Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 131 (2018) 866–872

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Congress of Information and Communication Technology (ICICT-2018)

Evaluating The Performance Of Deep Neural Networks For Health Decision Making

Emelia Opoku Aboagye*, Gee C.James, Rajesh Kumar

UESTC, P.O. Box 611731, Chengdu, China

Abstract

In this study, ResNet and Google Image Nets models which are both instances of Deep Neural networks are used to detect cancer cells and subsequently make predictions. Deep neural Networks (DNNs) have yielded massive accomplishment in Recommendation Systems, The deep neural models; Google Image Nets and ResNet are investigated and tested. Notwithstanding, there is generally few works on utilizing DNNs for proposal in the health sector considering the huge measure of patients' information in delicate zones like the health sector. The motivation behind this paper is to make prediction recommendations on cancer patients using deep learning models. Two types of datasets were used to train and test the data. Two Deep Neural Network models; RESNET and Google Image Nets were embraced. We show a lingering learning system to facilitate the preparation of systems that are considerably more profound than those utilized already.

Keywords: Health, preprocessing; deep learning; recommendation system.

1. Introduction

Recommender frameworks can be comprehensively categorized into collaborative filtering (CF), content-based, and hybrid methodologies. Matrix factorization techniques are deemed to be currently CF state-of- art frameworks for recommendation prediction. Content-based methods are based upon Item characteristics, and suggestions are founded on likeness amid such traits. .Last but not the least is the fact that, hybrid frameworks put together both item content and item-user response (Luo, 2017) .Deep learning algorithms possess the ability to gain knowledge from different environments, it is very sturdy and advances performance by becoming accustomed to the

* Corresponding author. Tel +8618628333154.
E-mail address: eoaboagye@yahoo.co.uk

1877-0509 © 2018 The Authors. Published by Elsevier Ltd.

10.1016/j.procs.2018.04.288

transformations in the surroundings

Deep learning have additionally been observed to be fitting for big data analytics with several implementations to natural language processing computer vision, pattern recognition, speech recognition systems, pattern recognition just to mention a few (Shen et al., 2016). Deep neural network techniques have the ability to learn from different environments, it is very sturdy and improves performance by adapting to the changes in the environment but difficult to train to prevent overfitting (Ansong, Huang, Yeboah, Dun, & Yao, 2015). The idea of profound knowledge came about through the study on (Ahmad, Ahmad, & Dar, 2016). To create a regular neural network (NN), activations of neurons are imperative to fine-tune the weights in order for the NNs to meet expectations. However, depending on the problems, the process of training a NN may cause computational complexities at the various stages. Backpropagation is an effectual gradient descent algorithm which has played a significant function which has played an important role in NNs since 1980 [5]. The ANNs is trained with a controlled learning method. Performance of backpropagation when harnessed to testing statistics might be unsatisfactory, even though the training accuracy is high as its basis is on local gradient data with an unsystematic primary point, the algorithm time and again gets cornered in local optima. Besides, the size of the training information if not adequately large, NNs will encounter the difficulty of outliers [6]. The unique characteristic of DNNs lies in its' efficient operation large scale data processing and its ability to generalize the outcome that accounts to its application in images where the data is massive. Since the 2012 ImageNet competition [7] achieved by Alex Net, which is the initial entry that used a Deep Neural Network, more than a few additional DNNs with rising intricacy have been investigated to the challenge so as to attain improved outcome. During the ImageNet categorization contest, the eventual aim is to gain the topmost precision in a multi-class grouping setback structure, irrespective of the definite presumption time (He, Zhang, Ren, & Sun, 2016). This has brought about several problems in the field. This paper in response to the challenge aspires to judge against most acceptable DNN functions. Selection algorithms are considered as having better-quality as compared to backward schemes as regards computational effectiveness for cancer prediction. This is because such processes have quite a lot of key shortcomings such as being computationally too costly or occasionally not possible to put into practice. There is comparatively minute effort on utilizing DNNs for suggestions in the health sector. Others have used backward methods but we think such methods have several major disadvantages such as being computationally too expensive or sometimes impossible to implement. We believe that, there is relatively little work on employing DNNs for recommendation in the health sector.

This study deals with the earlier mentioned research problems by formalizing a neural network modelling strategy for cancer collaborative filtering. We treat the image records as implicit feedback of patients' observed records.

1.1. Motivation

Cancer prediction is vital as the survival of the patient depends on timely interventions. This presupposes that, the prediction of such phenomenon requires fast and efficient models, and researching into such situations are worthwhile. It is well noting that, currently, academic research is skewing from system analysis of small-scale data processing. Deep Multi-Layer Perceptron with enhanced optimization characteristics that is simple and efficient to enhance such operations is critical.

