

Р.А. Ивченко, А.И. Купин

**ИССЛЕДОВАНИЕ И РАЗРАБОТКА АКТУАЛЬНЫХ ПРИЕМОВ,
ТЕХНОЛОГИЙ И МЕТОДИК, ПРИМЕНЯЕМЫХ ПРИ РЕШЕНИИ
ПРИКЛАДНЫХ ЗАДАЧ МАШИННОГО ОБУЧЕНИЯ В ЦЕЛЯХ
ПРЕДИКТИВНОГО АНАЛИЗА**

Аннотация. Произведено исследование актуальных приемов, технологий и методик, применяемых при решении прикладных задач машинного обучения, по материалам научных статей в высокорейтинговых журналах зарубежных исследователей, аналитических и обзорных заметок из открытых источников, а также технической документации и пресс-релизов технических и программных решений. Поиск новых методов выбора моделей, перекрестной проверки, эволюционного и аналитического подбора алгоритмов обучения представляет как научный, так и чисто практический интерес. Развитие технологий машинного обучения только ускорится в ближайшем будущем. В настоящее время мы наблюдаем прогресс в развитии методик автоматизированного поиска путей построения эффективных обучающихся моделей анализа данных, применимых ко многим практическим задачам интеллектуального анализа данных. В ходе обзора современных тенденций в машинном обучении нами выделены перспективные направления фундаментальных и прикладных исследований в данной области.

Ключевые слова: искусственные нейронные сети, аппроксимация нейронными сетями, нейросетевое моделирование, регрессионный анализ, Data Mining.

Введение. Многолетний прогресс в разработке методов машинного обучения породил не только разнообразные математические, программные и даже аппаратные решения, предназначенные для задач предиктивного и генеративного анализа данных в самых разных областях, но встретил на своем пути немало трудностей и препятствий. Основная трудность, с которой сталкивается человек в процессе знакомства с областью машинного обучения, – огромное количество разрозненных методов, каждый из которых обладает своими особенностями, областью использования и преимуществами. Однако такое разнообразие подчас ста-

вит в тупик и искусственных исследователей. С развитием математических и алгоритмических методов становится все труднее хорошо ориентироваться во всех нюансах применяемых алгоритмов. К сожалению, методологическая база значительно отстает от быстрого процесса разработки новых алгоритмов обучения, и процесс выбора обучаемой модели подчас сводится к простому перебору. В индустрии машинного обучения давно назрела необходимость создания более простых в использовании подходов, которые можно применять в широком кругу неэкспертов. Часто процесс использования систем машинного обучения предполагает выполнение более или менее полного цикла прикладных исследовательских работ по обработке (препроцессингу) данных, выделения признаков, выбора вида модели, обучения параметров, гиперпараметров, и т. д.

Исходные предпосылки. Несмотря на бурное развитие машинного обучения в последнее десятилетие, искусственный интеллект остается весьма расплывчатым понятием. Оно включает множество предметных областей: от предсказания временных рядов до генерации правдоподобных изображений по определенной теме. Методы машинного обучения, составляющие вычислительную основу технологий искусственного интеллекта, все еще остаются узкоспециализированными под каждую конкретную задачу. В качестве математической и инструментальной основы машинного обучения сегодня центральное место занимают искусственные нейронные сети. Сейчас они стали своеобразным универсальным языком представления обучаемых моделей. И хотя искусственный интеллект и машинное обучение как области знаний гораздо шире и включают целые семейства других методов, нейросетевые модели в настоящий момент так или иначе фигурируют в 90 % научных публикаций в данных областях.

Наиболее существенной из них можно отметить принцип работы любой нейросети как «черного ящика» – даже должным образом обученная сеть не дает исследователям информации о внутренней структуре проблемы и выявленных зависимостях в данных. Обученная нейросеть является набором матриц весов, и смысловая интерпретация этих весов в общем случае не предполагается. С этой точки зрения нейросети пред-

ставляют собой лишь инструмент решения конкретной задачи машинного обучения, но не дают экспертам аналитической информации для исследования проблемы. Этот недостаток заставляет исследователей заниматься проблемами интерпретируемости нейросетевых моделей. В частности, довольно широкую известность приобрела опубликованная в 2017 г. работа о применении довольно старого метода интерпретации предсказаний к обученным нейронным сетям через использование функций влияния [1]. В начале 2018 г. появилась книга, посвященная интерпретации моделей «черного ящика». Интерпретируемость модели машинного обучения понимается как «степень, с которой человек может понять причины того или иного решения». Построение интерпретируемых моделей позволяет находить ответ на такие вопросы как: «Как алгоритм создает модель?», «Как обученная модель делает предсказания?», «Как составные части модели влияют на предсказание?», «Как модель принимает определенное решение для определенного объекта/группы объектов?». В настоящее время выделяют класс так называемых «естественно интерпретируемых моделей» – моделей, человеческая интерпретация которых не представляет труда в силу самой архитектуры модели. К ним относят, в частности: линейные и логистические регрессии, деревья решений, наивные байесовские классификаторы, k ближайших соседей, модели правил вывода и другие [6]. Объяснение внутреннего устройства других типов моделей, в том числе глубоких нейронных сетей, представляет собой открытую научную задачу. Существуют следующие основные методы для построения моделенезависимых объяснительных механизмов (англ. model-agnostic methods): – функции и графики частичных зависимостей, которые показывают предельное влияние признаков на результирующее значение; – механизм индивидуального условного ожидания – расширение метода функций частичных зависимостей на локальный анализ отдельных предсказаний; – исследование уверенности модели (англ. model class reliance) – моделенезависимый вариант алгоритма важности признаков, который позволяет оценить зависимость средней ошибки модели в зависимости от дисперсии отдельных признаков; – метод LIME (от англ. local interpretable model-agnostic

explanations) основан на обучении локальных интерпретируемых моделей по предсказаниям модели, исследуемой для объяснения конкретных предсказаний. В целом, область исследований интерпретируемости моделей машинного обучения можно назвать очень актуальной.

Цель – исследование актуальных приемов, технологий и методик, применяемых при решении прикладных задач машинного обучения, по материалам научных статей в высокорейтинговых журналах зарубежных исследователей, аналитических и обзорных заметок из открытых источников а также технической документации и пресс-релизов технических и программных решений.

