

ОБЗОРНАЯ СТАТЬЯ

УДК 004.8

doi: 10.26907/2541-7746.2022.2-3.244-265

ПЕРСПЕКТИВЫ БИОИНСПИРИРОВАННОГО ПОДХОДА В РАЗРАБОТКЕ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА (обзор тенденций)

А.Р. Нурутдинов¹, Р.Х. Латыпов²

¹ПАО «Таттелеком», г. Казань, 420061, Россия

²Казанский (Приволжский) федеральный университет, г. Казань, 420008, Россия

Аннотация

Искусственный интеллект способен не только эффективно строить прогнозные модели в инженерных, политических, экономических и научных сферах, но и предлагать оптимальные стратегии для решения различных задач. Однако текущие реализации искусственного интеллекта оказались не настолько эффективными, как ожидалось десятилетия назад. В результате проблемы, которые выявляются при его широком использовании, все больше начинают нивелировать положительные эффекты от решаемых им задач. В статье рассмотрены трудности и ограничения, возникающие в системах искусственного интеллекта, и представлены возможные пути их преодоления.

Ключевые слова: искусственный интеллект, машинное обучение, биоинспирированный подход, модель мозжечка

Введение

Идея искусственного интеллекта (ИИ), предложенная в 1955 г., основана на гипотезе о том, что «каждый аспект обучения или любую другую характеристику интеллекта в принципе можно описать настолько точно, что можно создать машину для их моделирования» [1]. Пионеры направления верили, что создание общего ИИ человеческого уровня возможно в течение нескольких десятилетий. Эта гипотеза стала руководящей для большинства исследований в области ИИ, сфокусированных на конкретных научных и инженерных проблемах.

ИИ, в особенности машинное обучение (МО), прошел долгий путь от первых дней концептуальных теорий до того, чтобы стать неотъемлемой частью современного технологического общества. Быстрый рост ИИ/МО и их проникновение во множество гражданских и военных приложений, несмотря на успех, выявили также новые проблемы и трудности. Поскольку для некоторых новых систем ИИ/МО практически не требуется участие человека, возникла острая необходимость в более глубоком понимании того, как принимаются решения, предлагаемые системами ИИ. Это привело к появлению таких областей исследований ИИ, как, например, объяснимый ИИ (explainable AI) [2], а также гибридных интеллектуальных систем, сочетающих естественный интеллект и ИИ [3].

Как отмечают многие исследователи, не существует общепринятого определения ИИ [4, 5]. Множество проектов в области ИИ не направлены на прямое моделирование человека. В частности, известный тест Тьюринга [6], получивший широкое

распространение, некоторыми влиятельными исследователями ИИ был раскритикован как отвлекающий или даже вредный [7, 8]. Человеческое поведение зависит не только от интеллектуальной компетентности и механизмов межвидового взаимодействия, но и от биологических, эволюционных и культурных факторов, которые являются уникальными для людей. Подчеркиваемое многими авторами отсутствие глубокого понимания принципов работы интеллекта не позволяет сегодня дать исчерпывающее и универсальное определение. Отсюда и постоянные отсылки к эталону, то есть к естественному интеллекту.

В настоящей статье дан обзор различных трудностей и ограничений, возникших при развитии систем ИИ, и рассмотрены возможные пути их преодоления.

1. Тенденции в сфере искусственного интеллекта

В настоящее время ИИ представлен как набор технологий и функциональных приложений, закрепленных либо в виде научных публикаций, либо в форме зарегистрированных патентов. В отчете за 2019 г. Всемирной организации интеллектуальной собственности (ВОИС, World Intellectual Property Organization, WIPO) «Технологические тренды 2019. ВОИС. Искусственный Интеллект» [9] опубликован патентный анализ в области ИИ. Патенты являются объективным средством оценки тенденций в исследованиях, поскольку раскрывают области инноваций, на которых сосредоточены изобретатели, и приводят к осязаемым результатам.

Согласно отчету, с момента появления ИИ в 50-х годах XX в. новаторы и исследователи подали заявки на регистрацию почти 340000 изобретений и опубликовали более 1.6 млн научных работ, связанных с ИИ. Бум научных публикаций на эту тему начался в 2001 г., примерно за 12 лет до всплеска количества регистрируемых патентов. При этом соотношение научных публикаций к изобретениям уменьшилось с 8:1 в 2010 г. до 3:1 в 2016 г., что является показателем перехода от теоретических исследований к использованию технологий ИИ в коммерческих продуктах и услугах.

Подход, используемый в отчете, опирается на изучение предмета в трех направлениях:

- 1) методы ИИ – расширенные формы статистических и математических моделей, позволяющих производить вычисления для задач, обычно выполняемых людьми;
- 2) функциональные приложения ИИ – практические реализации с использованием одного или нескольких методов ИИ, обеспечивающие интеллектуальные функции;
- 3) области применения ИИ или отрасли науки и производства.

Методы ИИ подразделяют на пять категорий. Это нечеткие логики, методы логического программирования (включая экспертные системы), вероятностные выводы, проектирование онтологий и машинное обучение.

В области функциональных приложений ИИ рассмотрены девять категорий: компьютерное зрение, обработка речи, робототехника, обработка естественного языка, управление, планирование, прогнозирование, распределенный ИИ, представление знаний. Здесь наиболее популярным является компьютерное зрение, включающее распознавание изображений, а также обработка естественного языка и обработка речи.

Наконец, в областях применения ИИ ведущими отраслями являются транспорт, телекоммуникации, а также медико-биологические науки.

