SMART BRIDGE INTERNSHIP 2019

Project Name:

Malaria Prediction Recognisation

Team Members:

D.Manasa

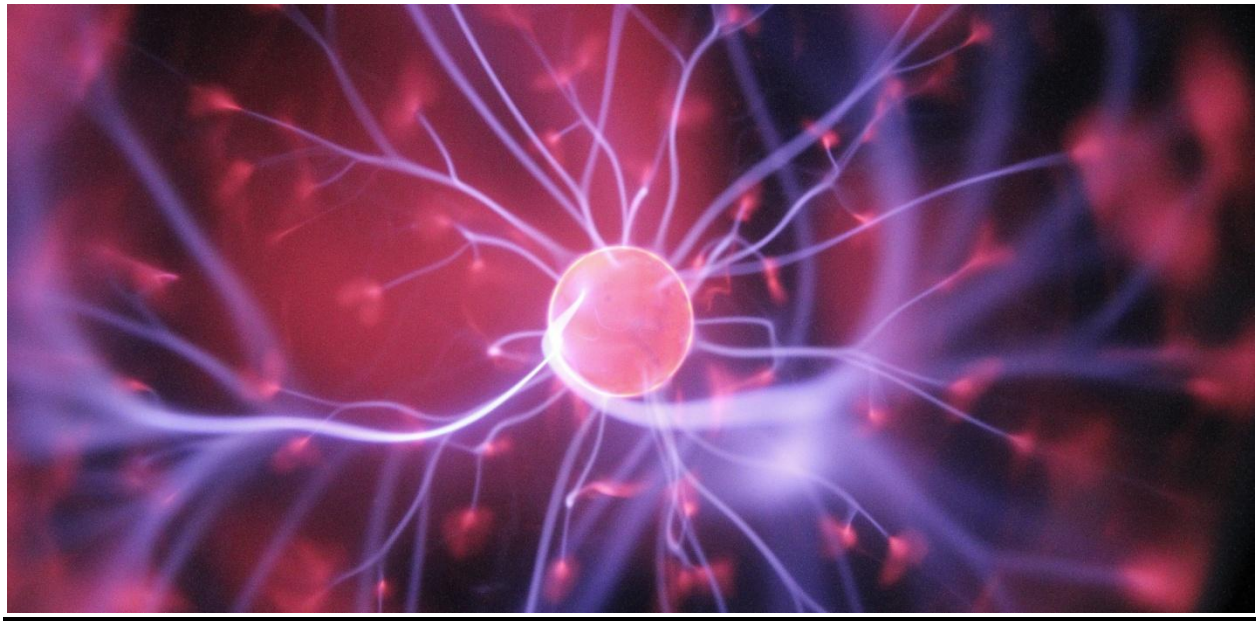
L.Srichandana

R.Akshaya

P.Snehitha

Malaria Prediction Recognisation

Introduction

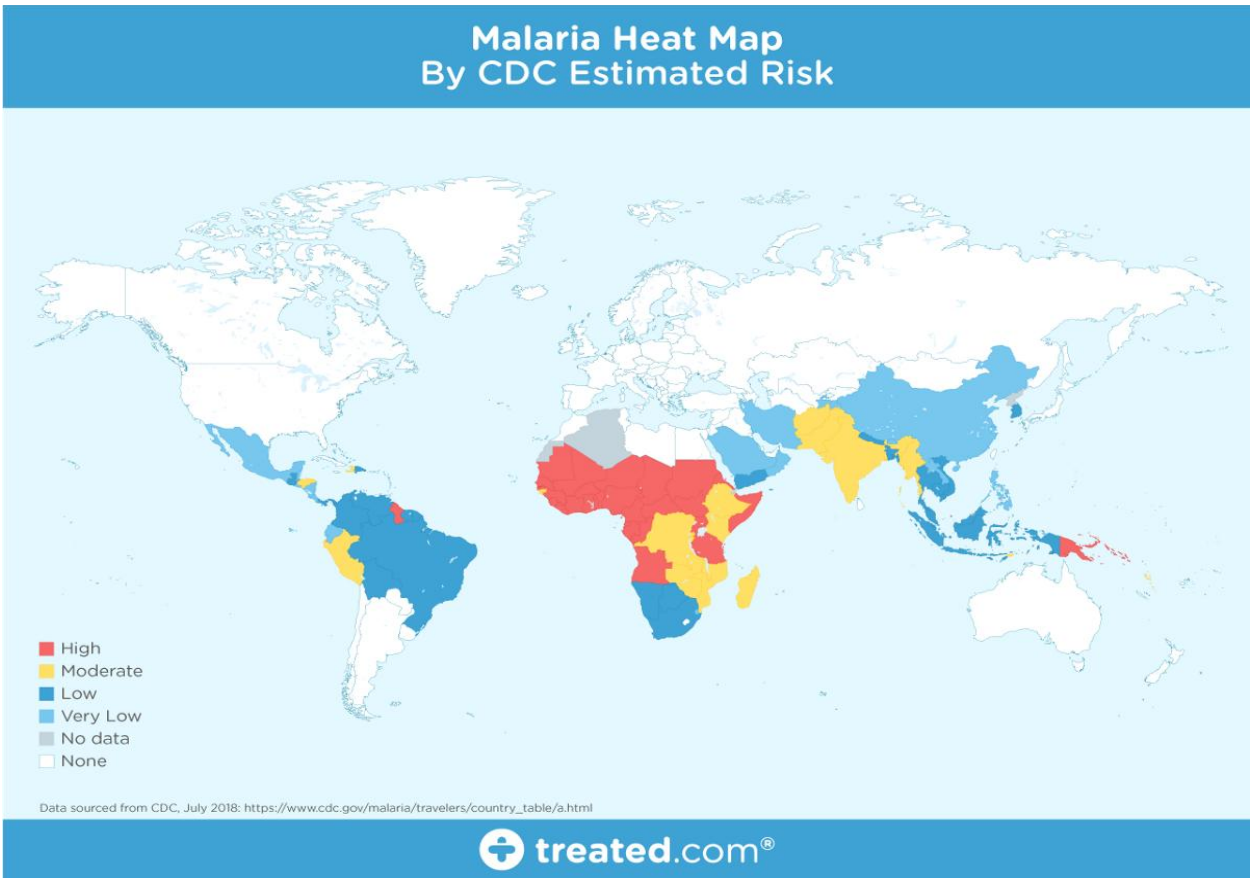


Welcome to the AI for Social Good Series, where we will be focusing on different aspects of how Artificial Intelligence (AI) coupled with popular open-source tools, technologies and frameworks are being used for development and betterment of our society. *"Health is Wealth"* is perhaps a quote yet very true! In this particular article, we will look at how AI can be leveraged for detecting malaria, a deadly disease and the promise of building a low-cost, yet effective and accurate open-source solution. The intent of the article is two-fold—understanding the motivation and importance of the deadly disease malaria and the effectiveness of deep learning in detecting malaria. We will be covering the following major topics in this article.

- **Motivation for this project**
- **Methods for Malaria Detection**
- **Deep Learning for Malaria Detection**
- **Convolutional Neural Networks (CNNs) trained from scratch**
- **Transfer Learning with Pre-trained Models**

Motivation

Malaria is a deadly, infectious mosquito-borne disease caused by Plasmodium parasites. These parasites are transmitted by the bites of infected female Anopheles mosquitoes. While we won't get into details about the disease, there are five main types of malaria. Let's now look at the significance of how deadly this disease can be in the following plot.



OBJECTIVES:

Objective 1: To prevent re-introduction and re-establishment of malaria.

- Maintain intensified case and entomological surveillance.
