DEEP LEARNING TECHNIQUES FOR BREAST CANCER RISK PREDICTION USING IBM CLOUD

MINI PROJECT

Submitted to

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY,

HYDERABAD

In partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

REPALA PRAVALIKA	**19UK1A0531**
MAMIDALA KRISHNA SRI	**19UK1A0529**
SYED ABDUL MUJEEB	**19UK1A0569**
CHIDIRALA SAI KIRAN	**19UK1A0566**

Under the esteemed guidance of

Ms. A.Swathi

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
VAAGDEVI ENGINEERING COLLEGE
(Affiliated to JNTUH, Hyderabad)
Bollikunta, Warangal - 506005

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
VAAGDEVI ENGINEERING COLLEGE
BOLLIKUNTA, WARANGAL – 506005

2019 – 2023

CERTIFICATE OF COMPLETION

UG PROJECT PHASE-1

This is to certify that the UG Project Phase-1 entitled **"DEEP LEARNING TECHNIQUES FOR BREAST CANCER RISK PREDICTION USING IBM CLOUD"** is being submitted by***R.PRAVALIKA(*H.NO:19UK1A0531),*M.KRISHNA,SRI*(H.NO:19UK1A0529),*S.A.MUJEEB*(H.NO:19UK1A0569),*CH.SAI KIRAN*(H.NO:19UK1A0566)** inpartialfulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering** to **Jawaharlal Nehru Technological University Hyderabad** during the academic year **2022-23**, is a record of work carried out by them under the guidance and supervision.

ProjectGuide
Ms. A. Swathi.
(AssistantProfessor)

Head of theDepartment
Dr. R. NaveenKumar
(Professor)

External

ACKNOWLEDGEMENT

We wish to take this opportunity to express our sincere gratitude and deep sense of respect to our beloved **Dr.P.PRASAD RAO,** Principal, Vaagdevi Engineering College for making us available all the required assistance and for his support and inspiration to carry out this UG Project Phase-1 in the institute.

We extend our heartfelt thanks to **Dr.R.NAVEEN KUMAR**, Head of the Department of CSE, Vaagdevi Engineering College for providing us necessary infrastructure and thereby giving us freedom to carry out the UG Project Phase-1.

We express heartfelt thanks to Smart Bridge Educational Services Private Limited, for their constant supervision as well as for providing necessary information regarding the UG Project Phase-1 and for their support in completing the UG Project Phase-1.

We express heartfelt thanks to the guide, **Ms. A. Swathi** Assistant professor, Department of CSE for her constant support and giving necessary guidance for completion of this UG Project Phase-1.

Finally, we express our sincere thanks and gratitude to my family members, friends for their encouragement and outpouring their knowledge and experience throughout the thesis.

PRAVALIKA (19UK1A0531)
KRISHNASRI (19UK1A0529)
ABDUL MUJEEB (19UK1A0569)
SAI KIRAN (19UK1A0566)

ABSTRACT

Breast Cancer is mostly identified among women and is a major reason for increasing the rate of mortality among women. Diagnosis of breast cancer is time consuming and due to the lesser availability of systems it is necessary to develop a system that can automatically diagnose breast cancer in its early stages. Various Machine Learning and Deep Learning Algorithms have been used for the classification of benign and malignant tumours. The Wisconsin Breast Cancer Dataset has been used which contains 569 samples and 30 features. The paper emphasises on various models that is implemented such as Logistic Regression, Support Vector Machine (SVM) and K Nearest Neighbour (KNN), Multi-Layer perceptron classifier, Artificial Neural Network(ANN)) etc. on the dataset taken from the repository of Kaggle. Each of these algorithms has been measured and compared with respect to accuracy and precision obtained. All the techniques are coded in python and executed in Google Colab, which is a Scientific Python Development Environment. The experiments have shown that SVM and Random Forest Classifier are the best for predictive analysis with an accuracy of 96.5%. To increase the accuracy of prediction, deep learning algorithms such as CNN and ANN have been implemented. The maximum accuracy obtained in the case of ANN and CNN are 99.3% and 97.3% respectively. Activation functions such as Relu and sigmoid have been used to predict the outcomes in terms of probabilities.

TABLE OF CONTENTS:-

1.INTRODUCTION

1.1 OVERVIEW

Breast cancer is one of the main causes of cancer death worldwide.Computer-aided diagnosis Systemsshowed the potential for improving diagnostic accuracy. But early detection and prevention cansignificantly reduce the chances of death.It is important to detect breast cancer as early as possible. Breast cancer is the global leading cause of cancer-related deaths in women, and the most commonly diagnosed cancer among women across the world (1). From our perspective, improved treatment options and earlier detection could have a positive impact on decreasing mortality, as this could offer more options for successful intervention and therapies when the disease is still in its early stages.Our team of IBM researchers published research in *Radiology* around a new AI model that can predict the development of malignant breast cancer in patients within the year, at rates comparable to human radiologists. As the first algorithm of its kind to learn and make decisions from both imaging data and a comprehensive patient's health history, our model was able to correctly predict the development of breast cancer in 87 percent of the cases it analyzed, and was also able to correctly interpret 77 percent of non-cancerous cases.When put to the test against 71 different cases that radiologists had originally determined as "non-malignant," but who ultimately ended up being diagnosed with breast cancer within the year, our AI system was able to correctly identify breast cancer in 48 percent of individuals (48 percent of the 71 cases) – which otherwise would not have been flagged.

1.2 PURPOSE:

The goal is to classify images into two classifications of malignant and benign. As early diagnostics significantly increases the chances of correct treatment and survival. In this application, we are helping the doctors and patients to classify the Type of Tumour for the specific image given with the help of Neural Networks.

One of the best ways to fight cancer is early detection, when it is still confined and can be fully excised surgically or treated pharmacologically. Cancer screening programs, that is, the practice of testing for the presence of cancer in people who have no symptoms, has been medicine's tool of choice for the earliest detection.

2. LITERATURE SURVEY

2.1 Existing problem

One of the main causes of cancer death worldwide is Breast Cancer. Computer-aided diagnosis systems showed the potential for improving diagnostic accuracy. But early detection and prevention can significantly reduce the chances of death. It is important to detect breast cancer as early as possible.

One of the best ways to fight cancer is early detection, when it is still confined and can be fully excised surgically or treated pharmacologically. Cancer screening programs, that is, the practice of testing for the presence of cancer in people who have no symptoms, has been medicine's tool of choice for the earliest detection.

