Evolution of Artificial Neural Networks

Alexey Averkin
Dorodnicyn Computing Centre
Federal Research Center "Computer Science and Control"
of Russian Academy of Sciences
Moscow, Russia
averkin2003@inbox.ru

Sergey Yarushev
Dubna State University
Dubna, Russia
Sergey.Yarushev@icloud.com

***Abstract*—The history of artificial neural networks originates from the distant 1943, when Warren McCulloch and Walter Pitts formalized the notion of a neural network in a fundamental article on the logical calculation of ideas and nervous activity. This was the starting point in the history of artificial neural networks. Later, a huge number of different architectures of neural networks, learning methods and optimization algorithms were proposed. There was a time when neural networks were left in oblivion, but after the 1980s they were talked about again when John Hopfield introduced his famous full-mesh network. Today, artificial neural networks are indispensable tools in a huge number of tasks, such as tasks of forecasting and analyzing time series, the task of recognizing images and even emotions, classification tasks, natural language recognition, neural networks are used in industry, defense and medicine; many enumerate areas of their application. All this tells us that artificial neural networks have become part of modern life.**

***Keywords*—artificial neural networks, forecasting, modular neural networks**

I. Introduction

What was once just a figment of the imagination of some our most famous science fiction writers, artificial intelligence (AI) is taking root in our everyday lives. We're still a few years away from having robots at our beck and call, but AI has already had a profound impact in more subtle ways. Weather forecasts, email spam filtering, Google's search predictions, and voice recognition, such Apple's Siri, are all examples. What these technologies have in common are machine-learning algorithms that enable them to react and respond in real time. There will be growing pains as AI technology evolves, but the positive effect it will have on society in terms of efficiency is immeasurable. In this paper we will talking about Artificial Neural Networks - great instrument of the Artificial Intelligence.

II. Early researches in Artificial Neural Networks

In 1943, neurophysiologist Warren McCulloch and mathematician Walter Pitts wrote a paper on how neurons might work [8]. In order to describe how neurons in the brain might work, they modeled a simple neural network using electrical circuits. In 1949, Donald Hebb wrote "The Organization of Behavior", a work which pointed out the fact that neural pathways are strengthened each time they are used, a concept fundamentally essential to the ways in which humans learn. If two nerves fire at the same time, he argued, the connection between them is enhanced. As computers became more advanced in the 1950's, it was finally possible to simulate a hypothetical neural network. The first step towards this was made by Nathanial Rochester from the IBM research laboratories. Unfortunately for him, the first attempt to do so failed.

A. First Model of Artificial Neural Network

American neurophysiologist Frank Rosenblatt. He proposed a diagram of a device that simulates the process of human perception, and called it a "perceptron". The perceptron transmitted signals from photocells, which are a sensory field, to blocks of electromechanical memory cells. These cells were connected randomly in accordance with the principles of connectivity. In 1957, the Cornell Laboratory of Aeronautics successfully completed the modeling of perceptron work on the IBM 704 computer, and two years later, on June 23, 1960 at Cornell University, the first neurocomputer Mark-1 was demonstrated, which was able to recognize certain letters of the English alphabet [4].

Frank Rosenblatt with his creation - "Mark-1." To "teach" the perceptron to classify images, a special iterative method of learning trial and error was developed, reminiscent of the process of human training - the method of error correction [3]. In addition, when recognizing a particular letter, the perceptron could distinguish the characteristic features of letters that are statistically more frequent than minor differences in individual cases. Thus, the perceptron was able to generalize letters written in different ways (handwriting) into one generalized image. However, the Perceptron's possibilities were limited: the machine could not reliably recognize partially closed letters, as well as letters of a different size, located with a shift or rotation, rather than those used at the stage of its training. The report on the first results appeared in 1958 - then Rosenblatt published an article "Perceptron: A Probable Model of Storage and Organization of Information in the Brain" [11]. But he describes his theories and assumptions about perception processes and perceptrons in more detail in 196 in his book "Principles of Neurodynamics: Perceptrons and Theory of Brain Mechanisms" [12]. In the book, he examines not only ready-made perceptron models with one hidden layer, but also multilayer perceptrons with cross-sections (the third chapter) and inverse (fourth chapter) connections. The book also introduces a number of important ideas and theorems, for example, the theorem of convergence of a perceptron is proved.

