

УДК 004.89

<https://doi.org/10.21869/2223-1560-2024-28-4-104-123>

Метод формирования многоярусной нейросетевой системы прогнозирования с возможностью реконфигурации

А.К. Крутиков¹ ✉, В.Ю. Мельцов¹ ✉

¹ Вятский государственный университет
ул. Московская, д. 36, г. Киров 610000, Российская Федерация

✉ e-mail: yadrodisk@yandex.ru

Резюме

Цель исследования. Повышение точности прогнозирования за счёт выявления логических связей в неструктурированных наборах данных и формирования многоярусной структуры специализированной нейросетевой вычислительной системы

Методы. Предложен параллельный алгоритм определения фрагментированной структуры обучающей выборки, используемый для выделения фрагментов, содержащих обучающие данные, на основе логических зависимостей выборки. На базе сгенерированной фрагментированной выборки разработан метод ассемблирования нейронных сетей, используемый для формирования эффективной структуры каскадной системы прогнозирования.

Результаты. В качестве основного эксперимента выбрано прогнозирование результатов неофициального командного зачета Международного фестиваля студенческого спорта 2023. Сформирована фрагментированная обучающая выборка на основе которой построен каскад нейросетевых модулей. В экспериментах протестировано четыре варианта каскада, показавших существенное повышение точности прогнозирования по сравнению с одномодульными аналогами. Для значительного повышения производительности нейросетевой системы при сверхкраткосрочных прогнозах рассмотрена аппаратная реализации каскадов на основе решающего поля ПЛИС. Предложена структура комплекса с возможностью его реконфигурации.

Заключение. Применение искусственных нейронных сетей в прогнозировании перспективно, однако может сталкиваться с проблемами неточности результатов из-за недостаточной вычислительной мощности и коллизий в обучающих выборках. Один из предложенных вариантов решения проблемы – каскадирование специализированных нейросетевых модулей. Положительные результаты продемонстрировали как программная, так и аппаратная реализация системы на основе предложенного каскада. Оценка аппаратной реализации демонстрирует возможность ускорения, по сравнению с программной реализацией, что может быть необходимо при проведении сверхкраткосрочных прогнозов. Предложенные методы и алгоритмы продемонстрировали свою корректность.

Ключевые слова: искусственная нейронная сеть; прогнозирование; обучающая выборка; каскадирование; ПЛИС; решающее поле.

Конфликт интересов: Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

© Крутиков А.К., Мельцов В.Ю., 2024

Для цитирования: Крутиков А.К., Мельцов В.Ю. Метод формирования многоярусной нейросетевой системы прогнозирования с возможностью реконфигурации // Известия Юго-Западного государственного университета. 2024. Т. 28, №4. С. 104-123. <https://doi.org/10.21869/2223-1560-2024-28-4-104-123>.

Поступила в редакцию 05.11.2024

Подписана в печать 28.11.2024

Опубликована 10.12.2024

A method for forming a multi-tiered neural network forecasting system with the possibility of reconfiguration

Alexander K. Krutikov ¹ ✉, Vasily Y. Meltsov ¹

¹ Vyatka State University
36, Moskovskaya str., Vyatka 610000, Russian Federation

✉ e-mail: yadrodisk@yandex.ru

Abstract

Purpose of research. Improving the accuracy of forecasting by identifying logical connections in unstructured datasets and forming a multi-tiered structure of a specialized neural network computing system.

Methods. A parallel algorithm for determining the fragmented structure of the training sample is proposed, which is used to isolate fragments containing training data based on the logical dependencies of the sample. Based on the generated fragmented sample, a method for assembling neural networks has been developed, which is used to form an effective structure of a cascade forecasting system.

Results. Forecasting the results of the unofficial team competition of the International Student Sports Festival 2023 was chosen as the main experiment. A fragmented training sample has been formed on the basis of which a cascade of neural network modules has been built. Four cascade variants were tested in experiments, which showed a significant increase in prediction accuracy compared to single-module analogues.

To significantly improve the performance of a neural network system with ultra-short-term forecasts, the hardware implementation of cascades based on the decisive field of FPGA is considered. The structure of the complex with the possibility of its reconfiguration is proposed.

Conclusion. The use of artificial neural networks in forecasting is promising, but it may face problems of inaccuracy of results due to insufficient computing power and collisions in training samples. One of the proposed solutions to the problem is cascading specialized neural network modules. Positive results were demonstrated by both the software and hardware implementation of the system based on the proposed cascade. The evaluation of the hardware implementation demonstrates the possibility of acceleration, compared with the software implementation, which may be necessary when conducting ultra-short-term forecasts. The proposed methods and algorithms have demonstrated their correctness.

Keywords: artificial neural network; forecasting; training sampling; cascading; FPGA; decision field.

Conflict of interest. The authors declare the absence of obvious and potential conflicts of interest related to the publication of this article.

For citation: Krutikov A. K., Meltsov V. Y. A method for forming a multi-tiered neural network forecasting system with the possibility of reconfiguration. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta = Proceedings of the Southwest State University*. 2024; 28(4): 104-123 (In Russ.). <https://doi.org/10.21869/2223-1560-2024-28-4-104-123>.

Received 05.11.2024

Accepted 28.11.2024

Published 10.12.2024

Введение

Прогнозирование играет ключевую роль во многих областях человеческой деятельности, поскольку оно позволяет предвидеть тенденции и развитие событий в будущем на основе анализа имеющихся данных и текущих трендов. Эти знания помогают принимать взвешенные решения и эффективно планировать свою деятельность. Прогнозирование широко используется в различных отраслях, таких как бизнес, финансы, транспорт, энергетика, здравоохранение и многих других. В финансовой сфере прогнозирование помогает анализировать риски и принимать обоснованные инвестиционные решения. В транспортной области – повышает качество логистических решений. В энергетике без прогнозирования и планирования затрат электроэнергии на краткосрочный и долгосрочный периоды невозможно эффективно организовать процесс выработки энергоресурсов. В здравоохранении оно используется и для прогнозирования развития сложных заболеваний, и для оценки вероятности возникновения и распространения эпидемий. Одной из областей, в которых прогнозирование получило второе дыхание благодаря мощному развитию суперкомпьютерных систем и информационных технологий, является область высоких спортивных достижений. Прогнозирование численных результатов позволяет планировать физическую

подготовку как отдельных спортсменов, так и сборных команд в целом. В случае сверхдолгосрочных прогнозов мы получаем возможность оценивать динамику развития атлетов, начиная с самого юного возраста.

Системы прогнозирования и планирования спортивных результатов используют различные методы и алгоритмы для анализа данных и составления прогнозов. Многие из них базируются на аппарате математической статистики [1, 2]. Статистические методы позволяют выявлять закономерности и тенденции, например, на основе регрессионного или корреляционного анализа [3]. Корреляционный анализ наиболее эффективен для прогнозирования результатов команд в игровых видах спорта. Он позволяет выявить взаимосвязь между различными факторами, такими как результаты последних матчей у каждой команды, результаты личных встреч, статистика игроков на данный период, погода в день матча и т. д. Как ни странно, достаточно точные результаты можно получить, используя метод «экспертной оценки». Эксперты в области спорта могут составить эффективный план подготовки индивидуального спортсмена к определённым соревнованиям в течение года, дать свою оценку возможных результатов призёров и победителей этих соревнований. Эти экспертные оценки основаны на опыте, знаниях и даже, в некоторых случаях, интуиции экспертов.

