

ВАРИАНТ ПРИМЕНЕНИЯ НЕЙРОСЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ИЗМЕНЕНИЯ ПАРАМЕТРОВ КАЧЕСТВА АРЕНДУЕМЫХ ИНФОРМАЦИОННЫХ ПОТОКОВ

Д.С. Горбуля, Академия ФСО России, gorbuly@ya.ru.

УДК 004.942

Аннотация. В статье проведена оценка применимости нейросетей для прогнозирования изменений значений параметров, характеризующих качество информационных потоков. Определена архитектура нейросети, объединяющая многослойные перцептроны и на основании сравнения результатов прогнозирования определена длина обучающей выборки, необходимая для адекватного автоматического построения нейросети. Также сравнение результатов прогнозирования изменения значений параметров с измеренными значениями позволило определить, что автоматически сформированная нейросеть превосходит результаты традиционных аналитических подходов.

Ключевые слова: мониторинг; прогнозирование технического состояния; нейросети.

A VARIANT OF USING NEURAL NETWORKS TO PREDICT CHANGES IN THE QUALITY PARAMETERS OF LEASED INFORMATION FLOWS

D.S. Gorbulya, Academy of the FSO of Russia.

Annotation. The article evaluates the applicability of neural networks to predict changes in the values of parameters characterizing the quality of information flows. The architecture of the neural network combining multilayer perceptron's is determined and, based on a comparison of the prediction results, the length of the training sample necessary for adequate automatic neural network construction is determined. Also, a comparison of the results of forecasting changes in parameter values with the measured values made it possible to determine that the automatically generated neural network surpasses the results of traditional analytical approaches.

Keywords: monitoring; forecasting of technical condition; neural networks.

Введение

Повышение роли инфотелекоммуникационных систем (ИТКС) в деятельности современного общества, усложнение алгоритмов их функционирования и критичность к дестабилизирующим факторам, способствовали развитию средств и систем мониторинга технического состояния. Однако получение возможности по измерению большего количества параметров, характеризующих функционирование систем, породило группу новых взаимоувязанных задач [1-3]:

- «сколько потоковых данных теряется из-за того, что их не успевают сохранить и/или обработать?»;
- «что из потерянных данных – существенно и что изменилось бы, если бы не были потеряны данные?»;
- «сколько (длительность), как часто (периодичность), что и где измерять?».

На разрешение указанных задач влияют две группы факторов:

- технические и технологические возможности средств обработки потока данных о состоянии контролируемой системы (чем выше производительность средств, тем больший объем входных данных

- возможно обработать);
- подходы, применяемые для обработки статистических данных (в том числе *bigdata*) обладают различной адекватностью, сложностью, условиями применимости, достоверностью.

Учитывая динамичность изменения значений параметров ИТКС, принятые подходы управления системой, основанные на оценке ее текущего состояния, не позволяют своевременно выявить дестабилизирующее воздействие, принимать решения о применении управляющих воздействий. Традиционные подходы описания функционирования систем с помощью дифференциальных уравнений, элементов математической статистики, имитационного моделирования из-за значительного усложнения ИТКС и протекающих в ней процессов становятся менее эффективны и востребованы в практической деятельности [4-8]. Для замены указанных математических подходов все в большем количестве отраслей деятельности применяют эвристические методы, основанные на применении технологий искусственного интеллекта: генетические алгоритмы, нечеткая логика и нейросети [9-11]. Также следует отметить, что из-за высокой скорости изменения состояния адаптивные системы, в том числе и ИТКС, переходят от оптимального к рациональному управлению, т.е. нахождению управляющих воздействий, удовлетворяющих заданным условиям в конкретный момент времени. Перечисленные факторы привели к коррекции требований, предъявляемых к достоверности результатов прогнозирования (моделирования), определяющих момент изменения тренда или достижения конкретных значений параметров, описывающих систему.

