This file is to provide a basic walkthrough of what each folder/file in this repo means/stands for. To make contributions easier for you

| FILE NAME | FILE CONTENT | FILE USAGE |
| --- | --- | --- |
|  |  |  |
| Dataset | The CSV file of the project.  A cleaned version is also available. | Can be used to train the prediction model of this project(ML models or neural networks) |
| Model | The models trained on the dataset. | Add your appropriately named model to this folder when you are done training. |
| Visualization | Contains the visualization notebooks and the plot images. | If you create new plots/visualizations, you should keep the notebook and plot images here. |
| \_\_pycache\_\_ | A directory that contains bytecode cache files that are automatically generated by python | Not needed by you. |
| static | Contains the frontend framework of the website. | Can be used to update the appearance of the website. |
| templates | Contain the HTML files of the website | Can be used to update the arrangement/format of the website. |
| .gitattributes | The .gitattributes file allows you to specify the files and path attributes that should be used by git when performing git actions | You don't need it. It just makes working with git via cmd possible |
| .gitignore | .gitignore file tells Git which files to ignore when committing your project to the GitHub repository | You don't need it but it ignores .vscode and .DS\_STORE files |
| Breast Cancer data.csv | Contains the raw CSV file of the project. | Can be used for data cleaning/preprocessing, feature engineering and data modelling |
| CODE\_OF\_CONDUCT.md | A policy lays out the project’s principles, standards, morals and ethical expectations | You need to know how to relate with other contributors and the repo’s manager. |
| Contributor’s guide | The guide each contributor is to follow as regards contributing to the project. | A must read for everybody interested in working on the project. |
| Contributor | A file that contains the name of everybody that has contributed to the project. | Kindly add your name after contributing to the project |
| LICENSE | Makes the software free to use, change and distribute. | You don't need it. Copy right protection of the owner works |
| Models Drive.txt | A notepad that contains the drive link to 2 heavy models. | If you find it hard to push your model to the github repo, you can always uplod it to this drive |
| README.md | Important information about your project | Helps to understand what the project is all about. |
| Transfer learning.py | Contains the notebook for a transfer learning approach |  |
| app.py | The backend file of the project. | You may decide to deploy more models or optimize the website/ Hence, the need for this |
| cnn\_app.py | The python script for a CNN model implemented with flask. |  |
| cnn\_model.h5 | The h5 version of the CNN model implemented with flask. |  |
| Ensemble\_modelling | The python script for an ensemble model implemented with |  |
| Implementation.py | The Random Forest implementation(i.e as a function) of this notebook. | You can decide to create an implementation file of any other model. |
| random\_forest.py | Random forest training of the dataset. | You might want to improve the accuracy of the random forest model. |
| requirement.txt | System requirement for the project. | The version for each package used in the project. Helps when trying to set the project up on your PC for real-time detection. |
| BCP\_Image\_Dataset | Image dataset | Should be used to train a deep learning model for image classification. |
| LOGO: | The logo image of the project. | Can be used on the website. |