DEEP LEARNING IN HEALTHCARE

Jyotsna Doonga

Date: 24-09-2022

Abstract

Deep learning is a sub discipline of machine learning and artificial intelligence that uses a machine learning technique called artificial neural networks to extract patterns and make predictions from large data sets. The increasing adoption of deep learning across healthcare domains together with the availability of highly characterised cancer datasets has accelerated research into the utility of deep learning in the analysis of the complex biology of cancer. While early results are promising, this is a rapidly evolving field with new knowledge emerging in both cancer biology and deep learning. In this review, if we focus on the deep learning applications for omics data types, including genomic, methylation and transcriptomic data, as well as histopathology-based genomic inference, and provide perspectives on how the different data types can be integrated to develop decision support tools. Deep Learning has helped to determine the best models for the cancer diagnosis and prognosis prediction tasks. Deep Learning is a generic model requiring minimal data manipulations and achieves better results while working with enormous volumes of data.

Problem statement

The problem statement is to apply. So the main objective is to use Artificial Neural Networks, which will not only give faster results but also demonstrate higher accuracy in Breast Cancer Prediction process.

Customer (Patient) Need Assessment

The patient should be able to get proper accurate reports and the error in prediction through the deep learning model should be minimalized. This can be done using Neural Networks i.e Deep Learning for faster and more accurate results This will reduce the time and will help the healthcare sector in analysing the results, reduce cost and save time added with giving accurate results reducing the overall error in reports.

Target Specification and characterization

A. To change the traditional pathology cancer detection process to an accurate process.

B. To identify and provide treatment in initial stage accurately.

C. Reducing frustration and death of patients due to delay in the prognosis process.

D. Predetermined dataset of cancer patients and normal patients is taken and based on that prediction is performed.

External Search (information and characterization)

Deep Learning is a branch of Artificial Intelligence that employs a variety of statistical, probabilistic and optimization techniques that allows computers to learn from past examples and to detect hard-to-discern patterns from large, noisy or complex data set. These datasets can be found in open-source platforms. I have taken this data from "UCI ML Breast Cancer Wisconsin (Diagnostic) dataset". Features are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe characteristics of the cell nuclei present in the image.

https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)

This data set characteristics are as below:

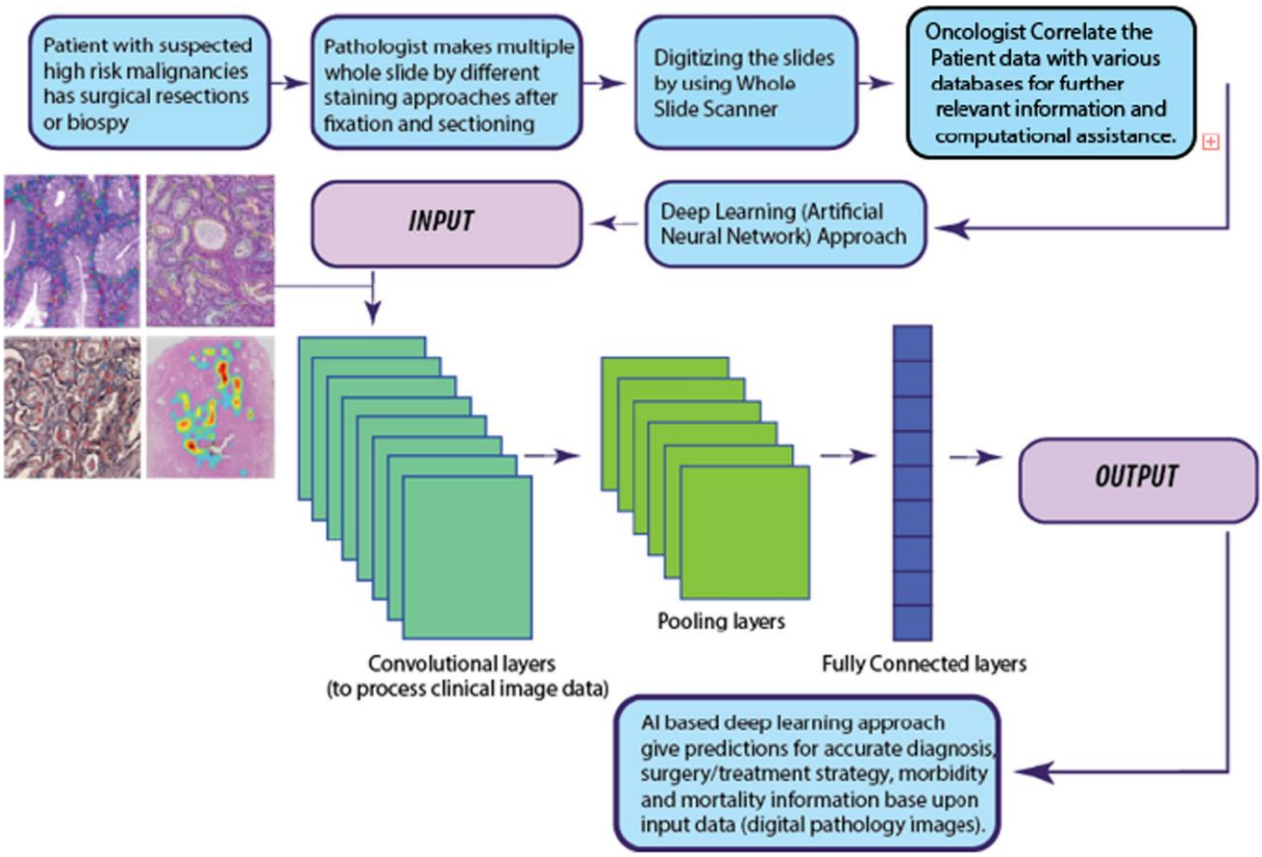
```
:Number of Instances: 569

:Number of Attributes: 30 numeric, predictive attributes and the class

:Attribute Information:
    - radius (mean of distances from center to points on the perimeter)
    - texture (standard deviation of gray-scale values)
    - perimeter
    - area
    - smoothness (local variation in radius lengths)
    - compactness (perimeter^2 / area - 1.0)
    - concavity (severity of concave portions of the contour)
    - concave points (number of concave portions of the contour)
    - symmetry
    - fractal dimension ("coastline approximation" - 1)

    The mean, standard error, and "worst" or largest (mean of the three
    largest values) of these features were computed for each image,
    resulting in 30 features.  For instance, field 3 is Mean Radius, field
    13 is Radius SE, field 23 is Worst Radius.

    - class:
            - WDBC-Malignant
            - WDBC-Benign
```

Information and Bench marking alternate products

Cancer is a complex and multifaceted disorder with thousands of genetic and epigenetic variations. AI-based algorithms hold great promise to pave the way to identify these genetic mutations and aberrant protein interactions at a very early stage. Modern biomedical research is also focused to bring AI technology to the clinics safely and ethically. AI-based assistance to pathologists and physicians could be the great leap forward towards prediction for disease risk, diagnosis, prognosis, and treatments. Clinical applications of AI and Machine Learning (ML) in cancer diagnosis and treatment are the future of medical guidance towards faster mapping of a new

treatment for every individual. By using AI base system approach, researchers can collaborate in real-time and share knowledge digitally to potentially heal millions.

Research on clinical oncology is now more focused to decode the molecular onset of cancer by understanding the complex biological architecture of cancer cell proliferation. It also focused to process the millions of relevant cases in big data and computational biology to tackle the current scenario of expanding number of cancer mortalities in the globe [6]. Moreover, the use of AI in clinical decision-making is believed to increase the chances of early disease prediction and diagnosis by NGS sequencing and high-resolution imaging techniques. It would also lead to introduce novel biomarkers for cancer diagnosis, designing novel personalized drugs, and delivery of potential treatment strategies by generating significant datasets and using specialized bioinformatics tools.