B. Next Researches

In 1959, Bernard Widrow and Marcian Hoff of Stanford developed models called "ADALINE" and "MADALINE." In a typical display of Stanford's love for acronymns, the names come from their use of Multiple ADAptive LINear Elements. ADALINE was developed to recognize binary patterns so that if it was reading streaming bits from a phone line, it could predict the next bit. MADALINE was the first neural network applied to a real world problem, using an adaptive filter that eliminates echoes on phone lines. While the system is as ancient as air traffic control systems, like air traffic control systems, it is still in commercial use. In 1962, Widrow & Hoff developed a learning procedure that examines the value before the weight adjusts it (i.e. 0 or 1) according to the rule: Weight Change = (Pre-Weight line value) * (Error / (Number of Inputs)). It is based on the idea that while one active perceptron may have a big error, one can adjust the weight values to distribute it across the network, or at least to adjacent perceptrons. Applying this rule still results in an error if the line before the weight is 0, although this will eventually correct itself. If the error is conserved so that all of it is distributed to all of the weights than the error is eliminated. Despite the later success of the neural network, traditional von Neumann architecture took over the computing scene, and neural research was left behind. Ironically, John von Neumann himself suggested the imitation of neural functions by using telegraph relays or vacuum tubes. In the same time period, a paper was written that suggested there could not be an extension from the single layered neural network to a multiple layered neural network. In addition, many people in the field were using a learning function that was fundamentally flawed because it was not differentiable across the entire line. As a result, research and funding went drastically down. This was coupled with the fact that the early successes of some neural networks led to an exaggeration of the potential of neural networks, especially considering the practical technology at the time. Promises went unfulfilled, and at times greater philosophical questions led to fear. Writers pondered the effect that the so-called "thinking machines" would have on humans, ideas which are still around today. The idea of a computer which programs itself is very appealing. If Microsoft's Windows 2000 could reprogram itself, it might be able to repair the thousands of bugs that the programming staff made. Such ideas were appealing but very difficult to implement. In addition, von Neumann architecture was gaining in popularity. There were a few advances in the field, but for the most part research was few and far between. In 1972, Kohonen and Anderson developed a similar network independently of one another, which we will discuss more about later. They both used matrix mathematics to describe their ideas but did not realize that what they were doing was creating an array of analog ADALINE circuits. The neurons are supposed to activate a set of outputs instead of just one. The first multilayered network was developed in 1975, an unsupervised network.

III. From 1980 to present

In 1982, interest in the field was renewed. John Hopfield of Caltech presented a paper to the National Academy of Sciences. His approach was to create more useful machines by using bidirectional lines. Previously, the connections between neurons was only one way. That same year, Reilly and Cooper used a "Hybrid network" with multiple layers, each layer using a different problem-solving strategy. Also in 1982, there was a joint US-Japan conference on Cooperative/Competitive Neural Networks. Japan announced a new Fifth Generation effort on neural networks, and US papers generated worry that the US could be left behind in the field. (Fifth generation computing involves artificial intelligence. First generation used switches and wires, second generation used the transister, third state used solid-state technology like integrated circuits and higher level programming languages, and the fourth generation is code generators.) As a result, there was more funding and thus more research in the field. In 1986, with multiple layered neural networks in the news, the problem was how to extend the Widrow-Hoff rule to multiple layers. Three independent groups of researchers, one of which included David Rumelhart, a former member of Stanford's psychology department, came up with similar ideas which are now called back propagation networks because it distributes pattern recognition errors throughout the network. Hybrid networks used just two layers, these back-propagation networks use many. The result is that back-propagation networks are "slow learners," needing possibly thousands of iterations to learn. Now, neural networks are used in several applications, some of which we will describe later in our presentation. The fundamental idea behind the nature of neural networks is that if it works in nature, it must be able to work in computers. The future of neural networks, though, lies in the development of hardware. Much like the advanced chess-playing machines like Deep Blue, fast, efficient neural networks depend on hardware being specified for its eventual use. Research that concentrates on developing neural networks is relatively slow. Due to the limitations of processors, neural networks take weeks to learn. Some companies are trying to create what is called a "silicon compiler" to generate a specific type of integrated circuit that is optimized for the application of neural networks. Digital, analog, and optical chips are the different types of chips being developed. One might immediately discount analog signals as a thing of the past. However neurons in the brain actually work more like analog signals than digital signals. While digital signals have two distinct states (1 or 0, on or off), analog signals vary between minimum and maximum values. It may be awhile, though, before optical chips can be used in commercial applications.