В последние годы одним из наиболее эффективных инструментов прогнозирования спортивных событий оказались искусственные нейронные сети (ИНС). Нейронные сети могут учитывать огромное количество факторов, иногда упускаемых из вида даже экспертами. Разработка специализированных систем прогнозирования [4], в том числе в спортивной сфере, одно из динамично развивающихся направлений в области IT-технологий.

Нейронные сети успешно решают задачи прогнозирования, так как они способны обрабатывать большие объемы данных и находить закономерности в них. Преимущества использования нейронных сетей в прогнозировании включают высокую точность предсказаний за счет обучения на больших объемах данных, возможность работы с нелинейными зависимостями и способность адаптироваться к изменяющимся условиям. Нейронные сети также могут автоматически извлекать признаки из данных, что делает их еще более эффективным инструментом. Однако стоит отметить, что некоторые архитектурные решения на базе нейронных сетей требуют больших объемов вычислительных ресурсов для обучения и работы. Кроме того, интерпретация результатов работы нейронных сетей может быть затруднительной, что усложняет их использование в некоторых областях прогнозирования [5, 6].

В работе [7] описывается специализированная вычислительная система, ко-

торая предназначена для решения задач прогнозирования на основе предварительно обработанных неструктурированных данных. Особенностью этой системы является то, что её операционная часть базируется на нейросетевых механизмах. Авторы работы подробно описывают процесс создания модели и структурных решений системы на базе искусственных нейронных сетей (ИНС). Также приводятся примеры использования системы для решения конкретных задач, в частности речь идет о прогнозе результатов в индивидуальной легкой атлетической дисциплине. Определение логических зависимостей производится экспертами в данном виде спорта [7]. Однако определение ряда логических зависимостей возможно автоматизировать, используя различные математические и статистические методы.

Для прогнозирования многопараметрических событий необходимо заранее подготовить релевантную обучающую выборку. Обучающая выборка представляет собой набор данных, содержащий некоторые известные факты за предыдущие временные периоды. Важность подготовки качественной обучающей выборки заключается в том, что для ИНС именно она определяет качество и точность модели обучения. Чем больше и разнообразнее данные в обучающей выборке, тем более точно модель сможет обобщать информацию и делать более точные прогнозы на основе поступающих данных. Если же обучающая выборка недостаточно объёмна или не ре-

презентативна, то модель может столкнуться с проблемой переобучения. Стандартная обучающая выборка имеет вид двумерного массива, где i -я строка соответствует одной из предикторных характеристик, а элементы строки заполняются реальными значениями этой характеристики в определённые моменты

времени. Тогда получаем, что каждая строка двумерного массива представляет собой вектор, содержащий набор численных значений выделенной характеристики за указанный период наблюдения [8]. Этому набору в следующем фрагменте выборки сопоставлен известный итоговый результат (рис. 1).

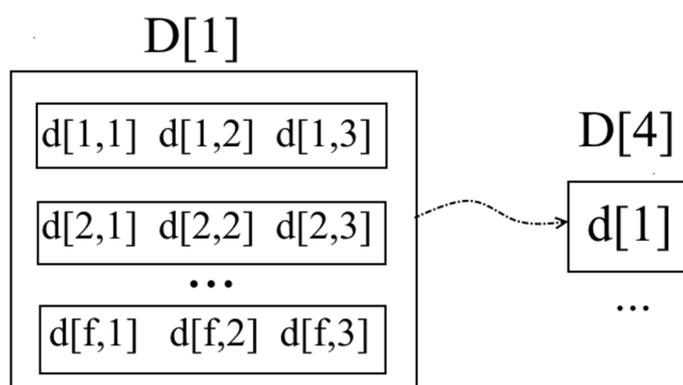


Рис.1. Логические зависимости в обучающей выборке

Fig. 1. Logical dependencies in the training sample

Материалы и методы

В данной работе при подготовке фрагментированной обучающей выборки предлагается исходный двумерный массив разбивать на меньшие двумерные массивы. Главный принцип разбиения – наличие логических связей между элементами, при этом элемент не должен быть итоговым результатом. Определение взаимозависимостей между фрагментами обучающей выборки (выбранными характеристиками наблюдения) основано на наличии выявленных ранее логических связей между указанным вектором исходного фрагмента и элементами столбца последующего фрагмента [7]. Количество «связываемых» фрагментов не ограниче-

но, а логические связи могут быть установлены либо с помощью специального анализа, выявляющего функциональные зависимости в структурированных наборах данных, либо на основе экспертного мнения специалиста-профессионала в данной предметной области [9]. Таким образом, фрагменты обучающей выборки выстраиваются в логическую последовательность или несколько последовательностей. Пример полной структуры обучающей выборки приведен на рис. 2.

Такая сложная структура выборок влияет и на детализированную структуру специализированной вычислительной системы прогнозирования. Эффективным решением будет использовать метод ассемблирования нейросетевых модулей в

составе комплекса. На основе сформированных последовательностей фрагментов выборки определяется количество уровней нейронного каскада и количество модулей на каждом ярусе. Тип нейронной сети, реализуемой в модуле, определяется задачей, решаемой на данном этапе прогнозирования. Начиная со второго уровня, в каждом модуле обрабатываются данные промежуточных прогнозов, поступающие от предыдущего яруса.

Выделяются два типа используемых переменных: предикторы, или независимые переменные, и целевые признаки - зависимые переменные, которые вычисляются на основе предикторов. Количество предикторов формирования прогнозов от уровня к уровню может как уменьшаться, так и увеличиваться. Сгенерированные фрагменты обучающей выборки подаются на соответствующие модули всех ярусов каскадной вычислительной системы.

