УДК 004.052

ПРОГРАММНО-АППАРАТНЫЙ КОМПЛЕКС ДЛЯ ОЦЕНКИ НАДЕЖНОСТИ С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Сметанина О.Н., Сазонова Е.Ю., Андрушко Д.Ю.

Уфимский государственный авиационный технический университет, Уфа, e-mail: smoljushka@mail.ru, rassadnikova_ekaterina@mail.ru, andrewrush@mpo14.ru

Обнаружение аномалий – широко применяемая концепция, которая используется во многих областях для обнаружения ошибок (в том числе на этапе проектирования) и предотвращения сбоев. Традиционные методы, которые основаны на базе теории обработки сигналов или сравнении поведения реальной системы с её моделью, успешно применяются для обнаружения неисправностей. Однако киберфизические системы (КФС) имеют сложную структуру и поведение. Они состоят из множества разнородных компонентов, которые формируют огромные объемы данных, обмениваются информацией друг с другом и обладают сложной моделью поведения. Это делает практически невозможным применение традиционных методов оценки надежности. В статье были рассмотрены основные модели машинного обучения и возможность их применения для решения проблемы надежности киберфизических систем. Предлагается использовать нейронные сети с архитектурой длинной кратковременной памяти (LSTM) для выявления аномалий в компонентах киберфизических систем. Разработана архитектура нейронной сети с тремя скрытыми, входным и выходным слоями. Также разработана архитектура тестового стенда для проведения исследований, связанных с применением аппарата нейронных сетей. Был проведен эксперимент на контрольной выборке (Tennessee Eastman Process) и выполнен анализ результатов.

Ключевые слова: тестовый стенд, киберфизические системы, надежность, машинное обучение, нейросети

SOFTWARE AND HARDWARE COMPLEX FOR ASSESSING RELIABILITY USING ARTIFICIAL INTELLIGENCE

Smetanina O.N., Sazonova E.Yu., Andrushko D.Yu.

Ufa State Aviation Technical University, Ufa, e-mail: smoljushka@mail.ru, rassadnikova_ekaterina@mail.ru, andrewrush@mpo14.ru

Anomaly detection is a widely used concept that is used in many areas to detect errors (including in the design phase) and prevent failures. Traditional methods, which are based on signal processing theory or comparing the behavior of a real system with its model, are successfully used for fault detection. However, cyber-physical systems (CPS) have complex structure and behavior. They consist of many heterogeneous components that form huge amounts of data, exchange information with each other and have a complex pattern of behaviour. This makes it almost impossible to use traditional methods of reliability assessment. Basic models of machine learning and possibility of their application for solving the problem of reliability of cyber-physical systems were considered in the article. It is suggested to use neural networks with Long short-term memory architecture (LSTM) to reveal anomalies in components of cyber-physical systems. Neural network architecture with 3 hidden, input and output layers has been developed. Test stand architecture has also been developed to conduct research related to the application of neural network apparatus. The experiment with testing data (Tennessee Eastman Process dataset) was conducted and the results were analyzed.

Keywords: test stand, cyber-physical systems, dependability, machine learning, neural networks

Компоненты робототехники и сенсорика являются одной из основных областей применения для программно-аппаратных средств, предназначенных для решения задач, определяемых понятием «киберфизические системы».

Автоматизированные системы управления технологическими процессами имеют в своем составе следующие функциональные уровни: уровень технологического объекта (полевой уровень), объединяющий датчики, исполнительные механизмы, промышленные компьютеры и контроллеры, и операторский уровень, на котором разворачиваются приложения.

Киберфизическая система (КФС) является интеграцией вычислительных, сетевых и физических процессов [1, 2].

Основными требованиями к подобным системам является надежность и предсказуемость поведения. Однако, согласно международному стандарту IEC 61508 [3], КФС подвержены различным типам неисправностей.