Важнейшим выводом отчета является отнесение машинного обучения к абсолютно доминирующему методу искусственного интеллекта, так или иначе упоминаемому в 89% патентов, при этом среди методов машинного обучения револю-

ционный прогресс наблюдается в области глубокого обучения и нейронных сетей. Сегодня между понятиями МО и ИИ часто ставят знак равенства, в итоге представление формируется об ИИ исходя из подходов, где человечество получило наиболее ощутимые результаты.

Похожие оценки получены и другими исследователями [10, 11].

2. Ограничения существующих подходов к машинному обучению

Математический аппарат классического МО сформировался еще в XX в. Например, метод главных компонент (principal component analysis, PCA) – один из основных способов уменьшения размерности данных, – изобретен Карлом Пирсоном в 1901 г. [12], а метод обратного распространения ошибки (backpropagation) – способ вычисления градиента, который используется при обновлении весов многослойного персептрона, – был неоднократно независимо открыт начиная с 60-х годов XX в. [13].

Прорыв в машинном обучении наступил после 2010 г., когда стали доступны высокопроизводительные вычислительные ресурсы, сформировалась культура оцифровки данных, возникло множество сообществ, распространивших идеи машинного интеллекта до уровня вирусного мема.

Однако, несмотря на огромные инвестиции в зарождающуюся отрасль, всестороннюю поддержку научного сообщества и прилагаемые усилия, машинное обучение достигло значимых результатов в ограниченном круге прикладных задач, таких как компьютерное зрение, распознавание речи и обработка естественного языка, то есть в областях хоть и с крупными, но ограниченными множествами в задачах классификации.

В чем причины таких ограничений?

Задача обучения с учителем (supervised learning) является наиболее распространенной формой машинного обучения и формулируется следующим образом [14].

Пусть задано множество X объектов, множество Y допустимых ответов и существует целевая функция $y^* : X \rightarrow Y$, значения которой $y_i = y^*(x_i)$ известны только на конечном подмножестве объектов $\{x_1, \dots, x_l\} \subset X$. Совокупность пар $X^l = \{(x_i, y_i) | i = 1, \dots, l\}$ называется обучающей выборкой. Задача обучения с учителем заключается в том, чтобы по выборке X^l восстановить зависимость y^* , то есть построить решающую функцию $g : X \rightarrow Y$, которая приближала бы целевую функцию $y^*(x)$, причем не только на объектах обучающей выборки, но и на всем множестве X . Функция g – элемент некоторого пространства возможных функций, обычно называемого пространством гипотез G .

Не углубляясь в теорию обучения с учителем, отметим, что в ее основе лежит модель, в качестве которой выступает та или иная функция, содержащая множество параметров. Поэтому задачей МО является выбор архитектуры и подбор параметров функции для максимального соответствия выходных значений набору, предоставленному для обучения.

Но текущая итоговая интерпретация ИИ в виде нахождения оптимальной функции не может служить полноценным аналогом реального процесса по следующим причинам.

1. Первую известную причину легко пояснить на изображении (рис. 1). Серая колеблющаяся кривая – это график эталонной функции

$$f(x) = \begin{cases} x \cdot \sin 2\pi x, & x < 0, \\ -x \cdot \sin \pi x, & x \geq 0, \end{cases}$$

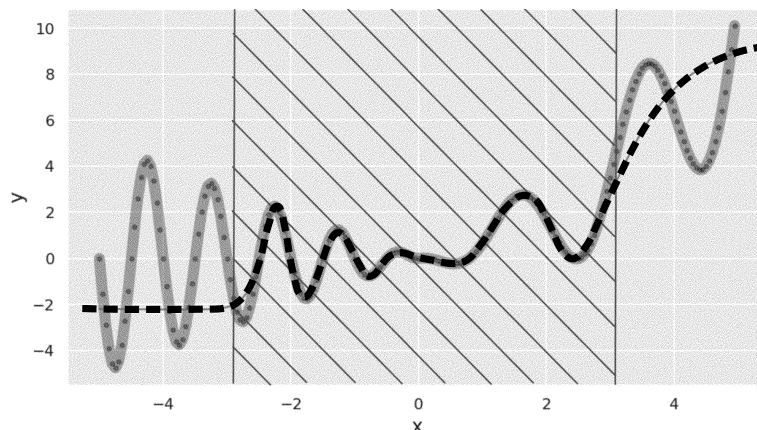


Рис. 1. Расхождение эталонной функции и функции, построенной нейронной сетью, за пределами обучающей выборки

служащей основой для генерации размеченной выборки. Черная прерывистая линия изображает функцию, построенную в результате обучения нейронной сети, которая имеет следующую архитектуру:

```
def baseline_model():
    model = Sequential()
    model.add(Dense(50, input_dim=1, activation='tanh',
                    kernel_initializer='he_normal'))
    model.add(Dense(1, input_dim=50, activation='linear',
                    kernel_initializer='he_normal'))
    sgd = SGD(lr=0.01, momentum=0.9, nesterov=True)
    model.compile(loss='mean_squared_error', optimizer=sgd)
    return model.
```

Обучение выполнено на выборке в диапазоне $(-3; 3)$, то есть в заштрихованной области. По итогам настройки аппроксимирующей функции в пределах обучающей выборки может показаться, что мы получили очень точное приближение. Однако при выходе за ее пределы видна несостоятельность построенной модели – несовпадение эталонной и обученной функций.

Причина состоит в том, что мы не пытаемся разобраться, понять и смоделировать реальный процесс, не находим и не используем глобальные закономерности, а просто пытаемся подобрать функцию с похожим поведением в известном диапазоне. Полный текст программ приведен в блокноте google colab (см. https://colab.research.google.com/drive/18UpWh9s4USmClUwqwHlVpuZrp_IsgL7n#scrollTo=ADUY-uA6Rekl).