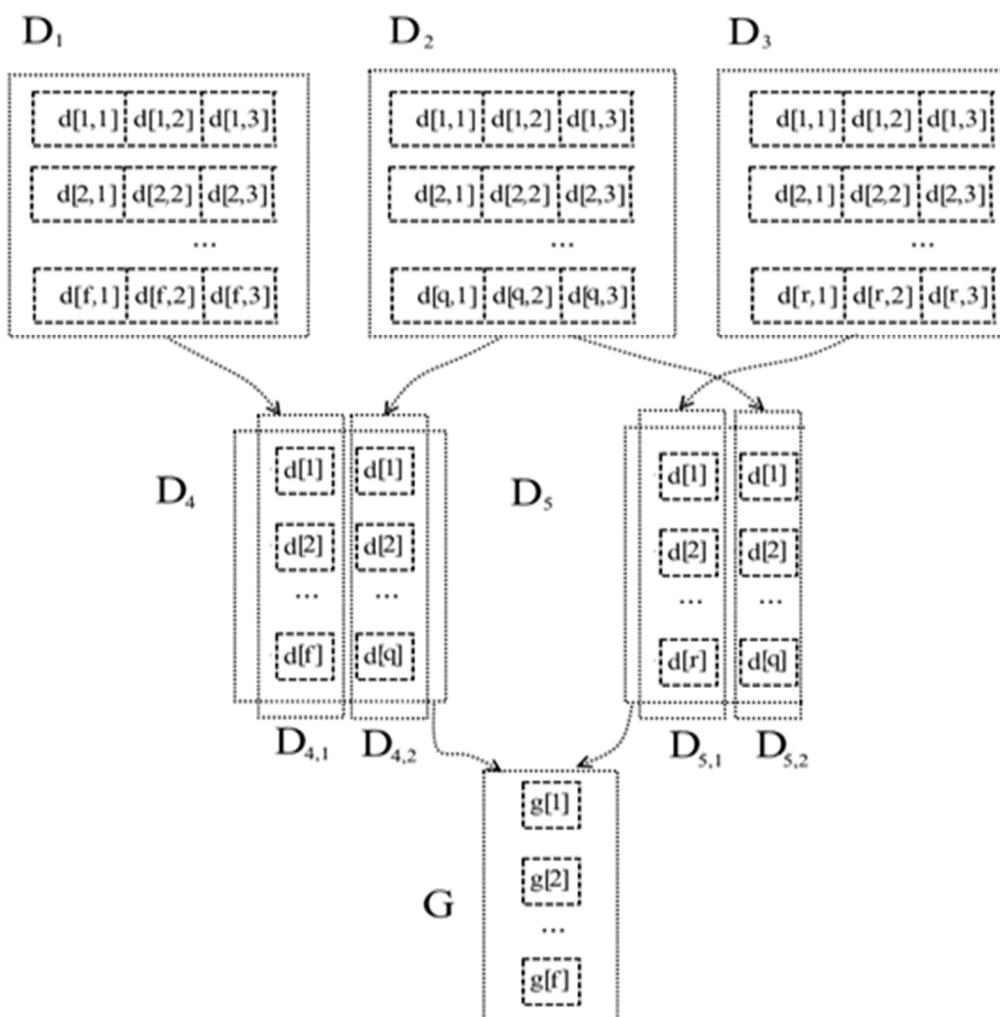


Рис. 2. Полная структура обучающей выборки

Fig. 2. The complete structure of the training sample

Некоторые параметры могут оказывать наибольшее влияние на промежуточные результаты прогнозирования и соответствующий итоговый результат. Такие параметры будем называть *ключевыми предикторами*. Предиктором, который является ключевым для формирования прогноза, называется тот, чей исключение из обучающей выборки приводит к увеличению ошибки обобщения. Величина ошибки обобщения определяет, насколько влияние этого предиктора влияет на конечный результат прогноза. В прикладных задачах, где существуют ограничения на время обучения или аппаратные затраты, выявление подобных характеристик позволит уменьшить размер обучающей выборки без значительных потерь в точности прогноза. В случае ассемблирования модулей, выявление ключевых предикторов на каждом ярусе позволяет сокращать столбцы и/или строки в фрагментах выборки и более «тонко» настраивать архитектуру нейросетевой системы для решения задач прогнозирования с указанными пользователем ограничениями.

На начальном этапе исследований были разработаны два прототипа системы: первый содержал только RBF-модуль [10], второй – только GRNN-модуль [11]. RBF-сети (нейронные сети с радиально-базисными функциями) строятся с использованием радиальных нейронов, которые имеют ненулевые значения только вблизи своего центра. Обучение RBF-сетей проходит в два этапа: настройка центров радиальных

функций и, затем, обучение выходного слоя. Проблема возникла при необходимости добавления некоторых параметров, которые также достаточно значимо могут влиять на прогнозируемый результат. Включение этих параметров в модель потребовало изменения структуры обучающей и исходной выборки, изменения архитектуры прототипа и переобучения соответствующих модулей нейросетевой системы.

GRNN-сеть, которая является разновидностью радиально-базисных нейронных сетей, обычно используется для непараметрической регрессии [10]. Обучение GRNN-модулей также проходило в два этапа: сначала настраивались центры радиальных функций, затем нейроны выходного слоя, с учётом фиксированных параметров радиальных нейронов. Основным недостатком GRNN-сети является ее размерность – для достижения хорошей точности потребовалось большое количество нейронов. Однако уменьшение обучающей выборки приводит и к уменьшению обобщенно-регрессионной нейронной сети.

В случае применения каскадной архитектуры предполагается формирование промежуточных прогнозных значений на каждом ярусе, передачу их на следующий ярус и формирование итогового прогноза. На первом ярусе все предикторы «собираются» в соответствующие фрагменты, количество которых определяет количество нейросетевых модулей на этом ярусе. Для каждого модуля системы обучающая выборка

формируется отдельно, и модули обучаются заранее. С учётом указанных выше особенностей был разработан третий прототип специализированной нейросетевой системы прогнозирования, содер-

жащий как GRNN, так и RBF-модули, организующие трёхъярусную архитектуру системы. Обобщённая структура каскада приведена на рис. 3.