Нейроэволюция. Искусственные нейронные сети доказали свою эффективность в решении многих задач, однако эта эффективность в большой мере зависит от умения подбирать правильную архитектуру сети под каждую конкретную задачу анализа данных. При проектировании архитектуры сети эксперт сталкивается с необходимостью принять множество решений, как количественных, так и качественных, от реализации которых напрямую зависит производительность результирующей модели. Среди них такие как: выбор количества слоев, количества нейронов в каждом слое, выбор функции активации, использование рекуррентных и сверточных слоев. В процессе такого проектирования задача эксперта – балансирование между вариативностью модели и склонностью к переобучению. Существуют еще и внешние факторы, которые необходимо брать в расчет: располагаемая вычислительная мощность, временные рамки решения задачи и т. д. Пространство возможных нейронных сетей, в котором ведется поиск, огромно. При проектировании нейронных сетей специалисты используют определенные эвристические правила и инструменты диагностики, однако полноценной методологией такой поиск назвать сложно, это скорее творческий процесс. Естественным направлением исследований в области искусственного интеллекта является построение методов и инструментов сокращения человеческого участия в построении систем машинного обучения, автоматизация данного процесса. В настоящее время существует два принципиально отличающихся подхода: автоматизация машинного обучения

(AutoML) и нейроэволюция (от англ. evolutionary artificial neural networks, EANN). Фундаментальной работой в области нейроэволюции стало исследование Яо [7], в которой исследована связь между процессами обучения искусственных нейронных сетей и эволюционными алгоритмами. Работа Яо дала старт исследованиям применения эволюционного программирования на различных этапах проектирования и обучения нейронных сетей. Дальнейшие исследования привели к созданию метода NEAT (от англ. neuroevolution through augmenting topologies), более оптимизированного по вычислительным ресурсам. Совсем недавно этот метод был адаптирован для эволюционного поиска структур глубоких нейронных сетей. На сегодняшний день эволюционное программирование дает возможность создавать сети, по производительности сравнимые с лучшими в своем классе при нулевом участии человека в процессе проектирования и обучения для задач классификации изображений и задач распознавания речи. Технологии искусственного интеллекта в менеджменте В настоящее время ведутся активные работы с использованием нейроэволюции в таких областях как: предсказание временных рядов ансамблевыми моделями, предсказание энергопотребления компьютерными кластерами, построение межпланетных траекторий, распознавание языка речи, предсказание цен на нефть. Главным недостатком нейроэволюционного подхода являются весьма высокие требования к вычислительной мощности для поддержки данного процесса до сходимости. Развитие глубоких нейронных сетей и распространение больших данных еще выше поднимает эту планку. Поэтому, в настоящее время мы наблюдаем спад интереса к эволюционному программированию искусственных нейронных сетей: время программиста и специалиста в области анализа данных сейчас стоит дешевле, нежели требуемое процессорное время. Однако большие корпорации, располагающие высокими вычислительными мощностями, могут позволить себе эксперименты с нейроэволюцией. По нашему мнению, с ростом производительности вычислительной техники интерес к эволюционным методам в машинном обучении будет возвращаться (возможно итерационно), пусть не в ближайший год, но в обозримом будущем.

Автоматизированное обучение. На протяжении последнего времени (примерно с 2013 г.) автоматизация машинного обучения стала широко обсуждаемой темой и одной из наиболее быстрорастущих областей теоретических и практических разработок. Р. Олсон, один из разработчиков библиотеки автоматизированного машинного обучения, выделяет три главных процесса, автоматизация которых возможна и высвобождает большое количество временных ресурсов экспертов: – подбор гиперпараметров моделей; – испытание большого количества разных моделей; – использование разных признаков, выделенных из данных [8]. Таким образом, автоматизированное машинное обучение (от англ. *automated machine learning, AutoML*) можно охарактеризовать как набор технологий и методов алгоритмического выбора, оценки эффективности моделей машинного обучения и итеративного моделирования. Сложность традиционного подхода к построению систем машинного обучения состоит в необходимости знания всех существующих алгоритмов искусственного интеллекта, умения их правильно применить и настроить. Исторически первым механизмом автоматизации процесса машинного обучения является основанный на байесовской оптимизации метод AutoWEKA [9]. Предлагаемые методы автоматизации машинного обучения связаны с существующими инструментальными средствами моделирования. Большинство из них ориентируется на популярную библиотеку *scikit-learn* языка Python. Например, довольно многообещающая система Auto-Sklearn, развивающая идеи байесовской оптимизации. Эта библиотека использует 15 классификаторов, 14 методов обработки признаков, 4 метода препроцессинга, позволяя вести автоматизированный поиск в пространстве моделей с более сотни гиперпараметров. Однако байесовская оптимизация не единственная методологическая основа автоматизации машинного обучения. Существуют решения, основанные на генетическом подходе [10]. Такие системы способны строить сложные процессы машинного обучения без какого-либо вмешательства человека в процесс проектирования. Современные рыночные системы AutoML, построенные в виде облачного сервиса могут строить модели, превосходящие созданные человеком в 30 % случаев. Однако экономия временных

ресурсов колоссальна – до двух порядков. Активным направлением исследований в области автоматизации машинного обучения является проработка вопроса использования сложных нелинейных конвейеров обработки данных. Преимущественным методом нахождения таких конвейеров является генетическое программирование. В середине прошлого года исследовательский отдел Google представил архитектуру AutoML, основанную на обучении с подкреплением [11]. Эта система строит рекуррентные сети, схожие по своей архитектуре с построенными человеком, но более сложные. В конце 2017 г. усовершенствованная система AutoML смогла построить сеть, превосходящую все существующие в задачах распознавания объектов на изображении (см. рис). Конечно, инструменты автоматизации машинного обучения не способны полностью вытеснить человека как участника процесса обработки данных. В конце концов, функцию целеполагания никакой инструмент обеспечить не в состоянии. Также одни из барьеров перед использованием систем AutoML – повышенные требования к вычислительным ресурсам. Однако с учетом экономии времени проектирования преимущество использования автоматизированного машинного обучения неоспоримо. Таким образом, можно с уверенностью назвать автоматизацию машинного обучения актуальной и активно развивающейся областью исследований. Поиск новых методов выбора моделей, перекрестной проверки, эволюционного и аналитического подбора алгоритмов обучения представляет как научный, так и чисто практический интерес. Развитие технологий машинного обучения только ускорится в ближайшем будущем. В настоящее время мы наблюдаем прогресс в развитии методик автоматизированного поиска путей построения эффективных обучающихся моделей анализа данных, применимых ко многим практическим задачам интеллектуального анализа данных. В ходе обзора современных тенденций в машинном обучении нами выделены следующие перспективные направления фундаментальных и прикладных исследований в данной области: 1. Теоретические исследования в области интерпретируемости моделей искусственного интеллекта в сочетании с анализом автоматически построенных моделей. 2. Практические исследования мультизадачных, генера-

тивных (порождающих) моделей. 3. Более широкое распространение автоматизированных средств машинного обучения. 4. Развитие и унификация инструментальных средств, в том числе облачных средств и сервисов интеллектуального анализа данных. 5. Разработка новых интеллектуальных продуктов пользовательского уровня, основанных на вышеперечисленных достижениях методологии машинного обучения.