The first cancer screening test to be widely used for cancer was the Pap test for finding cervical cancer. Since its introduction as a widely used test in the 1960's, cervical cancer death rate in the United States has declined by about 70%[1]. Similarly, breast cancer screening started to be widely used in the 1970's and has been shown to decrease mortality in multiple randomized controlled trials[1]. Screening for breast cancer is done using mammography exams in which radiologists scrutinize x-ray pictures of the breast for the possible presence of cancer. Mammography screenings, on average only find 7 out of 8 asymptomatic breast cancers [2], and this sensitivity has been increasing over the past years. On the other end of the spectrum there are the false positives. Out of 1,000 women, about 100 are recalled for additional diagnostic imaging, and of these 100 women, 4 or 5 are ultimately diagnosed with breast cancer[2] . These false positive exams lead to preventable harms, including patient anxiety, benign biopsies, and unnecessary intervention or treatment. Furthermore, high false-positive rates significantly contribute to the annual $7.8 billion mammography screening costs in the U.S[3].

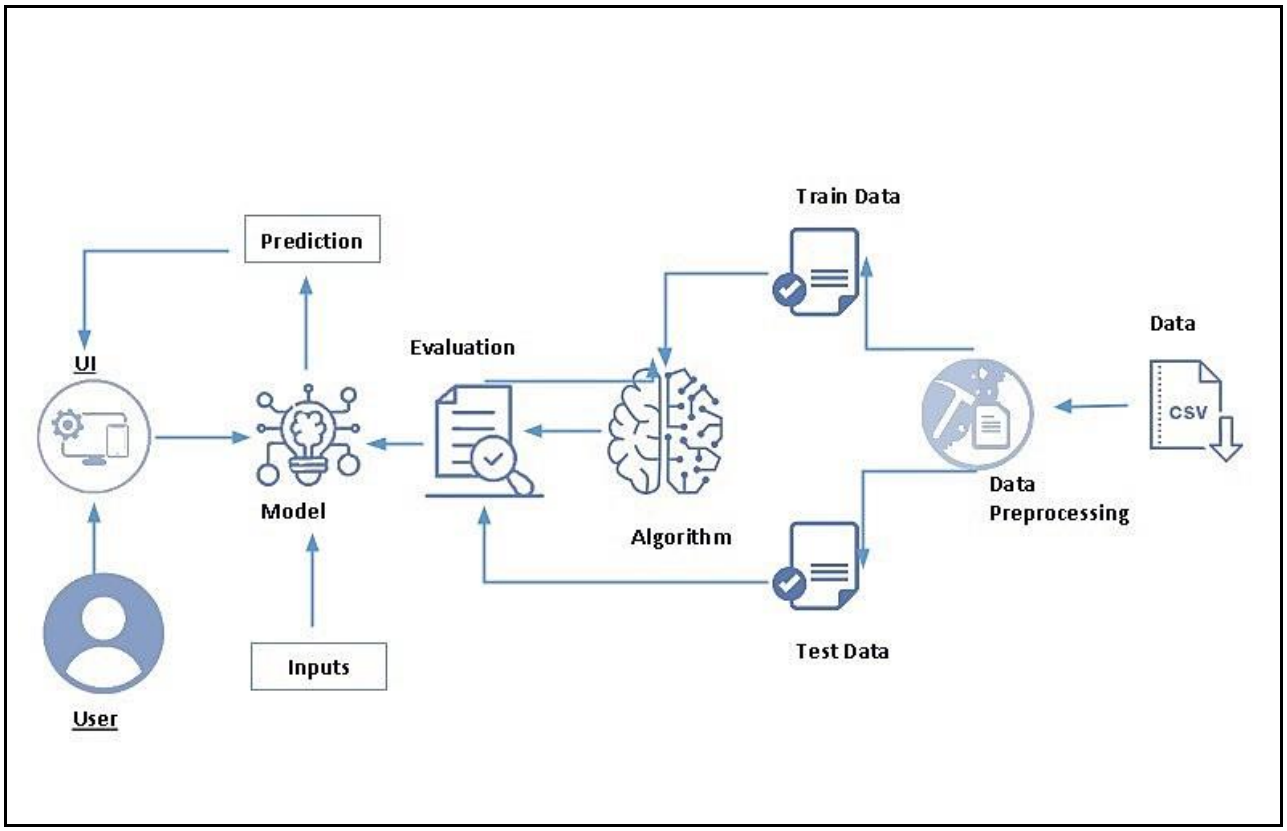
2.2 Proposed Solution

This project goal is to classify images into two classifications of malignant and benign. As early diagnostics significantly increases the chances of correct treatment and survival. In this application, we are helping the doctors and patients to classify the Type of Tumour for the specific image given with the help of Neural Networks.

Breast cancer is the global leading cause of cancer-related deaths in women, and the most commonly diagnosed cancer among women across the world (1). From our perspective, improved treatment options and earlier detection could have a positive impact on decreasing mortality, as this could offer more options for successful intervention and therapies when the disease is still in its early stages.

3. THEORITICAL ANALYSIS

3.1BlockDiagram



3.2 Hardware / Software designing

Software Requirements:

- Ananconda Navigator
- Tensor flow
- Keras
- Flask

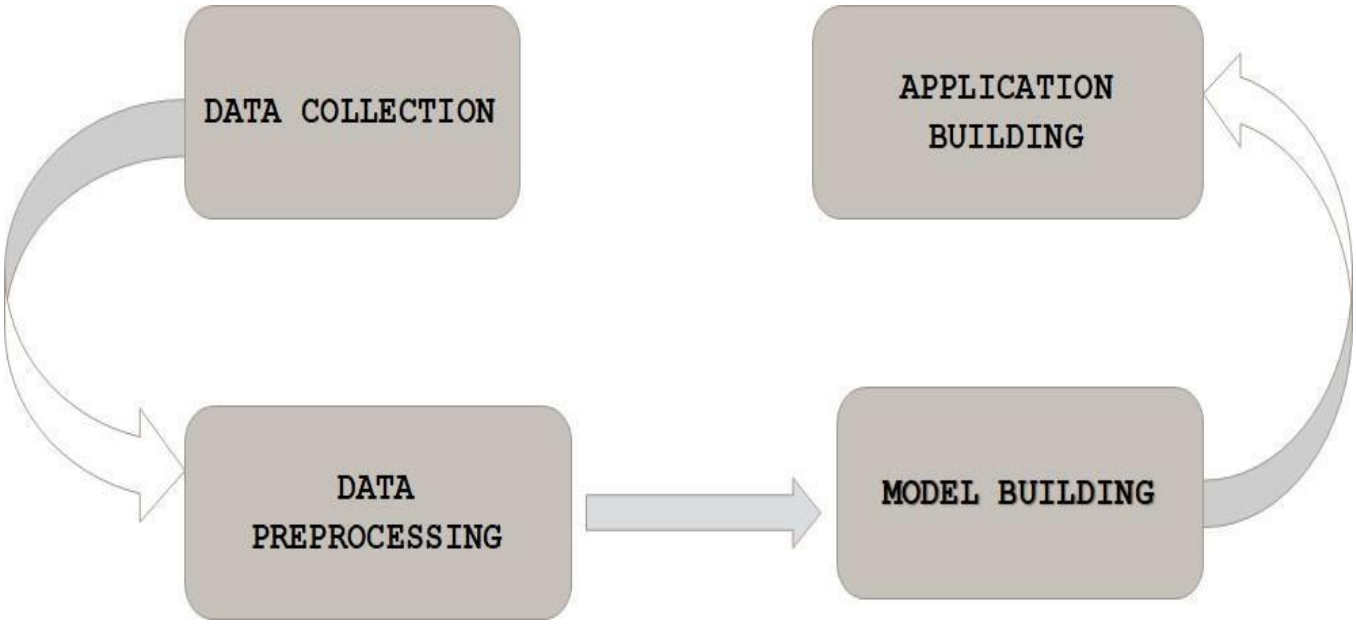
Hardware Requirements:

- Processor : Intel Core i3
- Hard Disk Space : Min 100GB
- Ram : 4GB
- Display : 14.1 “Color Monitor(LCD, CRT or LED)
- Clock Speed : 1.67 GHz

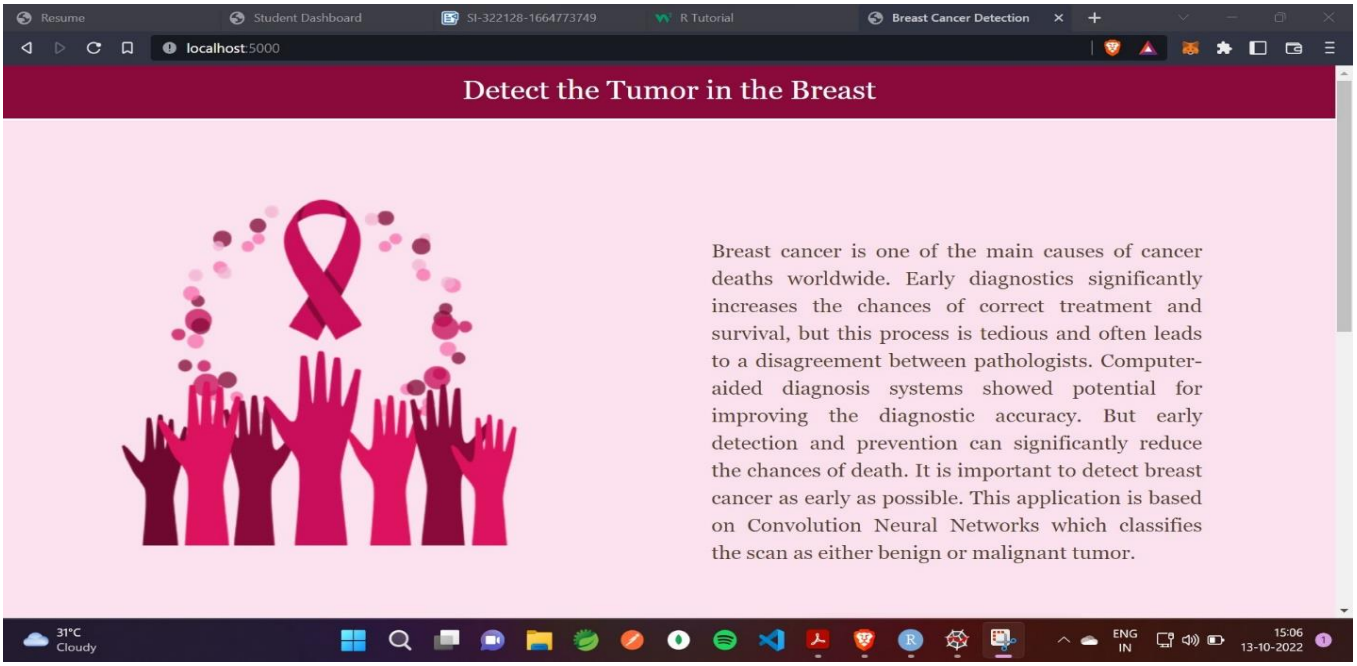
4. EXPERIMENTAL INVESTIGATIONS

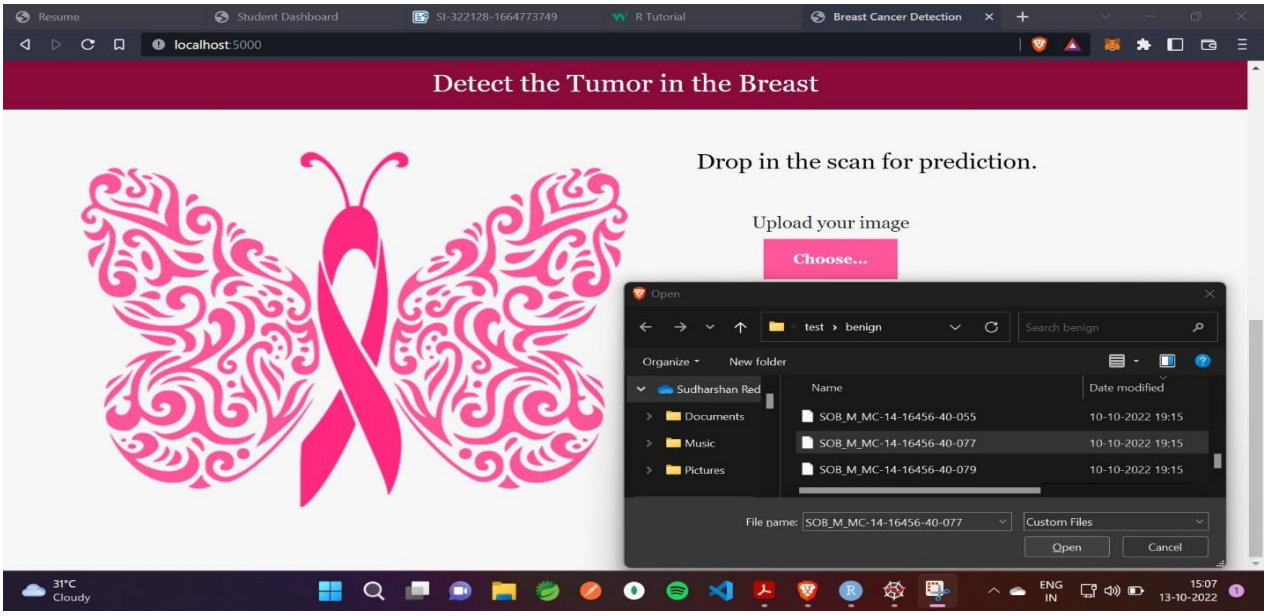
It shows that a model provided with different breast cancer scanning report images will predict the tumor and dispalys the output. Choose the image and click on upload, then it will predict the output.