1.2. Contribution

In this paper,

- We propose a criteria for data preparation
- We use the sub-models of RESNET and Google net for cancer detection based on…
- We analyse and detected unknown or new type of cancers successfully

The performance of ResNet and Google ImageNet are compared in terms of prediction accuracy and timeliness

1.3. Problem Formulation

We treat cancer prediction as an emergency issue and so formulate our problem with some initial considerations; Several important management methods have been suggested to lessen the restrictions (Ramachandra, McGrew, Baxter, Howard, & Elmslie, 2013) . in case of cancer diagnosis, let R be the probability of cancer cells growing, then, the matrix N would be a ected by (1-R) where R is any random value within the prediction range. If A is the matrix whose constituents show the degree of destruction to the cells caused by the disease and is defined as;

A=(1-R)N (1)

Eventually, the damaged disease cells would result in the failure of the affected cells and other life organs and matrix M is brought in to demonstrate the signal condition of the cancer cells employed as trendy produced relationship matrix

$$M_{k,t} = \begin{cases} 1, z_{i,j} \le \alpha_L \\ 0, y, t \ge \alpha_H \\ \dfrac{\alpha_L}{\alpha_{k,t}}, \alpha_L < y_{k,t} < \alpha H \end{cases} \tag{2}$$

Where α_L and α_H respectively depicts the lower and higher cancer regions. To discover the matrix with the utmost continued existence probability, a fresh array of matrix is described subsequent to the destruction. To get the most out of prediction rate following the harm initiated by the cancer cells, the anticipation matrix t is instituted whose components signify the position of expectation of survival If t is the anticipated matrix, and is the risky matrix, then, $Er = (C_\rho - \phi_?)\,t$ where matrix N and F as already defined above, is the overall quantity of protected cells on hand for is the total number of safe cells available for treatment and $\phi_?$ is the possibility of added inaccuracy which the model possibly will place in.

2. Proposed Deep Neural Cancer Prediction Model

For a real neural two-way filtering modelling, we take up a multi-layer algorithm such as (Parkhi, Vedaldi, Zisserman, & others, 2015) to model a patient–image interaction c_{pi} , here, the produce of an individual stratum becomes the effort of the layer after that which is above it. The underlying input stratum comprises double dimensional vectors V_{p^P} and V_{i^I} that explains patient p and image i on top of the input stratum lies the implanting stratum which is a whole connected stratum. The dimensional feature of the very last concealed stratum Y establishes the model's efficacy. The last output stratum depicts the envisaged grade. We at the moment put together the NCF's prognostic model being referred to as cancer deep neural collaborative filtering layers that maps the latent vectors to the prediction scores. Each layer of the deep neural CF layers is customized to reveal the latent structures of user–item relationships. The dimensional feature of the last hidden layer Y determines the model's efficacy. The last output. layer depicts the predicted score $\widehat{C}_{pi}$, and training is implemented by minimizing the pointwise loss between $\widehat{C}_{pi}$ and its target value C_{pi} . We now formulate the NCF's predictive model as;

$$c_{ni} = f(\mathrm{A}^I \mathrm{v}_n^P, \mathrm{B}^I \mathrm{v}_i^I \mid \mathrm{A}, \mathrm{B}, \theta \mathrm{f}) \tag{3}$$

$A \in R^{M \times K}$ and $B \in R^{N \times K}$ signifying the latent matrix for patients and images, respectively; and θ represents the model parameters of the interaction function f. Since the function f is since function f is described as a multi-layer neural network, it can be formulated as;

$$f(\mathrm{A}^T \mathrm{v}_p^P, \mathrm{B}^T \mathrm{v}_i^I = \theta_{out}(\theta_y(...\theta_2(\theta_1(\mathrm{A}^T \mathrm{v}_p^P, \mathrm{B}^T \mathrm{v}_i^I))...)) \quad (4)$$

Where θ_{out} and θ_v stands for the interaction function for the final phase and a neural collaborative filtering phase, where Y neural CF phases in .Length is available.