Разработка моделей экспресс-анализа бизнес процессов на основе нейросетей и нечеткой логики. Схема применение аппарата нечеткой логики для построения моделей технологических процессов предложен вариант формального представления многоэтапного технологического процесса обогащения [2].

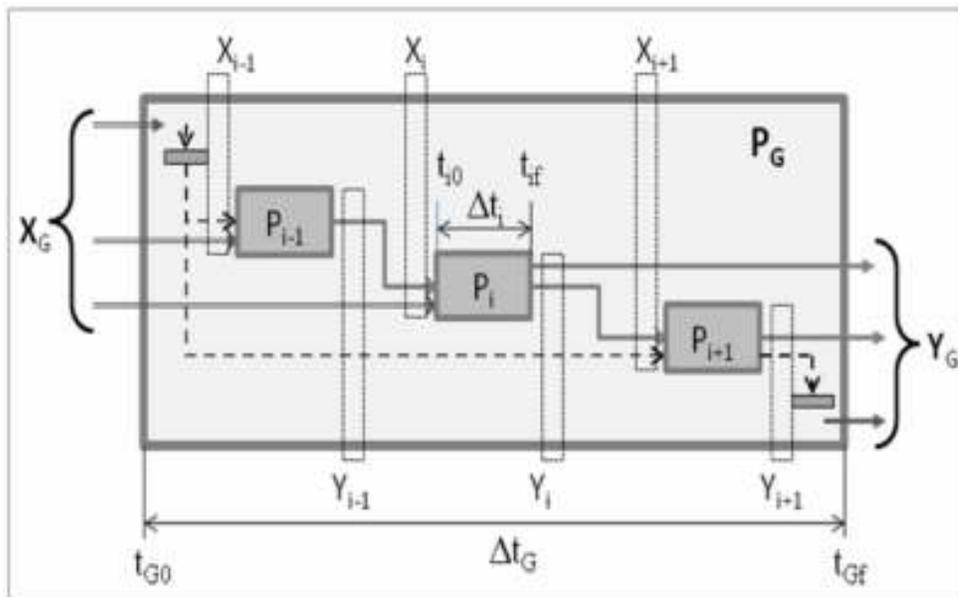


Рисунок 1 - Схема формального представления многоэтапного технологического процесса

На схеме использованы следующие обозначения: P_i – аппарат или этап технологической схемы P_G ; X – множество входных воздействий, влияющих на объект управления. В данное множество в качестве элементов включаются как внешние входы P_G , так и входы отдельных аппаратов P_i . Входами могут быть как контролируемые воздействия (управление), так и неуправляемые воздействия (возмущения). Возмущения в некоторых случаях измеримы, а в других - нет, и могут быть оценены

лишь в словесной форме (большое, маленькое и т.д.); Y – множество выходов, которое, аналогично множеству X , объединяет выходы всей схемы PG и выходы отдельных аппаратов P_i . Все выходы считаются потенциально контролируруемыми, хотя на практике, как правило, контролируется только часть из них. Далеко не все связи между входными сигналами и выходами могут определяться в виде классических передаточных функций $W(p)$. Это обусловлено размерностью задачи и степенью изученности исследуемого объекта. Зная о наличии некоторого параметра (фактора) не всегда удастся его измерить или оценить достаточно точно. Однако если предполагается его значимое влияние на динамику процесса, данный параметр должен быть включен в рассмотрение. При невозможности точной (количественной) оценки параметру может быть задана нечеткая (лингвистическая) оценка. Априорно принятым положением реализуемого подхода к разработке технологии оперативного прогнозирования обогатительных процессов является наличие неявного взаимного влияния технологических параметров процесса на характеристики продуктов разделения. Поэтому анализ начинается с рассмотрения ситуации, когда предполагается значимое влияние не только характеристик исходного сырья, но и всех технологических параметров на характеристики выходов процесса. Модель процесса представляется в виде направленного графа, узлами которого являются параметры, а дуги обозначают их влияние друг на друга. Подобное представление технологического процесса является неким подобием нейросети, в узлах которой находятся функции, осуществляющие преобразование сигналов. При этом часть функций представляется в виде «классических» передаточных функций $W(p)$, а часть – в виде «нечетких передаточных функций». Важной особенностью «нечетких передаточных функций» является их способность не только преобразовывать сигналы, но и осуществлять настройку передаточных функций узлов описанных с помощью $W(p)$. Данная особенность реализует свойство адаптивности модели, которое является весьма полезным с учетом динамичности свойств процесса. Для реализации нечетких передаточных функций для соответствующих параметров вводятся нечеткие функций принадлежности (ФП). Начальная

картина функций принадлежности формируется путем разбиения наблюдаемого (или - теоретического) интервала варибельности параметра на конечное число небольших отрезков и сопоставление каждому из них нечеткой ФП. Следует отметить, что использование большого количества нечетких ФП может сильно замедлить процесс вычисления. Поэтому изначально заданные интервалы разбиения диапазонов варьирования параметров следует, по возможности, укрупнить для снижения размерности задачи. Наличие больших объемов данных мониторинга обогатительного процесса позволяет установить соответствие между большинством интервалов значений выходного параметра и интервалами значений входов, при которых получен данный выход. Сокращение числа интервалов может осуществляться как на основе анализа имеющихся данных мониторинга, так и на основе экспертных рекомендаций. Формирование базы и верификация лингвистических правил отображения комбинаций нечетких значений входов в значения выходов также формируется с использованием данных мониторинга. В режиме прогнозирования в модель подается вектор, определяющий текущую ситуацию, который фаззифицируется и передается на обработку системе правил. Результаты нечеткого вывода дефаззифицируются, в результате чего определяются значения вектора выходов. Существует несколько алгоритмов осуществления операции фаззификации и заполнения базы знаний. Однако вопрос выбора формы ФП в литературных источниках практически не освещен. В качестве перспективного подхода к обоснованию выбора формы ФП планируется реализовать алгоритмы оценки скорости и ускорения (т.е. 1-й и 2-й производных) изменения значений параметров технологического процесса обогащения.

