Generative Adversarial Networks in Disease Gene Drug Relationships

Dr. S. Vijaya*

1Assistant Professor, Department of Information Technology, KG College of Arts and Science, Coimbatore, India

*E-mail: s.vijaya@kgcas.com

Abstract: The swift growth in the form of digital information stored in the biomedical databases in this digital era has activated a prototype shift in the models in the Deep learning approaches which have used in several contests in Machine Learning and in the domain of pattern recognition. Finding relationship between entities like genes, diseases, proteins and drugs is tedious task due the ambiguity in the terms used in biomedical domain. Treating cancer with the drug based on the gene that is associated with the disease increases the survival rate. Hence, the deep learning method “Generative Adversarial Networks” is proposed to find the relationship between Genes, Diseases and Drugs from biomedical abstracts in this work.

Keywords: Biomedical database, Genes, Diseases, Drugs, Deep learning, Generative Adversarial Networks

1. Introduction

Deep learning method is one of a sub part of machine learning methods which try to design complex abstractions in huge amount of dataset. Deep learning approaches use various layers for processing with composite structures. The approaches also poised of several non-linear transformations accomplished to handle huge amount of multidimensional data[1] [2-7]. When the data set is huge in size with multiple dimensions Deep learning methods are used to handle such data.

Major benefit of using this deep learning efficient algorithm is, it replaces handcrafted features and make better representations starting large-scale and unlabeled data as well as it creates the models to learn those representations. Different approaches of deep learning domain such as convolution neural networks, deep belief networks, reinforcement neural networks and Classic neural networks that have been useful in many fields like Relation Extraction, Natural language processing(NLP), Audio, Text recognition and bioinformatics modern results have been obtained on diverse tasks.

Deep Learning domain is considered one of the prominent field to train, analyze and generate patterns from large amounts of data. Tai Sing Lee and David Mumford [8] proposed a hierarchical Bayesian inference framework to be used in the visual cortex. They have used recurrent feed forward /feedback loop concept in visual cortex. They have recommended the methods of particle filtering and Bayesian-belief propagation for top-down and bottom-up observations.

A language modeling (LM) is a ability that catches the notable salient statistical uniqueness of the successions of terms in natural language [9] used temporally factored Restricted Boltzmann Machine Language Model for the words being predicted when compared with the context words.
Neural network is used to get promising results in speech recognition since many years. Mohamed et al. (2009) \[10\] used a five layer DBN as an alternate of the Gaussian mixture component of the GMM-HMM. Geoffrey E. Hinton et al. (2006) \[11\] presented “complementary priors” method to overcome the complexity of the effects that make explaining implication complicated in convolution nets with various hidden layers. They also have derived greedy algorithm that is fast and fine-tunes the weights with the help of a new version of the novice algorithm with slower learning procedure. From the studies of various research papers it is decided that deep learning methods can be effectively used to handle multidimensional data especially on the biomedical dataset. Hence, the deep learning method Generative Adversarial Networks has been proposed to identify the relationship between gene-disease-drug from biomedical abstracts.

2. Related Works

Deep learning models have a capability to produce appropriate results without any human intervention and it makes the results very promising for solving real time problems. A modified Deep Belief Networks is introduced by Nair and Hinton (2009) \[12\] where they have exposed the deployment of a third-order Boltzmann machine by top-layer. By using this method the error rate is reported extremely close to best result available on this task and the same was proved the better accuracy of DBN when compared with the other models like SVMs.

Tang and Eliasmith (2010)\[13\] proposed two different method to improve the results obtained by DBN. In the area of speech recognition, a new ensemble NN-HMM that is based on CNN has been proposed by Abdel-Hamid Ossama et al. (2012) \[14\]. To adjust speaker, they suggested limited filtering as well as the max pooling method in the frequency field. Their experiment results confirmed 10% reduction in relative error when CNN method is used.

Kaiming He et al. (2015) \[15\] developed a framework using residual learning to provide easiness to train deeper networks. Even though there is a presence of learning unreferenced functions, the layer is reformulated such as function of learning residual in position to the various layer inputs. They have got very promising results while evaluating residual nets on the ImageNet dataset with a have used depth of up to 152 layers and from the results obtained it is concluded that the method achieved 3.57% error on the ImageNet test set.

A new method GO using ‘value networks’ is used by David Silver et al. (2016) \[16\] to assess the positions on the slat and ‘policy network’ to decide moves. With the introduction of new search algorithm AlphaGo which is a combination of Monte Carlo simulation and policy network they got 99.8% winning chances when compared with Go programs. Yuxi Dong et al. (2017) \[17\] proposed CNN approach to find carotid plaques which may cause strokes.