2. Вторая причина требует более глубокого пояснения. Большую часть процессов и задач реального мира можно описать алгоритмами или системами дифференциальных уравнений, но невозможно описать функцией с разумным количеством параметров из-за ограниченности функций, вызванной ее экзистенциальной сущностью. Так, согласно классическому определению функции $y = f(x)$, предложенному Н.И. Лобачевским и П.Г.Л. Дирихле, каждому x соответствует уникальное значение y . Можно также определить многозначные функции, у которых область значений есть семейство множеств. По этой причине существует принципиальное отличие алгоритма и системы дифференциальных уравнений от функции – это зависимость не только от текущих значений входных параметров, но и от предыдущих своих состояний, то есть модель должна обладать своего рода памятью.

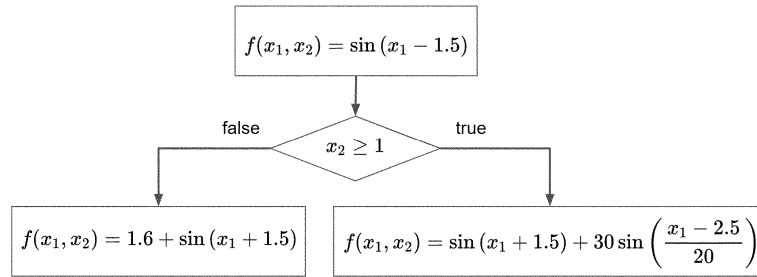


Рис. 2. Фрагмент алгоритма

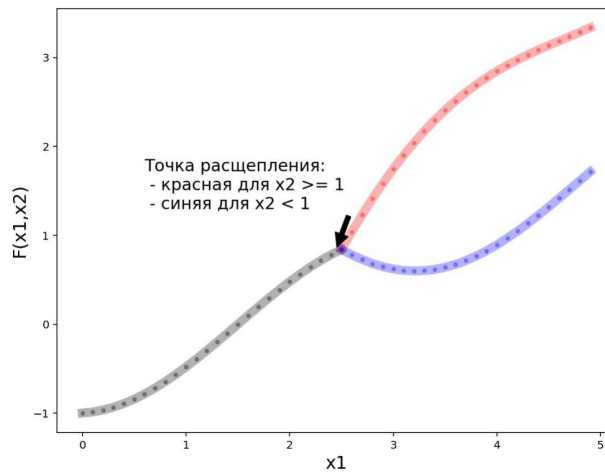


Рис. 3. Расщепление функции в точке ветвления

В задачах реального мира новое состояние системы во времени всегда учитывает ее предыдущие параметры. В алгоритмах зависимость от предыдущих значений возникает при ветвлении, являющейся основой алгоритмического подхода.

В компьютерных программах часто выполняются разные действия в зависимости от выполнения условия. Для примера на рис. 2 приведен фрагмент алгоритма (текст программ приведен в блокноте google colab, см. <https://colab.research.google.com/drive/1PdR0eovgCJfyJ7awFtmh9q28EdwiQ86#scrollTo=K53UiH8Nbxpa>).

На первом шаге выполнения алгоритма функция зависит только от x_1 . На следующем шаге в зависимости от значения параметра x_2 выполняются разные действия. Такое решение в программном коде называется ветвлением и получается с помощью if-теста, или в более общем смысле, блока if-else. График функции расщепляется в точке ветвления по крайней мере на две самостоятельные ветви. На рис. 3 представлен график функции – алгоритма процесса, который в точке $x_1 = 2.5$ развивается в зависимости от параметра, расщепление носит ступенчатый характер и происходит при достижении x_2 значения, равного 1. Безусловно, в функцию легко можно добавить новые параметры, которые учитывали бы предыдущие состояния и таким образом снова возвращали к однозначному отображению.

Для учета этого параметра и сохранения представления через функцию мы можем увеличить количество входных параметров до двух. В этом случае изображение функции $f(x_1, x_2)$ будет в виде разрезанной плоскости в трехмерном пространстве, как это показано на рис. 4.

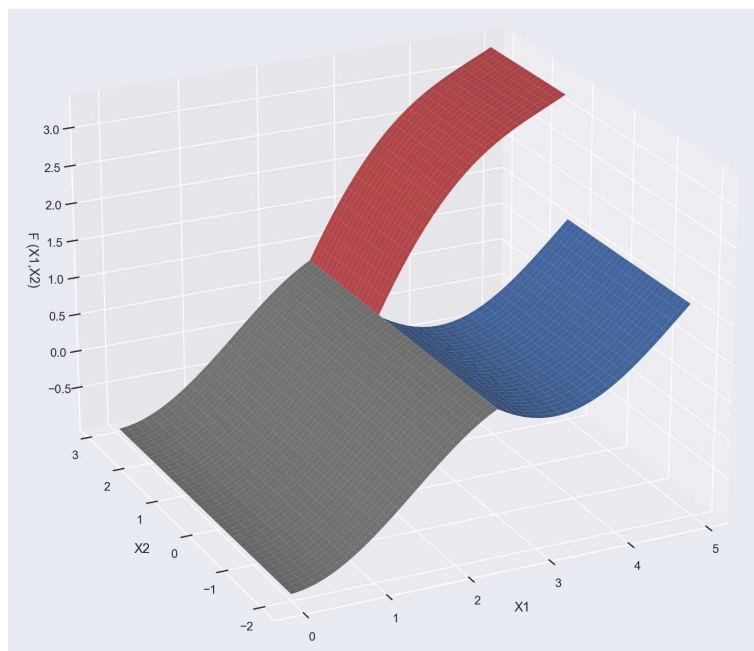


Рис. 4. Изменение размерности функции при увеличении числа параметров

На первый взгляд, функция подобным образом может отражать любой алгоритм, какое бы ветвление процессов он ни подразумевал. Однако у этого подхода есть как минимум два существенных недостатка.