Текущий этап развития технологии нейросетей показывает их высокую эффективность. Успешным примером применения нейросетей для анализа временных рядов, описывающих изменения экономических индексов на крипторынке, может служить нейросеть «*Mirocana*» (отечественная разработка), учитывающая в составлении прогноза более 100 валютных пар, множество технических индикаторов, а также разнообразные экономические новостей. Средняя ошибка прогнозирования нейросети «*Mirocana*» для индекса МосБиржи составила 6 % [12, 13].

Целью исследования является определение применимости использования нейросетей (обученной и использующей исходные данные от системы мониторинга *Zabbix*), как источника исходных данных для системы управления компьютерной сети связи (КСС), использующей арендуемые информационные потоки.

Задачами исследования является определение:

- структуры нейросети, применяемой для прогнозирования изменения параметров, характеризующих качество информационного потока;
- объема выборки, необходимого для обучения нейросети, позволяющей прогнозировать изменения параметров, характеризующих качество информационного потока с требуемой достоверностью;
- длительности обучения нейросети.

Выбор и обоснование структуры нейросети для прогнозирования изменения качества информационных потоков

Цель исследования достигается на основе обработки статистических данных созданного полунатурного стенда (рис. 1), синтеза группы нейросетей с использованием *STATISTICA* и анализа результатов прогнозирования.

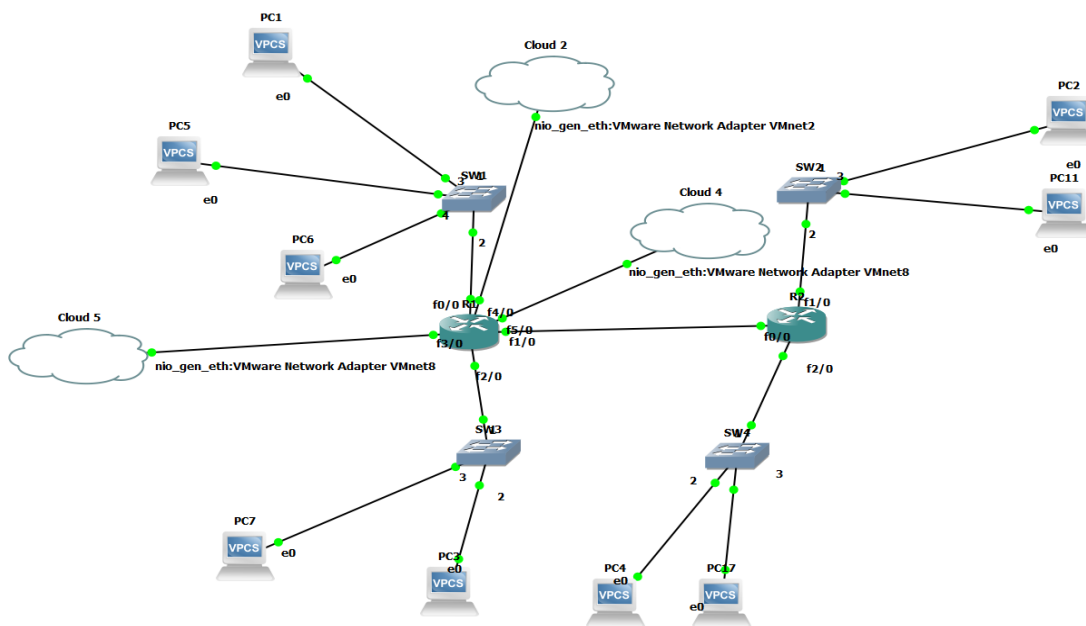


Рисунок 1

Текущий этап развития транспортной инфраструктуры ИТКС с коммутацией пакетов показывает, что увеличение информационных потоков, создаваемых элементами сети, является причиной роста коэффициента потери пакетов (*IPLR*) и коэффициента ошибок пакетов (*IPER*) и приводит к ухудшению качества предоставляемых услуг связи. Применение полунатурного стенда позволяет имитировать функционирование КСС с возможностью повторения экспериментов при одних и тех же условиях. В рамках исследования в качестве основы стенда использована среда *GNS3* (ver. 2.1.0). Сформированный стенд позволяет имитировать работу телекоммуникационного оборудования и измерять значения параметров отдельных элементов при изменении информационной нагрузки, формируемой всеми элементами транспортной сети.