Обнаружение аномалий – это хорошо изученная концепция, применяемая во многих областях, включая проектирование инженерных систем, где она помогает обнаруживать ошибки и предотвращать отказы. Во время работы системы детектор ошибок определяет, произойдет ли отказ системы в ближайшем будущем, основываясь на оценке текущего и серии прошлых состояний системы. Традиционные методы обнаружения аномалий, основанные либо на сравнении поведения реальной системы с ее моделью, либо на различных методах

обработки сигналов, успешно применяются при обнаружении неисправности и их изоляции (Fault Detection and Isolation, FDI) в мехатронных системах [4, 5].

Киберфизические системы (КФС) являются сложными как в структурном, так и поведенческом планах. Они состоят из многочисленных гетерогенных компонентов, генерирующих большие объемы данных, обменивающихся информацией и формирующих чрезвычайно сложные паттерны поведения. Это делает практически невозможным эффективную настройку и применение классических методов оценки надежности.

Современное состояние проблемы оценки надежности системы и постановка задачи

Популярные на сегодняшний день методы, основанные на глубоком обучении, используют классификатор, нейронную сеть, обученную отличать нормальное поведение системы от ненормального (табл. 1). Эти методы были предложены десятилетия назад, но только недавнее бурное развитие методов искусственного интеллекта [6, 7] позволило создавать эффективные средства выявления ошибок на основе методов глубокого обучения.

Подходы на основе классификации используют нейросеть для того, чтобы распознать нормальное и ошибочное состояния системы. Этот подход требует обучения с учителем, достаточного числа промаркированных экземпляров данных, как нормальных, так и ошибочных.

В подходах на основе предсказания текущие и предыдущие значения используются для предсказания/прогноза следующих нескольких шагов временных рядов. Следующее реальное значение сигнала сравнивается с предсказанным значением для обнаружения ошибки. Этот метод широко используется, когда ошибочные экземпляры трудно получить, при условии, что нормальный временной ряд поддаётся предсказанию на некоторое число шагов вперёд. В отличие от первого подхода, он позволяет даже смягчить временные ошибки путём замены настоящих ошибочных значений предсказанными.

Третий подход основан на специальной модели шифровщик-дешифровщик (Encode-decoder), которая реконструирует нормальный временной ряд. Этот метод основан на предположении, что такая модель будет плохо реконструировать ошибочные значения временного ряда.

Проблему обнаружения аномалий в КФС системах можно сформулировать в виде задачи множественной классификации.

Пусть дан многомерный временной ряд из размеченной обучающей выборки. Многомерный временной ряд состоит из нескольких переменных (признаков), например акселерометра, который выдает трехмерные данные в каждую единицу времени для каждой из трех осей [8].

Требуется разработать классификатор, позволяющий распознавать тип аномалий, то есть определить принадлежность текущего элемента контрольной выборки к определенному классу аномалий [9].

В качестве аномалий в исследовании рассматриваются ошибки сигналов в КФС, которые представляют собой данные временных рядов, записываемых непрерывно во времени.

Моделирование и реализация тестового стенда. Методика проведения эксперимента

Для проведения эксперимента необходимо разработать архитектуру тестового стенда. Он представляет собой программно-аппаратный комплекс: мобильное рабочее место, обеспечивающее возможность Machine и Deep Learning (машинного и глубокого обучения). Рабочее место будет использоваться при проведении исследований, связанных с применением аппарата нейронных сетей. Предполагаемая спецификация: видеокарта с поддержкой технологии CUDA; оперативная память 12–32 Gb; постоянное запоминающее устройство (SSD); процессор Intel серии Core i7 или Xeon.

Для определения характера процедуры обмена данным между элементами системы была построена модель тестового стенда по методологии DFD с помощью инструмента визуального моделирования Ramus. Модель (рис. 1) демонстрирует то, какие элементы тестового стенда обмениваются данными и какие данные необходимы для проведения эксперимента. Разработанная модель эксперимента включает процессы: формирование контрольной, обучающей и тестовой выборки; предварительная обработка полученных выборок; обработанные выборки; настройка нейросети; обучение нейросети на обучающей выборке; проверка результатов обучения нейросети на тестовой выборке.