1. С большой вероятностью функции, используемые в машинном обучении, не будут искать системные закономерности и упрощать получение релевантного результата. Все, как обычно, будет сводиться к прямому запоминанию через аппроксимацию, то есть к быстрому росту числа внутренних параметров, которые в пределе достигнут числа вариаций комбинаций и даже более. Например, оценочное число возможных шахматных позиций, вычисленное Клодом Шенноном, составляет 10^{43} [15].

Другим примером может служить GPT-3 (Generative Pre-trained Transformer 3) – модель прогнозирования языка, третье поколение от OpenAI. Полная версия GPT-3 имеет 175 млрд ($1.75 \cdot 10^{14}$) параметров машинного обучения [16]. Конечно, проблема комбинаторного взрыва может решаться за счет относительного неглубокого просчета или досчета шахматных позиций на каждом из ходов, что, в свою очередь, напоминает градиентный спуск, где каждый шаг – это оптимальный вектор с размерностью количества вариантов следующих позиций, а значит, по сути подбор огромного количества параметров, что аналогично настройке аппроксимирующей функции. Параллель с шахматами вполне уместна. GPT-3 – это не мыслящая машина, сопоставляющая значения слов с объектами или субъектами реального мира и прогнозирующая их взаимодействие, а просто сверхсложная функция в пространстве большой размерности, которая только поверхностно моделирует явления реального мира.

Сходная оценка прозвучала и от изобретателя сверточных нейронных сетей Яна Ле Куна: «многие ожидания по поводу способностей больших языковых моделей вроде GPT-3 являются совершенно нереалистичными», «пытаться построить умные машины, расширяя языковые модели, все равно что строить высотные само-

леты, чтобы отправиться на Луну. Вы можете побить рекорды высоты, но полет на Луну потребует совершенно другого подхода» (см. <https://news.ycombinator.com/item?id=24907318>).

2. Возникают серьезные вопросы и с обучением. Учитывая колоссальное количество параметров, необходимо подготовить огромные размеченные датасеты. В случае с GPT-3 использовались фильтрованные данные из проекта Common Crawl, Википедии, два датасета с книгами и один с текстами веб-страниц, всего объемом датасета 570 Гб, что примерно эквивалентно 400 млрд кодирований биграмм. В любом случае формирование крупных массивов размеченных данных – это огромные затраты. При этом по итогам их использования мы получаем или хорошо работающую узкоспециализированную функцию, как в случаях с распознаванием изображений, или сложную, но неглубокую симуляцию, как в случае с GPT-3. Кроме того, стоит учесть затраты на обучение трансформера с его миллиардами параметров.

Хотя языковые модели, такие как GPT-3, способны генерировать огромные объемы текста, которые трудно отличить от текста, созданного человеком, эти модели по-прежнему ненадежны и часто допускают ошибки, идущие против человеческих правил, соглашений и здравого смысла, особенно в длинных отрывках и в интерактивной обстановке.

3. Причины ограничений существующих моделей

Частично на этот вопрос проливает свет отчет исследовательской группы Стэнфордского университета «Исследования в области искусственного интеллекта за 100 лет (ИИ-100). Отчет Исследовательской группы за 2021 год» [17]. Переосмысление человеческого интеллекта за последние пять лет привело к появлению трех направлений исследований:

1) коллективный разум – понятие, согласно которому интеллект применим не только к отдельным людям, но и к коллективам;

2) когнитивная нейробиология – наука, изучающая, как устройство мозга участвует в реализации психологических и социальных процессов;

3) компьютерное моделирование – метод, позволивший разработать огромное количество моделей визуального распознавания, языковой обработки и других когнитивных процессов, основанных на машинном обучении.

Очевидно, что само понятие интеллекта начинает затрагивать очень разнообразные формы его проявления как в части базовых элементов для его построения, так и в части организационных подходов и архитектуры.

Данные обстоятельства подтверждают пригодность альтернативного определения ИИ как системы, обеспечивающей должную глубину и точность прогноза [18].

Из этого определения следует, что достоверные модели предсказания могут быть построены только при условии глубокого понимания причинно-следственных связей. Современные методы машинного обучения способны обнаруживать скрытые закономерности в данных, и эти открытия позволяют системам решать разнообразные задачи, но в основе большинства обучающих выборок лежит вторая сигнальная система – изложение событий и причинно-следственных связей через искусственную систему письменной речи, насыщенной культурными и эмоциональными аспектами. Это сдвигает выводы в сторону специфичных социальных аспектов, отдаляя от базовых для человека законов реального мира, не требующих дополнительного речевого подтверждения.

Языковые модели нейронных сетей, способные предсказывать последовательность слов, поражают своими возможностями исправлять грамматические ошибки,

отвечать на вопросы на естественном языке, писать компьютерный код, переводить с одного языка на другой и обобщать сложные или объемные специализированные тексты. Однако современные модели машинного обучения обладают лишь ограниченными возможностями для обнаружения причинно-следственных связей в мире, что подчеркивает лауреат премии Тьюринга Джуда Перл [19], в частности, в предсказании о том, как новые вмешательства будут влиять на мир, с которым они взаимодействуют. Роботы не знают, что возможно в этом мире, а что нет. Если мы хотим создать более мощные системы, то должны научить их понимать причинно-следственные связи. Остается открытым вопрос, возможно ли создавать системы ИИ, которые отлично понимали бы данные связи только из текста без реального взаимодействия.

