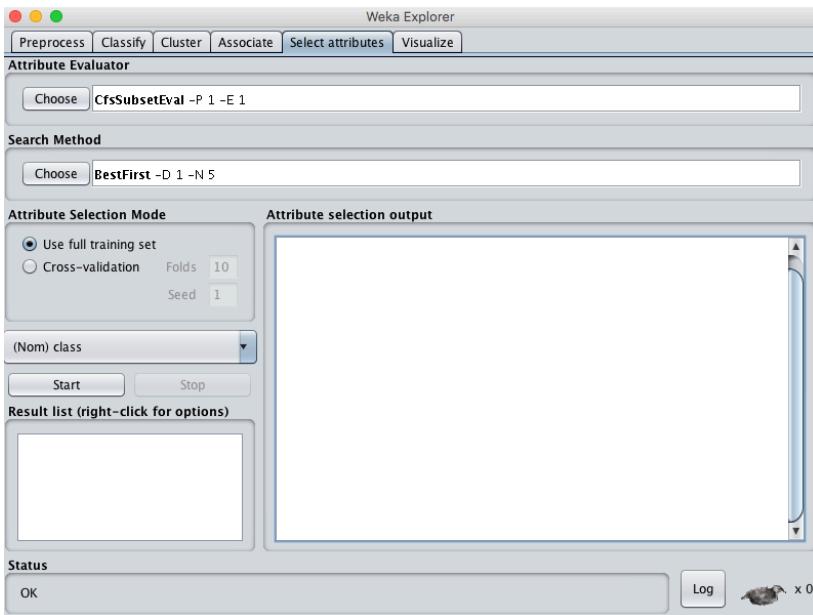
Pseudocode:

```
START
Generate the initial population
Compute fitness
REPEAT
    Selection
    Crossover
    Mutation
    Compute fitness
UNTIL population has converged
STOP
```

Feature Selection Algorithm (Using Weka tool):

A good place to get started exploring feature selection in Weka is in the Weka Explorer.

1. Open the Weka GUI Chooser.
2. Click the “Explorer” button to launch the Explorer.
3. Open the Pima Indians dataset.
4. Click the “Select attributes” tab to access the feature selection methods.



Weka Feature Selection

Feature selection is divided into two parts:

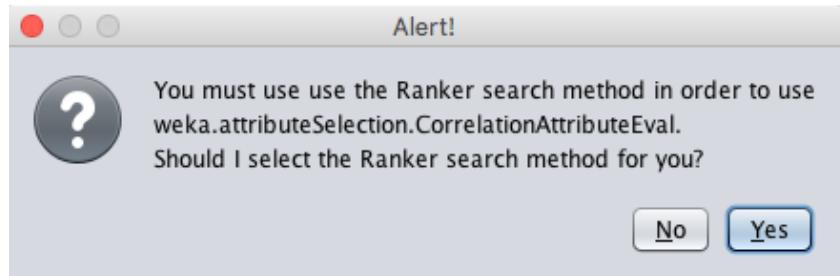
- Attribute Evaluator
- Search Method.

Each section has multiple techniques from which to choose.

The attribute evaluator is the technique by which each attribute in your dataset (also called a column or feature) is evaluated in the context of the output variable (e.g. the class). The search method is the technique by which to try or navigate different combinations of attributes in the dataset in order to arrive on a short list of chosen features.

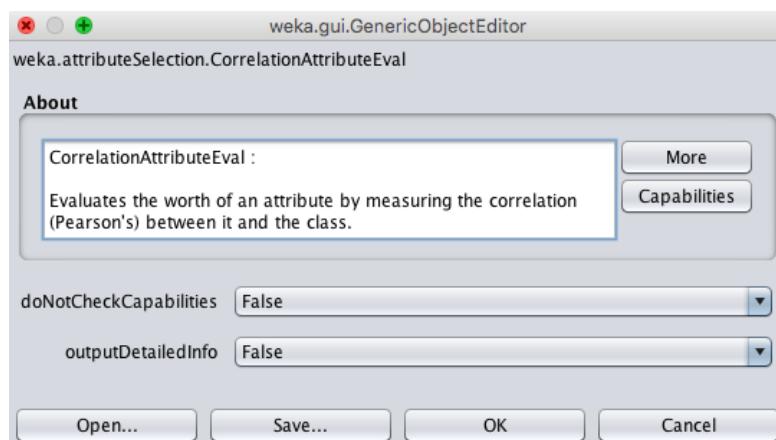
Some Attribute Evaluator techniques require the use of specific Search Methods. For example, the CorrelationAttributeEval technique used in the next section can only be used with a Ranker Search Method, that evaluates each attribute and lists the results in

a rank order. When selecting different Attribute Evaluators, the interface may ask you to change the Search Method to something compatible with the chosen technique.