- Ensure universal access to quality assured malaria diagnostic and treatment services free of charge.
- Liaise with the private sector to provide regulated quality diagnostic and treatment services.
- Detect all infections early and treat all patients with quality assured anti-malarial medicines based on national treatment guidelines to ensure radical cure and to prevent secondary transmission.
- Ensure all suspected cases are tested for malaria (microscopy/RDT).

- Establish web based real time surveillance system.
- Provide regular information to health care providers on early detection of imported malaria cases.
- Immediate notification of all patients strongly suspected of having malaria.
- Commence investigation of all cases and foci (including reactive parasitological surveillance and entomological surveillance) within 48 hours of notification.
- Respond to all cases within 7 days of notification according to the approved scope of work.
- Conduct entomological surveillance in accordance with the new national guidelines.
- Implement vector control measures as required.
- Ensure quality assurance in malaria diagnostic services.

Objective 2: To maintain zero mortality due to malaria.

- Provide universal access to malaria diagnostic and treatment services free of charge.
- Detect all infections early and to treat all patients with quality assured anti-malarias based on national treatment guidelines to prevent complications in both public and private health sectors.
- Ensure all suspected cases are tested for malaria (microscopy or RDT).
- Ensure all malaria cases are admitted to hospitals and managed.
- Make available anti-malarial medicines (including second line and injectable medicines) to diagnostic and treatment facilities when needed in a timely manner.