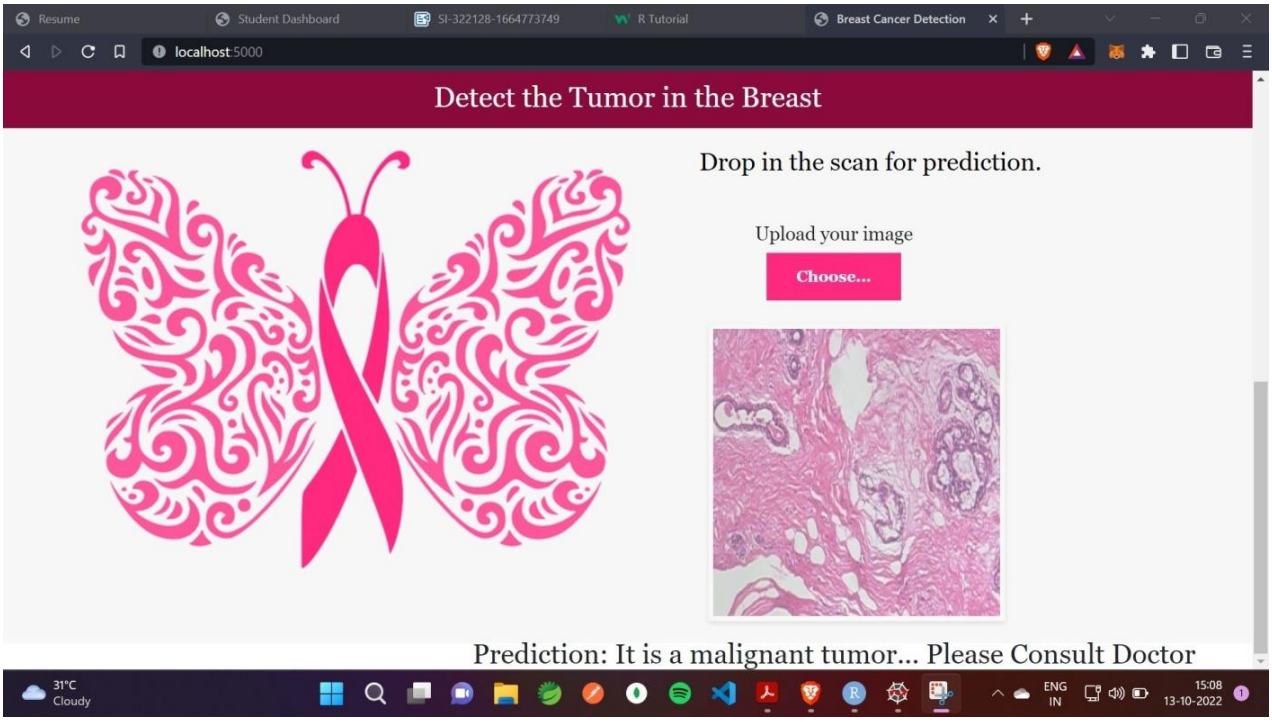
5. FLOWCHART



6. RESULT


Resume
Student Dashboard
SI-322128-1664773749
R Tutorial
Breast Cancer Detection
localhost:5000
Detect the Tumor in the Breast
Breast cancer is one of the main causes of cancer deaths worldwide. Early diagnostics significantly increases the chances of correct treatment and survival, but this process is tedious and often leads to a disagreement between pathologists. Computer-aided diagnosis systems showed potential for improving the diagnostic accuracy. But early detection and prevention can significantly reduce the chances of death. It is important to detect breast cancer as early as possible. This application is based on Convolution Neural Networks which classifies the scan as either benign or malignant tumor.
31°C
Cloudy
ENG
IN
15:06
13-10-2022


Resume
Student Dashboard
SI-322128-1664773749
R Tutorial
Breast Cancer Detection
localhost:5000
Detect the Tumor in the Breast
Drop in the scan for prediction.
Upload your image
Choose...
Open
test › benign
Search benign
Organize
New folder
Sudharshan Red
Documents
Music
Pictures
Name
Date modified
SOB_M_MC-14-16456-40-055
10-10-2022 19:15
SOB_M_MC-14-16456-40-077
10-10-2022 19:15
SOB_M_MC-14-16456-40-079
10-10-2022 19:15
File name: SOB_M_MC-14-16456-40-077
Custom Files
Open
Cancel
31°C
Cloudy
ENG
IN
15:07
13-10-2022


Resume
Student Dashboard
SI-322128-1664773749
R Tutorial
Breast Cancer Detection
localhost:5000
Detect the Tumor in the Breast
Drop in the scan for prediction.
Upload your image
Choose...
Prediction: It is a malignant tumor... Please Consult Doctor
31°C
Cloudy
ENG
IN
15:08
13-10-2022

7. ADVANTAGES AND DISADVANTAGES

Advantages:

- Early detection and prevention can significantly reduce the chances of death● Increased accuracy for Breast cancer risk prediction.
- Reduce the time complexity.

Disadvantages:

- Requires massive datasets to train on.
- Time consuming and more resources required.

8. APPLICATIONS

- Deep learning and Neural networks are key technologies used in the breast cancer risk prediction .
- It presents the results obtained by processing input from uploading image.

9. CONCLUSION

In this project, we have established the application to predict from uploaded image based on the IBM Cloud application. Breast cancer risk prediction can only use this web app to predict the breast cancer risk.

10.FUTURE SCOPE

The project can be further enhanced by deploying the deep learning model obtained using a web application and larger dataset could be used for prediction to give higher accuracy and produce the better result.

