DEEP LEARNING TECHNIQUES FOR BREAST CANCER RISK PREDICTION USING IBM CLOUD

1. INTRODUCTION

1.1 Overview

Breast cancer is one of the main causes of cancer death worldwide. 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.
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.

1.2 Purpose

To develop a mammography-based DL breast cancer risk model that is more accurate than established clinical breast cancer risk models.

2. LITERATURE SURVEY

2.1 Existing problem

It included interaction within the partner relationship as well as with family, friends, and colleagues. Data were obtained by individual and group interviews from 10 women with a diagnosis of breast cancer and 5 male partners. Both partnered and single women participated. There were four major findings seldom discussed in the literature, which have important implications for preventive intervention. First, partner relationships troubled before the diagnosis as well as those characterized by mutual caring faced challenges and negative changes. Second,

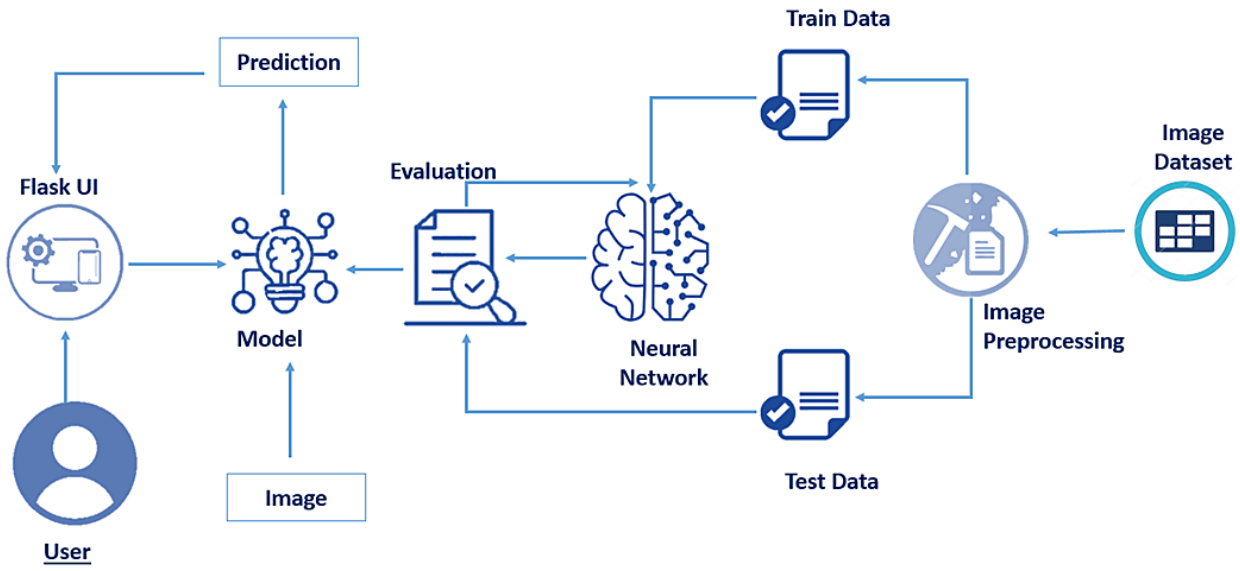
in an effort to protect each other, communication within the partner dyad became less open, and there were changes in the usual manner of conflict resolution. Third, unpartnered women appeared to be more vulnerable to problems of negative adjustment, largely because of relationship issues. Fourth, participants confirmed the need for a comprehensive intervention to facilitate coping with issues relative to relationships, intimacy, and sexuality. Although the sample was small, in-depth data were obtained that provide a basis for specific areas in which further empirical investigation is needed, and they indicate that preventive intervention may well be warranted.