Нильс Дж. Нильссон, пионер в области ИИ и автор одного из первых учебников, определил интеллект как способность надлежащим образом действовать в окружающей среде (“For me, artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment”) [20]. Когда в этой среде появляются люди, соответствующее поведение определяется сложными и динамичными паттернами.

В 2012 г. Левеск опубликовал статью [21], в которой предложил новый тест, так называемую схему Винограда, названную в честь ученого Терри Винограда. Проходя этот тест, ИИ отвечает на вопросы с двумя вариантами ответа. Вот один из вопросов: «Трофей не влезал в коричневый чемодан, потому что он был слишком маленьким (большим). Что было слишком маленьким (большим)?». Если в вопросе используется слово «маленьким», то ответ – «чемодан», если «большим», то «трофей». Чтобы ответить правильно, нужно обладать навыками, которые до сих пор недоступны ИИ, например, должен уметь представлять пространство, отношения между людьми, размер объектов.

Нынешнему ИИ не хватает здравого смысла, а он очень важен, поскольку обуславливает нашу связь с миром, заполняет пробелы и восполняет неявное. Мы знаем, что, если человек опрокинет стакан, вода выльется на стол, если не держать предмет в руках, то он упадет. Мы осознаем время и движение. Когда человек встает, мы знаем, что он больше не сидит, потому что человек не может находиться в этих двух состояниях одновременно. Модель мира, которую мы постепенно усваиваем, в течение первых месяцев и лет нашей жизни, позволяет нам сформировать эти многочисленные и весьма очевидные для нас законы физического мира [22].

4. Биоинспирированный подход в ИИ

Как ранее упоминалось, согласно отчету «Технологические тренды 2019. ВОИС. Искусственный Интеллект» [9], сегодня машинное обучение является доминирующим методом ИИ. На рис. 5 представлены первые пять наиболее крупных подкатегорий МО.

Из данной диаграммы можно сделать вывод, что биоинспирированный подход (Bio-inspired approach) находится на третьем месте по публикациям и не так сильно уступает позиции текущим лидерам. Кроме того, соотношение между научными публикациями и патентами косвенно свидетельствует о нахождении этого направления на стадии активного научного поиска и накопления теоретической базы в гораздо большей степени, чем этап практического внедрения у других лидеров, что говорит о вероятных перспективах указанного подхода.

Первые работы по биоинспирированным алгоритмам – симуляции эволюции – были проведены в 1954 г. Нильсом Баричелли на компьютере, установленном

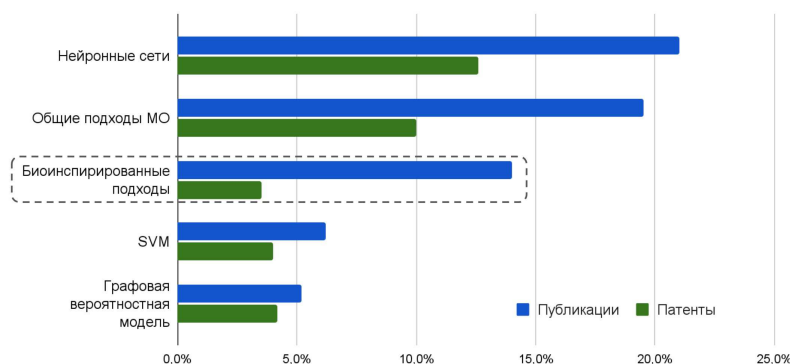


Рис. 5. Распределение публикаций по пяти крупнейшим направлениям машинного обучения

в Принстонском университете [23]. Биоинспирированный подход – это область исследования, направленная на решение проблем искусственного интеллекта с использованием биологических моделей, включая коннекционизм, искусственные нейронные сети, социальное поведение (например, роевой интеллект), а также эмерджентность (холизм) (см. [24]).

Традиционно биоинспирированные алгоритмы отличаются от классических методов искусственного интеллекта способами обучения, в которых в большинстве своем используют эволюционный подход. Как правило, биоинспирированные вычисления начинаются с набора простых правил и организмов (агентов), которые придерживаются этих правил. Со временем эти организмы развиваются в рамках простых ограничений. Этот метод можно считать восходящим или децентрализованным. В традиционном подходе интеллект часто программируется сверху вниз: программист, аналитики, эксперты являются творцами – они наполняют модель по сути своим видением, пониманием и вариативным прогнозным поведением (интеллектом).

Стоит отметить, что традиционный, исключительно «креационистский» подход к ИИ сталкивается с ограничениями быстро растущей сложности. Так, в 1956 г. У. Росс Эшби сформулировал «закон необходимого разнообразия», в котором внутренний порядок, создаваемый системой, определяется как ее реакция на внешние воздействия [25]. В последующем М. Бойсот и Б. МакКельви [26] развили принцип адекватности до «закона необходимой сложности», согласно которому для эффективной адаптации внутренняя сложность системы должна соответствовать внешней сложности, с которой она сталкивается.

Но есть еще один уровень, когда мы говорим об ИИ. Он является внутренней (управляющей) системой, а значит, система («творец»), которая проектирует и создает таковую, должна иметь еще большую сложность. Учитывая, что многие задачи реального мира имеют тенденцию к комбинаторным взрывам (например, кажущаяся простейшей задача коммивояжера), прямые алгоритмизированные взаимодействия с реальным миром быстро сталкиваются с ограничением глубины дерева решений или аналитических решений математическими методами.