Dirvanauskas et al. \[18\] planned with the images of embryo cell-human for 3 stages (1-cell or 2-cell or 3-cell) by a standard model known as GAN. Of these picture images which are synthesized that might be accustomed assist the event method in training the model, and analysis for the image process with taken embryo.

Zhang et al.\[19\] bestowed a new GAN-based approach to crack the information lack issues personally in the task of re-identification. For their work they took 2 read images (for example - cross read images) are framed by a novice conditional GAN in the existing original pictures and outline images. For person re-identification these pictures are sent as input to a discriminator.

Two variations of GAN models are utilized by Frid-Adar et al.\[20\] to come up with artificial liver lesions. From the experiment results they over that their model has improved artificial classification accuracy changed from 77.5 to 85.0% when it is compared to the classic knowledge extension. Rather than generating 2Dim HR MRIs, the authors Chen et al. \[21\] planned a tomography (HR Tomography) 3Dim HR high-resolution MRI image creation design as 3D networks fetch additional computing necessities.

3. Comparison of Various Deep Learning Techniques

Various Deep Learning Techniques had been utilized in totally different domains. To search out the acceptable technique to spot gene-disease-drug relationship distinct techniques has been studied and also the comparison is listed within the following table 1.
| Method Name                      | Description                                                                 | Works best in                                                                 |
|---------------------------------|------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Classic Neural Networks          | This technique is known by its multilayer perceptrons.                       | Any table dataset that is in the form of rows and columns format such as CSV.  |
|                                 |                                                                              | The Classification and Regression approaches with the real values.            |
|                                 |                                                                              | Any function with the very best suppleness, such as ANNS.                    |
| Convolutional Neural Networks    | It is a most sophisticated and high-tension variety of typical Artificial Neural Networks. | OCR – image dataset analysis.                                                 |
|                                 |                                                                              | Two – dimensional matrix format data file that might be additional remodeled to the format of one-dimensional matrix for quicker analysis |
|                                 |                                                                              | The representation can be concerned with its design to obtain output.         |
| Recurrent -Neural Networks (RNNs)| The RNN technique puts the information collected from its earlier state as associate degree of input data value for this calculation. | A single input connected to an input such as the classification of Image.     |
|                                 |                                                                              | Sequence of input data yielding only one output, like the output of Sentiment Analysis. |
|                                 |                                                                              | Sequence of input data yielding sequence of output data, as the classification of video. |
| Generative Adversarial Networks | Combination of a Generator and a Discriminator                              | Generation of Image and Text process.                                         |
|                                 |                                                                              | Enhancing the quality of Image New Drug finding methods.                     |
| Self-Organizing Maps            | Operate in addition with an unsupervised knowledge to reduce an amount of random data generation in an exceedingly model. | When the input data don’t escort a coordinate axis label values               |
|                                 |                                                                              | Project details for testing the dataset construction                          |
|                                 |                                                                              | Creative comes in Music, Text, and also in videos with the assistance of Artificial Intelligence. |
| Boltzmann Machines              | There is no lead for any predefined model.                                  | System observation                                                            |
|                                 |                                                                              | Setting of a recommendation of binary platform.                               |
|                                 |                                                                              | Testing specific kind of datasets.                                            |
| Deep Reinforcement Learning     | wherever associate degree agent intercommunicate with the surrounding to modify its condition. The agent will monitor and do actions based on the modification. | Board Games like Poker, online Chess, Automated Self-Drive Cars                |
|                                 |                                                                              | Inventory Management                                                          |
| Autoencoders                    | operates mechanically supported its data, before captivating associate degree creation operate and final result value decoding | Feature recognition                                                          |
|                                 |                                                                              | Recommendation model can be constructed. Include options to enormous datasets |
Back propagation

It is named as the essential mechanism for this neural networks to find out regarding any bugs in predicting knowledge.

Data Debugging

Gradient Descent

gradient means a kind of slop which contains a quantifiable angle and also that can be portrayed into an association between entities.

Updating parameters to an exceedingly given model

4. Proposed Methodology

The proposed methodology of this work, is Generative Adversarial Network. To improve scalability to find out relationship from biomedical abstracts such as gene, disease and drug various, a study on different deep learning technologies has been done and supported the study GAN technology has been proposed in this work.

4.1 Generative Adversarial Network

In this methodology the Generator can be referred because the reverse method of Convolutional Neural Network (CNN). Generally, CNN reads in a picture and outputs an inventory of sophistication possibilities that indicate whether or not the image contains an individual, animal, or object.