Разработка модели технологического процесса на основе использования нейросетей. Нейронные сети успешно применяются для синтеза систем управления динамическими объектами [3]. Нейросети обладают рядом свойств, которые определяют перспективность их использования в качестве аналитического аппарата систем управления. В контексте рассматриваемой задачи это, прежде всего, способность к обучению на примерах. Наличие больших объемов данных мониторинга, в которых

представлены взаимосвязанные измерения и входов, и выходов исследуемой системы, позволяет обеспечить нейросеть репрезентативными обучающими выборками. Другими важными свойствами является способность нейросети адаптироваться к изменению свойств объекта управления и внешней среды, а также высокая устойчивость к «сбоям» отдельных элементов сети в силу изначально заложенного в ее архитектуру параллелизма. Способности нейронной сети к прогнозированию напрямую следуют из ее способности к обобщению и выделению скрытых зависимостей между входными и выходными данными. После обучения сеть способна «предсказать» будущие значения выходов на основе нескольких предыдущих значений и текущих данных мониторинга. В рамках проводимых исследований наиболее перспективным представляется использование сетей встречного распространения. Сети данного типа имеют, в общем случае, существенно меньшее время обучения, чем сети обратного распространения. Поэтому такая сеть может более оперативно отреагировать на изменения условий протекания процесса обогащения, связанные с флуктуациями характеристик исходного сырья, технологических параметров или износом оборудования. В нейросети встречного распространения объединены два хорошо известных алгоритма: самоорганизующаяся карта Кохонена [4] и звезда Гроссберга [5]. Их объединение приводит к росту «обобщающих» способностей сети и позволяет получать правильный выход даже при неполных или незначительно искаженных входных данных. Анализ возможностей использования нейронных сетей для создания моделей экспресс-анализа производственных процессов. В результате анализа технологической схемы флотационного отделения, с учетом полученных ранее результатов исследования и уже принятыми в рамках данного проекта концептуальными принципами технологии моделирования производственных процессов, например, обогащения минеральных руд была определена структура нейросетевой модели. Используемые в модели параметры классифицированы на три группы: контрольные показатели, управляющие параметры и индикаторы. К контрольным показателям относятся характеристики входных и выходных продуктов технологической схемы. В модели рассматривается

15 контрольных параметров. К управляющим относятся те параметры, на которые можно воздействовать с целью изменения условий реализации технологического процесса и, как следствие, значений контрольных показателей. В качестве управляющих параметров в модели рассматривались температуры, влажности, давления, уровня жидкости, напряжения, тока и тд. Всего может рассматриваться порядка 27 индикативных параметров.

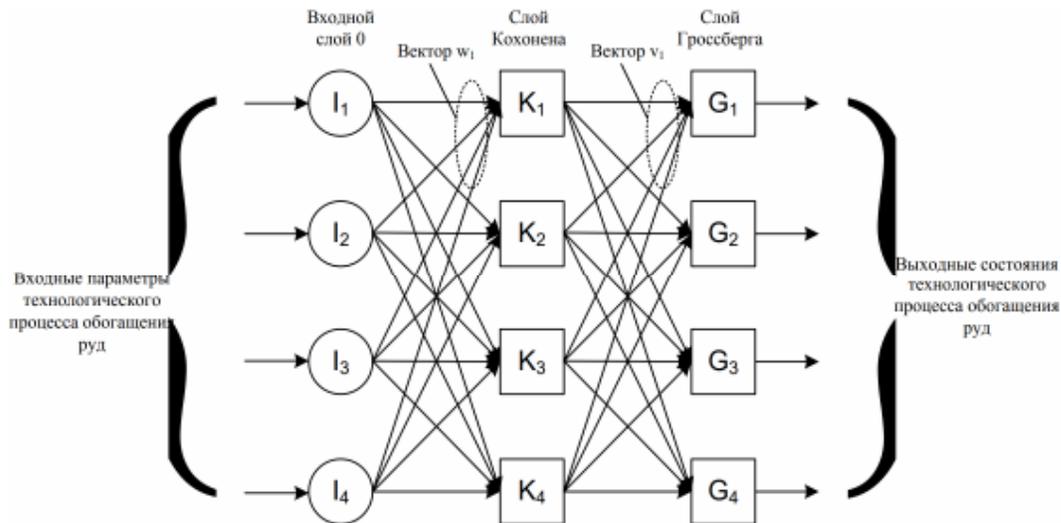


Рисунок 2 - Модель стандартной трехслойной нейросети встречного распространения

Нейроны слоя 0 служат точками разветвления и не выполняют вычислений. Каждый нейрон 0-го слоя связан с каждым нейроном слоя 1 (слой Кохонена). Аналогично, нейроны слоя 1 связаны с нейронами слоя 2 (слоем Гроссберга). С каждой связью ассоциирован собственный вес. Веса w_i связей слоев 0 и 1 образуют матрицу весов W , а веса V_j связей нейронов слоев 1 и 2 – матрицу весов V . Настройка значений весов производится в режиме обучения сети, когда в модель подаются априорно известные вектора входов X и выходов Y (рис. 1). В режиме прогнозирования в модель подается формируемый на основе текущих данных мониторинга входной вектор X , а выходной вектор Y генерируется сетью. Выход каждого нейрона слоя является просто суммой взвешенных входов. В результате сравнительной оценки взвешенных сумм входов нейронов Кохонена определяется «победитель» – нейрон, для которого это значе-

ние максимально. Выходу данного нейрона присваивается значение «1», а выходам остальных нейронов слоя Кохонена – «0». Слой Гроссберга функционирует по аналогичной схеме – его выходы определяются взвешенной суммой соответствующих входов от слоя Кохонена. Но, так как только у одного нейрона слоя Кохонена на выходе устанавливается значение «1», то фактически каждый нейрон слоя Гроссберга лишь выдает величину веса, который связывает этот нейрон с единственным ненулевым нейроном Кохонена. По сути, слой Кохонена классифицирует входные векторы в сходные группы, обеспечивая, тем самым, определение областей многомерного пространства входов, отображающихся в малую окрестность одной и той же «точки» пространства выходов. Это достигается с помощью настройки весов слоя Кохонена, обеспечивающей активацию одного и того же нейрона данного слоя соответствующими векторами входов. Перед началом обучения всем весам сети присваиваются некоторые случайные значения. В процессе обучения весовые векторы изменяются, «отслеживая» небольшую группу входных векторов. Обучение заканчивается, когда на выходе нейросети формируется требуемая картина выходов. Обучение слоя Гроссберга осуществляется путем корректировки лишь тех весов, которые ассоциированы с нейроном Кохонена, имеющим ненулевой выход. Величина коррекции веса пропорциональна разности между весом и требуемым выходом нейрона Гроссберга, с которым он соединен. Использование нейросетевой модели предполагает априорную классификацию состояний системы (обогачительного процесса) на конечное число вариантов. С каждым состоянием, при котором имеет место нарушение регламентных характеристик процесса, связан набор корректирующих воздействий, предполагающих конкретные изменения управляющих параметров. Для классификации могут быть использованы как экспертные оценки, так и формальные методы классификации из категории методов Data Mining, например – факторный и кластерный анализ. В качестве основного критерия классификации используются значения выходных векторов Y . Для определения текущего состояния процесса производится сравнение выхода нейросетевой модели и хранимых в информационной базе системы векторов, оп-

ределяющих выделенные состояния обогатительного процесса. Если в базе указано, что идентифицированному состоянию соответствует нарушение регламентных характеристик, то система извлекает из базы рекомендации по корректировке состояния. При наличии соответствующего исполнительного механизма запуск на выполнение корректирующих воздействий может быть автоматизирован. Разработанная нейросетевая модель процесса флотации была реализована и исследована в среде Matlab. На вход каждого из элементов нейросети подаются все контролируемые входные параметры. Весовые коэффициенты подбирались в процессе автоматического обучения на заранее заданных выборках реальных данных, полученных SCADA-системой в результате мониторинга производственного процесса. В ходе серии вычислительных экспериментов модель была настроена и обеспечила синтез выходных векторов, соответствующих контрольной выборке данных реального производственного процесса.

