**THEME: HEALTH CARE**

**Tech Stack**

1. Pytorch

2. Numpy

3. Pandas

4. FastAi

5. Google Colab

6. OpenCV

7. Matplotlib

8. Python

9.  Node.js

10. Express.js

11.  MongoDB

**Description of Proposed Idea**

**Tuberculosis** (**TB**), an infectious

disease usually caused by [Mycobacterium tuberculosis](https://en.wikipedia.org/wiki/Mycobacterium_tuberculosis)

(MTB) [bacteria](https://en.wikipedia.org/wiki/Bacteria), that affect many people in developing

countries. According to WHO, 1.7 billion people (about 23%) of the world

population are estimated to have latent TB infection, indicating a risk of

developing active TB during their lifetime. While treatment is possible, it

requires an accurate diagnosis first. Our idea aims at helping people detect

the onset early along with providing help to deal with it. We provide a deep

learning-based approach that can be used by people mainly in rural areas as

there are many cases where access to the X-Ray machine is possible (through

low-cost projects and donations), but often the radiological expertise is

missing for accurately asserting the images. The trained deep learning model(s)

will be hosted on a server and can be accessed by a front-end(web/mobile app).

An algorithm that could perform this task quickly and cheaply could drastically improve the ability to diagnose and ultimately treat the disease. In more developed countries, X-ray radiography is often used for screening new arrivals and determining eligibility for a work permit. The task of manually examining images is time-consuming and an algorithm could increase efficiency, improve performance, and ultimately reduce the cost of this screening. A well-trained model can perform like an expert practitioner, and our objective will be to reduce the trade-off between speed and accuracy.

**The Machine learning backend of our**

**Software**

We propose to implement a deep – learning framework that will take X-Ray scans as input and predict whether that particular X-Ray scan of lungs contains [Mycobacterium tuberculosis](https://en.wikipedia.org/wiki/Mycobacterium_tuberculosis) (MTB) [bacteria](https://en.wikipedia.org/wiki/Bacteria) hence predicting positive or negative case of tuberculosis. We will be using an EfficientNet/Resnet model, pre-trained on ImageNet, and further train this architecture.

This encoder-decoder model will extract features by downsampling the X-Ray image in the encoder block, while the decoder will use those features to generate a prediction for presence of MTB bacteria. Connecting links will be present between the encoder-decoder block to prevent the loss of information and complex features can be learned.

The loss function we will use is Binary Cross Entropy with logits Loss along with giving weights to corresponding classes such that Convolutional Neural Network not only focuses on easy examples and also carefully learn features from hard examples.

The output layer comprises of the score of the corresponding classes, and further applying the sigmoid function of the outputs we get probabilities of corresponding classes.

To evaluate the performance of our model, the metric we will use is ROC – AUC which is the area under the ROC curve of predicted probability and the observed target.

We will be training different pre-trained Efficientnet/ResNet model on Google Colab. The models will be trained on the metrics and loss functions discussed above.  The best performing model will be hosted on a server.

We will focus on a single-step approach, which is to get accurate predictions forTuberculosis disease.

The dataset used to train our model is created by the National Library of Medicine, Maryland, the USA in collaboration with Shenzhen No.3 People’s Hospital, Guangdong Medical College, Shenzhen, China. The Chest X-rays are from out-patient clinics and were captured as part of the daily routine using Philips DR Digital Diagnose systems.

**The Web App Description of our Software**

To deploy the ML model for production, there will be a web application. The frontend of the web application will be written in HTML, CSS, and javascript using bootstrap and jquery that would include two pages. The first page will be a landing page where users can sign in or register. The second page will have an uploader which will take X-Ray scans as input from the user. This input will be fed into the trained model which will be running on a virtual machine running on cloud.

The backend of the software will be implemented using node js and express. The templates will be rendered using Ejs. Python scripts will be run by using the child-process spawn method. The authentication would be done with passport js and user information would be added to MongoDB.

The output will be provided after passing through the trained machine learning model.

**Use Cases**

The solution we have provided may play an effective role in the early

diagnosis of TB disease that might be applied as a supportive tool. Modern

computer technologies should be trained in diagnostics for rapid disease

management. It should be noted that the only step needed in using our

approach in the real world is a medium to send requests to the server. Thus

once this step is done, there'll be little to no help needed in maintaining the

model. The X-ray scans that are sent to the server can be saved there for

increasing the size of the dataset which helps in updating the model. This can

be done by the admins of the server(us) without any trouble for the

organization using our solution. With help from expert practitioners, this

solution can be used even in places where specialized healthcare services are

not possible. The solution is aimed at areas with fewer healthcare facilities.

Depending on how high the organization is going to keep the price for their

services, the cost might increase or reduce for the patients, while it's definitely

a profit for the healthcare industry as a whole. We believe the government will

help in regulating the prices of such AI-based services keeping the healthcare

industry from exploiting the technology.

**Scope Of improvement**Currently, our aim is to make accurate predictions on X-Ray scans but if we get access to more data then our model can be improved quite easily and will also be able to segment the affected part of the lungs scan. This proposed idea will provide a huge help to practitioners, where theirs task would be analysing the segmented scans giving instructions to other helping doctors. Getting access to more amount of data will lead to further improvement in performance of our model.