В то же время эволюционный подход и роевое автопроектирование, не имеющие центра, осуществляющего управление, итерационно создают чрезвычайно сложные, гомеостазно устойчивые и эффективные системы. Мощь и немислимый для «творца» интеллект проявились в сложности и многообразии окружающего нас живого мира, в котором закон эволюции работает в полную силу, пусть и ценой

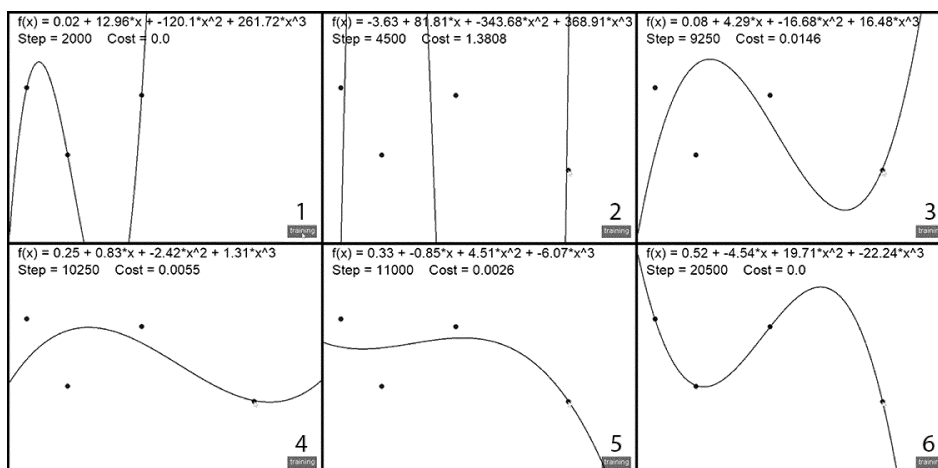


Рис. 6. Подбор коэффициентов многочлена (модели) с использованием градиентного спуска

чрезвычайных материальных и временных затрат. На первый взгляд, результаты биологической эволюции, то есть создания и управления сложнейшими экологическими системами в отсутствие не только соответствующего по сложности управляющего центра, но и вообще без такового, противоречат законам кибернетики. В действительности мы имеем дело с распределенной управляющей системой, основанной на обратной связи и прямом взаимодействии миллионов субъектов (участников эволюционного процесса) и миллиардов непрерывных итераций.

Проиллюстрируем отличие современных методов машинного обучения от биоинспирированных. На рис. 6 приведены результаты нескольких итераций процесса подбора коэффициентов многочлена (модели) с использованием градиентного спуска

$$f(x) = w_0 + w_1 \cdot x + w_2 \cdot x^2 + w_3 \cdot x^3.$$

Классические методы машинного обучения (*k*-nearest neighbors, линейная и логистическая регрессии, SVM, random forest и др.), как и искусственные нейронные сети, основаны на принципах обратной связи и итерационной эволюции модели через сравнение с эталонами и подстройкой весов. Однако в этих случаях мы имеем дело с одной моделью (пусть и сложной, состоящей из большого множества параметров) и одной петлей обратной связи, корректирующих параметры, уменьшая отклонение функции от эталонных значений. С математической точки зрения это одно сложное уравнение, в котором подбираются коэффициенты на основе сравнения с большим количеством готовых решений.

В основе биоинспирированного подхода, как правило, лежат эволюционные процессы большого количества взаимодействующих между собой интеллектуальных агентов. Показательным примером может служить рис. 7, на котором показаны этапы развития децентрализованной агентно-ориентированной модели сегрегации Томаса Шеллинга (Schelling Segregation Model, SSM) [27]. Данная модель широко используется для имитации большого количества явлений и процессов (сегрегация, ассимиляция, потребительское поведение, эволюция социальных сетей и др.) и позиционируется некоторыми исследователями как основной инструмент изучения институциональных трансформаций.

В случае с биоинспирированными системами, к которым относится модель Шеллинга, основанными на эволюционном развитии, мы имеем дело с множеством

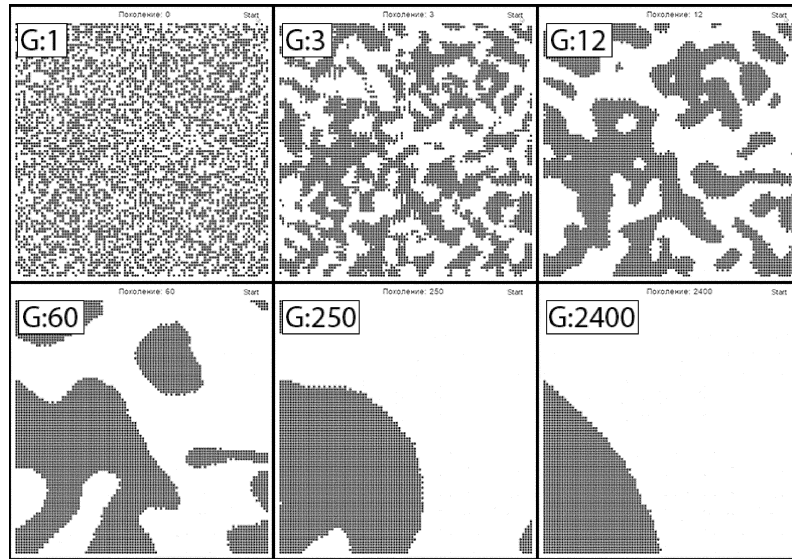


Рис. 7. Модель сегрегации Шеллинга

простых моделей/агентов, активно взаимодействующих друг с другом и имеющих:

- 1) собственную обратную связь;
- 2) обратные связи через различные цепочки взаимодействия друг с другом.

