

## MLP СЪС СТОХАСТИЧНО МАНИПУЛИРАН СКРИТ СЛОЙ

**Тодор Балабанов<sup>1</sup>, Румен Кетипов<sup>1</sup>, Зорница Атанасова<sup>1</sup>**

<sup>1</sup>  
*Институт по информационни и комуникационни технологии  
Българска академия на науките  
ул. „акад. Георги Бончев“, блок 2, кабинет 514, град София 1113, България  
todorb@iinf.bas.bg  
http://www.iict.bas.bg/*

## MLP WITH STOCHASTIC MANIPULATED HIDDEN LAYER

**Todor Balabanov<sup>1</sup>, Rumen Ketipov<sup>1</sup>, Zornitsa Atanassova<sup>1</sup>**

<sup>1</sup>  
*Institute of Information and Communication Technologies  
Bulgarian Academy of Sciences  
acad. Georgi Bonchev Str, block 2, office 514, 1113 Sofia, Bulgaria  
todorb@iinf.bas.bg  
http://www.iict.bas.bg/*

### Abstract

*Artificial neural networks are very popular in many areas for different problems solving. There are many different types of artificial neural networks, but one of the most popular is the multilayer perceptron. Invention of the multilayer perceptron was a keystone in the artificial neural networks evolution. It was used to prove that learning of XOR function is possible with adding layers and extending the idea of the regular perceptron. This research proposes training speed up of a regular multilayer perceptron by stochastic switching of some weights between the input and the hidden layer. All experiments are done with 2-10-1 topology of multilayer perceptron, which is used to learn the XOR function. The results of the experiments clearly show that the proposed stochastic modification speeds up the training.*

**Keywords:** artificial neural networks, backpropagation, machine learning.

### ВЪВЕДЕНИЕ

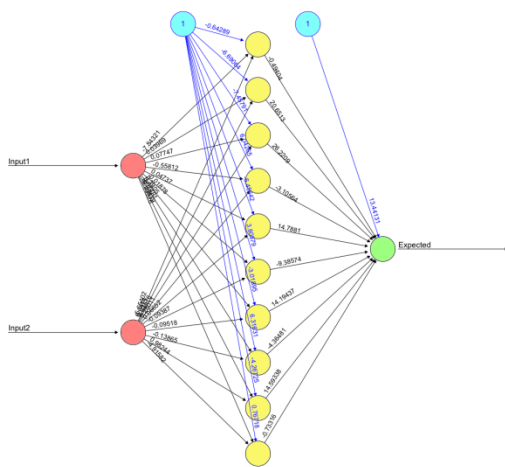
Изкуствените невронни мрежи са добре познат инструмент от областта на изкуствения интелект. Най-честата им употреба е в задачи за класификация или в задачи за прогнозиране [1,2]. Съществуват множество видове невронни мрежи [3], като най-популярният сред тях е многослойният перцептрон. Класическият перцептрон се оказва неприложим за обучение на изкуствени невронни мрежи, които трябва да работят с линейно неразделими множества (например функцията XOR). Тази трудност води до леко забавяне в развитието на изкуствените

невронни мрежи, докато не бива предложен многослойният перцептрон, който успешно решава проблемите с линейната неразделимост. Най-често използваният подход за обучение на многослойен перцептрон е точния числен метод с обратно разпространение на грешката [4]. Точните числени методи най-често са градиентни и това налага определени ограничения към вида на активационната функция [5]. Някои от ограниченията на точните числени методи биват преодолявани с помощта на глобални еволюционни евристики, каквато е еволюция на разликите (Differential Evolution) [6]. Из-

ползването на евристики за глобална оптимизация дава и допълнителна възможност изкуствените невронни мрежи да се обучават в разпределена среда на множество изчислителни машини [7,8].

## ФОРМУЛИРАНЕ НА ЗАДАЧАТА

Многослойният перцептрон се състои от неврони (възли на граф) и връзки (ребра в графа с тегла). Невроните са организирани в слоеве, като всеки неврон е свързан, чрез тегловна връзка с всеки неврон от следващия слой (Фиг.1). Информацията постъпва от външната среда във входния слой, предава се през вътрешния (може да е повече от един) слой и напуска невронната мрежа през изходния слой. Освен входни, изходни и обикновени неврони, в многослойния перцептрон се обособява по един допълнителен отместващ неврон (bias), който винаги емитира единична стойност и към него няма входни ребра.