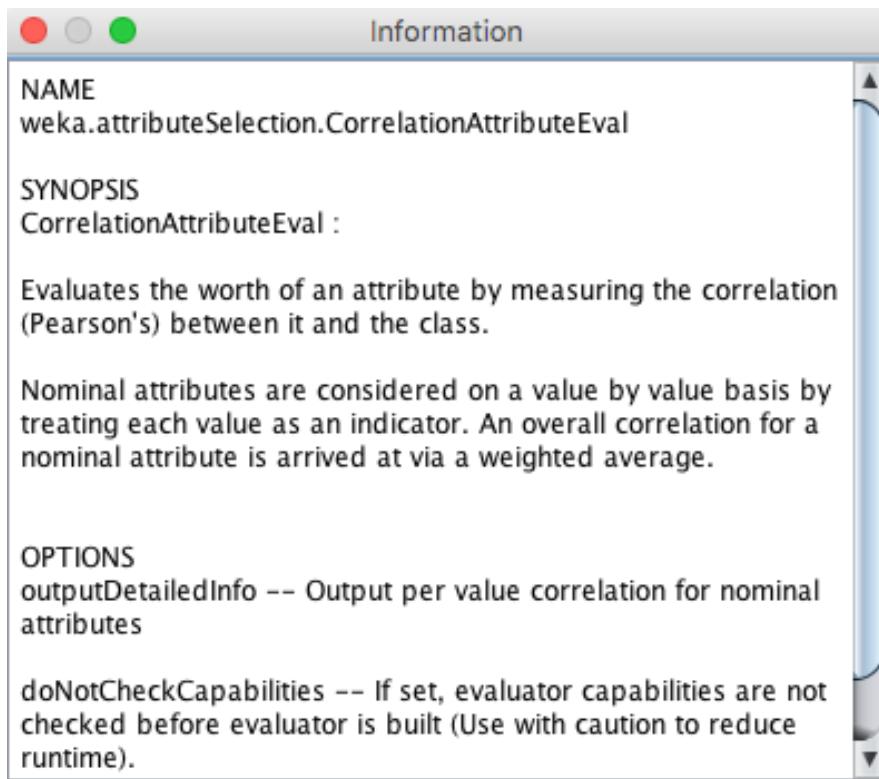
Weka Feature Selection Alert

Both the Attribute Evaluator and Search Method techniques can be configured. Once chosen, click on the name of the technique to get access to its configuration details.



Weka Feature Selection Configuration

Click the “More” button to get more documentation on the feature selection technique and configuration parameters. Hover your mouse cursor over a configuration parameter to get a tooltip containing more details.



Weka Feature Selection More Information

Now that we know how to access feature selection techniques in Weka, let's take a look at how to use some popular methods on our chosen standard dataset.

Results:

Genetic algorithm code:-

```
#include <iostream>
#include<cmath>
#include<stdio.h>
#include<string.h>
using namespace std;

int main(int argc, const char *argv[] )

{
    int dna[15];

    cout<<"\nPatient having chill? Yes or No\n";
    char ch[4]="YES";

    char chill[4];

    cin>>chill;

    if(strcmp(chill,ch)==0)

    {
        dna[0]=1;

    }
    else
    {
        dna[0]=0;
    }
    cout<<"\n\nPatient having Runny Nose? Yes or No\n";
    char rn[4]="YES";

    char runny[4];

    cin>>runny;

    if(strcmp(runny,rn)==0)

    {
```

```

    dna[1]=1;

}

else
{
    dna[1]=0;
}
cout<<"\n\nPatient having Headache? Yes or No\n";
char h[4]="YES";

char headache[4];

cin>>headache;

if(strcmp(headache,h)==0)

{

    dna[2]=1;

}

else
{
    dna[2]=0;
}
cout<<"\n\nPatient having Fever? Yes or No\n";
char f[4]="YES";

char fever[4];

cin>>fever;

if(strcmp(feaver,f)==0)

{

    dna[3]=1;

}

else
{
    dna[3]=0;
}

cout<<"\n\nEnter patient's AGE\n";

int age;

```

```
cin>>age;

if(age<15 && age>=0)

{

    dna[4]=0;
    dna[5]=0;

}

else if(age>=15 && age<22)

{

    dna[4]=0;
    dna[5]=1;

}

else if(age>=22 && age<=60)

{

    dna[4]=1;
    dna[5]=0;

}

else if(age>60)

{

    dna[4]=1;
    dna[5]=1;

}

cout<<"\n\nEnter patient's Blood Group\n";

char o[3]="O+";

char a[3]="A+";

char bg[3];

cin>>bg;

if(strcmp(bg,o)==0)
```

```

{

    dna[ 6]=0;

}

else if(strcmp(bg,a)==0)

{

    dna[ 6]=0;

}

else

{

    dna[ 6]=1;

}

cout<<"\n\nEnter RBC count range in million cells per
microliter (mCL)\n";
float rbc;
cin>>rbc;
if (rbc>=4.2&&rbc<=6.1)
{
    dna[ 7]=0;
}
else
{
    dna[ 7]=1;
}
cout<<"\n\nEnter WBC count range in cells per cubic
millimeter (cmm)\n";
float wbc;
cin>>wbc;
if (wbc>=4300&&wbc<=10800)
{
    dna[ 8]=0;
}
else
{
    dna[ 8]=1;
}
cout<<"\n\nEnter the number of occurrence of the
disease\n";

int occ;