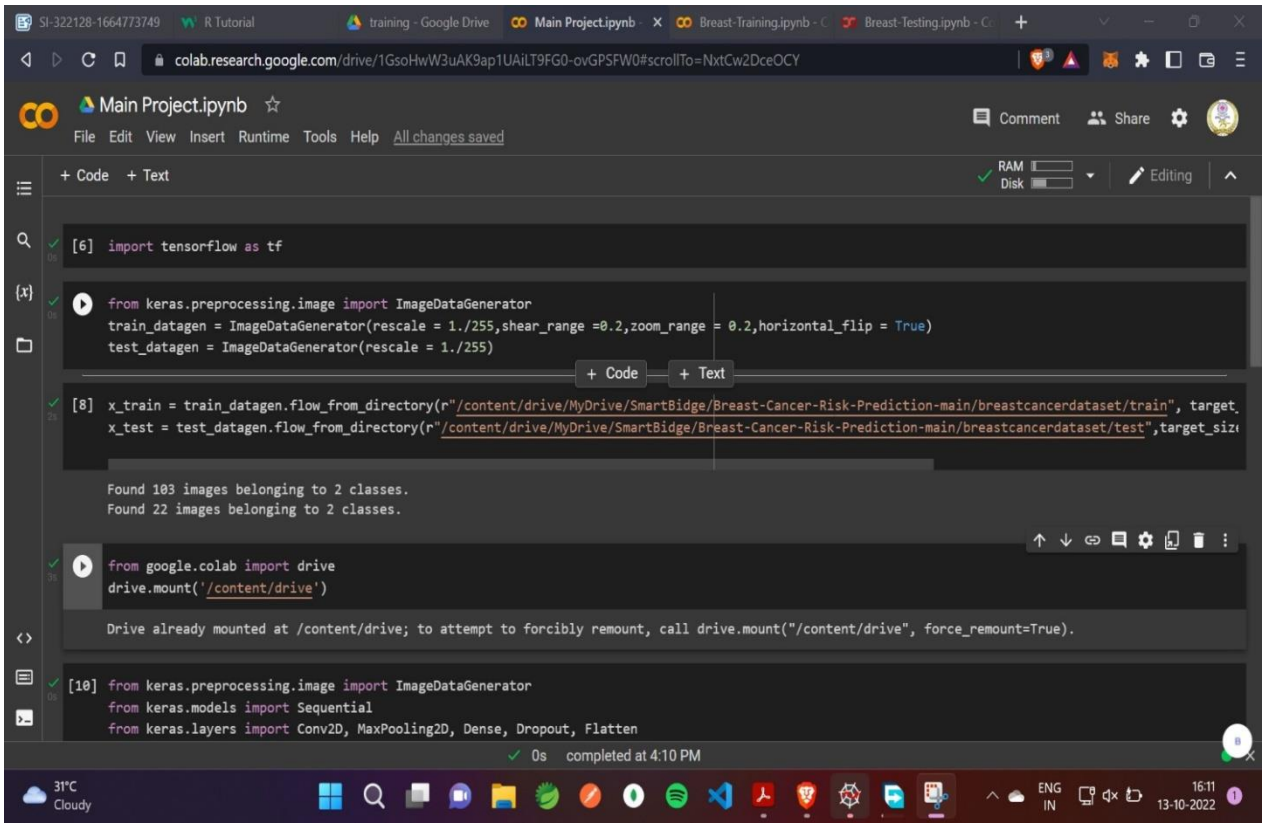
11. BIBILOGRAPHY

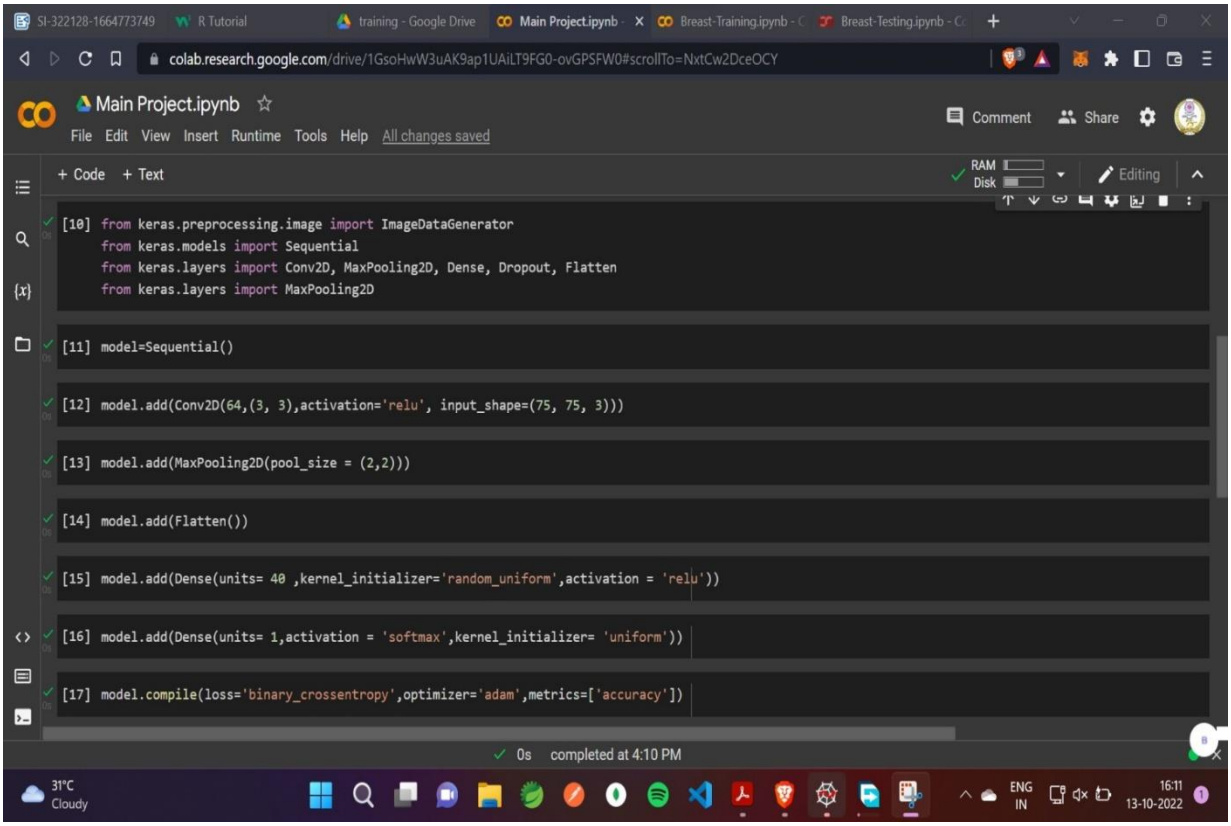
- Radiology , Breast cancer risk prediction using deep learning by *Min Sun Bae,MD,PhD and Hyug-Gi Kim, PhD.*
- Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics by *Le Lu, Xiasong Wang, Gustavo Carneiro and Lin Yang*