Фиг. 1 Трислойна мрежа

Като най-разпространен и един от най-ефективните алгоритми за обучение на изкуствени невронни мрежи, обратното разпространение на грешката, има за основна задача да изчисли такива стойности за теглата в мрежата, че мрежата успешно да научи функционалната зависимост между входните и изходните данни. Обратното разпространение на грешката е точен градиентен оптимизационен метод. Състои се от две основни фази – пас в права посока и

пас в обратна посока. При изчислението в права посока, сигналите се предават от входа на мрежата към изхода

□. На изх

мрежата се определя грешката, допусната при паса в права посока. На база изчислената грешка, следва пасът в обратна посока, при който грешката се разпространява обратно по слоевете. Благодарение на това обратно разпространение се изчислява частичната грешка, допусната от всеки неврон. С така изчислената частична грешка се взема решение с каква разлика да бъдат коригирани теглата, свързващи неврона с невроните от предходния слой.

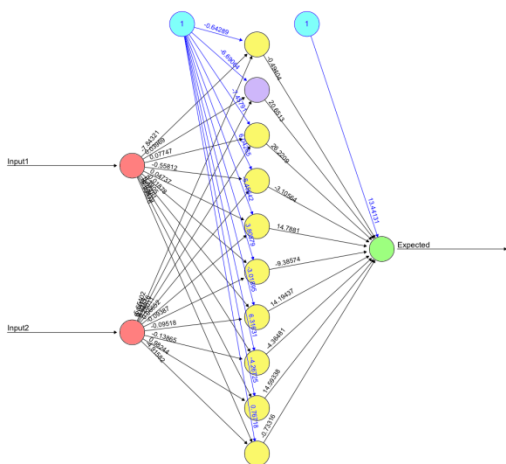
Основната задача при обучението на изкуствени невронни мрежи, по метода с обратно разпространение на грешката, е достигане до такива стойности на теглата в мрежата, че мрежата да допуска минимална обща грешка, когато се подават тренировъчните и проверочните примери. Основен стремеж на учените, работещи в областта на изкуствените невронни мрежи, е да търсят алгоритми, подходи и методи за максимално ускоряване на процеса по обучението на изкуствените невронни мрежи.

## МНОГОСЛОЕН ПЕРЦЕРПТРОН СЪС СТОХАСТИЧНО МАНИПУЛИРАН СКРИТ СЛОЙ

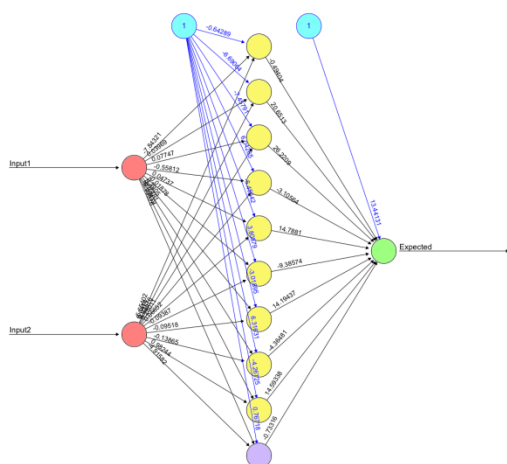
Стъпвайки на изследванията и резултатите публикувани от Занкински [9], настоящото изследване надгражда идеята с модификация на алгоритъма за обратно разпространение на грешката. Модификацията се състои в изключването на случайно избран неврон при корекцията на теглата (Фиг. 2-4), по време на обратния пас в алгоритъма за обратно разпространение на грешката.

При корекцията на теглата невронът, който е избран по случаен принцип не участва в обратния пас. Това позволява част от теглата в мрежата да останат непроменени до изпълнението на обратния пас от следващата епоха.

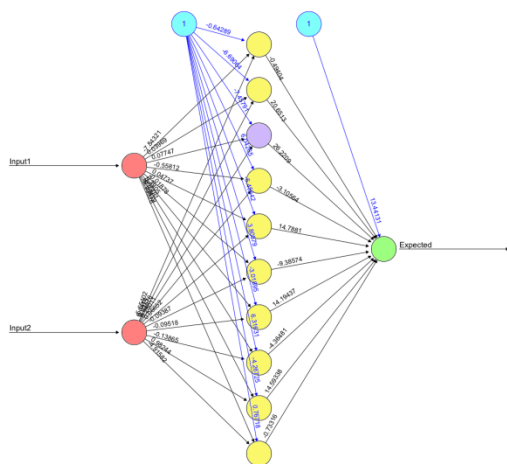
Съществено е да се отбележи, че всички неврони участват в правия пас и формират изходните сигнали на мрежата, дори случайно избраният, който не участва при обратното разпространение.