2.2 Proposed solution

The accurate classification of the histopathological images of breast cancer diagnosis may face a huge challenge due to the complexity of the pathologist images. Currently, computer-aided diagnosis is implemented to get sound and error-less diagnosis of this lethal disease. However, the classification accuracy and processing time can be further improved. This study was designed to control diagnosis error via enhancing image accuracy and reducing processing time by applying several algorithms such as deep learning, *K*-means, autoencoder in clustering and enhanced loss function (ELF) in classification. Histopathological images were obtained from five datasets and pre-processed by using stain normalisation and linear transformation filter. These images were patched in sizes of 512 × 512 and 128 × 128 and extracted to preserve the tissue and cell levels to have important information of these images. The patches were further pre-trained by ResNet50-128 and ResNet512. Meanwhile, the 128 × 128 were clustered and autoencoder was employed with *K*-means which used latent feature of image to obtain better clustering result. Classification algorithm is used in current proposed system to ELF. This was achieved by combining SVM loss function and optimisation problem. The current study has shown that the deep learning algorithm has increased the accuracy of breast cancer classification up to 97% compared to state-of-the-art model which gave a percentage of 95%, and the time was decreased to vary from 30 to 40 s. Also, this work has enhanced system performance via improving clustering by employing *K*-means with autoencoder for the nonlinear transformation of histopathological image.

3. THEORITICAL ANALYSIS

3.1 Block Diagram



3.2 Hardware / Software designing

Software Requirements:

- Anaconda Navigator
- Keras
- Flask

Hardware Requirements:

- Processor : Intel Core i3
- Hard Disk Space : Min 100 GB
- Ram : 8 GB
- Display : 14.1 “Color Monitor(LCD, CRT or LED)

