

Cancer Prevention and Early Detection during the COVID-19 Pandemic

Subash Kumar

Bachelor of Computer Science Engineering, Anna University.

To Cite this Article

Subash Kumar, "Cancer Prevention and Early Detection during the COVID-19 Pandemic", *International Journal for Modern Trends in Science and Technology*, 6(8): 187-190, 2020.

Article Info

Received on 13-July-2020, Revised on 03-August-2020, Accepted on 06-August-2020, Published on 15-August-2020.

ABSTRACT

Currently the governments across the world attempt to halt the spread of COVID-19, many of their efforts are being hampered by shortages of crucial testing kits. Even in many developed countries, the health system has come to the point of standstill due to the increasing demand for intensive care units. Any hospital in this world will be having a multidisciplinary team of experts involving radiologists, oncologists, surgeons, and other respiratory physicians and consultants. This multidisciplinary team uses a wide variety of data taken from CT Scans, X-Rays, Pet Scans, Biopsies, Clinical Examination, Imaging Mammography, Ultrasound (Sonography), Pathology Biopsy and Fine Needle Aspiration to identify whether a patient has cancer. Individuals with cancer, particularly those who are receiving systemic anticancer treatments, have been postulated to be at increased risk of mortality from COVID-19. Therefore, it is the need of an hour to implement an early detection of cancer that will prevent patients from COVID-19 and due treatment will be provided.

KEYWORDS: Covid-19 Pandemic, Corona Virus, Machine Learning, Neural Network, Deep Learning, CNN, Artificial Neural Network, Statistics, Artificial Intelligence, Cancer, Medicine

I. INTRODUCTION

Thousands of our grandmothers, mothers, and daughters fall victim to breast cancer every year. The human body comprises millions of cells each with its own unique function. When there is the unregulated growth of any of these cells it is termed as Cancer. In this, cells divide and grow uncontrollably, forming an abnormal mass of tissue called a tumor. Tumor cells grow and invade digestive, nervous and circulatory systems disrupting the bodies' normal functioning. Though every single tumor is not cancerous.

Breast cancer represents the second primary cause of cancer deaths in women today, the first one is Lung cancer. Breast cancer and has become the most common cancer among women both in

the developed and the developing world in the last years. 40,000 women die in a year from this disease, which is one woman every 13 minutes dying from this disease every day. Early detection of breast cancer is far easier to cure.

Breast cancer diagnosis differentiates benign (lacks the ability to invade neighboring tissue) from malignant (ability to invade neighboring tissue) breast tumors. Breast cancer is a type of cancer originating from the breast tissue, commonly from the inner lining of the milk ducts or lobules supplying the ducts with milk. It remains the number one form of cancer that woman is diagnosed with around the world. Even with enhanced treatment, the lack of early detection has

put women at even higher risk of dying from this disease.

Benign Breast Cancer (Non-Invasive): It is also known as carcinoma in situ. This type of cancer doesn't spread to neighboring tissue regions and hence is rarely a threat to life. These cells remain entirely in-situ (in their place of origin) because they have not yet developed the ability to spread outside of these ducts, either within the breast or elsewhere in the body.

Malignant Breast Cancer (Invasive): Malignant or Invasive is the type in which cancer has the potential to spread from the breast to other parts of the body and is a threat to life. Often they can be removed but sometimes grow back. The most common type of invasive breast cancer is invasive ductal cancer. This accounts for 80 % of all cases of breast cancer.

In the current world, there are various diagnostics tests for breast cancer such as Clinical Examination that helps us to understand about women above a certain age and showing symptoms should have a physical exam to check for breast cancer. The breast is examined for lumps or suspicious areas (change in texture, size).

Imaging Mammography, which is an X-Ray of the breast. Ultrasound (Sonography), a test uses sound waves to outline a part of the breast. It is usually helpful in women with dense breasts and is used to target a specific area found on the mammogram. Pathology Biopsy, this test involves taking a sample of tissue cells from the breast and testing to see whether it is cancerous or not.

What if the Computer is trained with unlabeled data, the algorithm is trained and left to find the commonalities among its input data and computer might be able to teach new things after it learns patterns from the data about tumors are so small less than 4 mm that it is very difficult to diagnose them via CT scan images that makes these algorithms particularly very helpful in cases where the human expert doesn't know what to look for in the data

II. METHODOLOGY

Deep learning is a sub-branch of the machine learning field comes from Artificial Intelligence, inspired by the structure of the human brain, a human brain consists of nearly 10 billion neurons with 60 trillion connections between each other. A neuron consists of a cell body called the soma where the nucleus is found, many dendrites where input signals are received and transmitted and a

synapse which is connections between neuron to neuron.