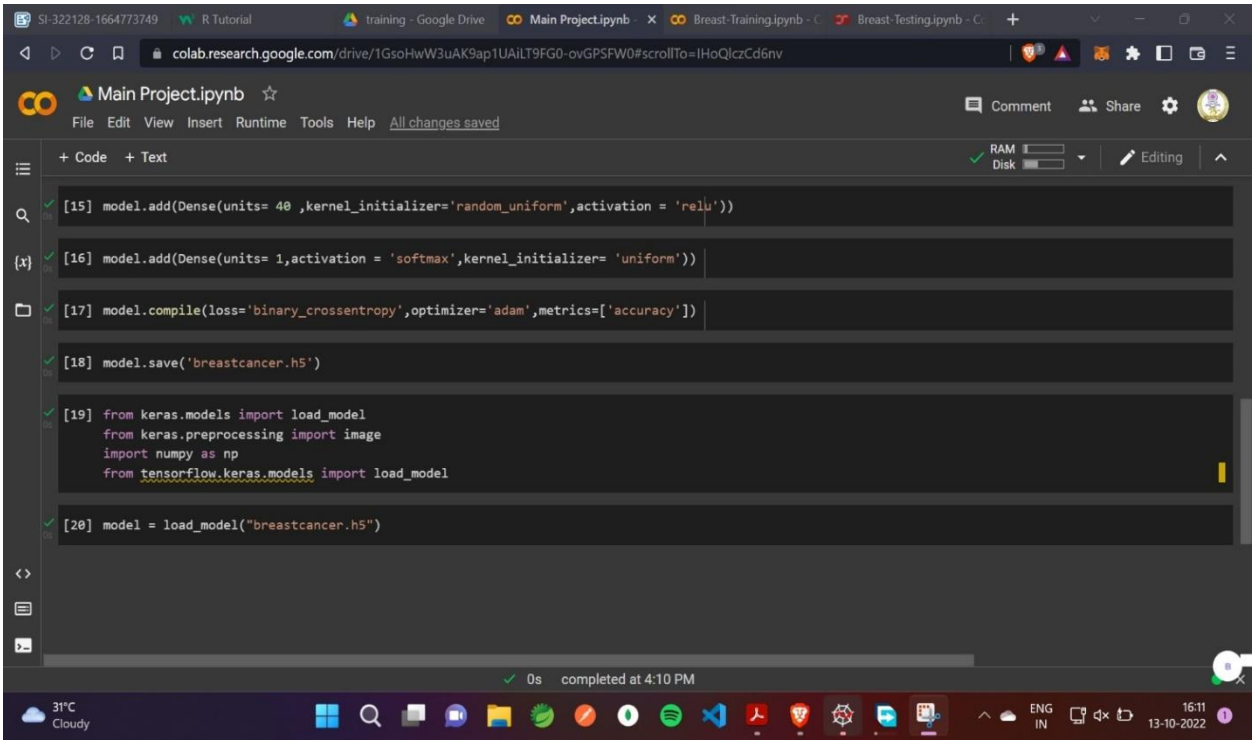
APPENDIX

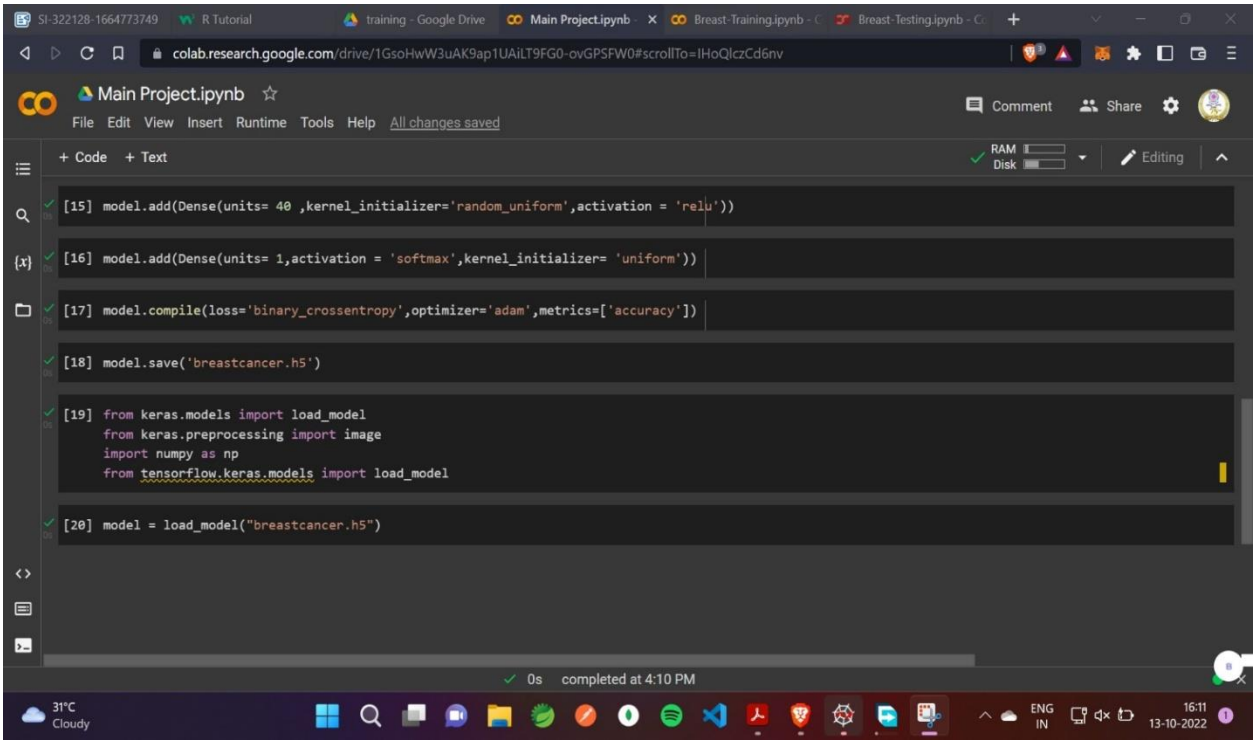
Source Code

- Main_Project.ipynb

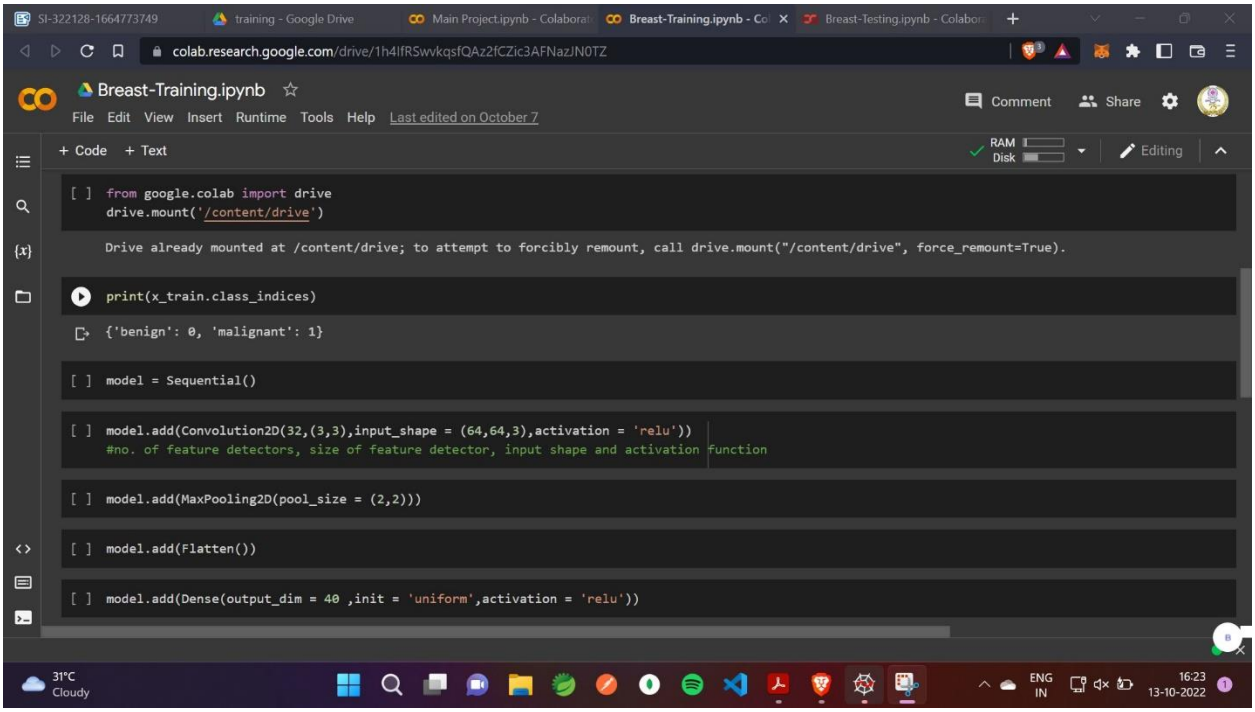


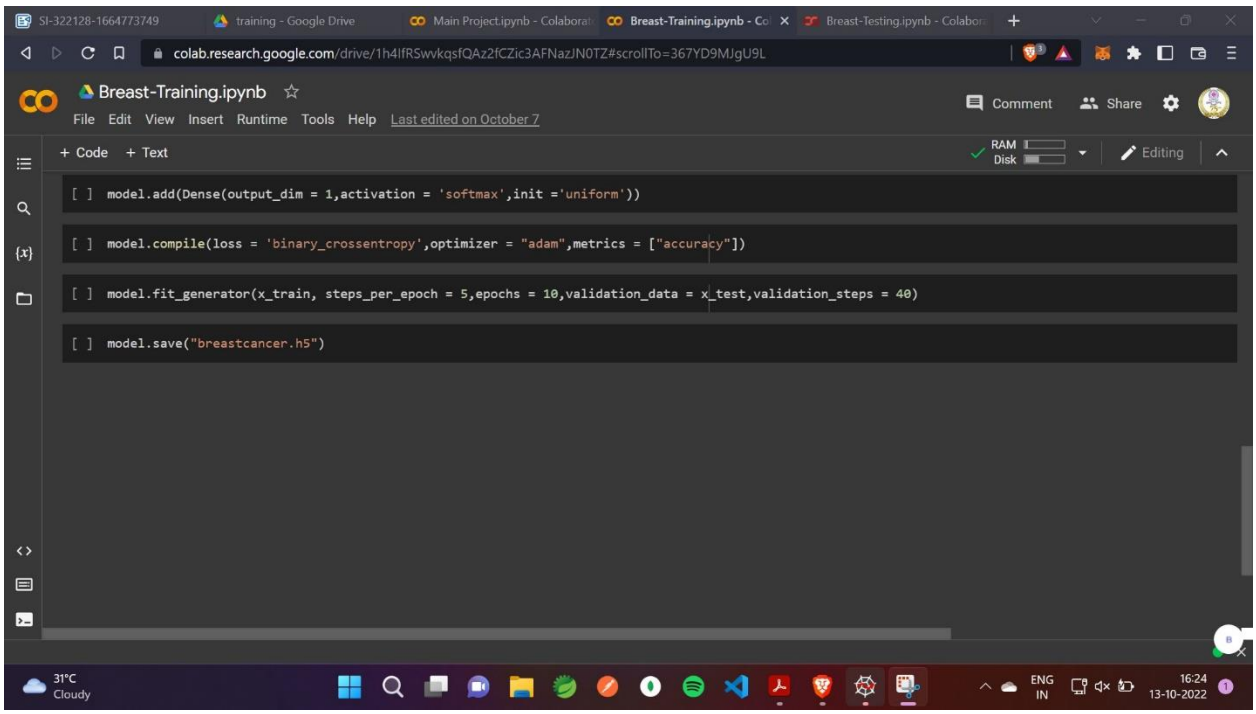

Main Project.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
+ Code + Text
RAM Disk
Editing
[10] from keras.preprocessing.image import ImageDataGenerator
from keras.models import Sequential
from keras.layers import Conv2D, MaxPooling2D, Dense, Dropout, Flatten
from keras.layers import MaxPooling2D
[11] model=Sequential()
[12] model.add(Conv2D(64,(3, 3),activation='relu', input_shape=(75, 75, 3)))
[13] model.add(MaxPooling2D(pool_size = (2,2)))
[14] model.add(Flatten())
[15] model.add(Dense(units= 40 ,kernel_initializer='random_uniform',activation = 'relu'))
[16] model.add(Dense(units= 1,activation = 'softmax',kernel_initializer= 'uniform'))
[17] model.compile(loss='binary_crossentropy',optimizer='adam',metrics=['accuracy'])
0s completed at 4:10 PM


Main Project.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
+ Code + Text
RAM Disk
Editing
[15] model.add(Dense(units= 40 ,kernel_initializer='random_uniform',activation = 'relu'))
[16] model.add(Dense(units= 1,activation = 'softmax',kernel_initializer= 'uniform'))
[17] model.compile(loss='binary_crossentropy',optimizer='adam',metrics=['accuracy'])
[18] model.save('breastcancer.h5')
[19] from keras.models import load_model
from keras.preprocessing import image
import numpy as np
from tensorflow.keras.models import load_model
[20] model = load_model("breastcancer.h5")
0s completed at 4:10 PM