![Generative Adversarial Network Methodology](image)

Figure 1 Generative Adversarial Network Methodology [22]

In the reverse CNN the 1-dimensional category vector can convert this data to an entirely complete machine-generated color image. From the category vector, it can be able to train the network to get any quite image. Really the generator may attempt to generate an image of an object. To form the looks of the item to seem smart and real object the technology Discriminator is employed.
**Figure 2** Generative Adversarial Networks

**Figure 3.** An example of the Generative Adversarial Network Model Architecture [22]
The technology discriminator is however, a standard CNN that’s trained to spot pictures of individuals, animals, objects. It compares the objects generated the Generator with a dataset of real pictures, and properly determine each real and generate smart image.

When the generator and therefore the technology discriminator used against each other, the generator produce a stream of fake pictures as result. The fake pictures then fed into the portion of discriminators. Then from the outputs discriminator can classify all the fakes made by the generator:

5. Proposed Principles of GAN

The vital Generative Adversarial Network approach consists of an input data vector, a generator, implicit perform with the technique discriminator perform. Each generator and the technique discriminator are implicit functions enforced by deep learning networks. To clarify basic principles of this methodology abstract mathematical language is employed for doing this work.

Assuming ‘x’ variable as the training sample, the data distribution $P_{data}$, then the mounted distribution is $P_{data}(x)$, the standard strategies suppose to facilitate the distribution $P_{data}(x)$ takes a mathematical Gaussian mixture-distribution and utilizes the utmost probability because of the resolution. The ensuring performance is a smaller amount once the model is complex. To beat the limitations, neural networks methods have been planned to outline the distribution $P_{g}(x)$. Here, generator is used during this neural network in addition with the parameter $\theta$ earns the variant ‘z’ from the sooner division and maps this to the prevailing sample with pseudo distribution using the neural network. Hence, the formulated information will be considered as $G(z)$. The input ‘z’ refers mathematical Gaussian noise. With reference to $\theta$, a straight forward input distribution is accustomed generate totally different and complicated distributions. The value of $P_{g}(x)$ is formulated by the generator and $P_{data}(x)$ the exact actual image distribution ought to take as same as possible. So, the objective is to search out the value of $G^*$ as below.

$$G^* = \arg \min DivG(P_{g}, P_{data})$$  

(1)

Authors Goodfellow et al. [26] have used the function called binary cross-entropy for outlining discriminator.

$$Loss = -(y \log(\hat{y}) + (1 - y)\log(1 - \hat{y}))$$  

(2)

Wherever the value of $\hat{y}$ is that the likelihood to the approach that predicts the example as a positive example, and the value of ‘y’ is that actual example label. If the tested label is positive example it assumes the worth as one(1), else the value assumed as zero (0).

The entire object perform for the discriminator can be calculated as:

$$V(G,D) = \mathbb{E}_{x \sim Pdata}[\log D(x)] + \mathbb{E}_{x \sim P_{g}}[\log(1 - D(x))]$$

(3)
By combining Equation (1) into (3), the target method performance of the fundamental model GAN is outlined by the Equation (4):

$$\min_G \max_D V(G, D) = \min_G \max_D E_{\tilde{x} \sim P_{\text{data}}} \log D(\tilde{x}) + E_{z \sim P_z} \log (1 - D(G(z)))$$

The GAN perform will be arrived by optimizing this equation. The technique discriminator maximizes $V(G, D)$ and generator perform practices to attenuate $V(G, D)$, by this it forms the minmax relationship. The parameters of the value $G(\theta_g)$ and the value $D(\theta_d)$ are perpetually updated.

6. Experimental Results

The classification performance of using various deep learning methods compared based on Precision, Recall and Accuracy measures. The sentences with gene, drug and disease names collected and trained using bi-grams and tri-grams. The results values obtained by Classical Neural Networks, Convolutional Neural Network, Recurrent Neural Networks and Generative Adversarial Networks are presented in Table 2 The graphical representation of Comparison results presented in Figure 5.

Table 2 Precision, Recall Accuracy of Deep learning methods

| Method                  | Precision | Recall | Accuracy |
|-------------------------|-----------|--------|----------|
| Classical Neural Networks | 90.9      | 83.33  | 83.33    |
| Convolutional Neural Networks | 86.95    | 83     | 80       |
| Recurrent Neural Networks | 85.71     | 81.81  | 78.79    |
| Generative Adversarial Networks | 96       | 92     | 90.32    |