Согласно диаграмме для проведения эксперимента необходимы данные временных рядов, которые в дальнейшем проходят обработку. Полученные контрольная и обучающая выборки необходимы для настройки, обучения нейросети. Обученная нейросеть проверяется с помощью тестовой выборки.

Функциональная модель тестового стенда построена по методологии IDEF0 с помощью инструмента визуального моделирования Ramus.

Таблица 1

Спектр методов машинного обучения

Модель	Лучшее применение	Худшее применение и побочные эффекты	Требования к ресурсам	Обучение
Случайные леса (статистические модели)	1. Обнаружение аномалий. 2. Системы с тысячами точек выбора и сотнями входов. 3. Регрессия и классификация. 4. Обрабатывает смешанные типы данных. 5. Игнорирует потерянные данные. 6. Линейно масштабируется вместе с вводом	1. Извлечение свойств. 2. Анализ с учетом времени и порядка следования	Низкие	1. Обучение на основе методов агрегации для максимальной эффективности. 2. Обучение без предвзятости с небольшим количеством ресурсов. 3. В основном под надзором
RNN (рекуррентная нейронная сеть)	1. Прогнозирование событий на основе последовательности. 2. Шаблоны в поточных данных. 3. Временные ряды. 4. Хранит предыдущие состояния для прогнозирования последующих (электрические сигналы, аудио, распознавание речи). 5. Неструктурированные данные. 6. Входящие переменные могут зависеть друг от друга	1. Анализ изображений и видео. 2. Системы, требующие применения тысяч свойств	1. Очень высокие при обучении. 2. Высокие при вычислении логических выводов	1. Обучение может быть более громоздким, чем в CNN. 2. Очень сложные в обучении. 3. Обучение с учителем
CNN (глубокое обучение)	1. Прогнозирование объекта на основе окружающих значений. 2. Распознавание шаблонов и свойств. 3. Распознавание двумерных изображений. 4. Неструктурированные данные. 5. Входящие переменные могут зависеть друг от друга	1. Прогнозирование на основе времени и порядка следования. 2. Системы, требующие применения тысяч свойств	1. Очень высокие при обучении (точность вычислений с плавающей точкой, большие обучающие наборы, серьезные требования к памяти). 2. Высокие при вычислении логических выводов	С учителем и без
Байесовские сети (вероятностные модели)	1. Неполные наборы данных, возможно, с шумом. 2. Шаблоны в поточных данных. 3. Временные ряды. 4. Структурированные данные. 5. Анализ сигналов. 6. Быстрое создание моделей	1. Предполагается, что все входящие переменные являются независимыми. 2. Плохо работает с многоуровневыми данными	Низкие	Требуется небольшое количество обучающих данных по сравнению с другими нейронными сетями