IV. Examples of Nowdays Artificial Neural Networks

A. Modular Neural Networks

The core of the modular neural networks is based on the principle of decomposition of complex tasks into simpler ones. Separate modules make simple tasks. More simple subtasks are then carried through a series of special models. Each local model performs its own version of the problem according to its characteristics. The decision of the integrated object is achieved by combining the individual results of specialized local computer systems in a dependent task. The expansion of the overall problem into simpler subtasks can be either soft or hard unit subdivision. In the first case, two or more subtasks of local computer systems can simultaneously assigned while in the latter case, only one local computing model is responsible for each of the tasks crushed. Each modular system has a number of special modules that are working in small main tasks. Each module has the following characteristics:

1) The domain modules are specific and have specialized computational architectures to rec-ognize and respond to certain subsets of the overall task;
2) Each module is typically independent of other modules in its functioning and does not in-fluence or become influenced by other modules;
3) The modules generally have a simpler architecture as compared to the system as a whole. Thus, a module can respond to given input faster than a complex monolithic system;
4) The responses of the individual modules are simple and have to combine by some inte-grating mechanism in order to generate the complex overall system response.

The best example of modular system is human visual system. In this system, different mod-ules are responsible for special tasks, like a motion detection, color recognition and shape. The central nervous system, upon receiving responses of the individual modules, develops a complete realization of the object which was processed by the visual system.

B. ANN in Time Series Forecasting

The recent surge in research of artificial neural networks (ANN) showed that neural networks have a strong capability in predicting and classification problems. ANN successfully used for various tasks in many areas of business, industry and science [14]. Such high interest in neural networks caused by the rapid growth in the number of articles published in scientific journals in various disciplines. It suffices to consider several large databases to understand the huge number of articles published during the year on the theme of the study of neural networks, it is thousands of articles. A neural network is able to work parallel with input variables and consequently handle large sets of data quickly. The main advantage of neural networks is the ability to find patterns [2]. ANN is a promising alternative in the toolbox professionals involved in forecasting. In fact, the nonlinear structure of the neural networks is partially useful to identify complex relationships in most real world problems. Neural networks are perhaps the universal method of forecasting in connection with those that they cannot only find the non linear structures in problems, they can also simulate the processes of linear processes. For example, the possibility of neural networks in the modeling of linear time series line were studied and confirmed by a number of researchers [5]. One of the main applications of ANN is forecasting. In recent years, it was seen increasing interest in forecasting using neural networks. Forecasting has a long history, and its importance reflected in the application in a variety of disciplines from business to engineering. The ability to accurately predict the future is fundamental to many decision making processes in planning, developing strategies, building policy, as well as in the management of supply and stock prices. As such, forecasting is an area in which a lot of effort has invested in the past. In addition, it remains an important and active area of human activity in the present and will continue to evolve in the future. Review of the research needs in the prediction presented to the Armstrong [1]. For several decades in forecasting dominated by linear methods. Linear methods are simple in design and use, and they are easy to understand and interpret. However, linear models have significant limitations, owing to which they cannot discern any nonlinear relationships in data. Approximation of linear models to complex are not linear relationships do not always give a positive result. Earlier in 1980, there have been large scale competition for forecasting, in which most widely used linear methods were tested on more than 1,000 real time series [7]. Mixed results showed that none of the linear model did not show the best results worldwide, which can be interpreted as a failure of the linear models in the field of accounting with a certain degree of non linearity, which is common for the real world. Predicting financial markets is one of the most important trends in research due to their commercial appeal [6]. Unfortunately, the financial markets are dynamic, nonlinear, complex, nonparametric and chaotic by nature [13]. Time series of multistationary, noisy, casual, and have frequent structural breaks. In addition, the financial markets also affects a large number of macroeconomic factors [9], such as political developments, global economic developments, bank rating, the policy of large corporations, exchange rates, investment expectations, and events in other stock markets, and even psychological factors. Artificial neural networks are one of the technologies that have received significant progress in the study of stock markets. In general, the value of the shares is a random sequence with some noise, in turn, artificial neural networks are powerful parallel processors nonlinear systems depending on their internal relations. Development of techniques and methods that can approximate any nonlinear continuous function without a priori notions about the nature of the process itself seen in the work of P. Pino [10]. It is obvious that a number of factors demonstrate sufficient efficacy in the forecast prices, and most importantly a weak point in this is that they all contain a few limitations in forecasting stock prices and use linear methods, the relative of this fact, although previous investigation revealed the problem to some extent, none of them provides a comprehensive model for the valuation of shares. If we evaluate the cost and provide a model in order to remove the uncertainty, it is largely can help to increase the investment attractiveness of the stock exchanges. Conduct research to get the best method of forecasting financial time series is currently the most popular and promising task.