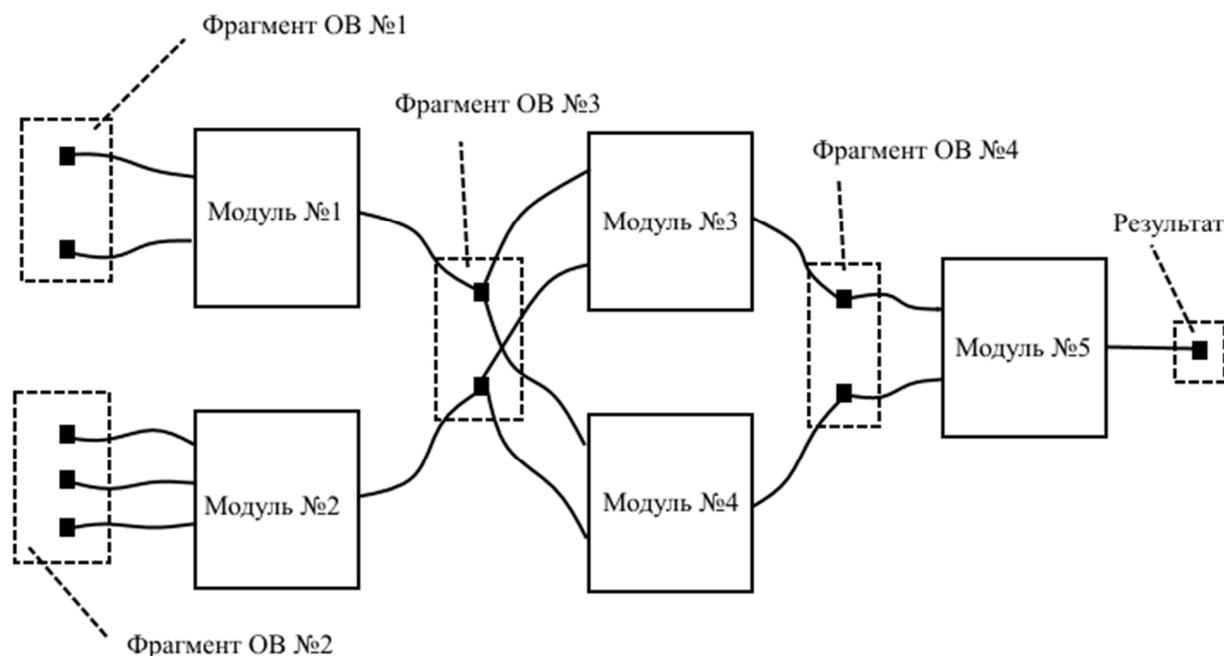


Рис.3. Обобщённая структура каскада нейросетевых модулей

Fig. 3. Generalized structure of the cascade of neural network modules

При использовании каскада, параметры, которые раньше не включались в выборку, динамически формируются в процессе прогнозирования на каждом ярусе и передаются между модулями каскада с одного уровня на другой. Эти значения являются промежуточными результатами прогнозирования и одновременно формируют предикторы для модулей последующих уровней. При обучении каскадной модели каждый модуль обучался по отдельности и включался в состав каскада после завершения обучения.

Результаты и их обсуждение

В качестве эксперимента рассмотрим прогнозирование результатов сборной команды Российской Федерации в неофициальном командном зачете на Международном фестивале студенческого спорта в 2023 году. Обычно, место в НКЗ определяется по основной системе оценки – количеству золотых медалей, завоеванных командой какой-либо страны. Однако, зачастую, применяется и более широкая рейтинговая система, которая учитывает общее количество медалей,

не разделяя их по достоинству. Для обучения нейросетевых прототипов была сформирована обучающая выборка, включающая результаты выступлений российских сборных с 1992 года, по всем видам спорта, входящим в программу Международного фестиваля. Анализ зарубежных [12] и отечественных [13] научных исследований в данной области показал, что для релевантного прогноза по НКЗ необходимо учитывать не только спортивные показатели команд, но и экономические и социальные факторы, влияющие на конечный результат выступления сборных команд. В частности, в указанных экспериментах в обучающие выборки были включены такие показатели, как численность населения страны, валовой внутренний продукт (ВВП) на душу населения, бюджеты Министерства спорта и отдельных Федераций, количество спортсменов, входящих в состав делегации на Фестиваль, удалённость места проведения соревнований от основных тренировочных баз (для учёта акклиматизации), фактор «родных стен» (на основе анализа выступлений стран-организаторов подобных спортивных мероприятий) и т.д.

При разделении обучающей выборки на фрагменты учитывались логические связи между исходными данными, выявленные на основе экспертных знаний [7]. Важно, что сформированные фрагменты выборок могут содержать как уникальные данные (включённые в состав только

указанного фрагмента), так и отдельные, повторяющиеся в нескольких фрагментах, данные, в зависимости от специфики каждого фрагмента.

Проанализируем результаты проведённых экспериментов на двух первых прототипах – на основе RBF-сетей и GRNN-сетей. Стоит отметить, что в обоих случаях проводилось по два отдельных эксперимента. Поскольку существуют разные системы оценок, то после прогнозирования места Российской команды по первой шкале (количество золотых медалей), необходимо было переобучать нейросетевые модули на другую шкалу (общее количество медалей). Настраиваемыми параметрами в RBF- и GRNN-прототипах являлись параметр влияния (SPREAD), определяющий крутизну функции radbas , и целевая ошибка обучения, соответственно. Параметр влияния – это параметр, который используется в радиально-базисных сетях для контроля степени влияния базисных функций на выходные значения. Он определяет ширину базисных функций и влияет на способность сети к обобщению и точности генерируемых прогнозов. Чем больше значение SPREAD, тем шире становятся базисные функции и тем меньше они реагируют на изменения входных данных. Это может привести к более обобщённым решениям, но также может снизить точность предсказаний для конкретных случаев. Результаты прогнозирования сведены в табл. 1 и табл. 2.

Таблица 1. Результаты экспериментов (RBF-сеть)**Table 1.** Experimental results (RBF network)

RBF-сеть / RBF network				
Золотые медали / Gold medals		Общее число медалей / Total number of medals		Параметр влияния / SPREAD
Прогноз (количество) / Forecast (quantity)	Прогноз (место) / Forecast (place)	Прогноз (количество) / Forecast (quantity)	Прогноз (место) / Forecast (place)	
167	1	507	1	0,0003
167	1	508	1	0,003
168	1	508	1	0,03
168	1	509	1	0,3

Таблица 2. Результаты экспериментов (GRNN-сеть)**Table 2.** Experimental results (GRNN network)

GRNN-сеть / GRNN network				
Золотые медали / Gold medals		Общее число медалей / Total number of medals		Целевая ошибка / GOAL
Прогноз (количество) / Forecast (quantity)	Прогноз (место) / Forecast (place)	Прогноз (количество) / Forecast (quantity)	Прогноз (место) / Forecast (place)	
128	2	458	2	0.001
128	2	458	2	0.01
128	2	459	2	0.1
131	1	459	2	1
131	1	459	2	2

Как видим, результаты по прогнозу количества золотых медалей если и различаются, то не так значительно. А вот прогнозирования общего количества

медалей имеют существенные отличия. Отложим детальный сравнительный анализ до окончания экспериментов с каскадной вычислительной системой.

Процессы обучения и прогнозирования при ассемблировании нейросетевых модулей имеют ряд положительных моментов. Во-первых, при наличии достаточно большого и репрезентативного набора исходных данных, обучение всех модулей системы можно проводить параллельно (одновременно). Таким образом, время подготовки каскадного прототипа к работе никак не изменится из-за наличия нескольких нейросетевых модулей.

Во-вторых, изменение структуры обучающей выборки привело к повышению эффективности прогнозирования за счёт существенной детализации полученных результатов. Обучающая выборка для первого уровня модулей разделена на два больших компонента. В первый компонент включены социальные (в том числе, экономические) параметры, а во второй – спортивные показатели, которые были описаны ранее при обучении отдельных ИНС. После первого уровня формируются фрагменты выборки, содержащие количество медалей различного достоинства,

завоёванных спортсменами указанной сборной. Стоит отметить, что во втором фрагменте объединены показатели количества как золотых, так и серебряных наград. Это сделано для повышения точности прогноза по главной шкале НКЗ, так как при равенстве у нескольких команд золотых медалей, учитывается количество серебряных наград. Соответственно, полученные на данном ярусе факторы являются промежуточными результатами для генерации итоговых прогнозов. А итоговых результатов прогноза можно получить сразу два, без переобучения нейросети. И это третье преимущество выбранного подхода с ассемблированием модулей. Первый результат, место сборной команды России по количеству золотых медалей, будет сформирован на выходе второго яруса. А второй, место российской сборной по общему количеству медалей, генерирует нейросетевой модуль третьего яруса. Общая фрагментированная структура выборки приведена на рис. 4.