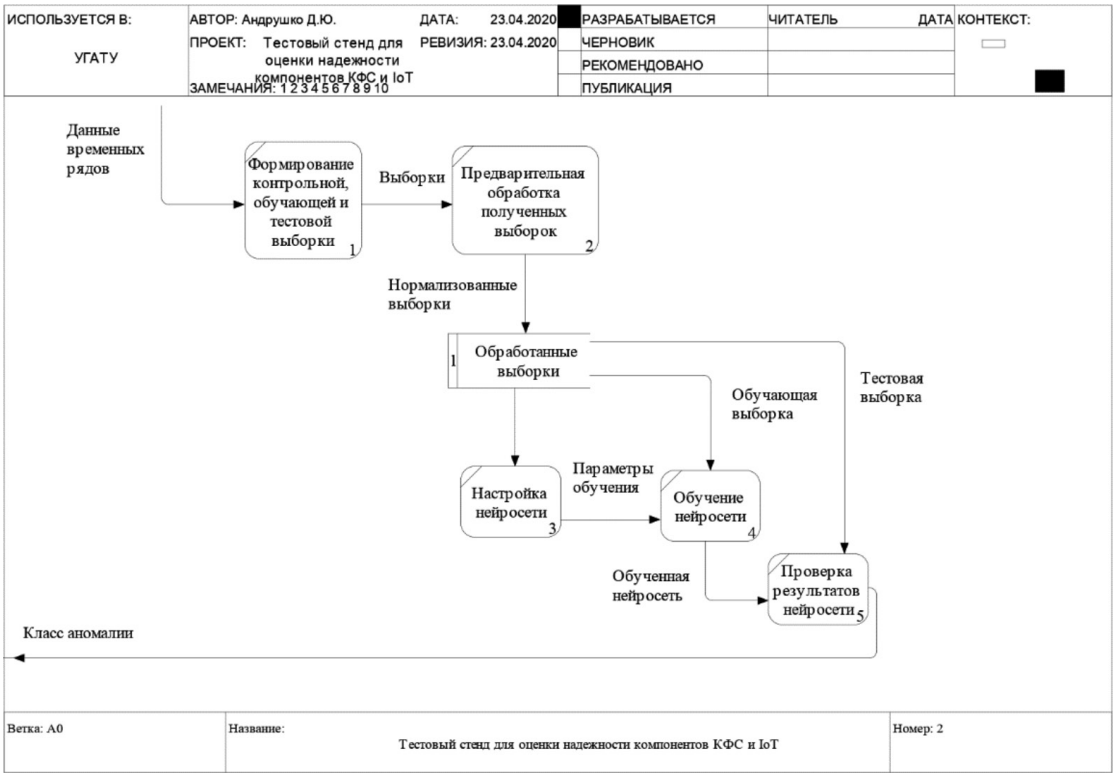

Рис. 1. Информационная модель эксперимента

Для формализации требований к системе построена функциональная модель с точки зрения разработчика. Моделирование позволяет определить действия, необходимые для проведения эксперимента на тестовом стенде. Модель способствует пониманию следующих вопросов: какая программная и аппаратная конфигурация используется для проведения эксперимента; как будет происходить оценка надежности компонентов КФС и IoT; какие функции должен реализовать разработчик.

Созданная с точки зрения разработчика модель должна реализовать следующие функции: формирование выборки на основе данных временных рядов, предварительная обработка выборки; настройка нейросети; обучение нейросети; проверка результатов обучения.

Согласно описанию исходными данными являются временные ряды. Правила оценки надежности компонентов КФС и IoT определяются методологией обучения нейросети. Выход представлен классом аномалии.

Функциональная модель эксперимента (рис. 2) включает следующие блоки: формирование контрольной, обучающей и тестовой выборки; предварительная обработка полученных выборок; определение параметров обучения; обучение нейросети на обучающей выборке; проверка результатов обучения нейросети на тестовой выборке.

На первом этапе происходит предварительная обработка исходных данных, в том числе путем нормализации. Затем осуществляется предварительная настройка параметров обучения нейросети, которая зависит от характера исходных данных. Длительность обучения нейросети определяется вычислительными мощностями тестового стенда. Результаты обучения необходимо проверить на тестовой выборке.

Выбор архитектуры нейронной сети в методах глубокого обнаружения аномалий в первую очередь зависит от характера входных данных. Входные данные могут быть классифицированы как последовательные (например, значения датчика) или непоследовательные (например, изображения).

В качестве примера были рассмотрены неисправности сенсорных сетей, которые также подвержены различным типам неисправностей. Для обучения была использована выборка, полученная в результате симуляции химических процессов в промышленности (выборка Tennessee Eastman Process или TEP dataset) [10, 11]. Набор данных состоит из четырех частей: обучающая (training) и тестовая (testing) выборка

для нормального (fault-free) и аномального (faulty) процессов. Наборы обучающей выборки содержат 500 временных измерений за 25 часов моделирования. Наборы тестовой выборки содержат 960 временных измерений за 48 часов моделирования.

Из обучающей выборки были извлечены контрольные данные (validation data) для её проверки. Контрольные данные составляют 20% от объема обучающей выборки. Использование контрольных данных позволяет оценить соответствие модели и обучающей выборки, что необходимо при подборе гиперпараметров модели.