2.1. Data Preparation and Setup

Real-world datasets are often incomplete that is, attribute values will be missing; noisy, have errors or outliers, and inconsistent, displaying a lot of discrepancies between the collected data. The unclean data can confuse the mining procedures and lead to unreliable and invalid outputs. Also, performing complex analysis and mining on a huge amount of such soiled data may take a very long time. Pre-processing and cleaning should improve the quality of data and mining results by enhancing the actual mining process. In this paper, we ensured that, the image dataset has a fixed characteristic vector size. Our objective is that, our images are preprocessed and scaled to have identical widths and heights using rescaling and ranking techniques, stretching from extra sophisticated processes that value the aspect proportion of the initial picture to the resized picture. Uncomplicated techniques that pay no attention to the aspect proportion and compress the distance across and height to the necessary measurements. Exactly which method one uses really depends on the complexity of your factors of variation. In some cases, ignoring the aspect ratio works just fine; in other cases, one'll want to preserve the aspect ratio. In this paper, we started with the basic solution of building an image preprocessor that resizes the image, ignoring the aspect ratio. We optionally pass in a list of image preprocessors that can be sequentially applied to the given input image. Specifying these preprocessors as a list rather than a single value is important: we first resized the image to a fixed size, then performed scaling (mean abstraction). The image array was converted to a format suitable for processing. Each of these preprocessors were novely implemented in an independent manner, allowing for sequential arrangements to the cancer images in an efficient manner. Based on this hierarchical directory structure, the datasets were freed from outliers and organized for efficient implementation of the neural networks

*2.2. Collection of Dataset*s

The cancer dataset is collected from two different sources one from Chengdu Sichuan Cancer Hospital, Chengdu, China. We collected 2000 patients' data that includes many samples which is shown in Figure1. This dataset comprises breast cancer, Chest cancer and other types of cancers. We spilt dataset into two parts one is testing and other for training. We spilt dataset into 20% for validation and 80% test for training.

3. Results and Discussion

We run our experiments on natural reliable real data sets from the Sichuan Hospital in Chengdu. The images contained the chest and other parts of patients diagnosed with cancer for a period of 10years. We adopted the two Deep Neural network models' ResNet and Google Net. It is observed that, Google net model has tremendous performance in terms of running time ref (table1), but performed poorly in terms of prediction and validation accuracies. We are tempted to believe that, if the data sets were larger than we have experimented, Google net would have performed better. However, ResNet (18, 50.101, and 152) really performed better in terms of prediction accuracy for recommendation (table1). It however performed poorly in terms of running time. We therefore conclude our study by proposing ResNet for image classification and prediction tasks and deduce that, ResNet is a good recommendation predictor for the survivability of cancer patients. The lower the Loss, the better a model. the deficiency is estimated on training and justification and its interpretation is dependent on whether the model is doing well or not for these two sets. Loss is not in proportion as an opposite of precision and it is a rundown of the miscalculations created for every case in point in training or justification sets. Loss is not in percentage as opposed to accuracy and it is a summation of the errors made for each example in training or validation sets. In terms of model loss 18layer ResNet model had the least loss though performed badly in terms of prediction accuracy.

Table1 Model Performance

Table I
MODEL PERFOMANCE

Models	Model Loss	Accuracy%	Prediction Time
Google ImageNet	N.A	0.95	5201secs
ResNet 18 layers	0.86	0.84	2131secs
ResNet 50 layers	4.16	0.98	1844secs
ResNet 101 layers	8.16	0.98	13964secs
ResNet 152 layers	11.56	0.98	1131secs