В качестве инструмента сбора статистических данных об изменении параметров, характеризующих качество информационного потока, применена система мониторинга оборудования транспортной сети с коммутацией пакетов *Zabbix* (ver. 3.4) (рис. 2).

Анализ работ в области синтеза нейросетей для анализа временных рядов [14-16] позволил сгруппировать известные архитектуры нейросетей в следующие группы:

- сети, объединяющие многослойные персептроны (*MLP NN*);
- рекуррентные нейронные сети (*RNN*);
- сети общей регрессии (*GRNN*);
- сети прямого распространения (*RBFN*);
- вероятностные нейронные сети (*PNN*).

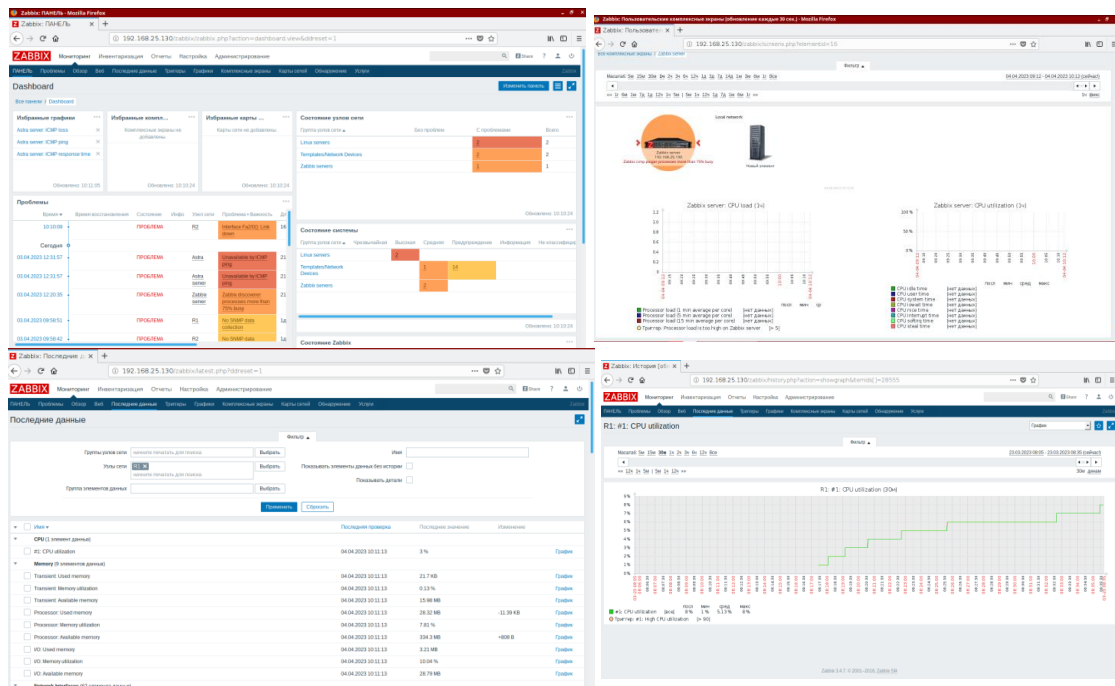


Рисунок 2

Сравнение результатов прогнозирования изменения временных рядов с использованием указанных архитектур нейросетей, проведенное в работах [15-17], показывает, что для решения заявленной в статье задачи возможно применение сети многослойных перцептронов. Данные сети позволяют осуществить прогнозирование изменения временного ряда с высокой степенью достоверности при незначительных временных затратах на их разработку (встроенные возможности *STATISTICA (ver. 10.0.1011.0)*), а также позволяют в автоматизированном режиме формировать группу различных нейросетей с общим принципом формирования топологии) на основе заданной длины обучающей выборки. Обучение нейросетей на основе многослойных перцептронов проводилось по выборкам длиной 50 75 и 100 отчетов (рис. 3).