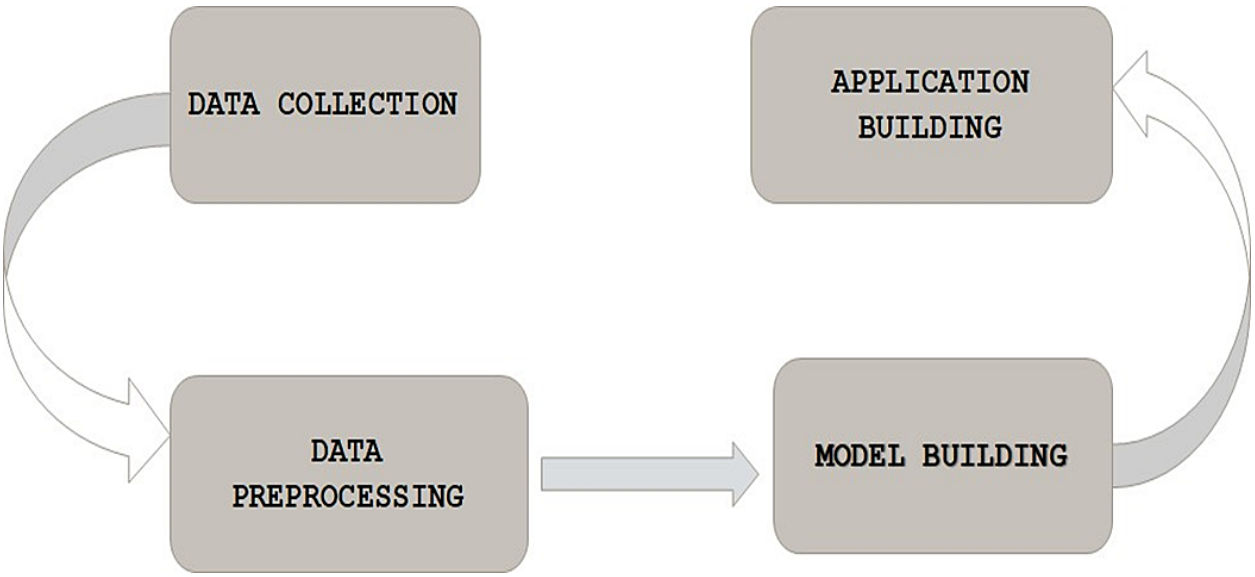
  Clock Speed : 1.67 GHz

4. EXPERIMENTAL INVESTIGATIONS

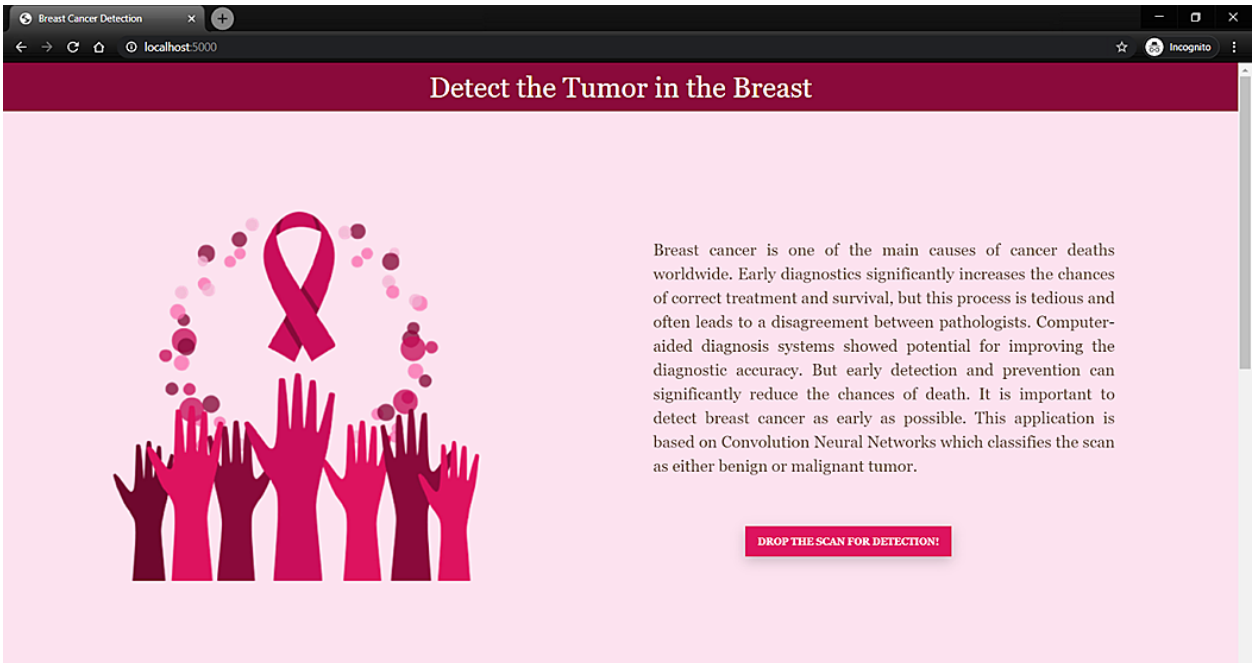
1. Install Anaconda.
2. Install Required Libraries.
3. Data Collection.
    a. Collect the dataset or Create the dataset ● Data Preprocessing.
        - Import the Libraries.
        - Importing the dataset.
        - Checking for Null Values.
        - Data Visualization.
        - Taking care of Missing Data.
        - Label encoding.
        - One Hot Encoding.
        - Feature Scaling.
        - Splitting Data into Train and Test.
4. Model Building
    a. Training and testing the model
        - Evaluation of Model
5. Application Building
    a. Create an HTML file