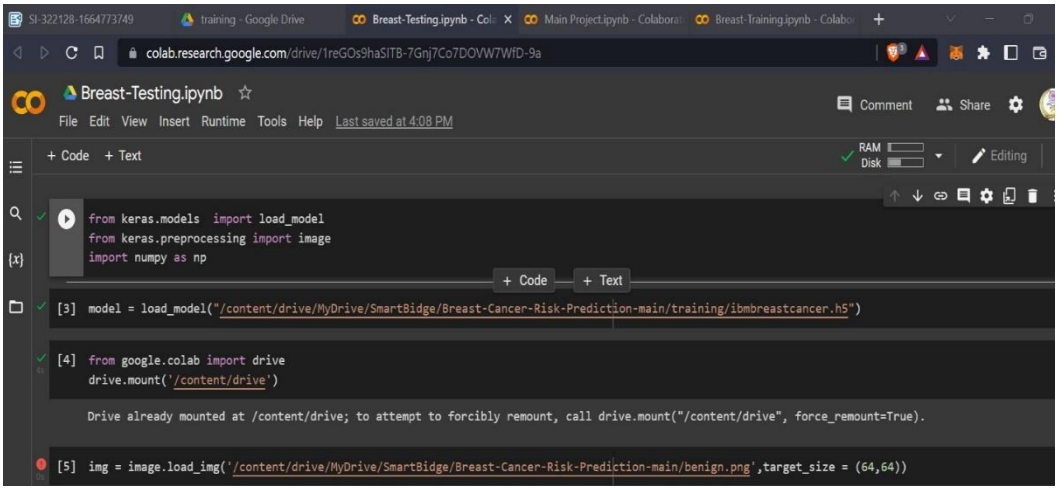


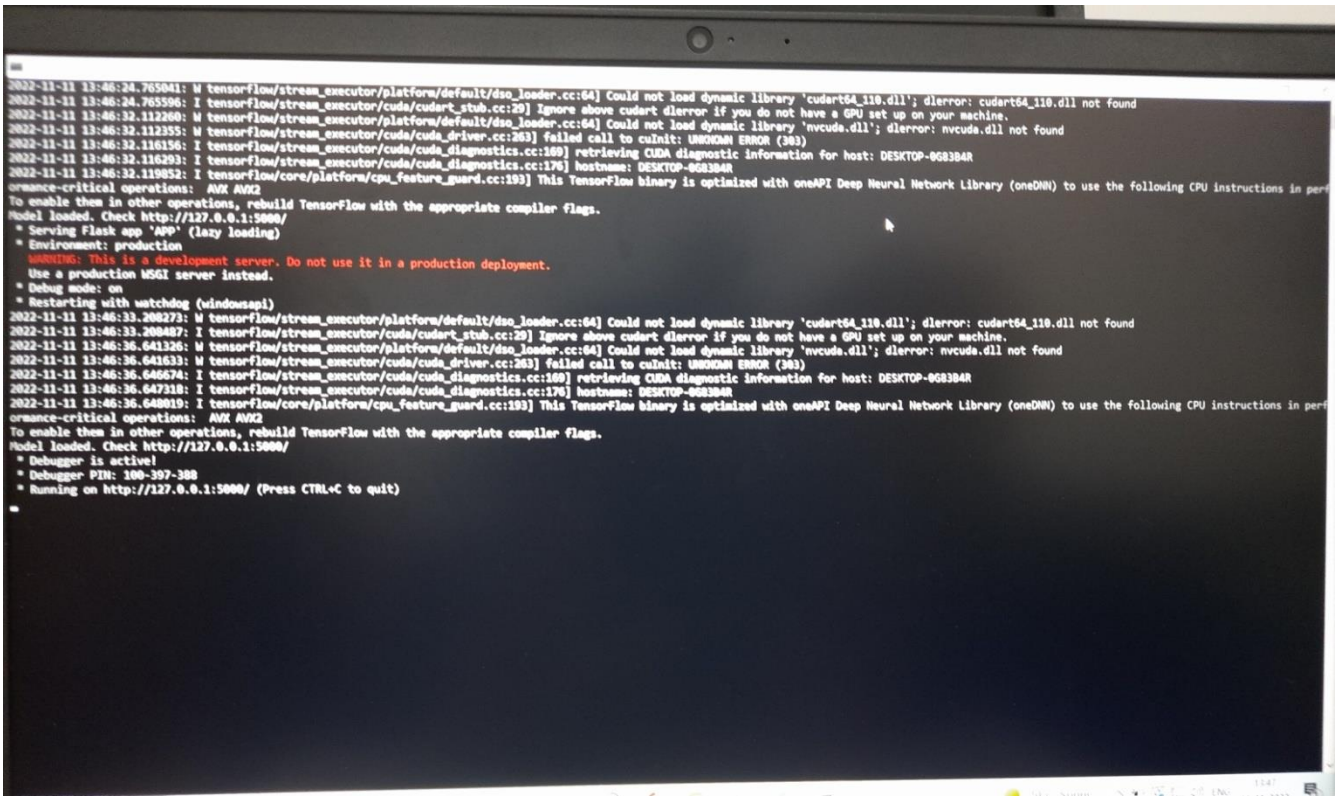
- Breast-Training.ipynb





- Breast-Testing.ipynb


SI-322128-1664773749
training - Google Drive
Breast-Testing.ipynb - Cola
Main Project.ipynb - Colaborat
Breast-Training.ipynb - Colabo
colab.research.google.com/drive/1reGOs9haSITB-7Gnj7Co7DOVW7WfD-9a
Breast-Testing.ipynb
File Edit View Insert Runtime Tools Help Last saved at 4:08 PM
Comment
Share
+ Code + Text
RAM
Disk
Editing
from keras.models import load_model
from keras.preprocessing import image
import numpy as np
+ Code + Text
[3] model = load_model("/content/drive/MyDrive/SmartBidge/Breast-Cancer-Risk-Prediction-main/training/ibmbreastcancer.h5")
[4] from google.colab import drive
drive.mount('/content/drive')
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
[5] img = image.load_img('/content/drive/MyDrive/SmartBidge/Breast-Cancer-Risk-Prediction-main/benign.png',target_size = (64,64))


2022-11-11 13:46:24.765041: W tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'cudart64_110.dll'; dlerror: cudart64_110.dll not found
2022-11-11 13:46:24.765596: I tensorflow/stream_executor/cuda/cudart_stub.cc:29] Ignore above cudart dlerror if you do not have a GPU set up on your machine.
2022-11-11 13:46:32.112260: W tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'nvcuda.dll'; dlerror: nvcuda.dll not found
2022-11-11 13:46:32.112355: W tensorflow/stream_executor/cuda/cuda_driver.cc:263] failed call to cuInit: UNKNOWN ERROR (303)
2022-11-11 13:46:32.116156: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:169] retrieving CUDA diagnostic information for host: DESKTOP-0GB3B4R
2022-11-11 13:46:32.116293: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:176] hostname: DESKTOP-0GB3B4R
2022-11-11 13:46:32.119052: I tensorflow/core/platform/cpu_feature_guard.cc:193] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library (oneDNN) to use the following CPU instructions in perf
ormance-critical operations: AVX AVX2
To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
Model loaded. Check http://127.0.0.1:5000/
* Serving Flask app 'APP' (lazy loading)
* Environment: production
WARNING: This is a development server. Do not use it in a production deployment.
Use a production WSGI server instead.
* Debug mode: on
* Restarting with watchdog (windowsapi)
2022-11-11 13:46:33.200273: W tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'cudart64_110.dll'; dlerror: cudart64_110.dll not found
2022-11-11 13:46:33.200487: I tensorflow/stream_executor/cuda/cudart_stub.cc:29] Ignore above cudart dlerror if you do not have a GPU set up on your machine.
2022-11-11 13:46:36.641326: W tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'nvcuda.dll'; dlerror: nvcuda.dll not found
2022-11-11 13:46:36.641633: W tensorflow/stream_executor/cuda/cuda_driver.cc:263] failed call to cuInit: UNKNOWN ERROR (303)
2022-11-11 13:46:36.646674: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:169] retrieving CUDA diagnostic information for host: DESKTOP-0GB3B4R
2022-11-11 13:46:36.647318: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:176] hostname: DESKTOP-0GB3B4R
2022-11-11 13:46:36.648019: I tensorflow/core/platform/cpu_feature_guard.cc:193] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library (oneDNN) to use the following CPU instructions in perf
ormance-critical operations: AVX AVX2
To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
Model loaded. Check http://127.0.0.1:5000/
* Debugger is active!
* Debugger PIN: 100-397-380
* Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)