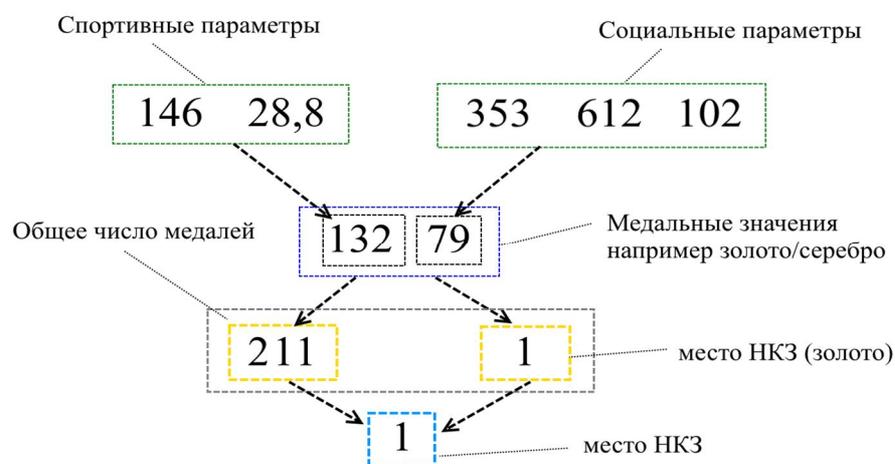


Рис. 4. Структура обучающей выборки в эксперименте

Fig. 4. The structure of the training sample in the experiment

Также важно, что практически половина значений обрабатываемых параметров могут отсутствовать в исходных выборках для начала процесса прогнозирования, так как динамически формируются в процессе прогнозирования в модулях каждого яруса и передаются с одного уровня на другой.

Программная симуляция каскадной вычислительной системы проводилась в

среде моделирования MATLAB с использованием пакета Deep Learning Toolbox [14]. Также, некоторые модули были написаны на языке программирования Python с использованием библиотеки Keras. Трёхъярусная структура была реализована в четырёх вариантах, при различном сочетании нейросетевых модулей (рис. 5).

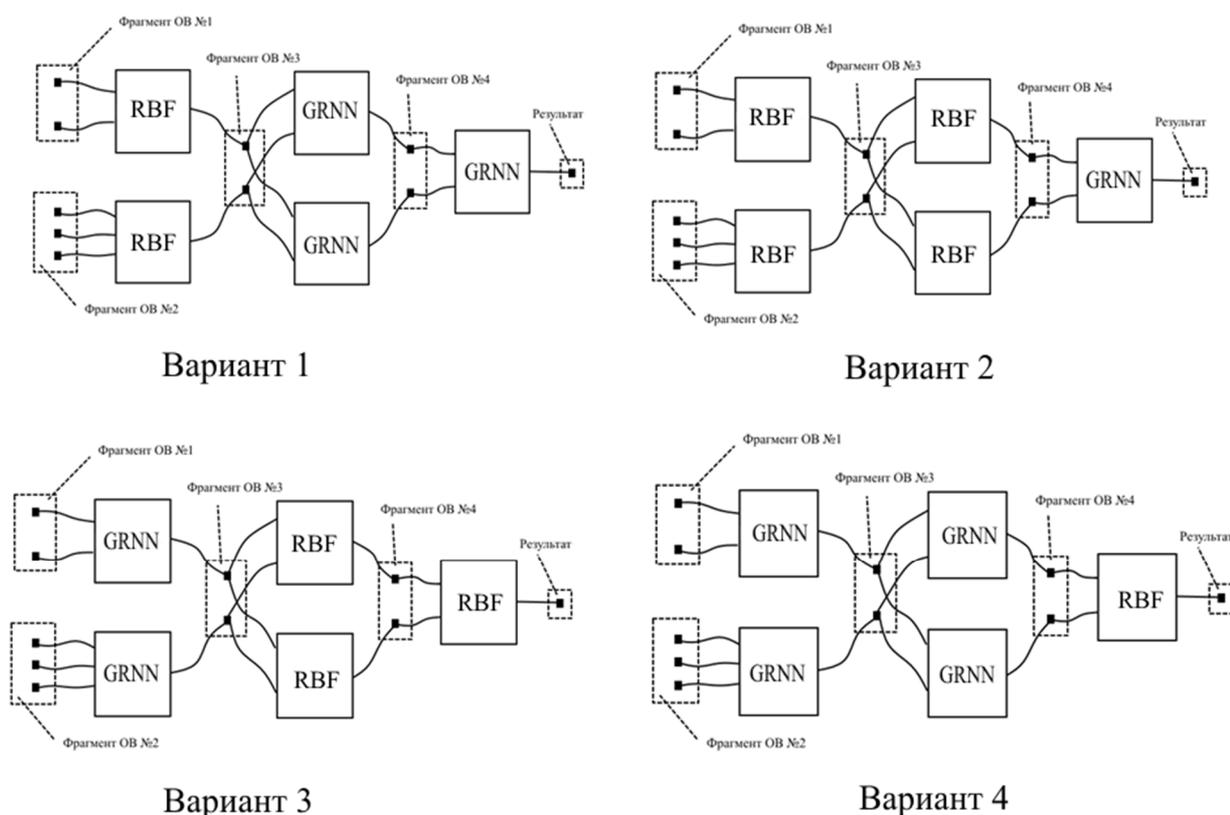


Рис. 5. Варианты реализации трёхъярусной нейросетевой вычислительной системы

Fig. 5. Implementation options for a three-tier neural network computing system

Результаты проведённых экспериментов на каскадной системе по прогнозированию места Российской сборной в НКЗ приведены в табл. 3.

Для анализа полученных результатов прогнозирования на всех трёх прототипах необходимо иметь фактические итоги участия Российской команды на

Международном фестивале студенческого спорта в 2023 году. А они следующие: золотых медалей – 144, серебряных медалей – 138, бронзовых медалей 199, НКЗ по золотым медалям (144) – 1 место, НКЗ по общему числу медалей (481) – 1 место.

Таблица 3. Результаты экспериментов (каскадная система)**Table 3.** Experimental results (cascade system)

Каскад / Cascade	Количество медалей (золото/серебро) / Number of medals (gold/silver)	Место в НКЗ (золото) / A place in the team standings (gold)	Количество медалей (общее) / Number of medals (total)	Место в НКЗ (общее) / Place in the team standings (general)
Вариант 1	153/128	1	494	1
Вариант 2	154/139	1	495	1
Вариант 3	136/142	1	472	1
Вариант 4	137/139	1	471	1

Таким образом, вернувшись к результатам экспериментов на первых двух прототипах с одномодульными вариантами, получаем отклонение по количеству прогнозируемых золотых медалей порядка 17% для RBF-сетей и 11% для GRNN-сетей. А по общему количеству медалей – порядка 5,5% и 5%, соответственно.

При анализе результатов на каскадной вычислительной системе видим, что точность прогнозов заметно увеличилась. Хотя варианты 1 и 2 (с RBF-модулем на первом ярусе) показывают всё ещё существенные отклонения от фактических показателей: по золотым медалям – около 7%, а по общему количеству медалей – около 6,5%, однако варианты 3 и 4 сформировали неплохие прогнозы: количество золотых медалей – отклонение примерно 5,5%, общее количество медалей – отклонение 2% и 2,5%, соответственно.