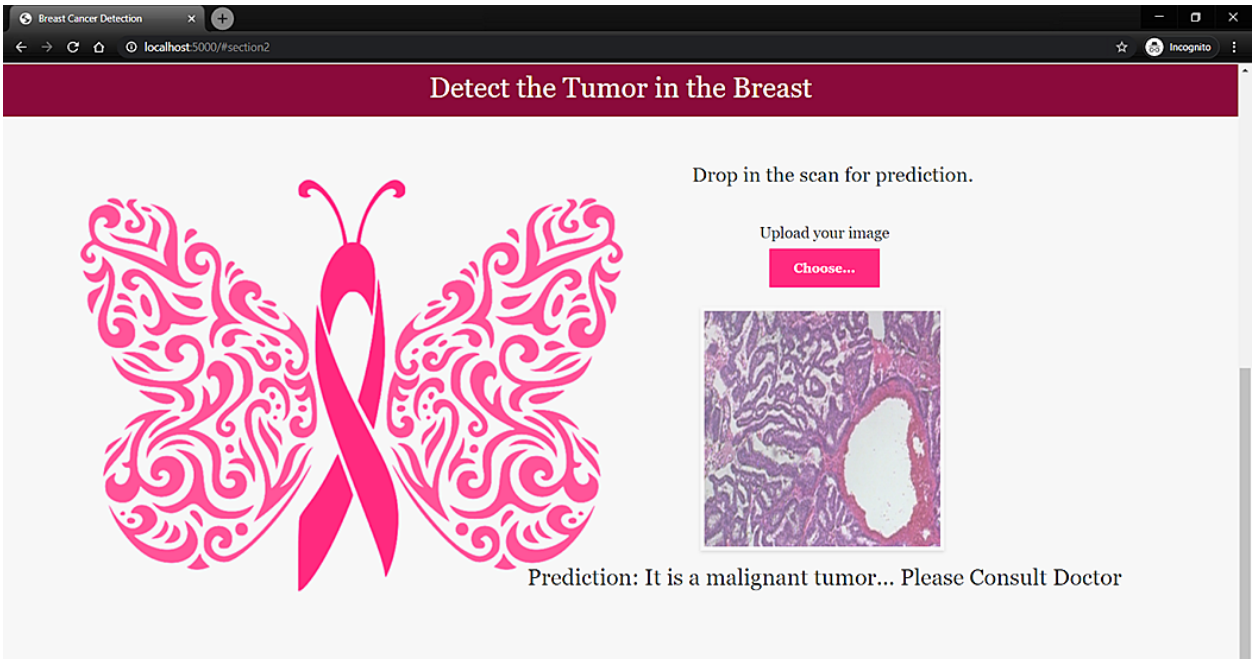
b. ○ Build a Python Code

5. FLOWCHART



6. RESULT





7. ADVANTAGES & DISADVANTAGES

Advantages:

Several studies have found underutilization of radiotherapy in patients with breast cancer; but

there are concerns about the completeness of various databases on radiotherapy. We used the linked Medicare-SEER (Surveillance, Epidemiology and End Results) database to compare information on receipt of radiotherapy after diagnosis of breast cancer. More than 18% of women identified by Medicare data as receiving radiotherapy were not so identified by SEER, and 7% of those identified as receiving radiotherapy by SEER were not identified by Medicare. Risk of discordance on radiotherapy information between the two data sets was especially high in women receiving breast-conserving surgery. The combined SEER-Medicare database gives a more complete picture on the use of radiotherapy. The previously reported geographic variations in the use of radiotherapy for breast cancer may be due in part to underreporting of radiotherapy in some areas.

Disadvantages:

Women who resided in the most socio-economically disadvantaged areas were significantly more likely (OR 1.21, 95% CI 1.07 to 1.37) than residents of the most advantaged areas to be diagnosed as having advanced breast cancer after adjustment for individual-level factors. When geographic remoteness and area-disadvantage (and all the individual-level factors) were simultaneously adjusted, the rates of advanced breast cancer were significantly higher for women residing in Outer Regional areas (OR 1.13, 95% CI 1.02 to 1.24) and those who lived in the most disadvantaged areas (OR 1.16, 95% CI 1.02 to 1.32). There was no statistically significant interaction between geographic remoteness and area disadvantage.

8. APPLICATIONS

- Better Power Output Wind power forecasts are important in efficiently using wind turbines for generating power output.
- Efficient Predicting features like wind speed and wind direction can greatly help one to make decisions on when to switch on the wind turbine and when to switch it off(when it is assumed to not get the suitable conditions for generating power)
- Environment friendly If we are able to achieve predicting the wind power output, then it will open up more avenues for efficient power production in this field. This will lower the

dependence on conventional sources of energy like coal which can cause harm to our environment.

9. CONCLUSION

Breast cancer if found at an early stage will help save lives of thousands of women or even men. These projects help the real world patients and doctors to gather as much information as they can. The research on nine papers has helped us gather the data for the project proposed by us. By using machine learning algorithms we will be able to classify and predict the cancer into being or malignant. Machine learning algorithms can be used for medical oriented research, it advances the system, reduces human errors and lowers manual mistakes.

10. FUTURE SCOPE

In future, these techniques may be implemented on datasets that consist of images. The system may also be integrated with an application or website. The accuracy of the model created may be increased in order to give better predictions

11. BIBILOGRAPHY

- Johns PC, Yaffe MJ: X-ray characterisation of normal and neoplastic breast tissues. Phys Med Biol. 1987, 32: 675-695. 10.1088/0031-9155/32/6/002.