```

```
cin>>occ;

if(occ==0)

{

    dna[9]=0;

}

else

{

    dna[9]=1;

}

cout<<"\n\nEnter the Time Period in days\n";

int days;

cin>>days;

if(days<=15)

{

    dna[10]=0;

}

else

{

    dna[10]=1;

}

cout<<"\n\nEnter Patient's status\n";

char pos[20]="Positive";

char neg[20]="Negative";

char status[20];

cin>>status;
```

```

if(strcmp(status,pos)==0)

{
    dna[11]=0;

}

else if(strcmp(status,neg)==0)

{
    dna[11]=1;

}
cout<<"\n\nEnter patient's food type\n";
char nv[20]="Non-Vegetarian";

char veg[20]="Vegetarian";

char type[20];

cin>>type;

if(strcmp(type,nv)==0)
{
    dna[12]=1;
}
else if(strcmp(type,veg)==0)
{
    dna[12]=0;
}
cout<<"\n\nEnter patient's region\n";
cout<<"\nEnter winter for states which experience
less summer\n";
cout<<"\nThese states include
JAMMU&KASHMIR, SIKKIM, HIMACHAL PRADESH, UTTARAKHAND\n";
cout<<"\nFor rest enter summer\n";
char w[20]="winter";
char s[20]="summer";
char region[20];
cin>>region;
if(strcmp(region,w)==0)
{
    dna[13]=0;
}
else if(strcmp(region,s)==0)
{
    dna[13]=1;
}

```

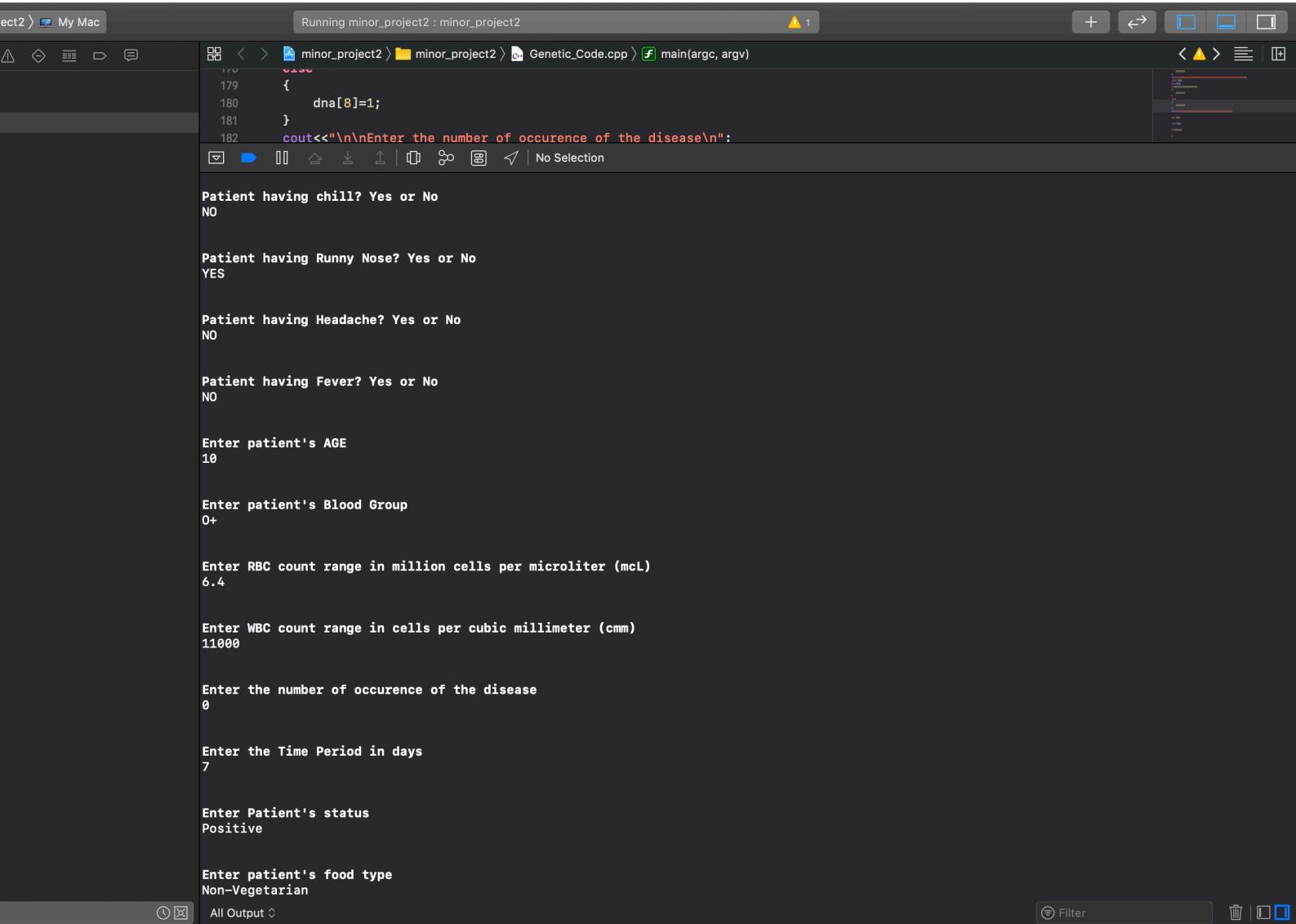
```

}

int odd=0,even=0;
for(int i=0;i<14;i++)
{
    if(dna[i]==1)
    {
        odd++;
    }
    else
    {
        even++;
    }
}
if(odd>=even)
{
    dna[14]=1;
}
for(int i=0;i<14;i++)
{
    cout<< dna[i];
}
cout<<"\n\nCheck with fittest string\n";
int JE=0,NJE=0;
for(int i=5;i<14;i++)
{
    if(dna[i]==1)
    {   JE++;      }
    else
    {   NJE++;      }
}
if(JE>=NJE)
{
    dna[14]=1;
}
cout<<"\n\nFittest string output\n";
for(int i=5;i<=14;i++)
{
    cout<< dna[i];
}
if(dna[14]==1)
{
    cout<<"\n\nJapanese Encephalitis
RESULT=Positive\n";
}
else
{
    cout<<"\n\nJapanese Encephalitis
RESULT=Negative\n";
}
```

```
        return 0;  
    }  
  
}
```