- Inform clinicians on management of malaria (both uncomplicated and severe).

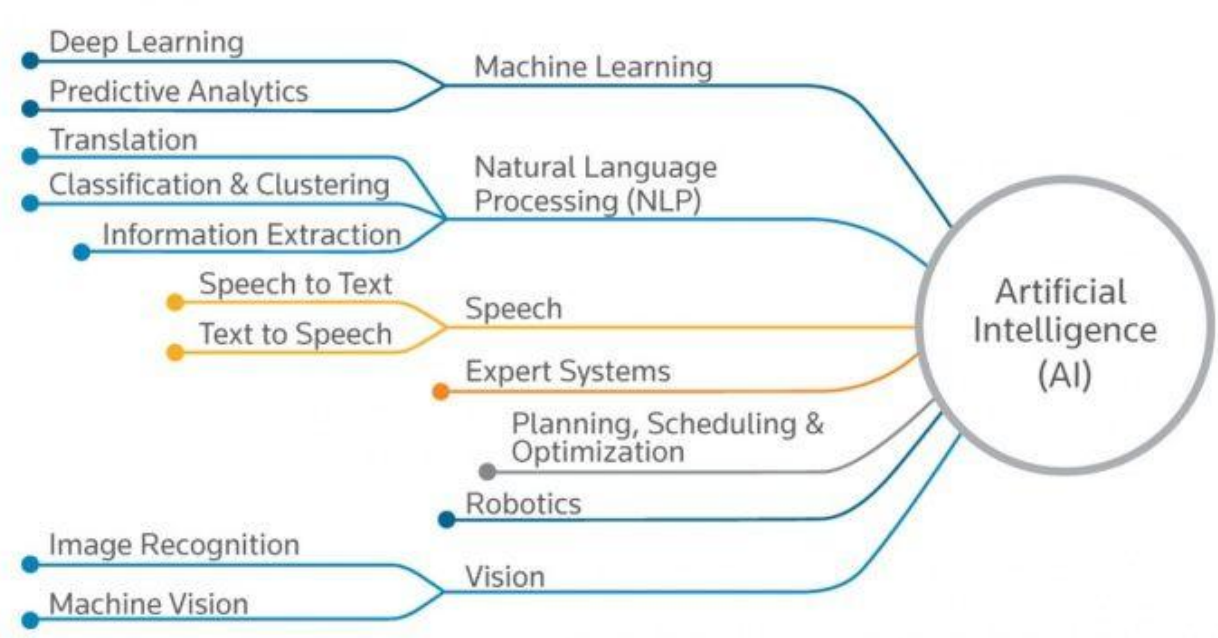
Carry out a medical audit of cases when recommended by the Case Review Committee

Problem Statement:

- Malaria is caused by *Plasmodium* parasites. The parasites are spread to people through the bites of infected female *Anopheles* mosquitoes, called "malaria vectors." There are 5 parasite species that cause malaria in humans, and 2 of these species

- Malaria is a life-threatening disease caused by parasites that are transmitted to people through the bites of infected female Anopheles mosquitoes. It is preventable and curable.

Review of Literature:



Artificial Intelligence:

Any technique which enables computers to mimic human behavior

- Machine Learning:

  Subset of AI technique which use statistical methods to enable machines to improve the experience

- Deep Learning:

  Subset of ML technique which make the computation of

Python:

Python is a high level programming language:

- Interpreted: python is processed at runtime by the interpreter
- Interactive: you can use a python prompt and interact with the interpreter directly to write your programs
- Object-oriented: python supports object-oriented technique of programming
- Beginner's Language: python is a great language for the beginner-level programmers and supports the development of a wide range of applications

Python basics:

- String: To identify the range of word using array
- List: A list in python is an ordered group of items or elements and these list elements don't have to be of the same type

  Python lists are "mutable" objects that can change their Values
- Tuples: python tuples are "Immutable" that cannot be changed
- Dictionaries: python dictionaries are kind of hash table type which consist of key-values pairs of unordered elements
- Sets: sets are similar to lists but duplicates are not allowed

Data Collection:

Let's talk about the dataset we would be using in our analysis. We are lucky to have researchers at the Lister Hill National Center for Biomedical Communications (LHNCBC), part of National Library of Medicine (NLM) who have carefully collected and annotated dataset

We can get further detail of the total number of images using the following code.