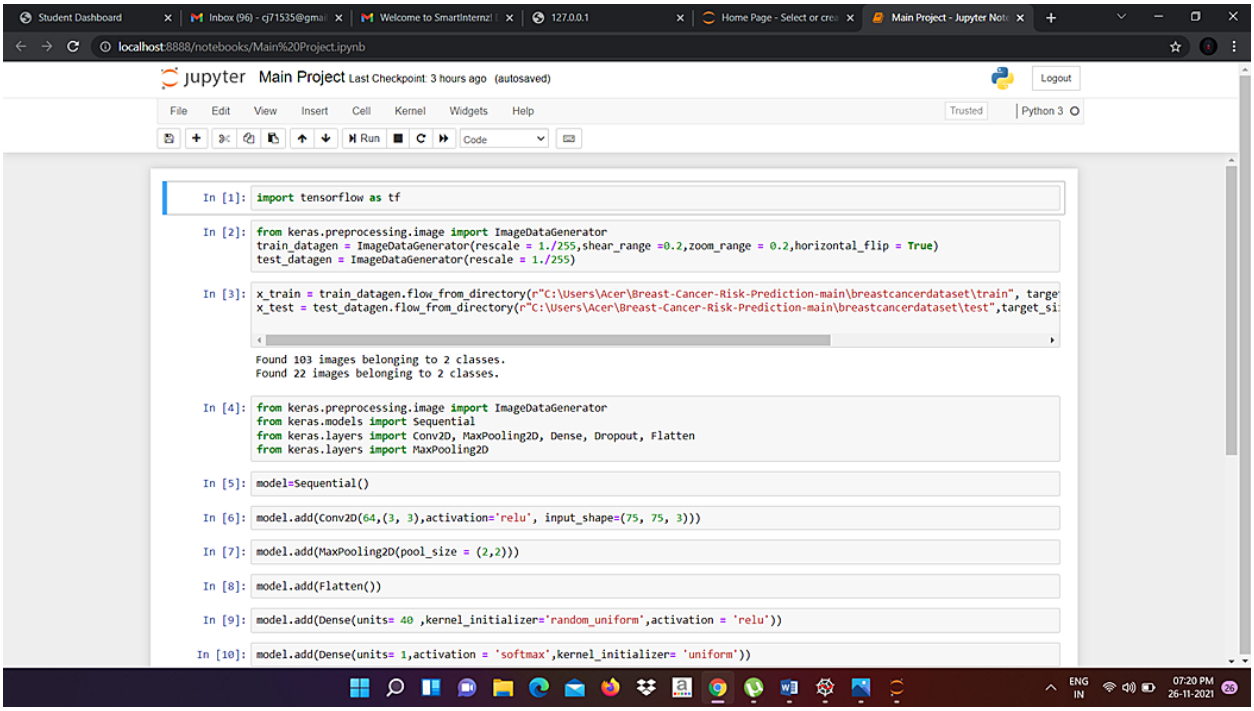
1. Egan RL: Breast Imaging: Diagnosis and Morphology of Breast Diseases. 1988, Philadelphia: WB Saunders Company
2. McCormack VA, dos Santos Silva I: Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev. 2006, 15: 1159-1169. 10.1158/1055-9965.EPI-06-0034.
3. Byrne C, Schairer C, Wolfe J, Parekh N, Salane M, Brinton LA, Hoover R, Haile R:

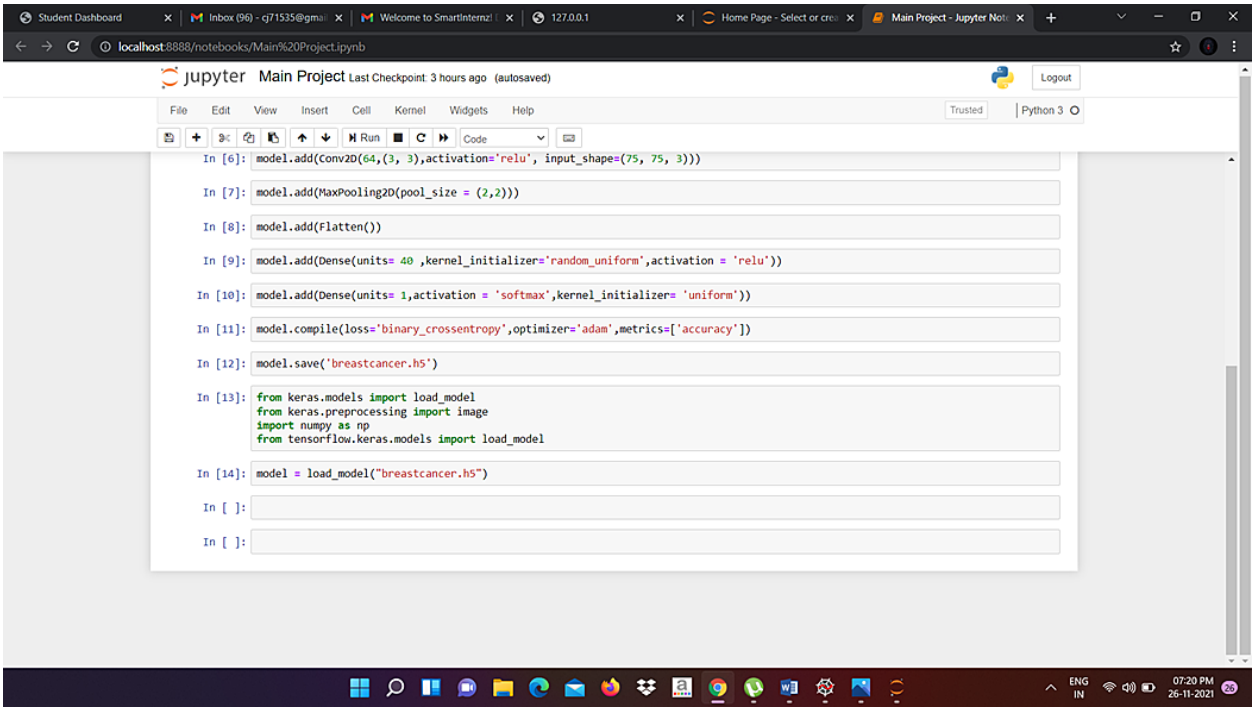
Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst. 1995, 87: 1622-1629. 10.1093/jnci/87.21.1622.

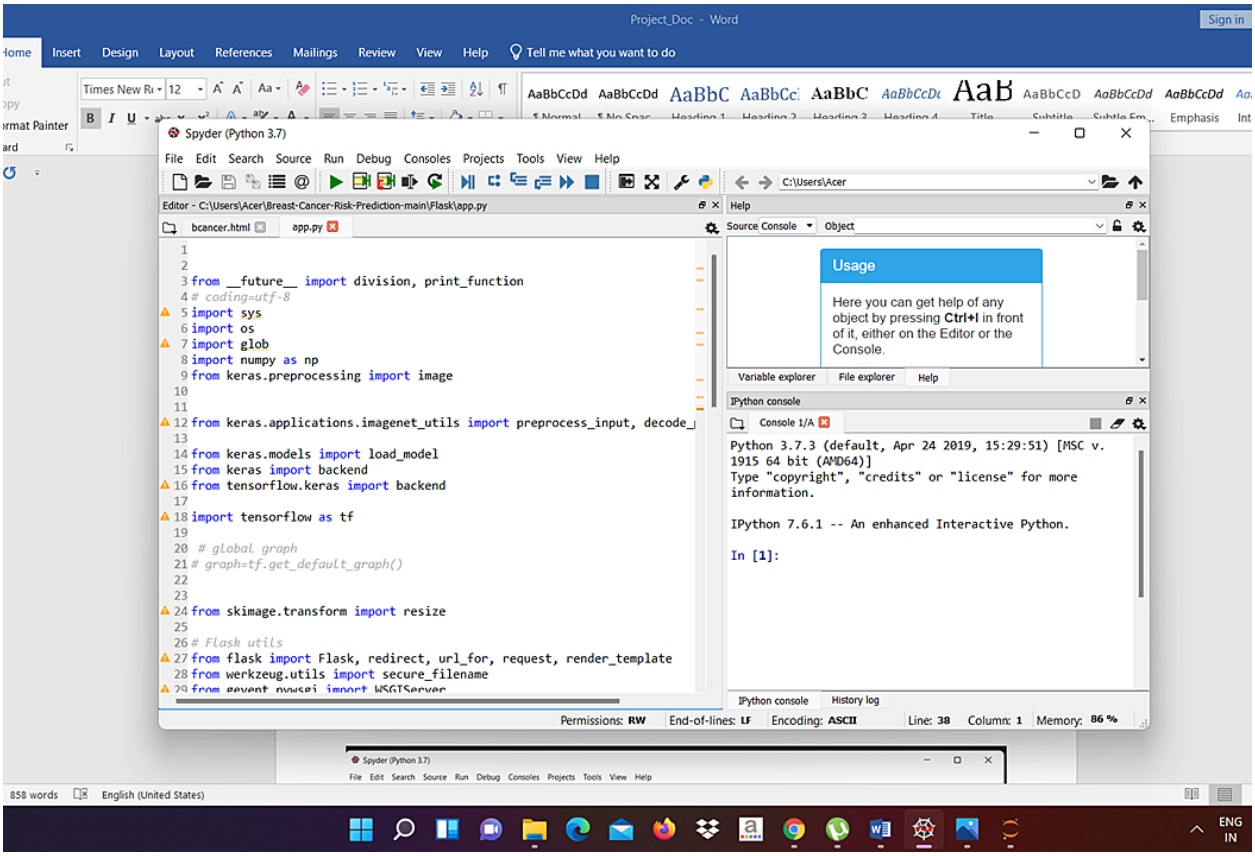
1. Yi-Sheng Sun, Zhao Zhao, Han-Ping-Zhu,"Risk factors and Preventions of Breast Cancer" International Journal of Biological Sciences.
2. D. Selvathi and A Aarthypoornila, "Performance analysis of various classifiers on deep learning network for breast cancer detection", International Conference on Signal Processing and Communication (ICSPC)