V. Conclusion

In this paper, we talked about evolution of artificial neural networks evolved from the first models of ANN, which presented Warren McCulloch and Walter Pitts. The history of ANN is not so long but evolution of architectures and methods is is impressive. Now we could use ANN for prediction and data analisys, for computer vision and emotions recognition. Neural networks are used in safety systems. We could see very wide fileds in our paper where ANN could be applied. In the development of new smartphones used artifitial intelligence, in automotive industry ANN is used for safety and assistant systems. All it demonstrate - Artificial Intelligence is a part of our nowdays life!

Acknowledgment

This work was supported by the Russian Foundation for Basic Research 17-07-01558 and 17-37-50013

References

[1] J Scott Armstrong. Research needs in forecasting. *International Journal of Forecasting*, 4(3):pp.449–465, 1988.

[2] Pei-Chann Chang, Yen-Wen Wang, and Chen-Hao Liu. The development of a weighted evolving fuzzy neural network for pcb sales forecasting. *Expert Systems with Applications*, 32(1):pp.86–96, 2007.

[3] Michele D Estebon. Perceptrons: An associative learning network. *Virginia Tech.—1997*, 1997.

[4] Donald Olding Hebb. *The organization of behavior: A neuropsychological theory*. Psychology Press, 2005.

[5] H Brian Hwarng. Insights into neural-network forecasting of time series corresponding to arma (p, q) structures. *Omega*, 29(3):pp.273–289, 2001.

[6] Ritanjah Majhi, Ganapati Panda, Gadadhar Sahoo, Pradipta K Dash, and Debi Prasad Das. Stock market prediction of s&p 500 and djia using bacterial foraging optimization technique. In *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*, pages pp.2569–2575. IEEE, 2007.

[7] Spyros Makridakis, A Andersen, Robert Carbone, Robert Fildes, Michele Hibon, Rudolf Lewandowski, Joseph Newton, Emanuel Parzen, and Robert Winkler. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of forecasting*, 1(2):pp.111–153, 1982.

[8] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):pp.115–133, 1943.

[9] Kai Miao, Fang Chen, and ZG Zhao. Stock price forecast based on bacterial colony rbf neural network [j]. *Journal of Qingdao University (Natural Science Edition)*, 20(2):pp.50–54, 2007.

[10] Raúl Pino, José Parreno, Alberto Gomez, and Paolo Priore. Forecasting next-day price of electricity in the spanish energy market using artificial neural networks. *Engineering Applications of Artificial Intelligence*, 21(1):pp.53–62, 2008.

[11] Frank Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):p.386, 1958.

[12] Frank Rosenblatt. Principles of neurodynamics. perceptrons and the theory of brain mechanisms. Technical report, CORNELL AERONAUTICAL LAB INC BUFFALO NY, 1961.

[13] Tuan Zea Tan, Chai Quek, and Geok See Ng. Brain-inspired genetic complementary learning for stock market prediction. In *Evolutionary Computation, 2005. The 2005 IEEE Congress on*, volume 3, pages pp.2653–2660. IEEE, 2005.

[14] Bernard Widrow, David E Rumelhart, and Michael A Lehr. Neural networks: applications in industry, business and science. *Communications of the ACM*, 37(3):pp.93–106, 1994.

ЭВОЛЮЦИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Аверкин А.Н.
Вычислительный центр им. А.А. Дородницына
Федерального исследовательского центра
«Информатика и управление» Российской академии наук
г. Москва, Россия
Ярушев С.А.
Государственный Университет "Дубна"
г. Дубна, Россия

История искусственных нейронных сетей берет свое начало с далекого 1943 года, когда Уоррен Мак-Каллок и Уолтер Питтс ормализовали понятие нейронной сети в фундаментальной статье о логическом исчислении идей и нервной активности. Это и стало отсчетной точкой в истории искусственных нейронных сетей. В последствии было предложено огромное количество разнообразных архитектур нейронных сетей, методов обучения и алгоритмов оптимизации. Было время, когда нейронные сети подвергались забвению, но после 1980х годов о них заговорили снова, когда Джон Хопфилд представил свою знаменитую полносвязную сеть сеть. Сегодня искусственные нейронные сети являются незаменимым инструментов в огромном количестве задач, таких, как задачи прогнозирования и анализа временных рядов, задачи распознавания образов, изображений и даже эмоций, задачи классификации, распознавания естесственного языка, нейронные сети применяются в промышленности, обороне и медицине, можно очень много перечислять области их применения. Все это говорит нам о том, что искусственные нейронные сети стали частью современной жизни.