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

In [4]:
```
malaria_dataset = pd.read_csv('malaria_prediction.csv')
malaria_dataset
```

Out[4]:

	age	sex	fever	cold	rigor	fatigue	headace	bitter_tongue	vomitting	diarrhea	Convulsion	Anemia	jundice	cocacola_urine	hypoglycemia	prostr
0	0	Male	yes	no	no	no	no	yes	no	no	no	no	yes	no	yes	
1	1	Male	no	no	yes	yes	yes	no	no	no	no	no	yes	no	no	
2	0	Female	yes	no	no	yes	no	yes	no	yes	yes	no	no	yes	yes	
3	1	Female	yes	yes	yes	no	yes	no	no	no	no	yes	yes	yes	yes	
4	1	Female	yes	yes	yes	no	yes	no	no	yes	yes	yes	yes	yes	no	
5	0	Female	yes	yes	no	yes	yes	no	yes	yes	no	no	yes	yes	no	
6	1	Male	yes	yes	yes	yes	yes	no	no	no	no	no	yes	yes	yes	
7	1	Male	yes	no	no	no	yes	yes	no	no	no	no	yes	no	yes	
8	0	Female	yes	yes	no	yes	yes	yes	no	no	yes	no	no	yes	yes	
9	0	Male	yes	no	yes	no	yes	yes	no	no	yes	yes	no	no	yes	
10	0	Female	no	yes	no	no	no	yes	no	yes	no	no	no	no	yes	
11	1	Female	yes	no	no	no	yes	no	no	yes	yes	no	yes	no	yes	

Methodology:

There are several methods and tests which can be used for malariadetection and diagnosis. The original paper on which our data and analysis is based on, *' Pre-trained convolutional neural networks as feature extractors toward improved Malaria parasite detection in thin blood smear images'* .

These include but are not limited to, thick and thin blood smear examinations, polymerase chain reaction (PCR) and rapid diagnostic tests (RDT). While we won't cover all the methods here in detail, an important point to remember is that the latter two tests are alternative methods typically used an alternative particularly where good quality microscopy services cannot be readily provided.

We will discuss briefly about a standard malaria diagnosis, based on a typical blood-smear workflow .

Statistical techniques and data analysis:

Exploratory Data Analysis:

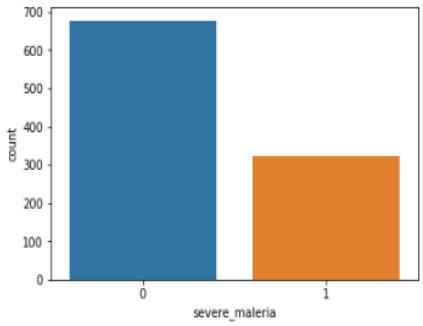
Matplotlib:

It is an excellent 2D and 3D Graphics library for generating scientific figures

Generally easy to get started for simple plots

```
In [19]: import seaborn as sns

         sns.countplot(malaria_dataset['severe_maleria'],label="Count")
```

```
Out[19]: <matplotlib.axes._subplots.AxesSubplot at 0x1e8e7ddda58>
```



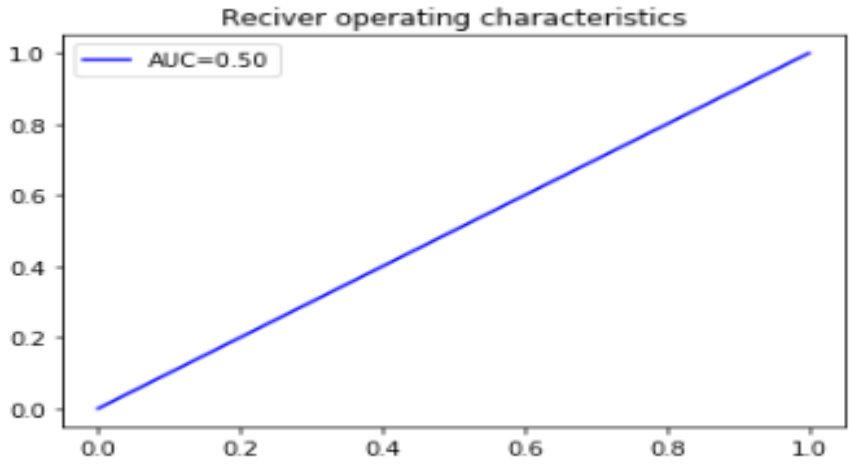

```
In [20]: from sklearn.model_selection import train_test_split

         X_train, X_test, y_train, y_test = train_test_split(new_malaria, y, test_size=0.3, random_state=0)
```