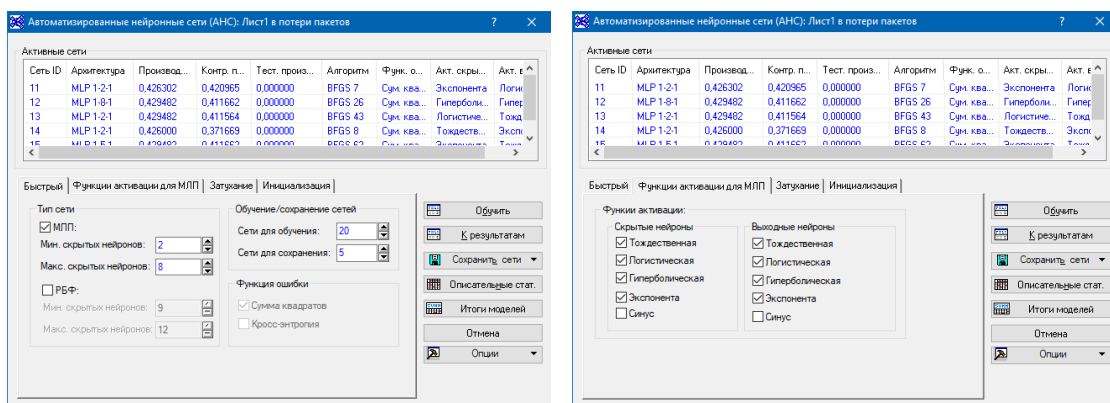


Рисунок 3

Для автоматического синтеза использовались следующие параметры:

- Минимальное количество скрытых нейронов – 2.
- Максимальное количество скрытых нейронов – 8.
- Количество обучаемых сетей – 20.

- Количество сохраняемых сетей – 5.

Функции активации скрытых нейронов – тождественная, логистическая, гиперболическая и экспоненциальная.

Функции активации выходных нейронов – тождественная, логистическая, гиперболическая и экспоненциальная.

На рис. 3 в окне «Активные окна» показаны архитектуры сгенерированных нейросетей.