Для приведенного примера обратная связь обеспечивается следующим алгоритмом:

```

for i in range(100):
    for j in range(100):
        if i>0 and i<99 and j>0 and j<99:
            if (Cell[i-1][j-1] + Cell[i][j-1] + Cell[i+1][j-1] +
                Cell[i-1][j+1] + Cell[i][j+1] + Cell[i+1][j+1] +
                Cell[i-1][j] + Cell[i+1][j]) > np.random.choice([3,4]):
                Cell_copy[i][j] = 1
            else:
                Cell_copy[i][j] = 0

```

По сути, каждая клетка (cell) является клеточным автоматом, состояние которого зависит от параметров соседних клеток, то есть имеет восемь взаимосвязанных петель обратной связи.

В данном случае с математической точки зрения мы имеем дело с системой многочисленных уравнений. При этом следует иметь в виду, что очень часто агенты описываются дифференциальными уравнениями, а системы даже простых дифференциальных уравнений часто не имеют аналитического решения (например, уравнения движения трех тел), и единственным способом их решения является вычислительный метод.

Ряд прорывов в системах ИИ инициирован именно открытиями, совершенными в нейробиологии. Сначала в 1957 г. американский психолог Фрэнк Розенблатт, изобретатель перцептрона, вдохновился открытиями в области нейробиологии своего времени [28]. Впоследствии Кунихико Фукусима, а потом и Ян ЛеКун были вдохновлены достижениями Дэвида Хьюбела и Торстен Н. Визеля, получившими в 1981 г. Нобелевскую премию по физиологии за работу над зрительной системой кошек.

Они обнаружили, что зрение возникает в результате прохождения визуального сигнала через несколько слоев нейронов: от сетчатки до первичной зрительной коры. При этом нейроны в каждом из этих слоев выполняют особые функции. В первичной зрительной коре каждый нейрон связан только с небольшой областью поля зрения, а именно со своим рецепторным полем [22]. Каждый из нейронов в пучке реагирует на разные линии и ориентацию контура, присутствующего в воспринимающем поле, с которым связан данный пучок. Эти нейроны реагируют также и на размер элементов. Такой механизм напоминает принцип работы экстрактора признаков. Д. Хьюбел и Т.Н. Визель как раз и объясняют это тем, что области первичной зрительной коры выступают в роли экстракторов признаков. Основываясь на их работах, Ян ЛеКун предложил многослойную архитектуру сети, которая сочетает в себе чередование простых и сложных клеток. Позднее он назовет данную архитектуру сверточной сетью, или CNN.

5. Мозжечковая архитектура как один из перспективных подходов к биоинспирированным вычислениям

На наш взгляд, биологические аналогии с мозжечком, одной из древнейших структур мозга, которая есть у представителей всех классов позвоночных, может служить основой нового биоинспирированного подхода. Мозжечок является одной из основных областей мозга и участвует в регуляции двигательных систем, речи, эмоций и других когнитивных функций организма [29].

На протяжении более двух десятилетий накапливались доказательства существенной роли мозжечка в немоторных функциях. Мозжечок играет важную роль не только в процессах обучения и сенсомоторики, все больше свидетельств показывает его роль в условном обучении и вознаграждении, что позволяет формировать наши ожидания в отношении результатов поведения. В более поздних работах, где показано, что мозжечок также необходим для чувства свободы действий, когнитивного процесса, который позволяет распознавать действие как наше собственное, высказывается предположение, что мозжечок может служить интерфейсом между сенсомоторной функцией и познанием. Единая модель, которая могла бы объяснить роль мозжечка в этих процессах, пока не создана.

Учитывая количество объектов, с которыми взаимодействуют позвоночные, мозжечок должен содержать большое количество нейронов, что подтверждается на практике. Например, мозжечок мыши, мелкой землеройки и игрунки содержит 60% всех нейронов мозга; 70% нейронов мозга – у крыс, морских свинок и макак; 80% – у агути, галаго и человека [30]. Кульминацией диспаритета может служить мозг африканского слона, который содержит 257 млрд нейронов, что в три раза больше, чем в среднем человеческом мозге, однако 97.5% нейронов в мозге слона (251 млрд) находятся в мозжечке. Это делает слона исключением по количеству мозжечковых нейронов в сравнении с другими млекопитающими, что может быть связано с сенсомоторными специализациями (по всей вероятности с наличием хобота). Напротив, кора головного мозга слона, которая имеет вдвое большую массу, чем кора головного мозга человека, содержит только 5.6 млрд нейронов, что примерно составляет только 1/3 от числа нейронов в коре головного мозга человека [31].

Мозжечок получает два вида информации. Первый вид, называемый афферентной информацией, передается из спинного мозга в кору полушарий головного мозга и сигнализирует о текущем состоянии регулируемой переменной (мышечный тонус, положение тела и конечностей в пространстве). Второй – эфферентная

информация – поступает от двигательных центров коры полушарий к спинному мозгу и дает представление о требуемом конечном состоянии. Сопоставляя первое и второе, кора мозжечка рассчитывает ошибку, о которой сообщает в двигательные центры. Таким образом непрерывно корректируются как произвольные, так и автоматические движения [29].

Но вопрос о функциях мозжечка до сих пор остается одним из открытых в неврологии. В теориях о функции мозжечка в течение как минимум 25 лет преобладали внутренние модели [32, 33].

В области моделирования управления движением разработана теория прямой модели [34]. Эта теория утверждает, что движения контролируются на основе постоянных взаимодействий между нашим организмом и окружающей средой через петли прямой и обратной связи. Петли прямой связи предсказывают, что произойдет, в то время как петли обратной связи сопоставляют предсказание с тем, что произошло, чтобы мы могли реагировать соответствующим образом [35, 36].

Другая теория, называемая *feedback error learning* (FEL), утверждает, что мозжечок предоставляет обратные модели [37], в которых усилена роль обратной связи для компенсации ошибки отслеживания при точном следовании по желаемой траектории. Однако большие коэффициенты усиления неприемлемы для автономных и биологических систем, поскольку они приводят к возникновению потенциально опасных несоответствующих движений. Кроме того, обучение ошибкам с обратной связью требует сложных эталонных структур.