Стоит отметить, что позже были проведены эксперименты на системе, в которой на всех трёх ярусах использовались только GRNN-сети. Результаты данного прототипа получились чуть менее точные, по сравнению с вариантом 4, на 0,4% и 0,2% соответственно.

Операционное устройство системы прогнозирования может быть реализовано как программно, так и аппаратно. Возможное ускорение при аппаратной реализации описано в работе [15]. При этом, необходимо учитывать целесообразность аппаратной реализации (например, при выполнении сверхкраткосрочных прогнозов), и наличие необходимых ресурсов.

Одним из наиболее эффективных вариантов аппаратной реализации искусственного нейрона и всей нейросетевой вычислительной системы является использование программируемых логических интегральных схем (ПЛИС). Требования к ПЛИС определяются дву-

мя основными характеристиками разрабатываемой системы прогнозирования: быстродействием и требуемой точностью [15]. Соответственно, с аппаратной стороны важнейшими параметрами являются тактовая частота ПЛИС и количество логических элементов на кристалле. Для оценки возможности реализации операционной части системы необходимо оценить количество нейронов, которое возможно разместить на одной ПЛИС.

Рассмотрим аппаратную реализацию второго прототипа каскадной вычислительной системы (GRNN-сеть) на базе FPGA. Расчёты показали, что нам потребуется: для всех регистров, включая буферные, – около 33 тысячи вентилях; для логических элементов “И” – почти 1.4 тысяч вентилях; для логических элементов “ИЛИ” – около 170 вентилях; для всех сумматоров – 3070 вентилях, плюс четыре вентиля для D-триггера. Также, на синтез всех счетчиков необходимо будет выделить примерно 1050 вентилях. Наиболее сложной частью является реализация мультиплексоров, для которых потребуется около 64 тысяч вентилях. Выходные комбинационные схемы требуют еще 6 тысяч вентилях, или 24 вентиля на разряд. В результате, получаем, что для реализации нейронной сети типа GRNN на ПЛИС потребуется примерно 119 тыс. вентилях. При ориентации авторов статьи на бюджетные ПЛИС [16] важно понимать, что при увеличении числа нейронов и размерности входного век-

тора данных количество вентилях значительно возрастет. Например, при увеличении числа нейронов вдвое (до 32) количество вентилях увеличится более чем в два раза (примерно в 3,7 раза). Это связано с тем, что добавления нейронов в GRNN-модель приводит к значительному увеличению числа связей. Следовательно, входные мультиплексоры для каждого нейрона, в том числе для вновь добавляемых нейронов, станут сложнее, и потребуются дополнительные вентилях для новых регистров весов.

Именно «гибкость» ПЛИС, при изменении структуры обучающей выборки, позволяет менять и структуру каскада. При этом, сохраняется возможность изменения структуры отдельного модуля.

При аппаратной реализации полноценных каскадов из нейросетевых модулей (многоярусные варианты системы прогнозирования) предлагается использовать специализированные структуры на основе решающих полей [17,18], содержащие от 4 до 8 ПЛИС. Определенную сложность при проектировании аппаратной части решающего поля вносит принципиальное различие архитектур нейросетевых модулей, применяемых при разработке системы. Одним из эффективных подходов к решению данной проблемы является формирование библиотеки базовых вариантов нейросетевых модулей в САПР и добавление инструментов для реконфигурации (модификации) как архитектуры отдельных модулей, так как и структурных решений

операционных устройств в целом. Пример структурной схемы нейросетевого модуля со встроенным блоком реконфигурации приведен на рис. 6. При этом, необходимо понимать, что количество вариантов структуры нейронной сети в одном блоке реконфигурации бу-

дет ограничено. Также в рамках всей каскадной вычислительной системы имеется модуль конфигурирования, позволяющий синтезировать необходимое количество ярусов из трёх возможных типов нейросетевых модулей.

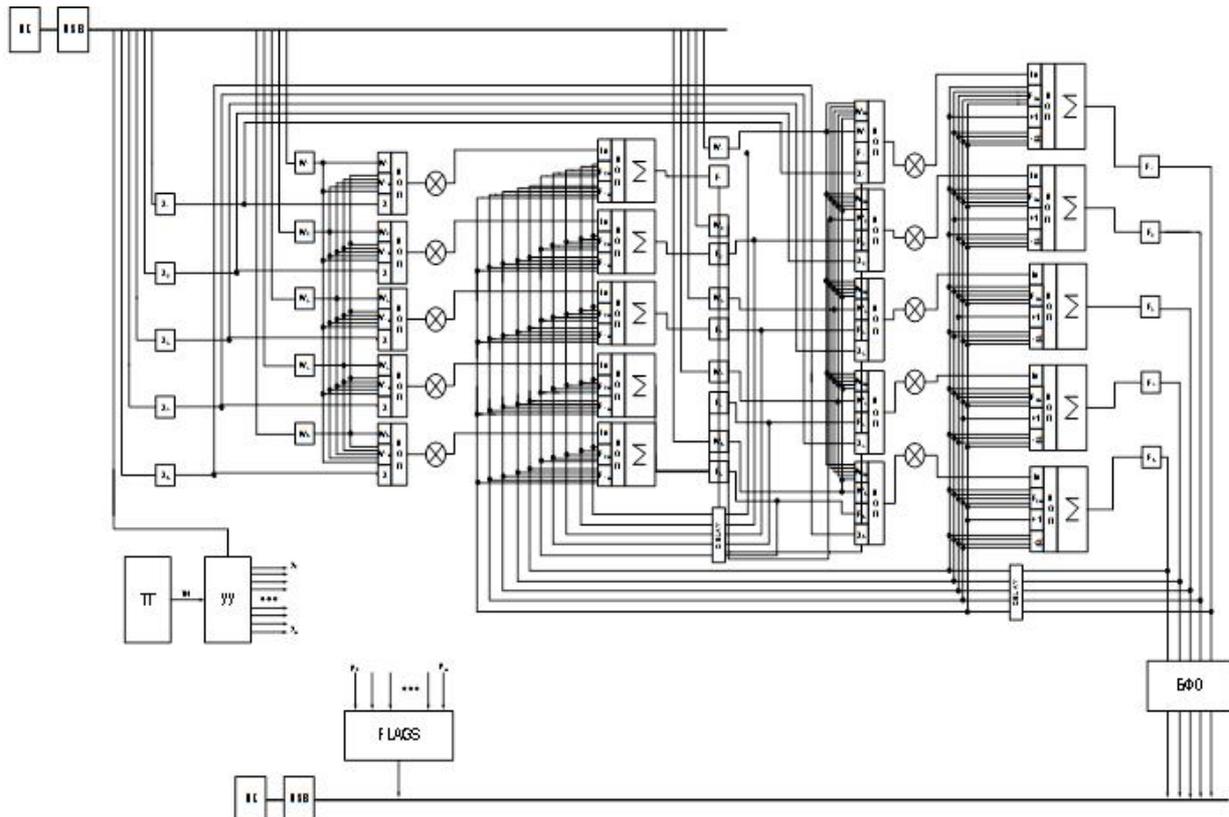


Рис. 6. Структура перестраиваемого нейросетевого модуля

Fig. 6. The structure of the reconfigurable neural network module

Выводы

Применение искусственных нейронных сетей в области прогнозирования является, несомненно, перспективным подходом. Однако в некоторых случаях, результаты могут быть недостаточно точным или некорректным из-за недостаточной вычислительной мощности системы и наличия коллизий и противоре-

чий в сформированных обучающих выборках. Эти проблемы можно решить, например, с помощью каскадирования специализированных нейросетевых модулей, выделения взаимосвязей между элементами фрагментированной обучающей выборки и передачи промежуточных прогнозных значений между ярусами каскада. Фрагментирование выборок также приводит к тому, что уменьшается

количество обрабатываемых параметров на первом уровне и, соответственно, возможно сокращение времени обучения при использовании параллельного метода обучения всех модулей системы одновременно.