Полученный набор данных, состоящий из контрольных данных, обучающей и тестовой выборки, содержит изменения 52 сигналов в течение 500 одинаковых временных интервалов. Каждому сигналу необходимо определить правильный тип ошибки, что является задачей классификации.

Для обучения предложена архитектура LSTM-сети (long short-term memory, *сеть долгой краткосрочной памяти*, разновидность рекуррентных нейронных сетей) с тремя скрытыми, входными и выходными слоями, которая позволяет классифицировать аномалии в работе компонентов киберфизической системы.

Обучение происходило в пакете математического моделирования Matlab R2019b Trial (с пакетами расширения Deep Learning Toolbox и Parallel Computing Toolbox) с использованием графического процессора NVIDIA GeForce MX150, который поддерживает аппаратное ускорение CUDA. Поскольку в данном примере обрабатывается большой объем данных, использование графического процессора значительно ускоряет время обучения.

Для оценки качества классификатора используется тестовая выборка, дополнительно для сравнения успешности классификации необходимо определиться с численной метрикой качества. Точность (accuracy) – метрика оценки качества классификатора, которая определяется отношением числа правильно классифицированных элементов тестовой выборки к общему размеру тестовой выборки:

$$Acc = \frac{P}{N},$$

где P – количество элементов тестовой выборки, которые были верно классифицированы, N – общее количество элементов тестовой выборки.