Исследование эффективности методов. Для оценки эффективности сетей радиального базиса и многослойного персептрона рассмотрим показанную в [8] задачу аппроксимации функции

$$d(x, y) = 3(1-x)^2 \exp(-x^2 - (y+1)^2) - 10\left(\frac{x}{5} - x^3 - y^5 \exp(-x^2 - y^2) - \frac{1}{2} \exp(-(x+1)^2 - y^2)\right) \quad (1)$$

при изменении переменных в пределах $-3 \leq x \leq 3$ и $-3 \leq y \leq 3$.

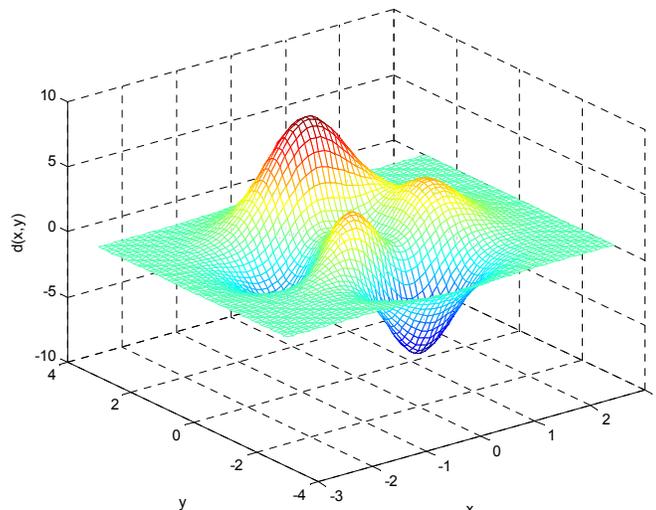


Рисунок 3 - График функции $d(x,y)$

График данной функции представлен на рис. 3. На основе обучающей выборки из 625 групп данных $([x, y], d)$, сгенерированных при равномерном распределении переменных x и y в областях их определения, в [8] построена сеть со структурой 2-36-1 (2 входных нейрона, 36 радиальных нейронов гауссовского типа и один выходной линейный нейрон). В [8] применялся гибридный алгоритм обучения, в результате максимальная погрешность аппроксимации после 200 итераций составила 0,06. Таким образом, проведенный вычислительный эксперимент показал, что нейронная сеть эффективно восстановила функцию $d(x, y)$ по ее табличным значениям. Однако в реальных условиях прикладные задачи, связанные с восстановлением функции, описывающей некоторое физическое явление, опираются на данные, содержащие различные помехи и погрешности измерения. По этой причине было решено повторить описанный вычислительный эксперимент, адаптировав его к прикладной области. Это значит, что данные, на которых строится нейросеть, должны быть не точными значениями функции, как в упомянутой работе, а содержать некоторый шум. Это позволит максимально приблизить эксперимент к реальным задачам.

Процесс построения нейросетевой модели можно условно разделить на 5 основных этапов (рис. 4).



Рисунок 4 - Основные этапы процесса построения нейросетевой модели

На первом этапе были сгенерированы 2 набора обучающих выборок $([x, y], d)$, содержащих случайные помехи (погрешности) с целью имитации данных, полученных при исследовании стохастического процесса или физического явления [12]. Первая группа, d' , содержала сильно зашумленные данные (помехи на уровне 20%), вторая, d'' , – слабо зашумленные (помехи на уровне 2%). Эти наборы данных были сгенерированы следующим образом. Функция $d(x,y)$ протабулирована в пределах $-3 \leq x \leq 3$ и $-3 \leq y \leq 3$ с шагом 0,25, и составлена таблица значений функции d . К каждому значению функции добавлена некоторая величина ε , полученная с помощью генератора случайных чисел с равномерным распределением.

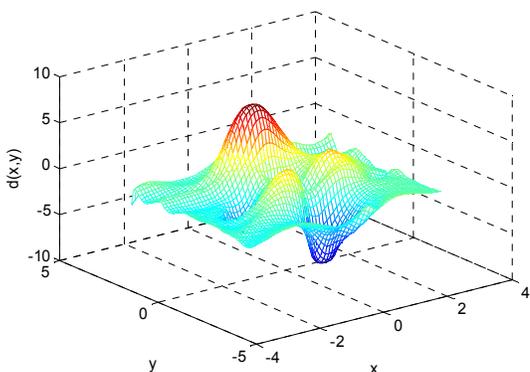
При этом
$$-\frac{p \max\{d\}}{2} \leq \varepsilon \leq \frac{p \max\{d\}}{2}$$
, где p – уровень помех в долях единицы. Таким образом, наборы точек $[x, y, d']$ и $[x, y, d'']$ имитируют результаты наблюдений за некоторым физическим процессом, содержащие погрешности измерения, на основе которых будет моделироваться сам процесс $d(x,y)$, в реальной задаче исследователю неизвестный.

На втором этапе проведено нормирование данных в диапазоне $[-1...1]$. Деление общей выборки на обучающую и тестовую в данном случае не требуется, так как для тестирования используется искомая функция $d(x,y)$.

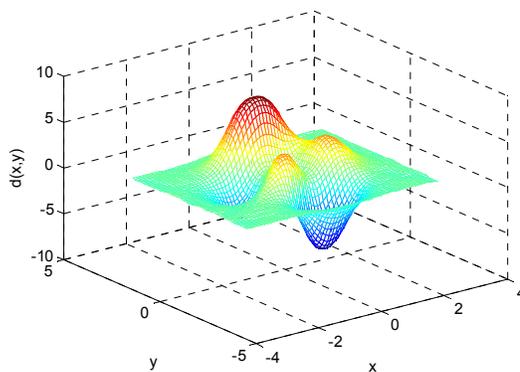
Третий этап подразумевает выбор типа нейросети из двух подходящих для решения задачи аппроксимации: многослойного персептрона и радиальной базисной сети. В данном эксперименте используются оба типа с целью сравнения их эффективности. Архитектура многослойного персептрона определялась следующим образом: сеть состояла из двух скрытых слоев по 8 нейронов в каждом. RBF-сеть содержит 1 скрытый слой, число нейронов в котором растет в процессе обучения.