Code Output:



The screenshot shows a terminal window titled "Running minor_project2 : minor_project2" with a yellow warning icon. The window displays the output of a C++ program. The code in the editor pane includes a section where dna[8] is set to 1, followed by a cout statement asking for the number of occurrences of a disease. The terminal output shows the user's responses to various prompts about patient symptoms, age, blood group, and other medical details.

```
cout<<"\n\nEnter the number of occurrence of the disease\n":  
  
Patient having chill? Yes or No  
NO  
  
Patient having Runny Nose? Yes or No  
YES  
  
Patient having Headache? Yes or No  
NO  
  
Patient having Fever? Yes or No  
NO  
  
Enter patient's AGE  
10  
  
Enter patient's Blood Group  
O+  
  
Enter RBC count range in million cells per microliter (mcl)  
6.4  
  
Enter WBC count range in cells per cubic millimeter (cmm)  
11000  
  
Enter the number of occurrence of the disease  
0  
  
Enter the Time Period in days  
7  
  
Enter Patient's status  
Positive  
  
Enter patient's food type  
Non-Vegetarian
```

To get the fittest string feature selection algorithm is used. It include and exclude attributes present in the data without changing them.

==== Run information ====

Evaluator: weka.attributeSelection.ClassifierAttributeEval -execution-slots 1 -B
weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --
Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1
Relation: JE_Report
Instances: 8
Attributes: 14

x1
x2
x3
x4
x5
x6
x7
x8
x9
x10
x11
x12
x13
x14

Evaluation mode: evaluate on all training data

==== Attribute Selection on all input data ====

Search Method:
Attribute ranking.
Attribute Evaluator (supervised, Class (numeric): 14 x14):
Classifier feature evaluator
Using Wrapper Subset Evaluator
Learning scheme: weka.classifiers.rules.ZeroR
Scheme options:
Subset evaluation: RMSE
Number of folds for accuracy estimation: 5

Ranked attributes:

0 13 x13
0 6 x6
0 4 x4
0 3 x3
0 2 x2

```
0 5 x5
0 7 x7
0 12 x12
0 8 x8
0 11 x11
0 10 x10
0 9 x9
0 1 x1
```

Selected attributes: 13,6,4,3,2,5,7,12,8,11,10,9,1 : 13

```
325     }
326     cout<<"\n\nCheck with fittest string\n";
327     int JE=0,NJE=0;
328     for(int i=5;i<14;i++)
329     {
330         if(dna[i]==1)
331         { JE++; }
332         else
333         { NJE++; }
334     }
335     if(JE>=NJE)
336     {
337         dna[14]=1;
338     }
339     cout<<"\n\nFittest string output\n";
340     for(int i=5;i<=14;i++)
341     {
342         cout<< dna[i];
343     }
344     if(dna[14]==1)
```

No Debug Session

```
Enter Patient's status
Positive

Enter patient's food type
Non-Vegetarian

Enter patient's region

Enter winter for states which experience less summer
These states include JAMMU&KASHMIR,SIKKIM,HIMACHAL PRADESH,UTTARAKHAND

For rest enter summer
winter
01000001100010

Check with fittest string

Fittest string output
0011000100

Japanese Encephalitis RESULT=Negative
Program ended with exit code: 0
```

All Output ◆ Filter

Dataset:

The screenshot shows a Microsoft Excel spreadsheet titled "JE_Report.csv". The interface includes a ribbon bar with tabs like Home, Layout, Tables, Charts, SmartArt, Formulas, Data, and Review. Below the ribbon is a toolbar with various icons for file operations, font selection, and alignment. The main area displays a data table with 41 rows and 18 columns. The columns are labeled A through P, and the first few rows show data points such as "Chill", "RunnyNose", "Headache", "Fever", "Age", "BloodGroup", "RBCcount", "WBCcount", "Occurrence", "TimePeriod", "Patient's Stat", "FoodChoice", "Region", and "Japanese Encephalitis ?". The last column is a question mark, indicating the final output or result of the dataset.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Chill	RunnyNose	Headache	Fever	Age	BloodGroup	RBCcount	WBCcount	Occurrence	TimePeriod	Patient's Stat	FoodChoice	Region	Japanese Encephalitis ?		
2	1	0	10	1	0	0	0	0	0	0	0	0	0	0	0	
3	0	1	0	0	10	0	1	1	0	0	0	1	0	0	0	
4	1	0	11	1	11	1	1	1	1	1	1	1	1	1	1	1
5	0	1	10	1	11	1	0	0	0	1	1	0	0	0	0	0
6	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
7	0	1	11	1	11	1	1	1	1	1	1	1	1	1	1	1
8	0	1	11	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	1	10	1	11	1	1	1	1	1	1	1	1	1	1	1
10	1	0	10	1	0	0	0	0	0	0	0	0	0	0	0	0
11	0	1	0	0	10	0	1	1	0	0	0	1	0	0	0	0
12	1	0	11	1	11	1	1	1	1	1	1	1	1	1	1	1
13	0	1	10	1	11	1	0	0	0	1	1	0	0	0	0	0
14	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
15	0	1	11	1	11	1	1	1	1	1	1	1	1	1	1	1
16	0	1	11	0	0	0	0	0	0	0	0	0	0	0	0	0
17	1	1	10	1	11	1	1	1	1	1	1	1	1	1	1	1
18	1	0	10	1	0	0	0	0	0	0	0	0	0	0	0	0
19	0	1	0	0	10	0	1	1	0	0	0	0	1	0	0	0
20	1	0	11	1	11	1	1	1	1	1	1	1	1	1	1	1
21	0	1	10	1	11	1	0	0	1	1	0	0	0	0	0	0
22	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
23	0	1	11	1	11	1	1	1	1	1	1	1	1	1	1	1
24	0	1	11	0	0	0	0	0	0	0	0	0	0	0	0	0
25	1	1	10	1	11	1	1	1	1	1	1	1	1	1	1	1
26	1	0	10	1	0	0	0	0	0	0	0	0	0	0	0	0
27	0	1	0	0	10	0	1	1	0	0	0	0	1	0	0	0
28	1	0	11	1	11	1	1	1	1	1	1	1	1	1	1	1
29	0	1	10	1	11	1	0	0	1	1	0	0	0	0	0	0
30	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
31	0	1	11	1	11	1	1	1	1	1	1	1	1	1	1	1
32	0	1	11	0	0	0	0	0	0	0	0	0	0	0	0	0
33	1	1	10	1	11	1	1	1	1	1	1	1	1	1	1	1
34	1	0	10	1	0	0	0	0	0	0	0	0	0	0	0	0
35	0	1	0	0	10	0	1	1	0	0	0	0	1	0	0	0
36	1	0	11	1	11	1	1	1	1	1	1	1	1	1	1	1
37	0	1	10	1	11	1	0	0	1	1	0	0	0	0	0	0
38	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
39	0	1	11	1	11	1	1	1	1	1	1	1	1	1	1	1
40	0	1	11	0	0	0	0	0	0	0	0	0	0	0	0	0
41	1	1	10	1	11	1	1	1	1	1	1	1	1	1	1	1

This code output is generated using the second row of the dataset. The last attribute in the dataset is the result(presence of JE virus) whereas the previous attributes are used as inputs for the code.