Applicable patents

- https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-021-00968-
  Deep learning in cancer diagnosis, prognosis and treatment selection
  Khoa A. Tran, Olga Kondrashova, Andrew Bradley, Elizabeth D. Williams, John V. Pearson & Nicola Waddell

- https://pubmed.ncbi.nlm.nih.gov/36119479/
  Applying artificial intelligence technology to assist with breast cancer diagnosis and prognosis prediction
  Meredith A Jones 1, Warid Islam 2, Rozwat Faiz 2, Xuxin Chen 2, Bin Zheng 2

Applicable Regulations (Government and environment)

- Patents on ML algorithms developed
- Laws related to privacy for collecting data from users
- Ensuring open-source, academic and research community for an audit of Algorithms.
- Review of existing work authority regulations.

Applicable Constraints

- Requires a whole lot of research to gather universal historical dataset in order to train available data
- Confidential health data to be obtained to train the model.
- In addition, it also requires a good deep knowledge to make models based on GANs that will provide us with better accuracy.
- Implementing the project by replacing the traditional method
- Continuous integration and development of the models

Business opportunity

Doctors are pretty good in diagnosing cancer while they are not so good in the prognosis of cancer. It takes more than two weeks to identify cancer in an individual. To overcome this hazardous circumstance, our main objective is to use Deep Learning, a subset of Machine Learning, which not only gives faster results but also demonstrates higher accuracy in the Cancer prediction process.

Concept Generation

Most types of cancer have four stages: stage I (1) to IV (4). Some cancers also have a stage 0 (zero). Here is a general description of cancer stage groupings.

- *Stage 0*: This stage describes cancer in situ. In situ means "in place." Stage 0 cancers are still located in the place they started. They have not spread to nearby tissues. This stage of cancer is often curable. Surgery can usually remove the entire tumor.

- *Stage I*: This stage is usually a cancer that has not grown deeply into nearby tissues. It also has not spread to the lymph nodes or other parts of the body. It is often called early-stage cancer.

- *Stage II and Stage III*: In general, these two stages are cancers that have grown more deeply into nearby tissue. They may have also spread to lymph nodes but not to other parts of the body.

- *Stage IV*: This stage means that the cancer has spread to other organs or parts of the body. It may be also called advanced or metastatic cancer.

So, for solving this issue, to generate a model, we use Deep Learning.

Concept Development

Initially we will analyze and perform Exploratory data analysis on the past data. Then we will design a neural network and complete the model and train the data. When the new data comes into the model, then the model will be able to predict cancer. Then we evaluate the model, check for loss and evaluate it by using classification_report and confusion_matrix.

Product Prototype

Model Development: A lot of manual supervised machine learning models can be used to perform the prediction forecasting.

1. Data collection and data pre-processing must be done.
2. Feature selection and feature engineering techniques should be performed.
3. Perform Exploratory Data Analysis to realize the dependent and independent features.
4. Algorithm training and optimization must be done to minimize overfitting of the model and hyperparameter tuning.
5. Python language and ML models is used for the integration of this product.

Product Details : Algorithms, frameworks, software needed

This product uses deep learning, machine learning algorithms. This requires python language for implementation. Here we have used Keras API for Tensorflow 2.0 .Visualizations can be done using some libraries like NumPy, pandas, matplotlib, seaborn, plotly, etc. Visual Studio Code and Google Colab is used as the editor.