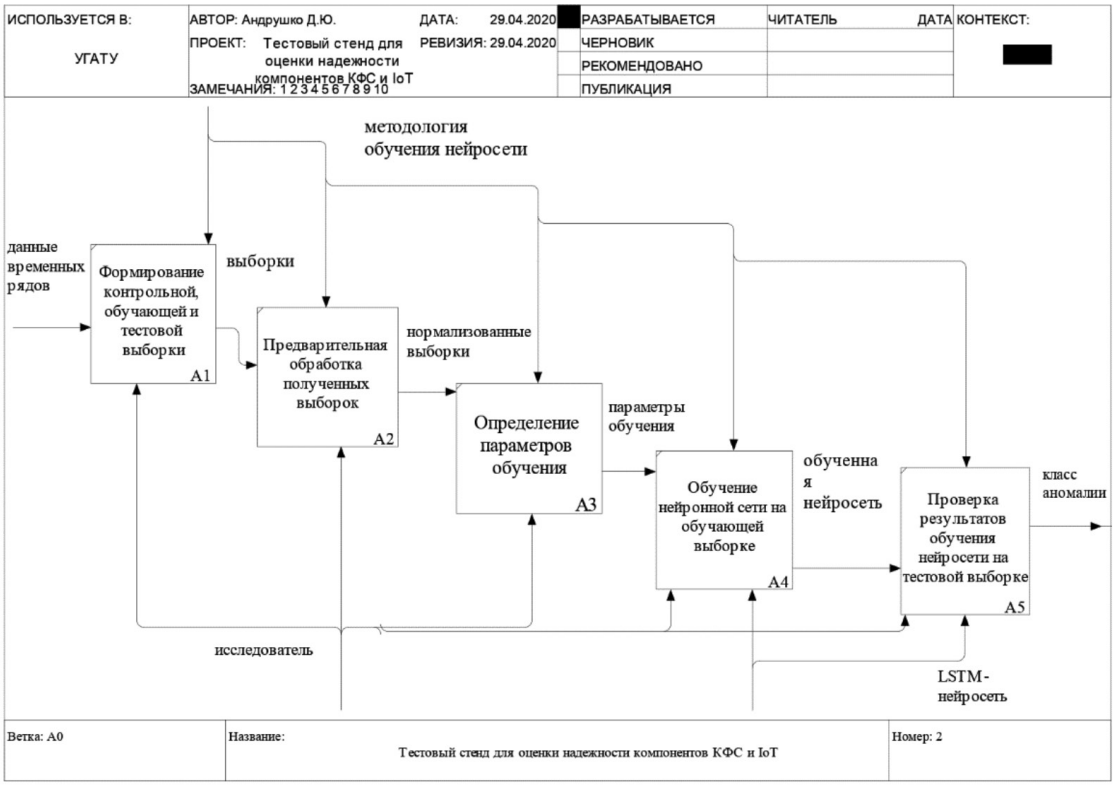

Рис. 2. Функциональная модель эксперимента

На практике значения точности удобнее и нагляднее оценить с помощью матрицы неточности (confusion matrix), которая представляет собой квадратную матрицу размерности $M \times M$, где M – количество классов. В пакете математического моделирования Matlab строки матрицы неточности резервируются за реальным классом элемента тестовой выборки, а столбцы – за решениями классификатора. Следовательно, на главной диагонали расположено количество правильно классифицированных элементов.

Процесс машинного обучения представляет собой алгоритм итеративной оптимизации – градиентный спуск. Градиент показывает скорость возрастания или убывания функции. Спуск подразумевает убывание. Процедура оптимизации повторяется до достижения оптимального результата. Для классификатора хорошим результатом обычно является высокая точность классификации (accuracy).

Функция потерь (cost function) показывает потери при неправильной классификации элементов выборки. Обычно в машинном обучении её называют Loss-функцией или просто Loss. Уменьшение значения Loss-функции показывает положительную динамику обучения.

На практике при обучении используют параметры (epoch, batch size, iterations per epoch). Эпоха (epoch) – полный проход выборки через нейросеть. Обычно одной эпохи недостаточно, поэтому при обучении задается количество эпох (epochs). При работе с выборками большого объема их разбивают на небольшие части (batch), чтобы хватило вычислительных мощностей оборудования. Batch size – количество элементов выборки, представленных в одном batch.

Итерации (iterations per epoch в Mathlab) – количество batch, которое необходимо обработать для завершения одной эпохи.

В табл. 2 показаны основные параметры обучения, время обучения и точность классификации. Как видно по таблице, с увеличением количества эпох точность классификации стремится к единице. Дальнейшее увеличение количества эпох не приведет к значительному приросту точности классификации, однако время обучения существенно увеличивается, что не имеет практического смысла. Поэтому было принято решение остановиться на 30 эпохах.

Результаты исследования и их обсуждение

По результатам проверки нейросети тестовой выборкой определена её точность – количество совпадений результатов классификации с реальными значениями типов неисправностей, деленное на общий размер тестовой выборки (0,9974). Высокая точность показывает, что нейросеть успешно классифицировала большинство элементов тестовой выборки (рис. 3).

Матрица неточностей показывает эффективность классификации. Она имеет числовые значения преимущественно на главной диагонали. Обученная сеть эффективна и правильно классифицирует более 99 % сигналов (рис. 4).

Заключение

В исследовании были рассмотрены основные модели машинного обучения: случайные леса, рекуррентные нейронные сети, сверточные нейронные сети, байесовские сети. Рекуррентные нейронные сети можно использовать при решения задач надежности киберфизических систем (КФС) на базе данных временных рядов.

Нейросети с долгой краткосрочной памятью (LSTM), которые являются разновидностью рекуррентных нейронных сетей, хорошо подходят для обнаружения аномалий в компонентах КФС. Архитектура нейросети зависит от характера исходных данных. В экспериментальных исследованиях для анализа точности классификации использовалась нейросеть с тремя скрытыми слоями.

Таблица 2

Параметры обучения, время обучения, точность классификации

	Количество эпох (Epochs)					
	5	10	15	20	25	30
Batch Size	50	50	50	50	50	50
Итерации (Iterations per epoch)	144	144	144	144	144	144
Общее число итераций (Total Iterations)	720	1440	2160	2880	3600	4320
Время обучения, (Training time, c)	354	749	1093	1434	1793	2140
Точность классификации (Accuracy)	0,8946	0,9594	0,9591	0,994	0,9884	0,9974