Figure 5 Comparison Results

7. Conclusion

Bioinformatics is rising and an emerging multidisciplinary field that’s accustomed perceive and organize data associated with the unit biomolecular through a mix of disciplines like Computer Science, Mathematics, Biology and Statistics. Recognizing entities from the collection of biomedical documents could be a tedious task. Although Machine Learning strategies perform well on recognizing these biomedical entities and finding relationships between the entities the measurability drawback is there once immense volume of information with multi dimensional attributes are concerned. Hence, the Generative Adversarial Networks has been planned for finding relationship between gene-disease and drugs from the corpus collected from biomedical abstracts and for this work to overcome the measurability problem occurred within the Machine Learning techniques.
REFERENCES

[1] Ben-Bright Benuwa et al., " A Review of Deep Machine Learning", International Journal of Engineering Research in Africa, ISSN: 1663-4144, Vol. 24, pp 124-136, doi:10.4028/www.scientific.net/JERA.24.124, 2016 Trans Tech Publications, Switzerland

[2] D. Li and D. Yu, 'Deep Learning: Methods and Applications,' Foundations and Trends in Signal Processing, Now Publishers, 2014.

[3] Y. Bengio, "Learning deep architectures for AI," Foundations and Trends© in Machine Learning, vol. 2, pp. 1-127, 2009.

[4] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 35, pp. 1798-1828, 2013.

[5] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85-117, 2015.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436-444, 2015.

[7] I. Arel, D. C. Rose, and T. P. Karnowski, "Deep machine learning—a new frontier in artificial intelligence research [research frontier]," Computational Intelligence Magazine, IEEE, vol. 5, pp. 13-18, 2010.

8. T. S. L. a. D. Mumford, "Hierarchical Bayesian inference in the visual cortex," J. Opt. Soc. Am. A, vol. Vol. 20, July 2003.

9. A. M. a. G. Hinton, "Three new graphical models for statistical language modelling," ICML ’07 Proceedings of the 24th international conference on Machine learning, pp. 641-648, 20- 24 June 2007.

10. G. D. a. G. H. Abdel-rahman Mohamed, "Deep Belief Networks for phone recognition," in Proceedings of the NIPS Workshop on Deep Learning for Speech Recognition and Related Applications, 2009.

11. S. O. Y.-W. T. Geoffrey E. Hinton, "A Fast Learning Algorithm for Deep Belief Nets," Neural Computation, pp. Pages 1527-1554, July 2006.

12. V. N. a. G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," ICML’10 Proceedings of the 27th International Conference on International Conference on Machine Learning, pp. 807-814, June 2010.

13. Y. T. a. C. Elia smith, "Deep networks for robust visual recognition," in Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 2010.

14. M. H. J. G. P. Ossama Abdel-Hamid, "Applying Convolutional Neural Networks concepts to hybrid NN-HMM model for speech recognition," in 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012.

15. K. H. X. S. R. J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2015.

16. A. H. a. David Silver, "Mastering the game of Go with deep neural networks and tree search," in Nature 529(7587):484-489, 2016.

17. Y. P. J. Z. a. W. X. Y. Dong, "Learning to Read Chest X-Ray Images from 16000+ Examples Using CNN," in 2017 IEEE/ACM International Conference on Connected Health: Applications , Systems and Engineering Technologies (CHASE), Philadelphia, 2017.

18. Dirvanaukas D, Maskeliunas R, Raudonis V, Damasevicius R, Scherer R. HEMIGEN: human embryo image generator based on Generative Adversarial Networks. Sensors (Basel). (2019) 19:3578. doi: 10.3390/s19163578

19. Zhang C, Zhu L, Zhang S, Yu W. PAC-GAN: an effective pose augmentation scheme for unsupervised cross-view person re-identification. Neurocomputing. (2019) 387:22–39. doi: 10.1016/j.neucom.2019.12.094

20. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing. (2018) 321:321–31. doi: 10.1016/j.neucom.2018.09.013

21. Chen Y, Feng S, Christodoulou AG, Zhou Z, Li D. Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network. arXiv:1803.01417. (2018). doi: 10.1007/978-3-030-00928-1_11

22. https://machinelearningmastery.com/

23. Creswell A, White T, Dumoulin V, Kai A, Sengupta B, Bharath AA. Generative adversarial networks: an overview. IEEE Signal Proc Mag. (2018) 35:53-65. doi: 10.1109/MSP.2017.2765202.

24. Spall JC. A stochastic approximation technique for generating maximum likelihood parameter estimates. In: American Control Conference. St. Louis, MO (2009).

25. Guo Y, Yu L, Oerlemans A, Lao S, Song W, Lew MS. Deep learning for visual understanding: a review. Neurocomputing. (2016) 187:27–48. doi: 10.1016/j.neucom.2015.09.116.

26. Goodfellow IJ, Pouget-Abadie J, Mirza M, Bing X, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets. In: International Conference on Neural Information Processing Systems Montréal, QC (2014).