Фиг. 2 Първа епоха на обучение



Фиг. 3 Втора епоха на обучение

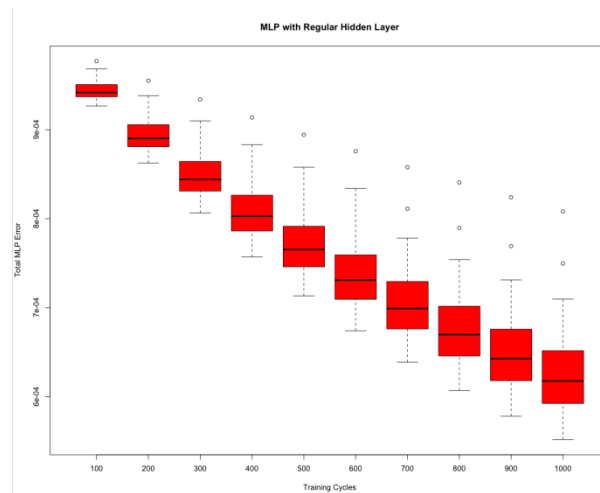


Фиг. 4 Трета епоха на обучение

В основата на това предложение стои факта, че в реалните нервни системи, част от нервните клетки умират или биват претоварени със сигнали.

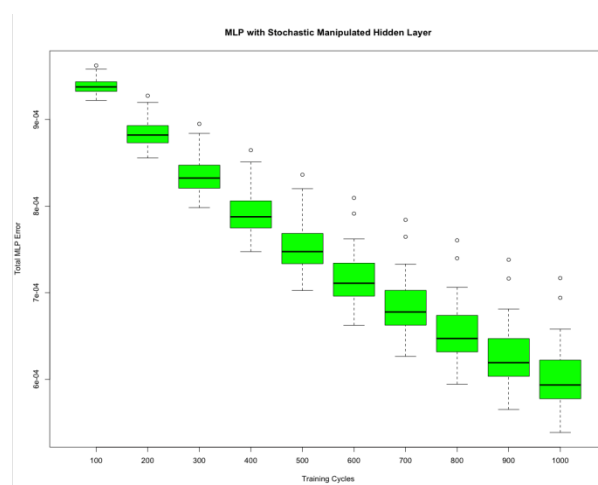
## ЕКСПЕРИМЕНТИ И РЕЗУЛТАТИ

Всички проведени експерименти са достъпни под формата на програмен код с лиценз за свободно разпространение [9]. Експериментите са реализирани с програмния език Java, като за основа е използван съществуващ вече проект за работа с многослоен перцептрон.

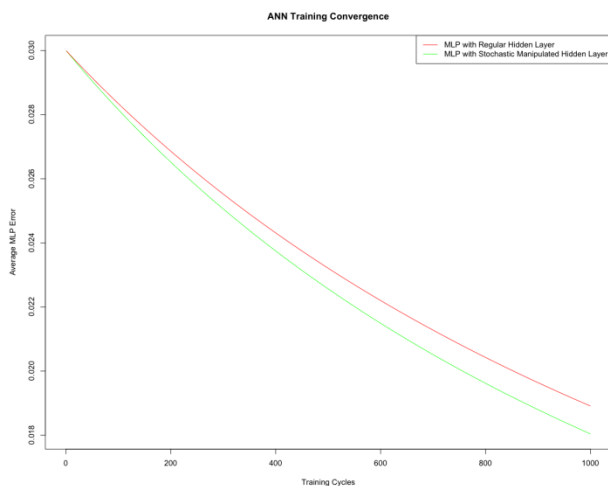


Фиг. 5 Сходимост при обучението на трислойна мрежа

Изследването е проведено с трислойна мрежа, чиято топология е от тип 2-10-1. Задачата на мрежата е да научи функцията XOR, по зададени входно-изходни примери. Сравняват се две мрежи – класическа трислойна и трислойна с манипулиран скрит слой.