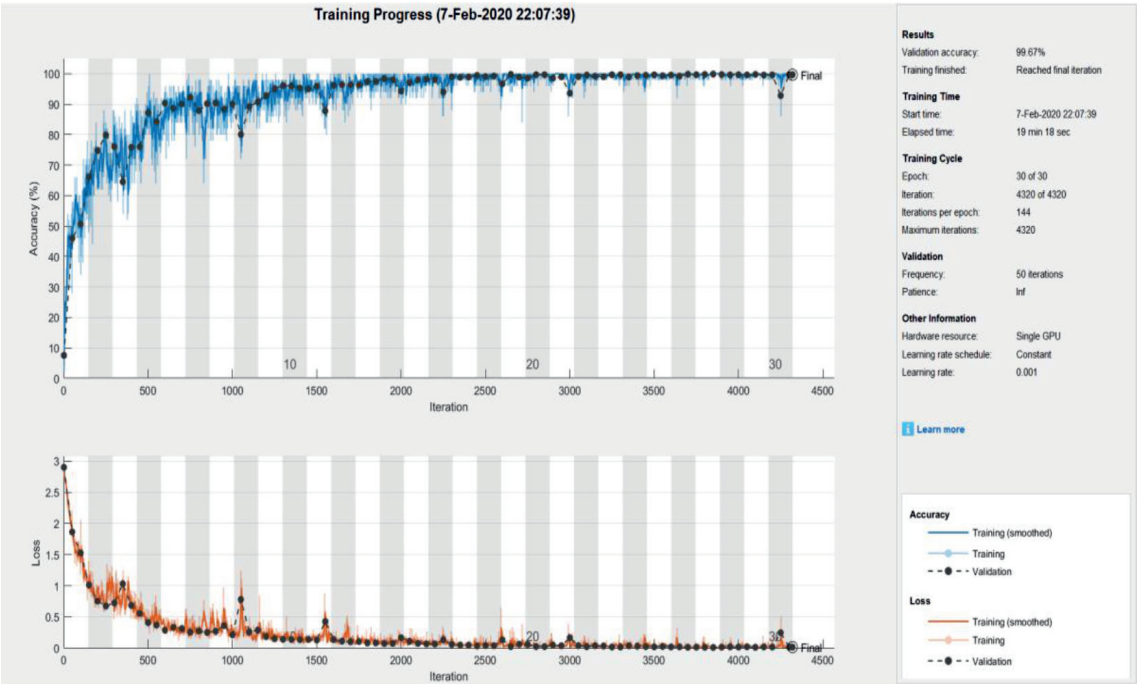

Рис. 3. Результаты обучения

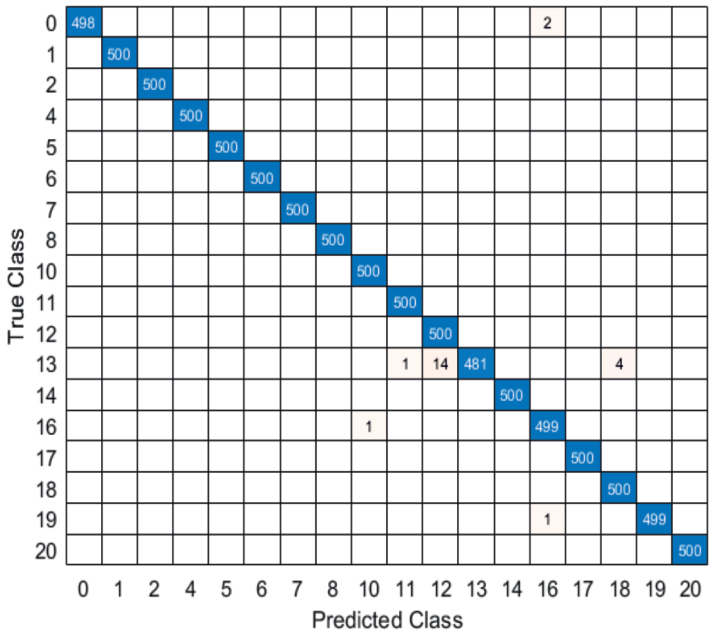

Рис. 4. Матрица неточностей

Архитектура тестового стенда представляет собой программно-аппаратный комплекс, который используется при проведении исследований, связанных с применением аппарата нейронных сетей, и представляет собой мобильное рабочее место. Спецификация тестового стенда выбрана с учетом применения в процессе исследований машинного и глубокого обучения: видеокарта с поддержкой технологии CUDA;

оперативная память 12–32 Gb; постоянное запоминающее устройство (SSD); процессор Intel серии Core i7 или Xeon.