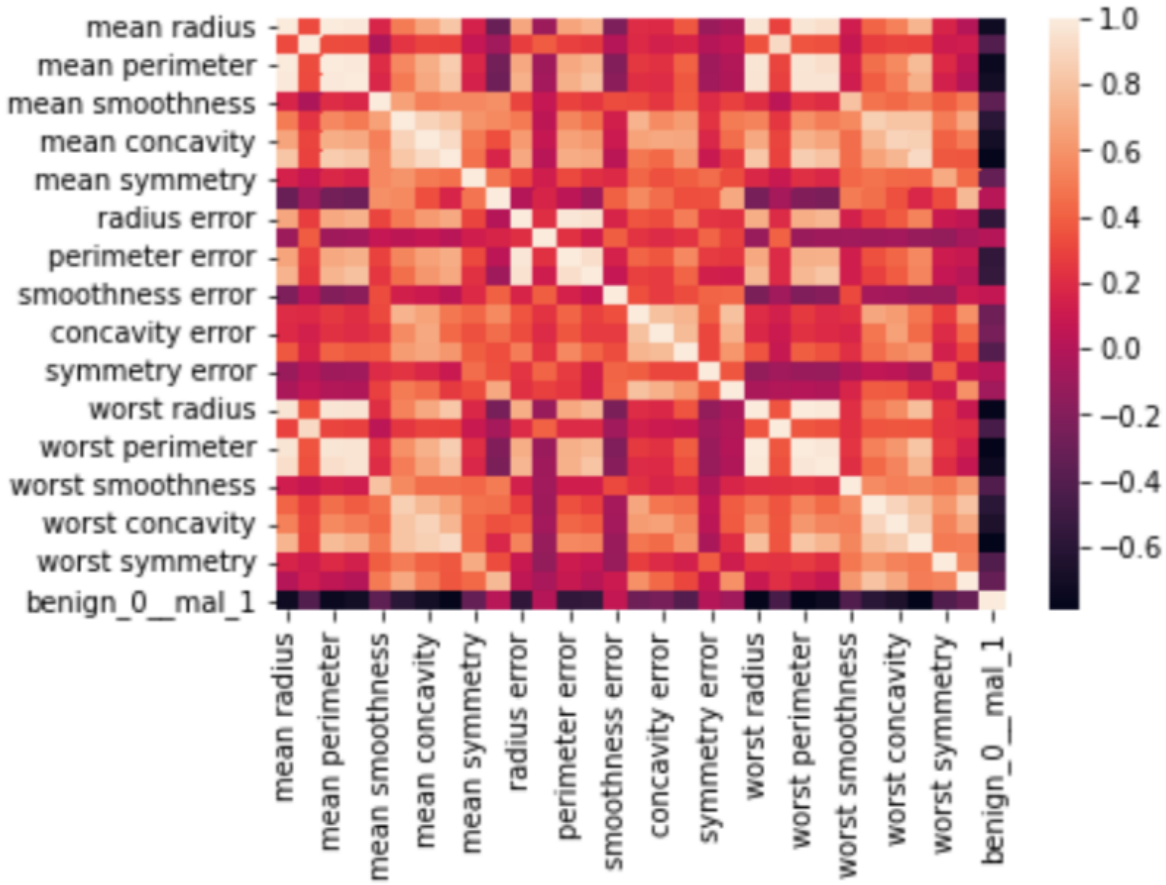
Code Implementation

Link to Code: https://github.com/jyotsnad246/Feynn-Labs/blob/main/Task-0/Cancer_Classification.ipynb

Visualizations:

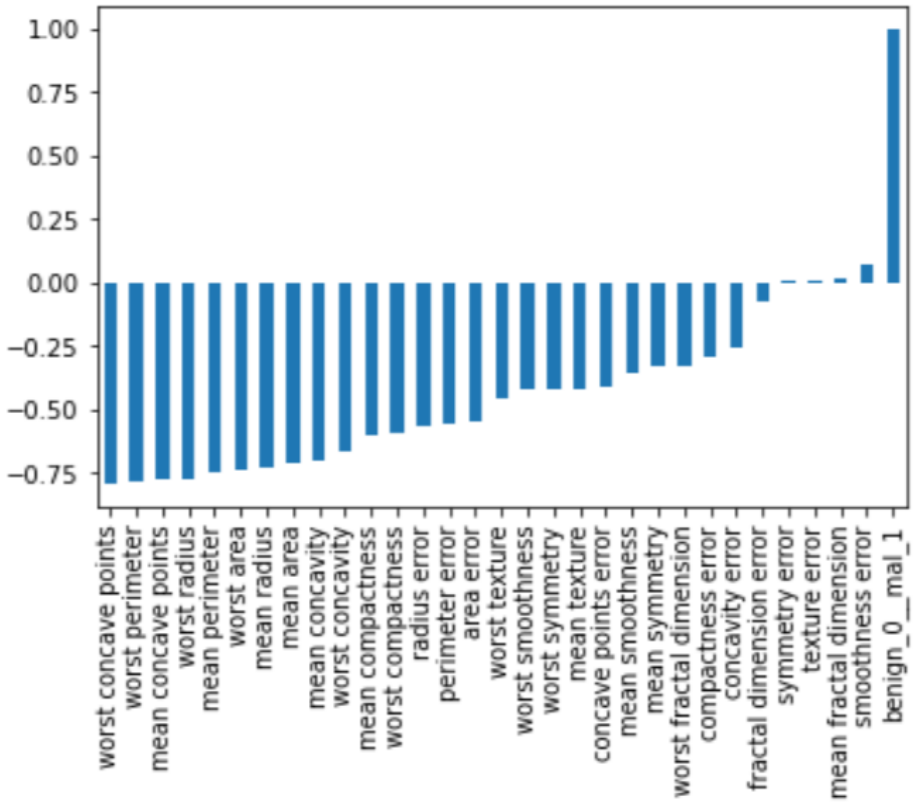
```
sns.heatmap(df.corr())
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fa01e831ed0>
```