Использование предобработанных данных на последующих уровнях повышает точность прогнозирования. Проблема предобработки обучающих данных может быть решена стандартными методами, такими как замена переменных или отбрасывание неполных векторов [19].

Рассмотренная возможность аппаратной реализации каскадной вычислительной системы на базе современных ПЛИС позволяет значительно, почти в 40 раз, сократить время прогнозирования, что крайне важно при проведении краткосрочных и сверхкраткосрочных прогнозов. Отличительной особенностью, предлагаемой авторами аппаратной реализации, является введение в состав системы, и каждого модуля в отдельности, блоков реконфигурации, для подбора наиболее эффективных архитектурно-структурных решений нейросетевого устройства.

В свою очередь использование решающих полей на базе ПЛИС, реконфигурируемых полей [20], позволит разместить каскад и произвести полную его аппаратную реализацию, без программной эмуляции отдельных блоков или модулей. Целесообразность аппаратной реализации оценивается исходя из материально-технических возможностей, необходимости получения кратко-

срочных и сверхкраткосрочных прогнозов, а также оценки потенциальных изменений структуры обучающих данных и структуры самой системы.

Ещё одним важным результатом данных исследований является определение ключевых предикторов, наиболее сильно влияющих на точность прогнозирования. При необходимости, сокращения аппаратных затрат на ПЛИС или сокращения общего времени решения задачи прогнозирования, некоторые «неключевые» параметры могут быть исключены из исходных выборок. Влияние указанных действий на точность прогнозирования может быть оценено предельной ошибкой обобщения. В разработанном варианте специализированной нейросетевой системы пользователь, после проведённых экспериментов, имеет возможность вручную исключить некоторые параметры выборки. В свою очередь, выделение ключевых предикторов, уменьшает объем выборки. Это позволяет сократить время обучения и модулей каскада. При необходимости можно перестроить каскад на основе обучающей выборки, содержащей только ключевые предикторы, что приведет к снижению затрат на программное и аппаратное обеспечение. В дальнейшем алгоритм определения ключевых предикторов будет автоматизирован и в состав системы будет добавлен блок для автоматической реконфигурации модулей и блоков операционной части с учетом анализа значений предельной ошибки.

Описанный в работе подход позволяет улучшить производительность и эффективность работы системы за счет оптимизации структуры нейросетевого каскада под конкретные задачи прогнозирования, а использование современных ПЛИС, в свою очередь, позволяет ускорить вычисления. Использование

предобработанных данных и аппаратной реализации на базе ПЛИС в каскадных вычислительных системах может значительно улучшить точность и скорость прогнозирования, что делает этот подход перспективным для решения задач прогнозирования в различных областях.

Список литературы

1. Baltagi B.H., Li D. Prediction in the panel data model with spatial correlation // Springer, Berlin: *Advances in Spatial Econometrics: Methodology, Tools and Applications*. 2004. Chapter 13. P. 283–295.
2. Kholodilin K.A., Siliverstovs B., Kooths S. A dynamic panel data approach to the forecasting of the GDP of German Länder // *Spatial Economic Analysis*. 2008. Vol. 3. P. 195–207.
3. Белоущенко Я. А. Статистические методы прогнозирования в экономике. Керчь: ФГБОУ ВО КГМТУ, 2016. 44 с.
4. Реуцкая Е.А. Комплексная система диагностики лыжников-гонщиков с целью определения перспективности и прогнозирования предраположенности к высоким спортивным результатам // *Наука и спорт: Современные тенденции*. 2022. Т. 10, №4. С. 79-87.
5. Жуков С.В., Зеленский А.В. Использование нейронных сетей в построении оптимальной тренировочной траектории в биатлоне // *Учёные записки университета им. П.Ф. Лесгафта*. 2017. № 5 (147). С. 81–88.
6. Юшкин В.Н. Проблемы спортивного прогнозирования // *Ученые записки Уиверситета им. П.Ф. Лесгафта*. 2021. №4(194). С. 473-478.
7. Крутиков А.К., Параллельный алгоритм формирования фрагментированной обучающей выборки для специализированной системы прогнозирования на основе ассемблирования нейросетевых модулей // *Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления*. 2024. № 49. С. 61-81. DOI 10.15593/2224-9397/2024.1.04.
8. Розыходжаева Г.А., Розыходжаева Д.А. Особенности формирования обучающей выборки и обучения нейронной сети с неполными входными данными при решении частных медицинских задач // *Научное обозрение. Биологические науки*. 2017. № 5. С. 28-32.
9. Дорохова Ю.В., Гончарук Н.С. Использование метода экспертных оценок при прогнозировании социальных процессов // *Среднерусский вестник общественных наук*. 2014. №3(33). С. 76-81.

10. Жуков В.Г., Бухтояров В.В. О применении искусственных нейронных сетей с радиальными базисными функциями в задачах обнаружения аномалий в сетевом трафике // Решетневские чтения. 2013. Т. 2, № 17. С. 285-286.

11. Konakoglu B., Cakir L. Generalized Regression Neural Network for Coordinate Transformation // Montenegro, International Symposium on Advancements in Information Sciences and Technologies (AIST). 2018. P. 66-74.

12. Малютин А. Не счесть золота? // Forbes. URL: <http://www.forbes.ru/forbes/issue/2004-08/20514-ne-schest-zolota> (дата обращения: 21.05.2024).

13. Ясницкий Л.Н., Павлов И.В., Черепанов Ф.М. Прогнозирование Олимпийских Игр 2014 года в неофициальном командном зачете методами искусственного интеллекта // Современные проблемы науки и образования. 2013. № 6.

14. Deep Learning Toolbox. URL: <https://exponenta.ru/neural-network-toolbox> (дата обращения: 02.04.2024).

15. Krutikov A.K., Meltsov V.Y., Strabykin D.A. Evaluation the Efficiency of Forecasting Sports Events Using a Cascade of Artificial Neural Networks Based on FPGA // Proceedings of ElConRus-2022. St. Petersburg: ETU LETI, 2022. P. 355–360.

