

Research/Technical Note

Tuning the Training of Neural Networks by Using the Perturbation Technique

Huseyin Murat Cekirge* 

Mechanical Engineering, The City College of New York, CUNY, New York, USA

Abstract

The calculation of biases and weights in neural network are being calculated or trained by using stochastic random descent method at the layers of the neural networks. For increasing efficiency and performance, the perturbation scheme is introduced for fine tuning of these calculations. It is aimed at introducing the perturbation techniques into training of artificial neural networks. Perturbation methods are for obtaining approximate solutions with a small parameter ϵ . The perturbation technique could be used in several combination with other training methods for minimization of data used, training time and energy. The introduced perturbation parameter ϵ can be selected due nature of training of the data. The determination of ϵ can be found through several trials. The application of the stochastic random descent method will increase training time and energy. The proper combined use with the perturbation will shorten training time. There exists abundance of usage of both methods, however the combined use will lead optimal solutions. A proper cost function can be used for optimum use of the perturbation parameter ϵ . The shortening the training time will lead determination of dominant inputs of the out values. One of the essential problems of training is the energy consuming will be decreased by using hybrid training methods.

Keywords

Neural Networks, Stochastic Random Steepest Descent, Perturbation Techniques

1. Introduction

The neural networks with the concept of “a thinking machine” can be traced of the antiquity. The calculation of neural networks Hardesty [1] and Yang [2] are being performed by the stochastic random descent method Bouwmeesteretal [3]. An efficient method, the perturbation method can be used as combined, Boccara [4] and Nayfeh [5]. The focus on the important events that led to the development of thinking around neural networks Shao and Shen [6], Levitan and Kaczmarek [7] and Rosenblatt [8]. Human brain functions and structures are leading to the design of Artificial Neural Networks (ANNs). The human brain is being created to pro-

cessing formation and inputs by using neural networks. This concept is becoming a base of artificial intelligence (AI). This methodology is used many AI applications. These applications are prediction, language and image recognition models. The ANN (Artificial Neural Network) is a numerical and mathematical model by using the methodology that using methodology similar to the human brain. The methodology is collecting data to reach the conclusion by using a layered structure. An ANN contain in a input layer, hidden and output layers. The layers are presented by nodes which connected each other by thresholds and weights. It should though that

*Corresponding author: hmcekirge@usa.net (Huseyin Murat Cekirge)

Received: 5 June 2025; **Accepted:** 21 June 2025; **Published:** 6 July 2025



Copyright: © The Author(s), 2025. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

each node has own individual linear regression model,

$$y_i = b_i + \sum_j w_{ij} \quad (1)$$

Here,

y_j, x_j, b_i and w_{ij} are outputs, inputs, biases and weights, respectively.

i, j and k are the indices of outputs, inputs and perturbation order.

$i = 1$ to n_1 , $j=1$ to n_2 ;

n_1 = number of inputs and n_2 = number of outputs.

If input layers are considered, biases and weights are chosen according to the importance of the input variables. All input values are multiplied by the related weights. This process is passing data from one layer to forward layer finally output layer. The selected nodal values, biases and weight values may result desired values. The stochastic random steepest descent method, or stochastic gradient descent, is an optimization algorithm used to find the minimum of the error (cost) function for determining the biases and weights, Bishop [9], Needell, et al. [10] and Bottou [11]. It's an iterative method and makes efficient for high dimensional problems and large datasets. The solution obtained from stochastic random steepest descent method is tuned or extended by us-

ing the perturbation technique Bender [12], Holmes [13] and Wiesel [14]. The perturbation theory provides methods for obtaining an acceptable solution to many mathematical problems; the essential feature of the technique is dividing the problem into "solvable" problem. In perturbation solutions, the solution is expressed as a power series in a small parameter ε . The first solution is the obtainable solution 0^{th} (zeroth order) problem. Further terms in the perturbation algorithm at higher powers of ε usually are smaller. A perturbation solution is achieved by truncating the series, most likely the first two terms are sufficient, or the solution to the known problem and the 'first order' perturbation correction. The perturbation provides an expression for the expected solution in terms of a power series which are known as a perturbation series in parameter ε that presents the deviation from the essential problem. The leading term in the perturbation series is the solution of the essential problem, while further terms present the deviation in the solution the approximation to the solution in a series in the small parameter ε .

2. Analysis

The equation for neural net work is for the i^{th} output;

$$y_i = \varepsilon^0 (b_i^0 + \sum_j w_{ij}^0 x_j) + \varepsilon^1 (b_i^1 + \sum_j w_{ij}^1 x_j) + \varepsilon^2 (b_i^2 + \sum_j w_{ij}^2 x_j) + \dots \quad (2)$$

The appropriate b_i^0 and w_{ij}^0 could be found from Equation (1), Vapnik [15] and Goodfellow [16],

$$y_i = b_i^0 + \sum_j w_{ij}^0 x_j, \quad (3)$$

where b_i^k and w_{ij}^k are biases and weights of perturbation at k^{th} level.