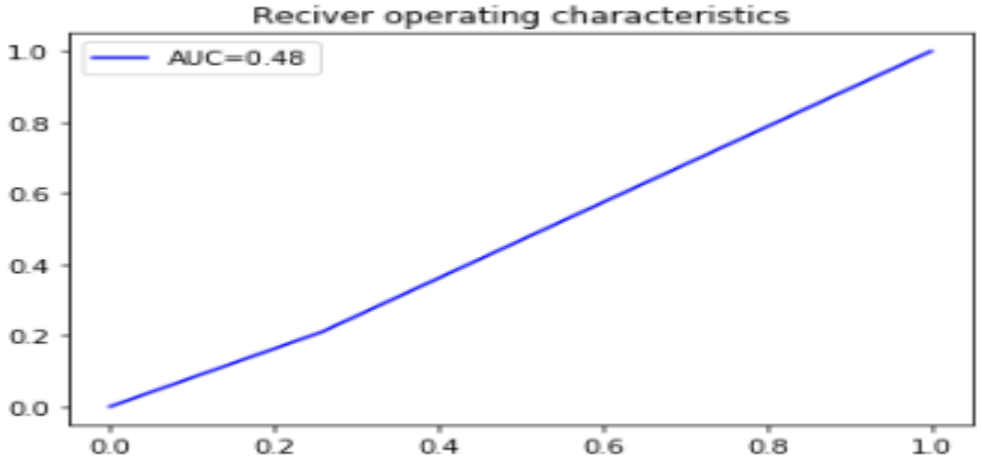
```
In [21]: X_train.values
```

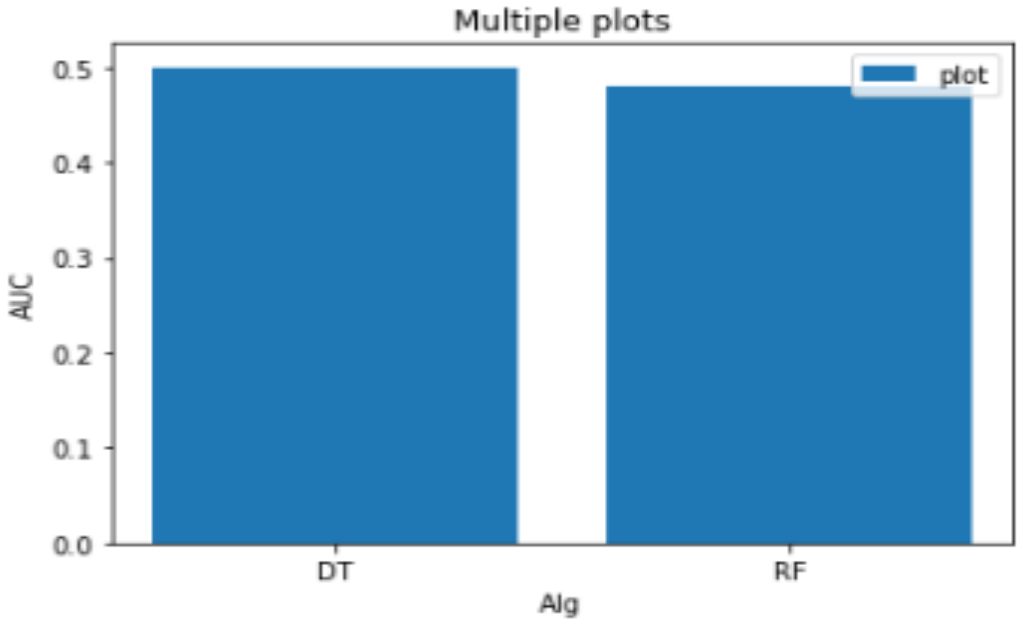
Figures And Tables:

Take ROC_AUC Values

- Decision Tree
- Random Forest


Reciver operating characteristics
AUC=0.50
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


Reciver operating characteristics
AUC=0.48
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


Multiple plots
plot
0.5
0.4
0.3
0.2
0.1
0.0
AUC
DT
RF
Alg

Conclusion

We looked at an interesting real-world medical imaging case study of malaria detection in this article. Malaria detection by itself is not an easy procedure and the availability of the right personnel across the globe is also a serious concern. We looked at easy to build open-source techniques leveraging AI which can give us state-of-the-art accuracy in detecting malaria thus enabling AI for social good. I encourage everyone to check out the articles and research papers mentioned in this article, without which it would have been impossible for me to conceptualize and write this article. Let's hope for more adoption of open-source AI capabilities across healthcare making it cheaper and accessible for everyone across the world!