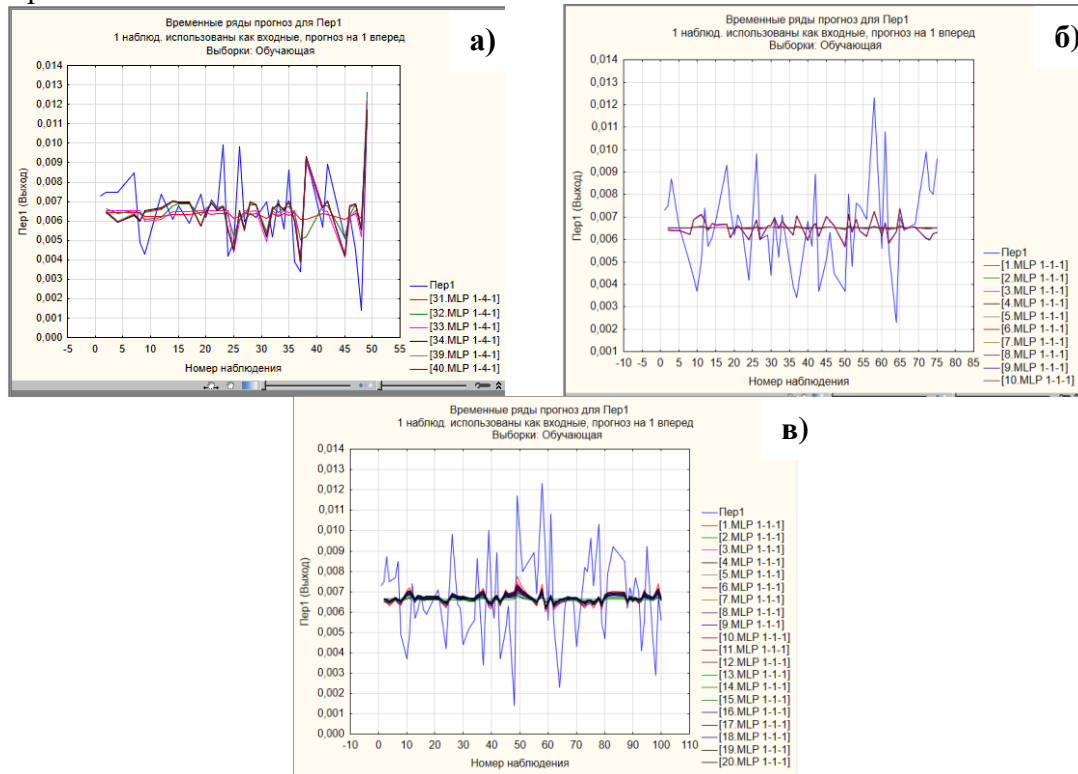


Рисунок 4

- а) результаты прогнозирования нейросетей, обученных на выборке – 50 отсчетов
- б) результаты прогнозирования нейросетей, обученных на выборке – 75 отсчетов
- в) результаты прогнозирования нейросетей, обученных на выборке – 100 отсчетов

Результаты прогнозирования изменения параметров качества арендуемых информационных потоков с применением автоматически сформированных и обученных нейросетей показаны на рис. 4 (а-в).

Длительность обучения указанных групп нейросетей составило от 3 до 10 с. Время выбора нейросети, обладающей наивысшей достоверностью прогнозирования из сгенерированных сетей (по критерию распределения остатков), составляет до 10 с (при выборе из 5 лучших сгенерированных сетей). Процесс автоматического формирования данных для прогнозирования данных соизмерим со временем измерения. Таким образом, процесс формирования набора исходных данных, обучения и выбора наилучшей нейросети для дальнейшего применения не превышает 20 с. Исходя из того, что при обучении сетей один отсчет равен 1 с., нейросеть должна формировать прогноз на 20 отсчетов вперед.

Выбор нейросети, обладающей наибольшей достоверностью результатов проводился в два этапа: на первом этапе выбиралась лучшая сеть среди одной выборки обучения (50, 75 и 100 отсчетов); на втором – из трех сетей, обученных на указанных выборках. В качестве критерия принятия решения приняты данные о наименьшем распределении остатков обучения (рис. 5).