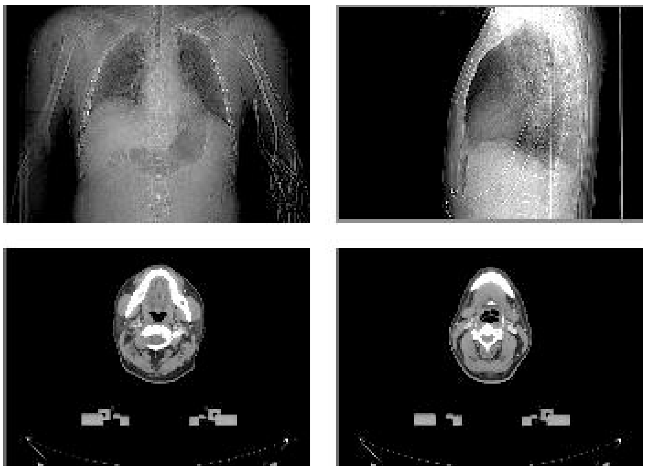

Fig 1. cancer datasets

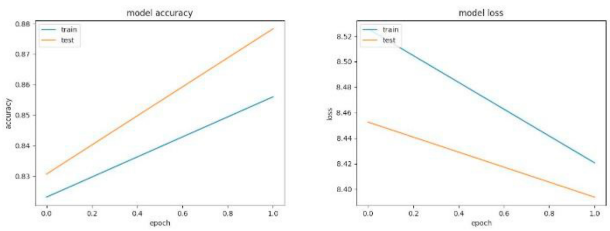

Fig 2. 18 Layers Model Accuracy

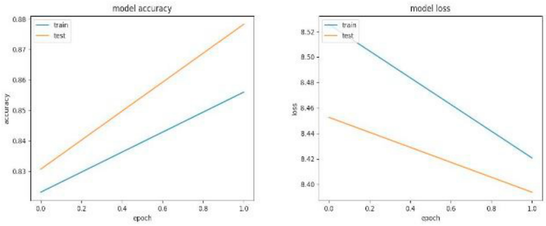

Fig 3. 50 Layers Model Accuracy

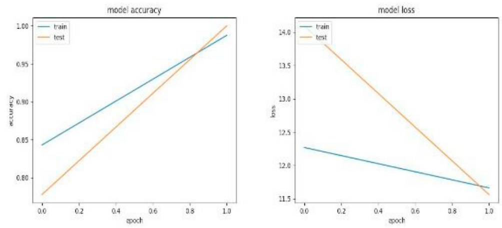

Fig 4. 101 Layers Model Accuracy

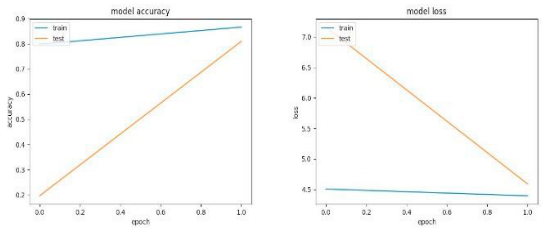

Fig 5. 152 Layers Model Accuracy

4. Conclusion

We have performed both experimental and theoretical study of neural networks for recommendation prediction of the cancer disease**.** With the development of big data analysis, deep learning have been used for scenarios where massive amounts of unsupervised data are involved. As an efficient tool for big data analysis, the deep learning technique have achieved great success with huge amounts of unlabeled training data. However, when only a limited amount of training data is available, more powerful models are required to achieve an enhanced learning ability. It is therefore of great significance to consider how to design deep models to learn from fewer training data, especially for speech and visual recognition systems. This was evident in our experiment with Google ImageNets. Uses of optimization algorithms to adjust the network parameters:. In DNNs, a large number of parameters need to be adjusted. Moreover, with an increasing number of hidden nodes, the algorithm is more likely get trapped in the local optimum. Optimization techniques, such as the PSO are therefore required to avoid this problem.

The deficiency is estimated on training and justification and its interpretation is dependent on whether the model is doing well or not for these two sets. The proposed training algorithm should be able to extract the features automatically and reduce the loss of information so as to mitigate both the curse of dimensionality (Vervliet, Debals, Sorber, & De Lathauwer, 2014)(Patel et al., 2017) and the local optimum

5. Acknowledgements

We express our profound gratitude to the Applied Basic Research programs of Sichuan Province (2015JY0043), the Fundamental Research Funds for the Central Universities (ZYGX2015J154, ZYGX2016J152,, ZYGX2016J170), Research and Development projects of high and new technology development and Industrialization of Sichuan Province(2017GZ0007Key

References

1. Ahmad, F., Ahmad, I., & Dar, W. M. (2016). Identification and classification of voxels of human brain for rewardless-related decision making using ANN technique. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-016-2413-6
2. Ansong, M. O., Huang, J. S., Yeboah, M. A., Dun, H., & Yao, H. (2015). Non-Gaussian Hybrid Transfer Functions : Memorizing Mine Survivability Calculations, *2015*.
3. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90
4. Luo, J. (2017). A Hybrid Recommendation Algorithm with LDA and SVD++ Considering the News Timeliness. *Revista de La Facultad de Ingenier{'\i}a*, *31*(12).
5. Parkhi, O. M., Vedaldi, A., Zisserman, A., & others. (2015). Deep Face Recognition. In *BMVC* (Vol. 1, p. 6).
6. Patel, D., Thompson, M. D., Manna, S. K., Krausz, K. W., Zhang, L., Nilbuol, N., … Kebebew, E. (2017). Unique and novel urinary metabolomic features in malignant versus benign adrenal neoplasms. *Clinical Cancer Research*, clincanres--3156.
7. Ramachandra, R., McGrew, S. Y., Baxter, J. C., Howard, J. R., & Elmslie, K. S. (2013). NaV 1.8 channels are expressed in large, as well as small, diameter sensory afferent neurons. *Channels (Austin, Tex.)*, *7*(1), 34–7. https://doi.org/10.4161/chan.22445
8. Shen, Y., Jin, R., Chen, J., He, X., Gao, J., & Deng, L. (2016). A Deep Embedding Model for Co-occurrence Learning. In *Proceedings - 15th IEEE International Conference on Data Mining Workshop, ICDMW 2015*. https://doi.org/10.1109/ICDMW.2015.32
9. Vervliet, N., Debals, O., Sorber, L., & De Lathauwer, L. (2014). Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis. *IEEE Signal Processing Magazine*, *31*(5), 71–79. https://doi.org/10.1109/MSP.2014.2329429