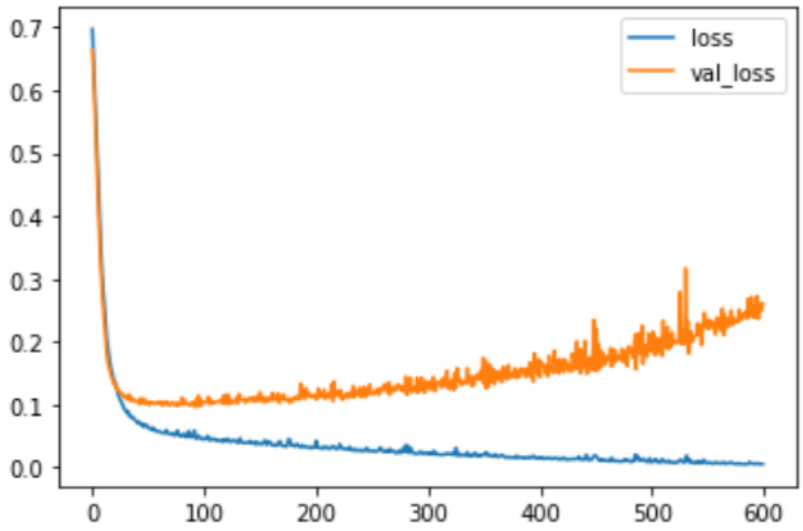


```
df.corr()['benign_0__mal_1'].sort_values().plot(kind='bar')
```