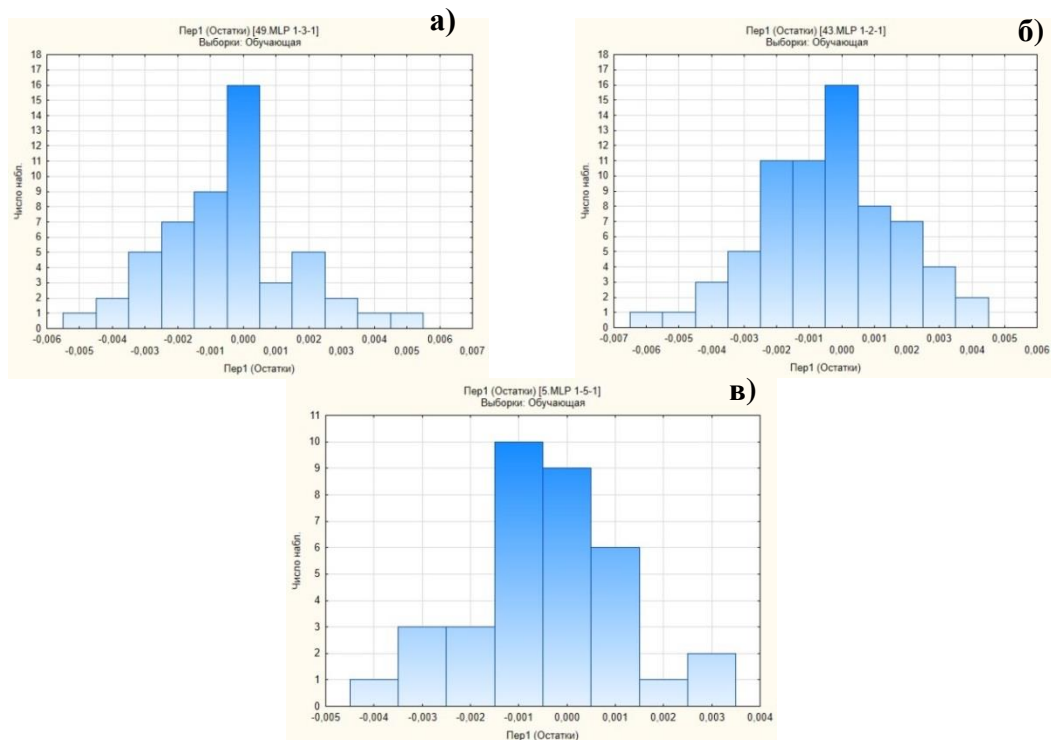


Рисунок 5

- а) для нейросети, обученной на выборке – 50 отсчетов;
- б) для нейросети, обученной на выборке – 75 отсчетов;
- в) для нейросети, обученной на выборке – 100 отсчетов;

Сравнение распределения остатков обучения показало, что наибольшей достоверностью обладает нейросеть, обученная на выборке 50 отсчетов (рис. 6).

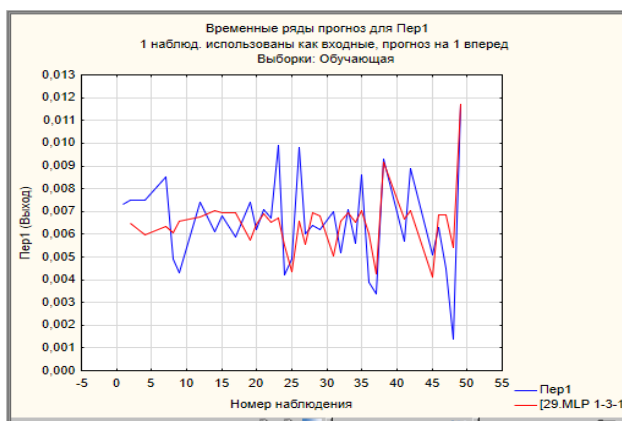


Рисунок 6

С целью оценки применимости нейростей для прогнозирования изменения параметров качества информационных потоков осуществлено сравнение полученных результатов с результатами расчетов аналогичного параметра с использованием известных методов математической статистики: простой скользящей средней, взвешенных скользящих средних, наименьших квадратов и описании в виде полинома и собранных статистических данных (рис. 7).



Рисунок 7

- а) для нейросети и метода скользящего среднего;
- б) для нейросети и метода взвешенного скользящего среднего;
- в) для нейросети и метода наименьших квадратов;
- г) для нейросети и описания ряда полиномом.

Численная оценка результатов прогнозирования проведена с помощью сравнения следующих параметров:

Среднеквадратичная ошибка (*Root Mean Squared Error, RMSE*) –
$$\left(RMSE = \sqrt{\frac{1}{n} \sum (R^{\text{прогноз}} - R^{\text{измер}})^2} \right);$$

Средняя абсолютная погрешность (*Mean Absolute Error, MAE*) –
$$\left(MAE = \frac{1}{n} \sum |R^{\text{прогноз}} - R^{\text{измер}}| \right);$$

Средняя абсолютная процентная погрешность (*Mean Absolute Percentage Error, MAPE*) –
$$\left(MAPE = \frac{1}{n} \sum \frac{|R^{\text{прогноз}} - R^{\text{измер}}|}{R^{\text{измер}}} \right).$$

Результаты расчетов и их сравнения представлены в табл. 1, 2.

Таблица 1.

Метод прогнозирования	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>
Нейросеть	0,001298	0,001077	0,0167778
Метод скользящего среднего	0,0776209	0,0625	0,096775968
Метод взвешенного скользящего среднего	0,002350644	0,01113	0,174441962
Метод наименьших квадратов	0,0025698	0,0024077	0,027548
Полином	0,0874514	0,07812	0,06741

Таблица 2.