Opencv python code:

```
#%import required libraries
import numpy as np # vector operation
#import cv2 # image processing
import os # to set and read from folder
import glob
from skimage import io
from skimage.transform import resize

#% Specify root folder location where images present
path = "/Users/apple/Desktop/cell_images/Uninfected"
names = os.listdir(path)
#% define two lists for images and labels
images = []
labels = []
#% fetch the root folder dataconda

for name in names:
    cur_path = path + "/" + name
    cur_label = name
    i = 1
    for file in glob.glob(cur_path + "/*.*"):
        image = io.imread(file) # read image using opencv

        image = resize(image, (32,32)) # resize the image
        images.append(image) # saving individual images in to a list
        labels.append(cur_label) # saving individual labels in to a list
        i += 1
#% feature extraction
images_array=np.array(images) # convert the list to array
labels_array=np.array(labels)
images_array = images_array.reshape((images_array.shape[0], 3072)) # reshape every image pixel into integer feature values
print("features matrix: {:.1f}MB".format( images_array .nbytes / (1024 * 1000.0))) #memory size of feature matrix

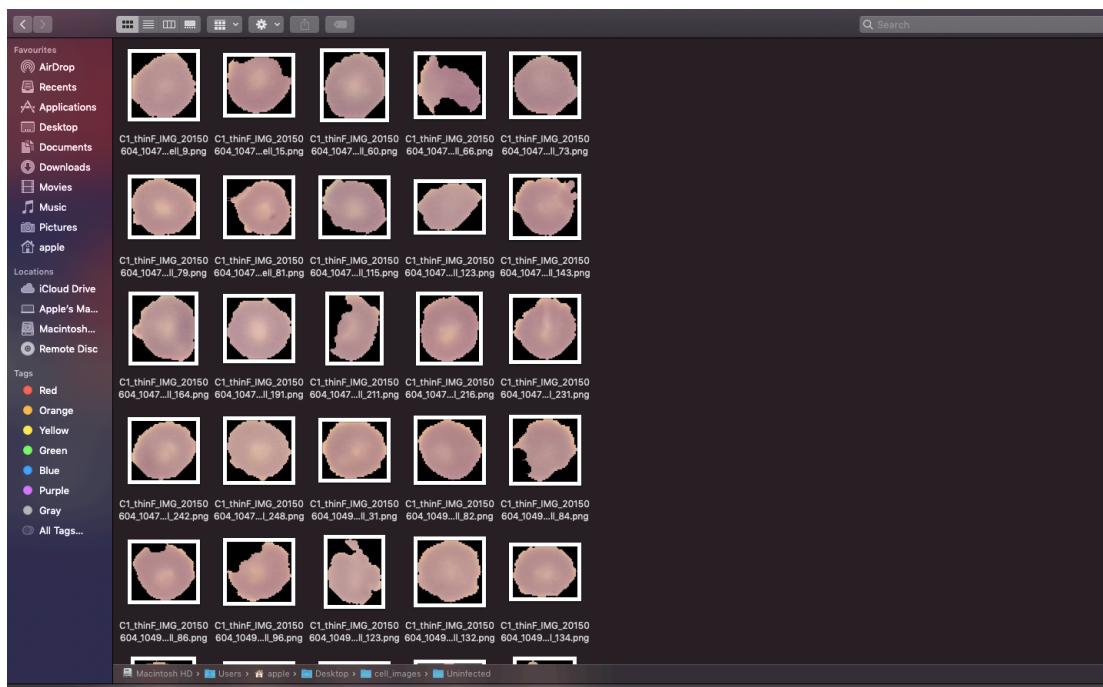
#% converting string labels to integer
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
labels = le.fit_transform(labels)
#% perform a training and testing split, using 70% of the data for
# training and 30% for evaluation

from sklearn.model_selection import train_test_split
(trainX, testX, trainY, testY) = train_test_split(images_array ,labels_array, test_size=0.25)
#% creating the simple KNN model
from sklearn.neighbors import KNeighborsClassifier
model = KNeighborsClassifier(n_neighbors = 5, metric = 'minkowski', p = 2)
model.fit(trainX, trainY)
#% evaluation parameter analysis
predictions = model.predict(testX)
from sklearn.metrics import confusion_matrix
print(confusion_matrix(testY,predictions))
from sklearn.metrics import accuracy_score
accuracy_score(testY, predictions)
from sklearn.metrics import classification_report
print(classification_report(testY, predictions,target_names=le.classes_)) # classification report
```