16. FPGA. URL: <https://alt.ru/Mikroshemi.html> (дата обращения: 01.02.2024).

17. Современные высокопроизводительные вычислительные системы с реконфигурируемой архитектурой на основе ПЛИС Xilinx Virtex-7 и Virtex UltraScale / И.И. Левин, А.И. Дордопуло, И.А. Каляев, Ю.И. Доронченко, М.К. Раскладкин // Russian Supercomputing Days 2015. 2015. С. 435-447.

18. Высокопроизводительные реконфигурируемые вычислительные системы нового поколения / А.И. Дордопуло, И.А. Каляев, И.И. Левин, Е.А. Семерников // Труды Международной суперкомпьютерной конференции с элементами научной школы для молодежи «Научный сервис в сети Интернет: экзафлопсное будущее». М.: Изд-во МГУ, 2011. С. 42-49.

19. Предобработка данных для обучения нейронной сети / Ю.П. Качановский, Ю.П. Качановский, Е.А. Коротков, Е.А. Коротков // Фундаментальные исследования. 2011. № 12-1. С. 117-120.

20. Гузик В.Ф., Каляев И.А., Левин И.И. Реконфигурируемые вычислительные системы. Ростов-на-Дону: изд-во Южного федерального университета, 2016. 471 с.

References

1. Baltagi B.H., Li D. Prediction in the panel data model with spatial correlation. Springer, Berlin: Advances in Spatial Econometrics: Methodology, Tools and Applications. 2004; 13: 283-295.

2. Kholodilin, K.A., Siliverstovs, B., Kooths, S. A dynamic panel data approach to the forecasting of the GDP of German Länder. *Spatial Economic Analysis*. 2008; 3: 195-207.

3. Belouschenko Ya. A. Statistical methods of forecasting in economics. Kerch; 2016. 44 p. (In Russ.).
4. Reutskaya E.A. A comprehensive diagnostic system for ski racers in order to determine the prospects and predict predisposition to high sports results. *Nauka i sport: Sovremennye tendentsii = Science and sport: Modern trends*. 2022. 10(4): 79-87. (In Russ.).
5. Zhukov S.V., Zelensky A.V. The use of neural networks in building optimal training trajectory in biathlon. *Uchenye zapiski universiteta im. P.F. Lesgafta = Scientific notes of the P.F. Lesgaft University*. 2017; (5): 81-88 (In Russ.).
6. Yushkin V.N. Problems of sports forecasting. *Uchenye zapiski universiteta im. P.F. Lesgafta = Scientific notes of the P.F. Lesgaft University*. 2021; (4): 473-478. (In Russ.).
7. Krutikov A.K., A parallel algorithm for the formation of a fragmented training sample for a specialized forecasting system based on the assembly of neural network modules. *Vestnik Permskogo natsional'nogo issledovatel'skogo politekhnicheskogo universiteta. Elektrotehnika, informatsionnye tekhnologii, sistemy upravleniya = Bulletin of the Perm National Research Polytechnic University. Electrical engineering, information technology, control systems*. 2024; (49): 61-81. DOI 10.15593/2224-9397/2024.1.04 .
8. Rozyhodzhaeva G.A., Rozyhodzhaeva D.A. Features of the formation of a training sample and training of a neural network with incomplete input data in solving private medical problems. *Nauchnoe obozrenie. Biologicheskie nauki = Scientific Review. Biological sciences*. 2017; (5): 28-32. (In Russ.).
9. Dorokhova Yu.V., Goncharuk N.S. The use of the method of expert assessments in forecasting social processes. *Srednerusskii vestnik obshchestvennykh nauk = Central Russian Bulletin of Social Sciences*. 2014; (3): 76-81 (In Russ.).
10. Zhukov V.G., Bukhtoyarov V.V. On the use of artificial neural networks with radial basis functions in the tasks of detecting anomalies in network traffic. *Reshetnevskie chteniya = Reshetnev readings*. 2013. 2(17): 285-286 (In Russ.).
11. Konakoglu B., Cakir L. Generalized Regression Neural Network for Coordinate Transformation. *Montenegro, International Symposium on Advancements in Information Sciences and Technologies (AIST)*. 2018. P. 66-74.
12. Malyutin A. Can't count the gold? Forbes. Available at: <http://www.forbes.ru/forbes/issue/2004-08/20514-ne-schest-zolota> (accessed: 21.05.2024). (In Russ.).
13. Yasnitskiy L.N., Pavlov I.V., Cherepanov F.M. Forecasting the 2014 Olympic Games in the unofficial team competition using artificial intelligence methods. *Sovremennye problemy nauki i obrazovaniya = Modern problems of science and education*. 2013; (6). (In Russ.).
14. Deep Learning Toolbox. Available at: <https://exponenta.ru/neural-network-toolbox> (accessed: 02.04.2024).

15. Krutikov A.K., Meltsov V.Y., Strabykin D.A. Evaluation the Efficiency of Forecasting Sports Events Using a Cascade of Artificial Neural Networks Based on FPGA. *Proceedings of ElConRus-2022*. St. Petersburg: ETU LETI; 2022. P. 355-360.

16. FPGA. Available at: <https://alt.ru/Mikroshemi.html> (accessed: 01.02.2024).

17. Levin I.I., Dordopulo A.I., Kalyaev I.A., Doronchenko Yu.I., Razladkin M.K. Modern high-performance computing systems with reconfigurable architecture based on FPGAs Xilinx Virtex-7 and Virtex UltraScale. *Russian Supercomputing Days 2015*. 2015; 435-447 (In Russ.).

18. Dordopulo A.I., Kalyaev I.A., Levin I.I., Semernikov E.A. High-performance reconfigurable computing systems of a new generation. In: *Trudy Mezhdunarodnoi superkomp'yuternoï konferentsii s elementami nauchnoi shkoly dlya molodezhi «Nauchnyi servis v seti Internet: ekzaflopsnoe budushchee» = Proceedings of the International Supercomputer Conference with elements of a scientific school for young people "Scientific service on the Internet: an examflop future"*. Moscow: Izd. MGU; 2011. P. 42-49 (In Russ.).

19. Kachanovsky Yu.P., Kachanovsky Yu.P., Korotkov E.A., Korotkov E.A. Data pre-processing for neural network training. *Fundamental'nye issledovaniya = Fundamental research*. 2011; (12-1): 117-120. (In Russ.).

20. Guzik V.F., Kalyaev I.A., Levin I.I. Reconfigurable computing systems. Rostov-on-Don; 2016. 471 p. (In Russ.).

Информация об авторах / Information about the Authors

Крутиков Александр Константинович, старший преподаватель, Вятский государственный университет, г. Киров, Российская Федерация, e-mail: yadrodisk@yandex.ru

Alexander K. Krutikov, Senior Lecturer, Vyatka State University, Kirov, Russian Federation, e-mail: yadrodisk@yandex.ru

Мельцов Василий Юрьевич, кандидат технических наук, доцент, Вятский государственный университет, г. Киров, Российская Федерация, e-mail: yadrodisk@yandex.ru

Vasily Y. Meltsov, Cand. Sci. (Engineering), Associate Professor, Vyatka State University, Kirov, Russian Federation, e-mail: yadrodisk@yandex.ru