Эксперимент, проведенный на контрольной выборке (Tennessee Eastman Process), показал высокую точность классификации (0,9974). Матрица неточностей имеет числовые значения преимущественно на главной диагонали. Таким образом, используемая нейросеть показала высокую эффективность, правильная классификация происходит в более чем 99 % случаях.

Результаты исследований, связанные с программной реализацией, в основу которой положены модели и алгоритмы интеллектуальной обработки данных, частично получены в рамках выполнения гранта РФФИ 18-07-00193.

Решения, полученные в рамках выполнения гранта РФФИ 19-07-00709, связаны с выбором моделей и методов выявления закономерностей на больших данных.

Вопросы исследования и описания проблемы решения задачи анализа свойств объекта, подходы к ее решению и готовые программные реализации, а также проведения эксперимента с целью выбора наиболее эффективного метода классификации для программной реализации с учетом метрик качеств получены в рамках государственного задания № FEUE-2020-0007.

Список литературы

1. National Science Foundation. Cyber-Physical Systems (CPS). [Electronic resource]. URL: https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503286 (date of access: 16.06.2020).

2. Lvovich Y., Preobrazhenskiy A.P., Choporov O.N. Simulation of controlling alternative current actuator on neural network basis. 17th International Ural Conference on AC Electric Drives (ACED), Ekaterinburg, 2018. P. 1–5. DOI: 10.1109/ACED.2018.8341698.

3. IEC 61508. Functional Safety of Electrical/Electronic/Programmable Electronic. [Electronic resource]. URL: Safety-related Systems. https://webstore.iec.ch/preview/info_iec61508-1%7Bed2.0%7Db.pdf (date of access: 16.06.2020).

4. Thirumarimurugan M., Bagyalakshmi N., Paarkavi P. Comparison of fault detection and isolation methods: A review. 10th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, 2016. P. 1–6. DOI: 10.1109/ISCO.2016.7726957.

5. Mutzke T., Ding K., Morozov A., Janschek K., Braun J. Model-based analysis of timing errors for reliable design of mechatronic medical devices. 3rd Conference on Control and Fault-Tolerant Systems (SysTol), Barcelona, 2016. P. 233–238. DOI: 10.1109/SYSTOL.2016.7739756.

6. Юсупова Н.И., Богданова Д.Р., Бойко М.В. Обработка слабоструктурированной информации на основе методов искусственного интеллекта. М.: Инновационное машиностроение, 2016. 250 с.

7. Lvovich Y., Preobrazhenskiy A.P., Choporov O.N. The simulation of error-correcting communication channel for video transmission. Moscow Workshop on Electronic and Networking Technologies (MWENT), Moscow, 2018. P. 1–6. DOI: 10.1109/MWENT.2018.8337296.

8. Yusupova N.I., Hilbert A., Boyko M.V., Bogdanova D.R. Models and Methods for Quality Management Based on Artificial Intelligence Applications. ITIDS+RRS'2014 Proceedings of the 2 nd International Conference «Information Technologies for Intelligent Decision Making Support». 2014. P. 231–237.

9. Javaid A., Niyaz Q., Sun W., Alam M. A Deep Learning Approach for Network Intrusion Detection System. In Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS) (BICT'15). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Brussels, BEL, 2016. P. 21–26. DOI: 10.4108/eai.3-12-2015.2262516.

10. Rieth C.A., Amsel B.D., Tran R.., Maia B. Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation. Harvard Dataverse. Version 1. 2017. DOI: 10.7910/DVN/6C3JR1.

11. Heo S., Lee J.H. Fault Detection and Classification Using Artificial Neural Networks. Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology. P. 470–475. DOI: 10.1016/j.ifacol.2018.09.380.