Like a brain neuron, neural networks consist of an artificial neuron or node. An artificial neuron consists of a weight to determine the strength of a connection, a linear function that needs to be computed and an activation function that computes the weighted sum of the linear function that is then compared to a threshold value, with this inspiring technique, an artificial intelligence tool can quickly detect cancerous tumor It can classify information which would not normally be recognized with the naked

Deep learning techniques used in recent years continue to show an impressive performance in the field of medical image processing, as in many fields. By applying deep learning techniques to medical data, we are able to draw meaningful results from medical data. Deep learning models have been used successfully in many areas such as classification, segmentation and lesion detection of medical data. Analysis of image and signal data obtained with medical imaging techniques such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT) and X-ray with the help of deep learning models helps us to analyze, detect and diagnosis of diseases such as diabetes mellitus, brain tumor, skin cancer and breast cancer.

A timely diagnosis of any disease is critical in the medical field, with an increasing population of breast cancer patients, this paper is dedicated to all medical professionals who are trying to save many lives. Machine learning and Deep learning technique will help physicians to diagnose accurately, a family of algorithms known as neural networks has recently seen a revival under the name "deep learning." While deep learning shows great promise in many machine learning applications, deep learning algorithms are often tailored very carefully to a specific use case. Here, we will only discuss some relatively simple methods, namely multilayer perceptron's for classification and regression, which can serve as a starting point for more involved deep learning methods. Multilayer perceptron's (MLPs) are also known as (vanilla) feed-forward neural networks, or sometimes just neural networks.

Let us see the novelty of this research paper. In this paper, we will use 12 attributes, out of which 10 real-valued features from each cell nucleus obtained from the Wisconsin diagnostic breast cancer dataset that explains about the stage of breast cancer M (Malignant) and B (Benign) will be the input to the neural network

Features explanations:

1. ID: Patient id
2. Diagnosis (M = Malignant, B = Benign)
3. Radius (mean of distances from the center to points on the perimeter) (worst). Worst texture. Texture (standard deviation of gray-scale values) (worst). Worst perimeter. perimeter (worst)
4. Texture (Breast cancer can cause changes and inflammation in skin cells that can lead to texture changes), here we can take the standard deviation of grey scaled values
5. Perimeter: Size of the core tumor
6. Area: Area of the core tumor
7. Smoothness: Local variation in radius length
8. Compactness: $(\text{perimeter}^2 / \text{area} - 1.0)$
9. Concavity (severity of concave portions of the contour)
10. Concave points (Number of concave portions of the contour)
11. Symmetry
12. Fractal Dimension

Except for the ID and Diagnosis, all other features are divided into three parts, the first part is Mean that tells about the mean of all cells, the second part is Standard Error that tells about the standard the error of the cells and the third part is the worst mean of the worst cells. Now we have a total of 30 features (radius mean, texture mean, perimeter mean, area mean, smoothness mean, compactness mean, concavity mean, concave points mean, symmetry mean, fractal dimension mean, radius se, texture se, perimeter se, area se, smoothness se, compactness se, concavity se, concave points se, symmetry se, fractal dimension se, radius worst, texture worst, perimeter worst, area worst, smoothness worst, compactness worst, concavity worst, concave points worst, symmetry worst, fractal dimension worst) that will be the data for the Support Vector Machine Learning - Artificial Intelligence (Machine Learning & Deep Learning) For example, field 1 is the Mean Radius, field 11 Standard Error Radius, 21 Worst mean radius of the tumor. The data set contain 569 rows and 32 columns. 'Diagnosis' is the column which we are going to predict, which says if the cancer is M = malignant or B = benign. 1 means the cancer is malignant and 0 means benign. We can identify that out of the 569 persons, 357 are labeled as B (benign) and 212 as M (malignant).

Python programming language has the inbuilt library Sklearn Neural Network with MLP Classifier

which is a widely used library for implementing the machine learning algorithms. The neural network learned a very nonlinear but relatively smooth decision boundary. By default, the MLP uses 100 hidden nodes, which is quite a lot for this small dataset. We can reduce the number (which reduces the complexity of the model) and still get a good result. Finally, we can also control the complexity of a neural network by using an l2 penalty to shrink the weights toward zero, as we did in ridge regression and the linear classifiers. The parameter for this in the MLPClassifier is alpha (as in the linear regression models), and it's set to a very low value (little regularization). The accuracy of the MLP is quite good, but not as good as the other models, this is likely due to the scaling of the data. Neural networks also expect all input features to vary in a similar way, and ideally to have a mean of 0 and a variance of 1. We must rescale our data so that it fulfils these requirements.

One of the possible inference we can make is that features that have very small weights for all of the hidden units are "less important" to the model. We can see that "mean smoothness" and "mean compactness," in addition to the features found between "smoothness error" and "fractal dimension error," have relatively low weights compared to other features. This could mean that these are less important features or possibly that we didn't represent them in a way that the neural network could use.