Фиг. 6 Сходимост при обучението на трислойна мрежа с манипулиран скрит слой



**Фиг. 7** Средна стойност на допусната грешка от двата типа мрежи

Обучението с обратно разпространение на грешката е извършено за 1000 епохи, като измервания са направени на всеки 100 (Фиг. 5-6). Двете графики показват boxplot диаграми на резултатите от 30 независими стартирания на алгоритъма за обучение. Всяка кутия от двете графики показва грешката допусната от мрежата. На диаграмите са представени медианата и квантилите. На Фиг. 6 ясно се вижда, че трислойната мрежа със стохастично манипулиран скрит слой има много по малко разсейване в процеса на обучението.

На Фиг. 7 е представена графика на усреднените стойности (от 30 независими изпълнения) за общата грешка, която двата вида мрежи допускат. При тази графика измерване е направено на всяка епоха от обучението. Графиката ясно показва, че мрежата със стохастично манипулиран скрит слой допуска по-малка обща грешка.

## ЗАКЛЮЧЕНИЕ

От направените експерименти ясно се вижда, че процесът по обучение на многослоен перцептрон схожда по-бързо, когато бъде приложена стохастична манипулация на невроните в скрития слой. Като бъдещо развитие на изследването е интересно да се проучи възможността за сходна манипулация в изходния слой. Също така, предложената стохастична манипулация е интересно да се изследва при еври-

стични алгоритми за обучение и при обучение на невронни мрежи в разпределена среда.

## БЛАГОДАРНОСТИ

This work was supported by private funding of Velbazhd Software LLC.

## REFERENCE

- [1] Atanasova, T., Barova, M., Balabanov, T.: Use of neural patterns to analyze time lines in large amounts of data. Publishing complex Vasil Levski, 193--198, ISSN:1314-1937, (2016)
- [2] Tomov, P., Monov, V., Artificial Neural Networks and Differential Evolution Used for Time Series Forecasting in Distributed Environment, Proceedings of International conference Automatics and Informatics, ISSN 1313-1850, 129--132, Sofia, Bulgaria, (2016)
- [3] Tashev, T., Hristov, H., Modeling of synthesis of information processes with generalized nets. In: Drinov, M. (ed.) Cybernetics and Information Technologies, vol. 2, Academic Publishing House, Sofia, 92--104, (2003)
- [4] Zankinski, I., Tomov, P., Balabanov, T., Alternative Activation Function Derivative in Artificial Neural Networks, Proceedings of XXV International Symposium Management of Energy, Industrial and Environmental Systems, ISSN 1313-2237, Bankya, Bulgaria, 79--81, (2017)
- [5] Balabanov, T., Avoiding Local Optimums in Distributed Population based Heuristic Algorithms (in Bulgarian), Proceedings of XXIII International Symposium Management of energy, industrial and environmental systems, John Atanasoff Union of Automation and Informatics, Sofia, Bulgaria, 83--86, (2015)
- [6] Balabanov, T., Zankinski, I., Dobrinkova, D., Time Series Prediction by Artificial Neural Networks and Differential Evolution in Distributed Environment, International Conference on Large-Scale Scientific Computing, Sozopol, Bulgaria, Lecture Notes in Computer Science, vol. 7116, no. 1, 198--205, (2011)
- [7] Balabanov, T., Zankinski, I., Barova, M., Strategy for Individuals Distribution by Incident Nodes Participation in Star Topology of Distributed Evolutionary Algorithms, Cybernetics and Information Technologies, Sofia, Bulgaria, vol. 16, no. 1, 80--88, (2016)

- [8] Keremedchiev, D., Barova, M., Tomov, P., Mobile application as distributed computing system for artificial neural networks training used in perfect information games, Proceedings of 16th International scientific conference UNITECH16, Gabrovo, Bulgaria, vol. 2, 389--393, (2016)
- [9] Zankinski, I., Stoilov, T., Effect of the Neuron Permutation Problem on Training Artificial Neural Networks with Genetic Algorithms in Distributed Computing, Proceedings of XXIV International Symposium Management of Energy, Industrial and Environmental Systems, ISSN 1313-2237, Bankya, Bulgaria, 53--55, (2016)
- [10] Balabanov, T., Multi Layer Perceptron implemented in Java, <https://github.com/Coding-Sunday-Sofia/multi-layer-perceptron> , (2018)