Метод прогнозирования	Нейросеть		
	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>
Метод скользящего среднего	98,33	98,28	82,66
Метод взвешенного скользящего среднего	44,78	90,32	90,38
Метод наименьших квадратов	49,49	55,27	39,1
Полином	98,52	98,62	75,11

Заключение

Сравнение результатов оценки прогнозных моделей изменения временного ряда показывают, что сформированная нейросеть обладает наилучшими показателями и ее применение позволяет с большей обоснованностью принимать решения по управлению конфигурацией сети.

Таким образом, проведенные исследования показывают, что применение нейросетей, как источника исходных данных для системы управления компьютерной сети связи (КСС), использующей арендуемые информационные потоки, позволяет повысить достоверность результатов прогнозирования (по сравнению с выбранными методами прогнозирования) от 39 до 98 %. Численные результаты оценки достоверности результатов прогнозирования позволяют утверждать о применимости нейросетей, как источника исходных данных и достижении цели исследования.

Литература

- Новиков Д.А. Кибернетика: Навигатор. История кибернетики, современное состояние, перспективы развития. – М.: Ленанд, 2016. – 160 с.
- Добрышин М.М., Горбуля Д.С. Подходы оценки качества связи и предоставления услуг связи и задачи по их совершенствованию в рамках обеспечения информационной безопасности // Экономика и качество систем связи, 2023. – № 3 (29). – С. 60-71.
- Белов А.С., Добрышин М.М. Предложение по удаленному мониторингу программных средств автономных комплексов связи // Авиакосмическое приборостроение, 2021. – № 6. – С. 13-20.
- Юсупов Р.М. Элементы теории испытаний и контроля технических систем / В.И. Городецкий, А.К. Дмитриев, В. М. Марков и др. Под ред. Р. М. Юсупова. – Л.: Энергия, 1978. – 192 с.
- Грешилов А.А. и др. Математические методы построения прогнозов / А.А. Грешилов, В.А. Стакун, А.А. Стакун. – М.: Радио и связь, 1997. – 112 с.
- Четыркин Е.М. Статистические методы прогнозирования. Изд. 2-е, перераб. доп. / Е.М. Четыркин – М.: Статистика, 1977. – 200 с.
- Тихов М.С., Котельникова М.В. Современные методы статистического оценивания параметров: Учебно-методическое пособие. – Нижний Новгород, Нижегородский госуниверситет, 2016. – 120 с.
- Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы; [пер. с польск. И.Д. Рудинского]. – М.: Горячая линия – Телеком, 2006. – 452 с.
- Бураков М.В. Генетический алгоритм: теория и практика: учеб. пособие // – М. В. Бураков. – СПб.: ГУАП, 2008. – 264 с.
- Круглов В.В., Дли М.И., Годунов Р.Ю. Нечеткая логика и искусственные нейронные сети. – М.: Физматлит, 2001. – 224 с.
- Coinspot: электрон. ресурс. Mirosana – нейронная сеть для торговли на крипторынке. 2017 URL: <https://coinspot.io/technology/mirosana-nejronnaya-set-dlya-torgovli-na-kriptorynke/> (дата обращения 10.01.2024).

12. Сергеев В.А. Использование нейросетей в прогнозировании фондового рынка // Азимут научных исследований: экономика и управление, 2018. – Т. 7. – № 4 (25). – С. 280-282.
13. Fjodor van Veen. The neural network zoo. 2016. <https://www.asimovinstitute.org/neural-network-zoo>
14. Саймон Хайкин. Нейронные сети: полный курс = Neural Networks: A Comprehensive Foundation. 2-е изд. – М.: Вильямс, 2006.
15. Jahn M. Artificial neural network regression models in a panel setting // Predicting economic growth. September, 2020. – Vol. 91. – pp. 148-154.
16. Michael Štencl and Jiří Šťastný. Artificial Neural Networks Numerical Forecasting of Economic Time Series. Artificial Neural Networks – Application. Publisher InTech, 2011.
17. Wei huang, Kin Keung Lai, Yoshiteru Nakamori, Shouyang Wang, lean yu. neural networks in finance and economics forecasting // International Journal of Information Technology & Decision Making, 2007. – Vol. 6. – № 1. – pp. 113-140.