На четвертом этапе на каждом из наборов данных обучались два указанных типа нейросетей. Для построения и обучения нейросетей использовался инструмент NeuralToolbox пакета Matlab.

На заключительном этапе строились поверхности отклика полученных нейросетевых моделей (рис. 5, 6). Оценено среднеквадратическое отклонение S_2 как ошибка полученной модели.



а)

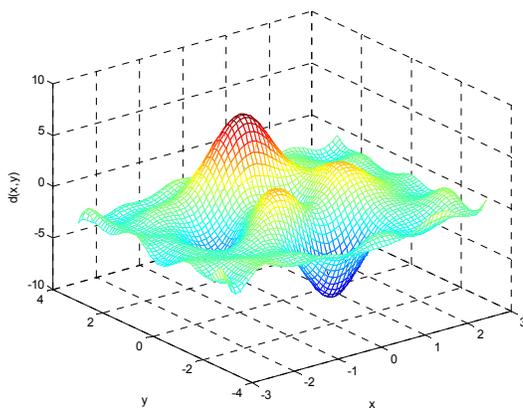


б)

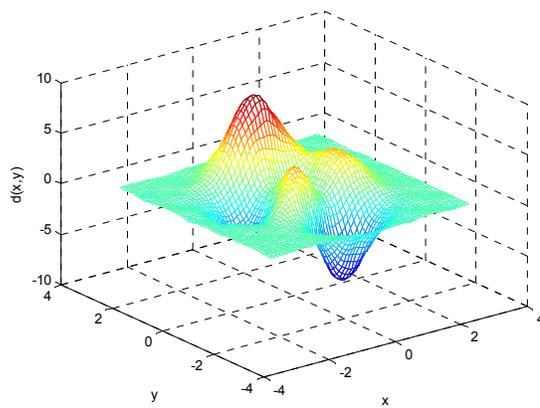
Рисунок 5 - Функция аппроксимирована MLP-сетью: а – на основе сильно зашумленных данных; б - на основе слабо зашумленных данных

На рис. 5а функция смоделирована на сильно зашумленных данных со среднеквадратическим отклонением $S^2=1,806$ (22%) многослойным перцептроном. S^2 нейросети – 1,63 (20%). Отклонение нейросети от зашумленного сигнала $S^2 =0,844$ (9%).

На рис. 5б функция содержит слабые помехи $S^2 =0,185$ (2%), аппроксимирована многослойным перцептроном. S^2 нейросети – 0,179 (2%). Отклонение нейросети от зашумленного сигнала $S^2 =0,11$ (1%).



а)



б)

Рисунок 6 - Функция аппроксимирована RBF-сетью: а – на основе сильно зашумленных данных; б - на основе слабо зашумленных данных

На рис. 6а функция содержит помехи $S^2 = 1,88$ (23%), аппроксимирована сетью радиального базиса. S^2 нейросети – 1,68 (20%). Отклонение нейросети от зашумленного сигнала $S^2 = 0,85$ (10%).

На рис. 6б функция содержит помехи $S^2 = 1,185$ (2%), аппроксимирована сетью радиального базиса. S^2 нейросети – 0,184 (2%). Отклонение нейросети от зашумленного сигнала $S^2 = 0,185$ (2%).

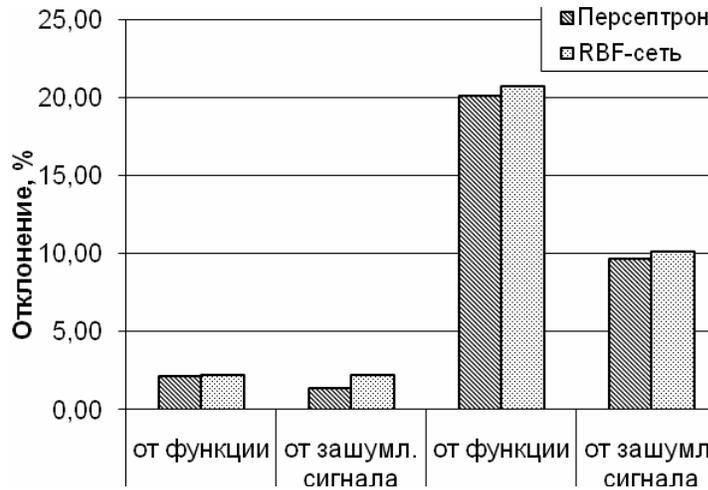
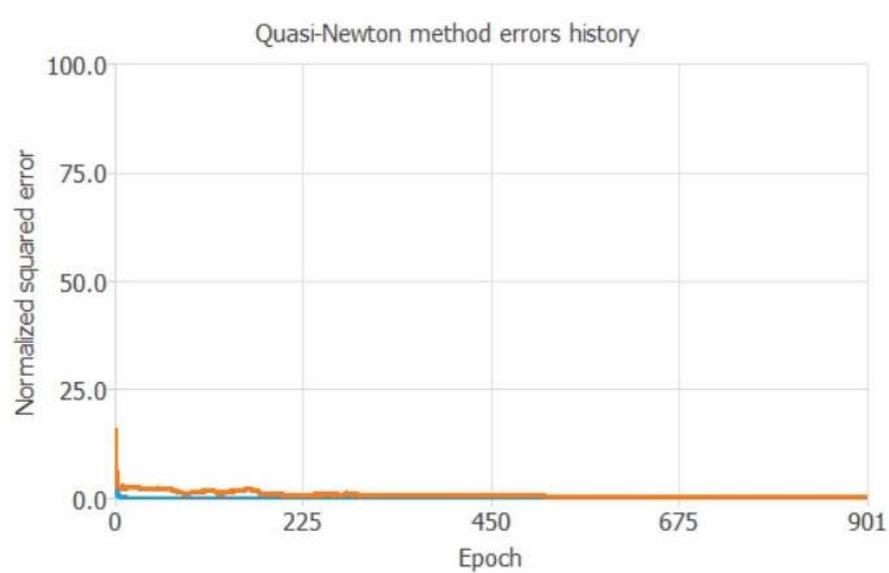