The indice k may change from 1 to desired accuracy level. The perturbation expansion through the powers ε for biases and weights are as follows:

$$b_i = \varepsilon^0 b_i^0 + \varepsilon^1 b_i^1 + \varepsilon^2 b_i^2 + \dots \text{ and} \quad (4)$$

$$w_{ij} = \varepsilon^0 w_{ij}^0 + \varepsilon^1 w_{ij}^1 + \varepsilon^2 w_{ij}^2 + \dots. \quad (5)$$

The stochastic random decent method or similar methods can be used for evaluation or training of the zeroth order terms.

For higher order perturbation terms, b_{ij}^k and w_{ij}^k which are tuning values of biases and weights, can be obtained from the following equation, for $k \geq 1$:

$$y_i - \sum_{l=0}^{k-1} \varepsilon^l (b_i^l + \sum_j w_{ij}^l x_j) = \varepsilon^k (b_i^k + \sum_j w_{ij}^k x_j). \quad (6)$$

The value of ε and Equation (6) are for tuning of the values obtained from the Equation (3). The further training of the data could be performed by using the stochastic random decent method. These alternatives are for improving and tuning of training of the neural network.

This methodology is a hybrid numerical method for faster and robust computations in AI problems.

3. Conclusion

The perturbation method which actively many fields and disciplines, is blended into the stochastic random steepest

descent method for tuning the results of the stochastic random steepest descent method. Perturbation theory is used in a wide range of problems. The aim is at improving stochastic random steepest descent method or similar methods. The field in general remains actively and heavily researched across multiple disciplines.

Author Contributions

Huseyin Murat Cekirge is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Hardesty L (14 April 2017). "Explained: Neural networks". MIT News Office. Archived from the original on 18 March 2024. Retrieved 2 June 2022.
- [2] Yang Z (2014). *Comprehensive Biomedical Physics*. Karolinska Institute, Stockholm, Sweden: Elsevier. p. 1. Archived from the original on 28 July 2022. Retrieved 28 July 2022.
- [3] Bouwmeester, Henricus; Dougherty, Andrew and Knyazev, Andrew V. (2015). "Nonsymmetric Preconditioning for Conjugate Gradient and Steepest Descent Methods". *Proceed. Computer Science*. 51: 2 76–285. ArXiv: 1212.6680. <https://doi.org/10.1016/j.procs.2015.05.241>
- [4] Boccara, N. *Essentials of Mathematica: With Applications to Mathematics and Physics*. Springer, New York, 2007.
- [5] Nayfeh, Ali H. (2004) *Perturbation Methods*, Wiley-VCH Verlag GmbH & Co. Kga A.
- [6] Shao, Feng; Shen, Zheng (January 9, 2022). "How can artificial neural networks approximate the brain?". *Front. Psychol.* 13: 970214. <https://doi.org/10.3389/fpsyg.2022.970214>
- [7] Levitan, Irwin; Kaczmarek, Leonard (August 19, 2015). "Inter cellular communication". *The Neuron: Cell and Molecular Biology* (4th ed.). New York, NY: Oxford University Press. pp. 153–328.
- [8] Rosenblatt, F. (1958). "The Perceptron: A Probabilistic Model For Information Storage and Organization In The Brain". *Psychological Review*. 65(6): 386–408. Cite Seer X 10.1.1.588.3775. <https://doi.org/10.1037/h0042519>
- [9] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning* (Information Science and Statistics). Springer.
- [10] Needell, D., Srebro, N. & Ward, R. (2015, January). Stochastic gradient descent weighted sampling, and the randomized Kaczmarz algorithm. <https://arxiv.org/pdf/1310.5715.pdf>
- [11] Bottou, L. (1991) Stochastic gradient learning in neural networks. *Proceedings of Neuro-Nimes*, 91. <https://leon.bottou.org/publications/pdf/nimes-1991.pdf>
- [12] Bender, Carl M. (1999). *Advanced mathematical methods for scientists and engineers I: asymptotic methods and perturbation theory*. Steven A. Orszag. New York, NY: Springer.
- [13] Holmes, Mark H. (2013). *Introduction to perturbation methods* (2nd ed.). New York: Springer.
- [14] Wiesel, William E. (2010). *Modern Astrodynamics*. Ohio: Aphelion Press p. 107.
- [15] Vapnik, V. N., (1998). *The nature of statistical learning theory* (Corrected 2nd print. ed.). New York Berlin Heidelberg: Springer.
- [16] Goodfellow, Ian, Bengio, Yoshua and Courville, Aaron (2016). *Deep Learning*. MIT Press. Archived from the original on 16 April 2016. Retrieved 1 June 2016.