Neural networks have re-emerged as state-of-the-art models in many applications of machine learning. One of their main advantages is that they are able to capture the information contained in large amounts of data and build incredibly complex models. Given enough computation time, data, and careful tuning of the parameters, neural networks often beat other machine learning algorithms (for classification and regression tasks). This brings us to the downsides. Neural networks—particularly the large and powerful ones—often take a long time to train. They also require careful pre-processing of the data, as we saw here. Similarly to SVMs, they work best with "homogeneous" data, where all the features have similar meanings. For data that has very different kinds of features, tree-based models might work better. Tuning neural network parameters is also an art unto itself. In our experiments, we barely scratched the surface of possible ways to adjust neural network models and how to train them. A helpful measure when thinking about the model complexity of a neural

network is the number of weights or coefficients that are learned. If we have a binary classification dataset with 100 features, and you have 100 hidden units, then there are $100 * 100 = 10,000$ weights between the input and the first hidden layer. There are also $100 * 1 = 100$ weights between the hidden layer and the output layer, for a total of around 10,100 weights. If we add a second hidden layer with 100 hidden units, there will be another $100 * 100 = 10,000$ weights from the first hidden layer to the second hidden layer, resulting in a total of 20,100 weights. If instead, we use one layer with 1,000 hidden units, you are learning $100 * 1,000 = 100,000$ weights from the input to the hidden layer and $1,000 * 1$ weights from the hidden layer to the output layer, for a total of 101,000. If we add a second hidden layer you add $1,000 * 1,000 = 1,000,000$ weights, for a whopping total of 1,101,000—50 times larger than the model with two hidden layers of size 100. A common way to adjust parameters in a neural network is to first create a network that is large enough to over fit, making sure that the task can actually be learned by the network. Then, once you know the training data can be learned, either shrink the network or increase alpha to add regularization, which will improve generalization performance.

III. RESULTS

The data we use is usually split into training data and test data. The training set contains a known output and the model learns on this data in order to be generalized to other data later on. We have the test dataset (or subset) which is about 25% of total sample size 569 is 143, in order to test our model's prediction on this subset.

The below criteria were used for the performances of deep learning transfer models

$$\text{Accuracy} = (\text{TN} + \text{TP}) / (\text{TN} + \text{TP} + \text{FN} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

True positives (TP): These are cases in which we predicted yes (they have the disease), and they do have the disease. , we predicted true positive for our model as 50

True negatives (TN): We predicted no, and they don't have the disease, we predicted true negative for our model as 87

False positives (FP): We predicted yes, but they don't have the disease. (Also known as a "Type I error."), we predicted the false positive for our model as 3

False negatives (FN): We predicted no, but they do have the disease. (Also known as a "Type II error."), we predicted the false negative for our model as 3

The results are much better after scaling, and already quite competitive. We got a warning from the model, though, that tells us that the maximum number of iterations has been reached. This is part of the algorithm for learning the model and tells us that we should increase the number of iterations. Increasing the number of iterations only increased the training set performance, not the generalization performance. Still, the model is performing quite well. As there is some gap between the training and the test performance, we might try to decrease the model's complexity to get better generalization performance. Here, we choose to increase the alpha parameter (quite aggressively, from 0.0001 to 1) to add stronger regularization of the weights. Accuracy on training set: 98.8% Accuracy on test set: 97.2%

IV. DISCUSSION

We got a warning from the model, though, that tells us that the maximum number of iterations has been reached. This is part of the algorithm for learning the model and tells us that we should increase the number of iterations. Increasing the number of iterations only increased the training set performance, not the generalization performance. Still, the model is performing quite well. As there is some gap between the training and the test performance, we might try to decrease the model's complexity to get better generalization performance

REFERENCES

- [1] Lee E.Y., Ng M.Y., Khong P.L. COVID-19 pneumonia: what has CT taught us? *Lancet Infect. Dis.* 2020;20(4):384–385. [PMC free article] [PubMed] [Google Scholar]
- [2] Bernheim A., Mei X. Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. *Radiology.* 2020 doi: 10.1148/radiol.2020200463. In press. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
- [3] Pan F., Ye T. Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. *Radiology.* 2020 doi: 10.1148/radiol.2020200370. In press. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
- [4] Abdel-Hamid O, Mohamed AR, Jiang H, Penn G (2012) Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. *ICASSP, IEEE Int Conf Acoust Speech Signal Process - Proc* 4277–4280. doi: 10.1007/978-3-319-96145-3_2
- [5] Abdeljaber O, Avci O, Kiranyaz S, et al (2017) Real-time vibration-based structural damage detection using one-dimensional