Рисунок 7 - Среднеквадратическое отклонение рассмотренных нейросетей

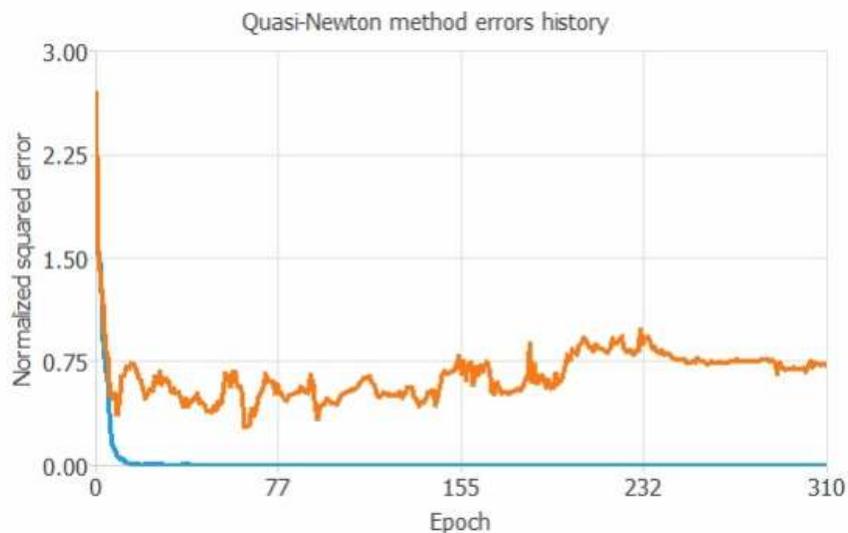
Оба рассмотренных типа нейросетей справились с задачей построения регрессионной модели зашумленного сигнала. Во всех случаях сети продемонстрировали способность к фильтрации шумов. RBF-сеть во всех случаях демонстрировала несколько большее отклонение (погрешность), однако, учитывая слишком малую величину разницы с многослойным персептроном (рис.7), следует признать это несущественным. Сеть радиального базиса выгодно отличается от многослойного персептрона, так как не требуется определение числа слоев и нейронов экспертом. В случае сети радиального базиса количество нейронов увеличивается в процессе обучения с целью достижения заданной точности модели. Однако число нейронов RBF-сети значительно (в данном исследовании - на порядок) больше, чем в персептроне, что снижает скорость работы с RBF-сетью. В задачах, которые требуют прогнозирования поведения функции за пределами диапазона обучения, выгоднее применять

многослойный перцептрон, так как он обладает способностью экстраполировать функцию.



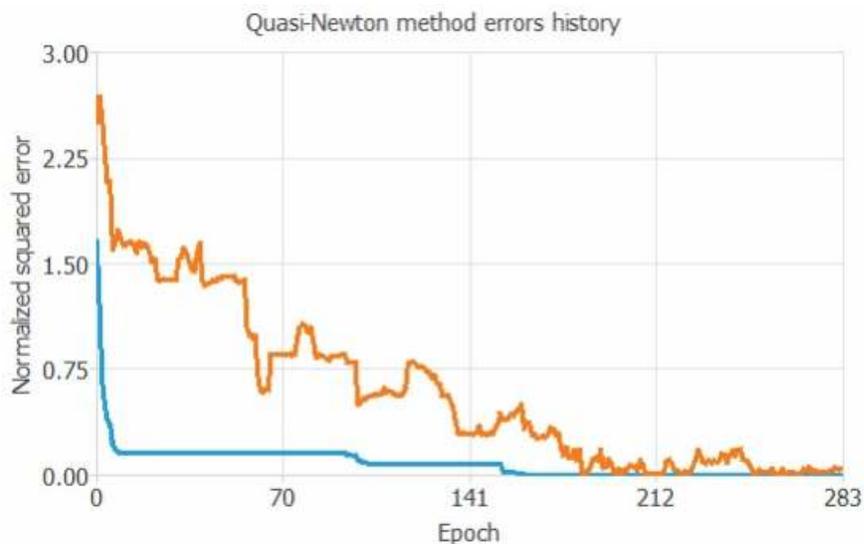
	Value
Final parameters norm	3.61
Final training error	6.18e-06
Final selection error	0.372
Final gradient norm	0.000984

Рисунок 8 - График обучения 3х слойного перцептрона на основе линейной функции активации выходного нейрона и скрытыми слоями с гиперболической функцией активации



	Value
Final parameters norm	12
Final training error	0.000481
Final selection error	0.723
Final gradient norm	0.000997

Рисунок 9 - Обучение 3х слойного перцептрона на основе логистической функции активации внутренних слоев



	Value
Final parameters norm	5.3
Final training error	0.000159
Final selection error	0.0398
Final gradient norm	0.000959

Рисунок 10 - Обучение 3х слойного перцептрона на основе гиперболической функции активации всех нейронов

Вывод. Среди рассмотренных архитектур нейросетей для решения задач аппроксимации и регрессионного анализа применимы следующие сети: многослойные перцептроны и сети радиального базиса. Оба типа нейросетей имеют свои достоинства и недостатки при использовании в задачах восстановления зависимостей. Каждая из рассмотренных сетей эффективно аппроксимирует сложные функции, обучаясь на зашумленных данных. Многослойные перцептроны показывают хорошие результаты при обработке экспериментальных данных, в том числе многомерных, позволяя моделировать скрытые в них закономерности. При обучении 3х слойный перцептрон на основе линейной функции активации выходного нейрона и скрытыми слоями с гиперболической функцией активации показал самый лучший результат в плане правильности обучения и точности прогнозирования.

ЛИТЕРАТУРА / ЛІТЕРАТУРА

1. Коротеев, М. В. Обзор некоторых современных тенденций в технологии машинного обучения Creative Commons «Attribution» («Атрибуция») 4.0. всемирная. (<http://creativecommons.org/licenses/by/4.0/>)
2. Олейник А.Г. Схема оперативного прогнозирования производственных процессов обогащения руд / А.Г. Олейник, Л.П. Ковалева // Труды Кольского научного центра РАН. Информационные технологии. – Апатиты: Изд-во КНЦ РАН. - 4/2011(7). - Вып. 2. - С.211-219.
3. Чернодуб А.Н. Обзор методов нейроуправления / А.Н. Чернодуб, Д.А. Дзюба // Проблемы программирования. -2011.-№ 2. - С.79 -94.
4. Kohonen, T. Self-organization and associative memory /Т. Kohonen // 2d ed. -New-York, Springer-Verlag, 1988. - 312 p.
5. Grossberg S. Some networks that can learn, remember and reproduce any number of complicated space-time patterns /S. Grossberg // Journal of Mathematics and Mechanics, 1969. -Vol. 19, № 1, - P.53-91.
6. Хохлова Д. (2016). Бум нейросетей: Кто делает нейронные сети, зачем они нужны и сколько денег могут приносить, 12.06.2016. Режим доступа: <https://vc.ru/16843-neural-networks> (дата обращения: 06.09.2018).
7. Molnar C. (2018), “Interpretable machine learning”, available at: <https://christophm.github.io/interpretable-ml-book/> (accessed September 6, 2018).
8. Olson R. (2016), “TPOT: A Python tool for automating data science”, available at: <https://www.kdnuggets.com/2016/05/tpot-python-automating-data-science.html/2> (accessed September 6, 2018).
9. Thornton C. et al. (2013), “Auto-WEKA: combined selection and hyperparameter optimization of classification algorithms ACM”, pp. 847–855.
10. Olson R.S. & Moore J.H. (2016), “TPOT: A tree-based pipeline optimization tool for automating machine learning”, pp. 66–74.
11. Zoph B. & Le Q.V. (2016), “Neural architecture search with reinforcement learning”, available at: <https://arxiv.org/abs/1611.01578>.
12. Гулаков К.В. Выбор архитектуры нейронной сети для решения задач аппроксимации и регрессионного анализа экспериментальных данных. Вестник Брянского государственного технического университета. 2013. № 2(38).