APPENDIX

Source Code




Student Dashboard
Inbox (96) - cj71535@gmai
Welcome to SmartInternz!
127.0.0.1
Home Page - Select or crea
Main Project - Jupyter Note
localhost:8888/notebooks/Main%20Project.ipynb
jupyter Main Project Last Checkpoint: 3 hours ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3
Run
Code
In [6]: model.add(Conv2D(64,(3, 3),activation='relu', input_shape=(75, 75, 3)))
In [7]: model.add(MaxPooling2D(pool_size = (2,2)))
In [8]: model.add(Flatten())
In [9]: model.add(Dense(units= 40 ,kernel_initializer='random_uniform',activation = 'relu'))
In [10]: model.add(Dense(units= 1,activation = 'softmax',kernel_initializer= 'uniform'))
In [11]: model.compile(loss='binary_crossentropy',optimizer='adam',metrics=['accuracy'])
In [12]: model.save('breastcancer.h5')
In [13]: from keras.models import load_model
from keras.preprocessing import image
import numpy as np
from tensorflow.keras.models import load_model
In [14]: model = load_model("breastcancer.h5")
In [ ]:
In [ ]:
ENG
IN
07:20 PM
26-11-2021


Project_Doc - Word
Spyder (Python 3.7)
File Edit Search Source Run Debug Consoles Projects Tools View Help
Editor - C:\Users\Acer\Breast-Cancer-Risk-Prediction-main\Flask\app.py
bcancer.html
app.py
from __future__ import division, print_function
# coding=utf-8
import sys
import os
import glob
import numpy as np
from keras.preprocessing import image
from keras.applications.imagenet_utils import preprocess_input, decode_
from keras.models import load_model
from keras import backend
from tensorflow.keras import backend
import tensorflow as tf
# global graph
# graph=tf.get_default_graph()
from skimage.transform import resize
# Flask utils
from flask import Flask, redirect, url_for, request, render_template
from werkzeug.utils import secure_filename
Help
Source Console
Object
Usage
Here you can get help of any object by pressing Ctrl+I in front of it, either on the Editor or the Console.
Variable explorer
File explorer
Help
IPython console
Console 1/A
Python 3.7.3 (default, Apr 24 2019, 15:29:51) [MSC v. 1915 64 bit (AMD64)]
Type "copyright", "credits" or "license" for more information.
IPython 7.6.1 -- An enhanced Interactive Python.
In [1]:
IPython console
History log
Permissions: RW
End-of-lines: LF
Encoding: ASCII
Line: 38
Column: 1
Memory: 86 %
858 words
English (United States)