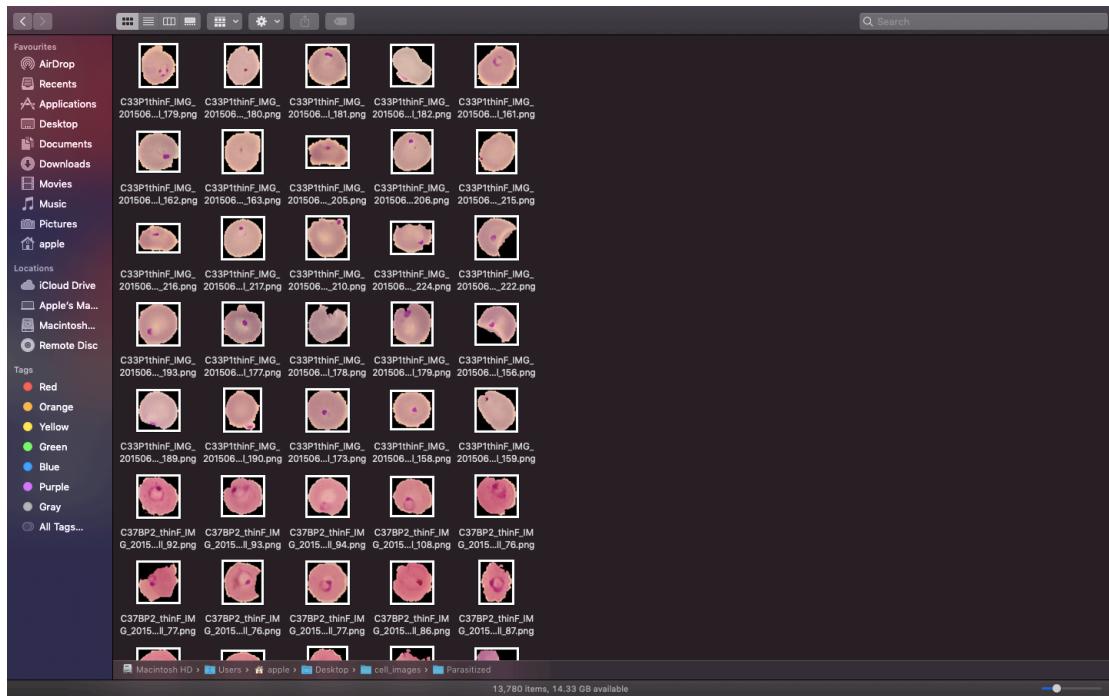
Ln: 39 Col: 4

Dataset Cell images):

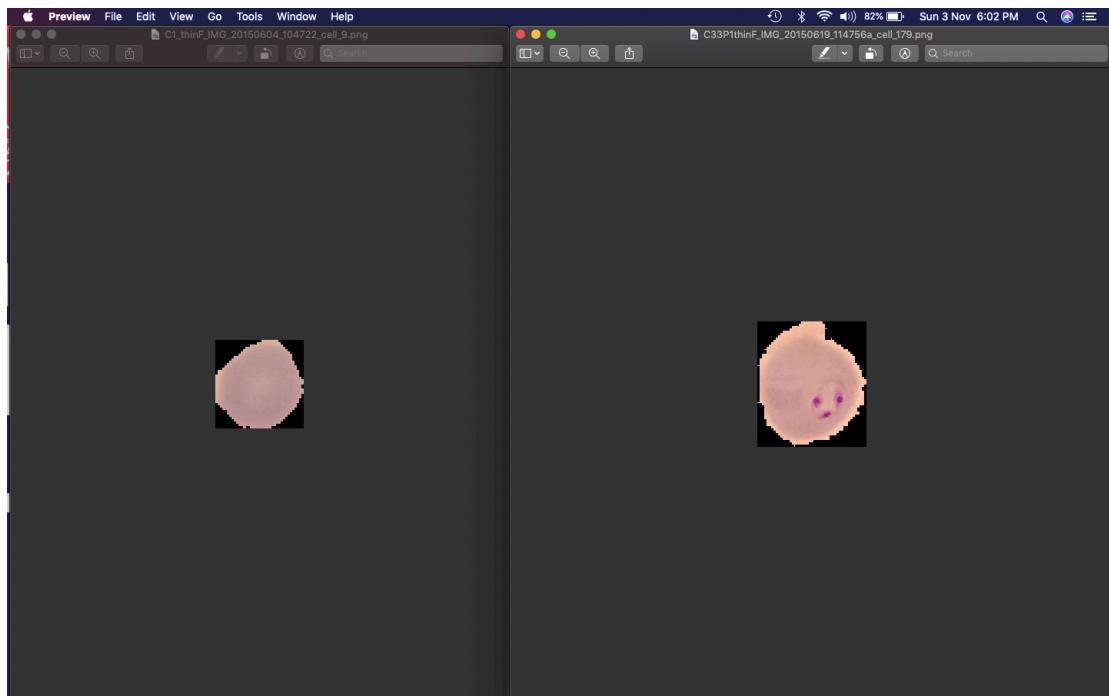
Uninfected:



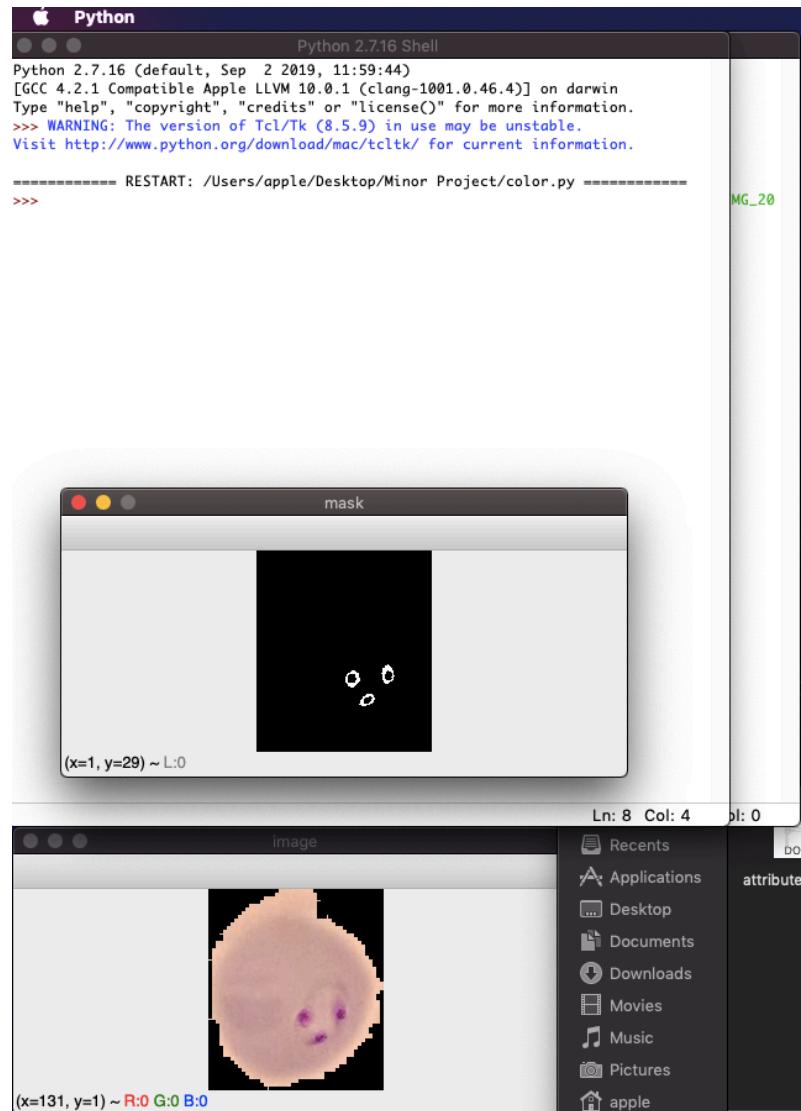
Infected:



Input:



Output:



COMPARISON

This project dealt with the affected patients data and analysed the result i.e whether the disease exists or not.

Since there are no existing methodologies in particular related to Japanese Encephalitis using the concept of Machine Learning, there is no possible comparison which could be done.

CONCLUSION

The outcome of the project is the prediction of whether a patient has Japanese encephalitis. This genetic code output is generated using the symptoms present in the dataset. The last attribute in the dataset is the result(presence of JE virus) whereas the previous attributes are used as inputs for the code. The result of Opencv colour change code gives us the idea of whether the cell is parasitised or uninfected. On the basis of above result we could predict whether a person is prone to disease or not.

This Project helped to add to our technical knowledge. We could learn about machine learning concepts and understand a lot about Bioinformatics. We would also like to thank our guide and faculties for their constant guidance and support. It was an enriching experience. We are really looking forward to such opportunities to widen our horizon.

REFERENCES

STUDY OF JAPANESE ENCEPHALITIS

- [1] Saluzzo JF, Dodet BTsai TF (1997) Factors in the changing epidemiology of Japanese encephalitis and West Nile fever. in *Factors in the emergence of arbovirus diseases*. eds Saluzzo JF, Dodet B (Elsevier, Paris), pp 179–189.
- [2] Solomon T (1997) Viral encephalitis in southeast Asia. *Neurological Infections and Epidemiology* **2**:191–199.
- [3] Miyake M (1964) The pathology of Japanese encephalitis. *Bull World Health Organ* **30**:153–160.
- [4]. Saluzzo JF, Dodet BGould EA, Zanotto PM, Holmes EC (1997) The genetic evolution of flaviviruses. in *Factors in the emergence of arbovirus diseases*. eds Saluzzo JF, Dodet B (Elsevier, Paris), pp 51–63.
- [5]. Tsai TF, Popovici F, Carnescu C, et al. (1998) West Nile encephalitis epidemic in southeastern Romania. *Lancet* **352**:767–771.
- [6]. Anonymous (1999) Outbreak of West Nile-like viral encephalitis: New York, 1999. *MMWR* **48**:845–849.
- [7]. Briese T, Jia X-Y, Huang C, Grady LJ, et al. (1999) Identification of a Kunjin/West Nile-like flavivirus in brains of patients with New York encephalitis. *Lancet* **354**:1261–1262.
- [8]. Kaisre R (1995) Tick-borne encephalitis in southern Germany. *Lancet* **345**:463.
- [9].Davidson MM, Williams H, Macleod J (1991) Louping ill virus in man: a forgotten disease. *J Infect* **23**:241–249.

[10]. "Japanese encephalitis". *World Health Organization*. December 2015. Archived from the original on 13 July 2017. Retrieved 29 October 2017.

[11]. https://en.wikipedia.org/wiki/Japanese_encephalitis

STUDY OF GENETIC ALGORITHM

[12] .<https://towardsdatascience.com/understanding-genetic-algorithms-cd556e9089cb>

[13].https://en.wikipedia.org/wiki/List_of_genetic_algorithm_applications

[14].https://en.wikipedia.org/wiki/Genetic_algorithm

[15].https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol1/hmw/article1.html

STUDY OF FEATURE SELECTION ALGORITHM

[16] .<https://machinelearningmastery.com/perform-feature-selection-machine-learning-data-weka/>

STUDY OF OPENCV PYTHON COLOUR CHANGE DETECTION

[17] .<https://www.pyimagesearch.com/2014/08/04/opencv-python-color-detection/>