REFERENCES

1. Koroteev, MV A review of some current trends in the technology of machine learning Creative Commons Attribution 4.0. the world. (<http://creativecommons.org/licenses/by/4.0/>)
2. Oleynik, A.G. Scheme of operative forecasting of production processes of ore enrichment / AG. Oleynik, L.P. Kovalev // Proceedings of the Kola Scientific Center of the Russian Academy of Sciences. Information technology. - Apatity: Publishing House of the KSC RAS. - 4/2011 (7). -No. 2. - P.211-219.
3. Chernodub, AN Review of Neuro-management Methods / AN Chernodub, D.A. Beetle // Programming problems. -2011.-№ 2. - P.79 -94.
4. Kohonen, T. Self-organization and associative memory / T. Kohonen // 2d ed. -New-York, Springer-Verlag, 1988. - 312 p.
5. Grossberg, S. Some networks that can learn, remember and reproduce any number of complicated space-time patterns / S. Grossberg // Journal of Mathematics and Mechanics, 1969. -Vol. 19, No. 1, - P.53-91.
6. Khokhlova D. (2016). Neural Network Boom: Who Makes Neural Networks, Why They Need It and How Much Money They Can Make, 06/12/2016. Access mode: <https://vc.ru/16843-neural-networks> (accessed: 06/09/2018).
7. Molnar C. (2018), “Interpretable machine learning”, available at: <https://christophm.github.io/interpretable-ml-book/> (accessed September 6, 2018).
8. Olson R. (2016), “TPOT: A Python tool for automating data science”, available at: <https://www.kdnuggets.com/2016/05/tpot-python-automating-data-science.html> / 2 (accessed September 6, 2018).
9. Thornton C. et al. (2013), “Auto-WEKA: combined selection and hyperparameter optimization of ACM classification algorithms”, pp. 847–855.
10. Olson R.S. & Moore J.H. (2016), “TPOT: A tree-based pipeline optimization tool for automating machine learning”, pp. 66–74
11. Zoph B. & Le Q.V. (2016), “Neural architecture search with reinforcement learning”, available at: <https://arxiv.org/abs/1611.01578>.

12. Gulakov K.V. The choice of neural network architecture for solving the problems of approximation and regression analysis of experimental data. Bulletin of the Bryansk State Technical University. 2013. No. 2 (38).

Received 10.02.2020.

Accepted 14.02.2020.

Дослідження і розробка актуальних прийомів, технологій і методик, що застосовуються при вирішенні прикладних задач машинного навчання з метою інтелектуального аналізу

Були розглянуті архітектури нейромереж для вирішення завдань апроксимації і регресійного аналізу, застосовні наступні мережі: багатосарові перцептрони і мережі радіального базису. Було вироблено дослідження на основі якого було виявлено що при навчанні 3х шарового перцептрону на основі лінійної функції активації вихідного нейрона і прихованими шарами з гіперболічної функцією активації показав більш оптимальний результат в плані правильності навчання і точності прогнозування.

Research and development of actual methods, technologies and methods applied to solution of applied machine training problems for predictive protection

A study was made of relevant techniques, technologies and techniques used to solve applied problems of machine learning, based on materials from scientific articles in highly rated journals of foreign researchers, analytical and review notes from open sources, as well as technical documentation and press releases of technical and software solutions. The search for new methods of model selection, cross-validation, evolutionary and analytical selection of training algorithms is of both scientific and purely practical interest. The development of machine learning technologies will only accelerate in the near future. Currently, we are witnessing progress in the development of automated search methods for constructing effective learning models for data analysis that are applicable to many practical problems of data mining. During the review of modern trends in machine learning, we identified promising areas of fundamental and applied research in this area.

Development of a process model based on the use of neural networks. Neural networks are successfully used for the synthesis of control systems for dynamic objects. Neural networks have a number of properties that determine the prospects of their use as an analytical apparatus of control systems. In the context of the problem under consideration, this is, above all, the ability to learn by example. The presence of large volumes of monitoring data, which presents interconnected measurements of both the inputs and outputs of the studied system, allows the neural network to be provided with representative training samples. Other important properties are the ability of the neural network to adapt to changes in the properties of the control object and the external environment, as well as high resistance to "failures" of individual network elements due to the parallelism originally built into its architecture. The ability of a neural network to predict directly follows from its ability to generalize and highlight hidden relationships between input and output data. After training, the network is able to "predict" future output values based on several previous values and current monitoring data. In the framework of ongoing research, the most promising is the use of counterpropagation networks

Among the considered neural network architectures, the following networks are applicable for solving the problems of approximation and regression analysis: multilayer perceptrons and

radial basis networks. Both types of neural networks have their advantages and disadvantages when used in dependency recovery tasks. Each of the networks considered effectively approximates complex functions, learning from noisy data. Multilayer perceptrons show good results in processing experimental data, including multidimensional ones, allowing us to simulate patterns hidden in them. When training a 3-layer perceptron based on the linear function of activation of the output neuron and hidden layers with a hyperbolic activation function, it showed the best result in terms of training accuracy and prediction accuracy.

Ивченко Родион Анатольевич - аспирант кафедры Автоматизации компьютерных наук и технологий, Криворожский национальный университет.

Купин Андрей Иванович - д.т.н., профессор, заведующий кафедрой компьютерных систем и сетей, Криворожский национальный университет.

Ивченко Родіон Анатолійович - аспірант кафедри Автоматизації комп'ютерних наук і технологій, Криворізький національний університет.

Купін Андрій Іванович - д.т.н., професор, завідувач кафедри комп'ютерних систем та мереж, Криворізький національний університет.

Ivchenko Rodion Anatoliyovych - Kryvyi Rih National University.

Kupin Andriy Ivanovich - Doctor of Technical Sciences, Professor, Kryvyi Rih National University.