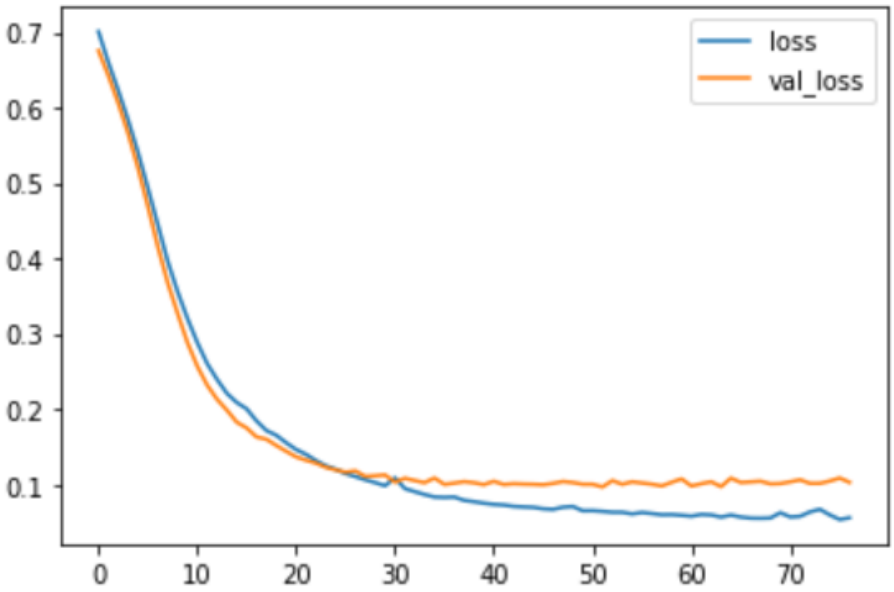
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fa01b14f3d0>
```



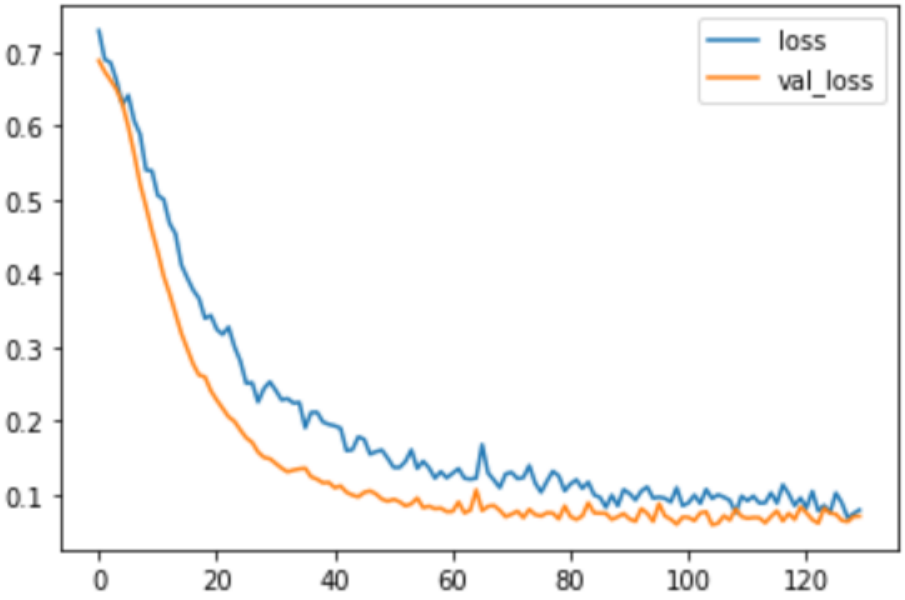
After training the model we notice that we have trained it way too much as it shows in loss plot



So we use early stopping to track val_loss and stop training once it begins increasing too much



This helps us in not overfitting the model. Adding of dropout layers also helped in reducing loss



Conclusion

AI is changing the way we are. So, we have to move forward to the advanced options to keep our pace in the world. Many healthcare organizations are switching to the AI module for their applications. With the right tools, software and programs, you can develop better models that will benefit healthcare sector such as for predicting Cancer and Tumour in suspected patients with much better accuracy.

Hence, I have explained the prototype and concept for a breast cancer prediction model using Artificial Neural Networks. It can further be improved by using Convolutional Neural Networks and Transfer Learning It has a lot of scope for the current world. This will definitely create a great opportunity to improve the healthcare industry. The implementation in the actual real-world scenarios will bring a huge impact in this sector.

References

1. https://www.nature.com/articles/s41568-020-00327-9
2. https://www.hindawi.com/journals/cmmm/2021/9025470/
3. W.N. Street, W.H. Wolberg and O.L. Mangasarian. Nuclear feature extraction for breast tumor diagnosis. IS&T/SPIE 1993 International Symposium on Electronic Imaging: Science and Technology, volume 1905, pages 861-870, San Jose, CA, 1993.
4. K. P. Bennett and O. L. Mangasarian: "Robust Linear Programming Discrimination of Two Linearly Inseparable Sets", Optimization Methods and Software 1, 1992, 23-34
5. https://cancerci.biomedcentral.com/articles/10.1186/s12935-021-01981-1/figures/2
6. https://www.udemy.com/course/complete-tensorflow-2-and-keras-deep-learning-bootcamp/
7. https://www.news-medical.net/news/20220812/AI-Model-That-Reduces-Cost-And-Time-And-Increasing-Accuracy-Of-Cancer-Diagnosis.aspx