Синтезом этих теорий является гипотеза о том, что в мозжечке находится множество прямых и обратных моделей [38]. Эти теории связаны с представлениями робототехники о прямой и обратной кинематике; FEL – это вариант метода вычисления крутящего момента в робототехнике.

Контроллер артикуляции на основе модели мозжечка (Cerebellar Model Articulation Controller (СМАС)) был предложен Дж.С. Альбусом в 1975 г. [39]. СМАС представляет собой простую и быструю нейронную сеть, основанную на локальных аппроксимациях. СМАС широко используется в управлении роботами, коррекции цвета и приближенного оценивания моделей систем.

Базовая структура СМАС показана на рис. 8.

Принцип работы СМАС можно включить следующие этапы:

- 1) на вход поступает m -разрядный вектор (x_1, x_2, \dots, x_m) ;
- 2) каждая входная переменная x_i делится на k_i дискретных блоков; таким образом входное пространство делится на k_i^m подпространств, которые называются гиперкубами. При этом если каждая входная переменная имеет k_j режимов деления на блоки, то имеется $k_j \cdot k_i^m$ гиперкубов. Это означает, что квантованное значение входного вектора задает состояние, которое будет покрыто k_j различными гиперкубами;
- 3) каждый гиперкуб связан с единицей физической памяти, в которой хранятся синаптические веса;
- 4) выходные данные СМАС будут получены путем вычисления взвешенной суммы соответствующих весов блоков памяти.

Было доказано, что СМАС обладает высокой скоростью сходимости при обучении и сильной способностью к обобщению, однако на производительность СМАС сильно влияет режим деления входных переменных.

Модель СМАС существовала одновременно с популярной концепцией перцептрона [40], хотя в то время только разрабатывались эффективные схемы обучения для настройки перцептронов. Кроме того, существовала теория, что мозжечок аналогичен перцептрону [41]. Однако 1969 г. М. Минский и С. Пейперт указали на ограничения, заключающиеся в том, что логика исключительного ИЛИ (сложение по модулю 2) не может быть решена с помощью модели перцептрона, в своей

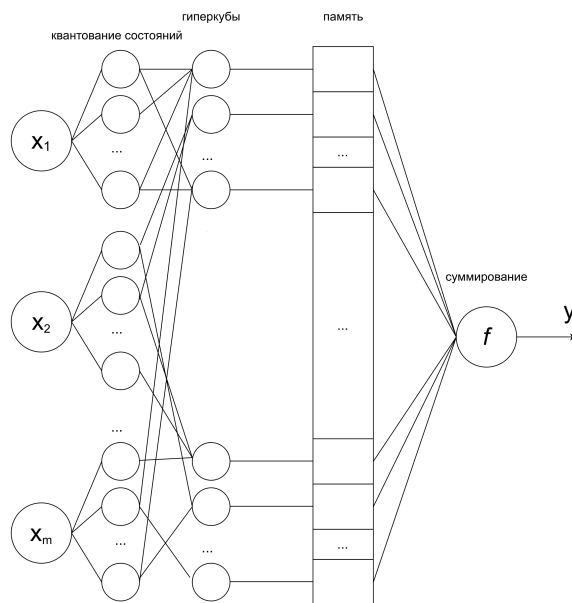


Рис. 8. Базовая схема СМАС

книге «Перцептроны: введение в вычислительную геометрию» [42]. Поэтому в последнее время подчеркивается, что СМАС следует понимать как метод обращения к ассоциативной таблице, который может адаптироваться к реальному времени.

СМАС показал свою эффективность во многих классических инженерных задачах управления. Были предложены различные обобщения структуры СМАС [43, 44]. Исследования последних лет в этой области сосредоточены на конкретных приложениях [45–47]. Однако эта модель в последние десятилетия также столкнулась с отсутствием фундаментального прорыва. По всей видимости, ограничения современных моделей СМАС связаны с чрезмерным упрощением структуры мозжечка.

Заключение

Биоинспирированные вычисления находятся на пересечении исследований в информатике, математике и биологии. В последние годы в области биоинспирированных методов в машинном обучении интенсивно проводятся исследования и разработки для поиска оптимальных решений сложных задач в науке и технике. Однако эти проблемы обычно нелинейны и также имеют нелинейные ограничения, что требует больших затрат времени вычислений и предполагает высокую размерность задач. Биоинспирированные алгоритмы дают многообещающий подход к построению эффективных технологий искусственного интеллекта.

Разработка нейронной сети следующего поколения должна учитывать последние научные результаты в нейробиологии, в частности недавно открытые новые функции и модели мозжечка. По всей видимости, недооценка значения и механизма работы мозжечка является причиной парадокса Моравека (см., например, [48]), суть которого, по наблюдениям исследователей искусственного интеллекта и робототехники, заключается в том, что, вопреки традиционным предположениям, для рассуждений требуется очень мало вычислений, а сенсомоторные навыки и навыки восприятия требуют огромных вычислительных ресурсов. Принцип был

сформулирован Хансом Моравеком, Родни Бруксом, Марвином Минским и другими исследователями в 80-х годах XX в. Х. Моравек писал в 1988 г.: «Сравнительно легко заставить компьютеры демонстрировать производительность взрослого уровня в тестах на интеллект или игре в шашки, и трудно или невозможно дать им навыки годовалого ребенка, когда дело доходит до восприятия и мобильности» [49].

До сих пор также актуально высказывание известного лингвиста и когнитивиста Стивена Пинкера. Он в 1996 г. отмечал, что «главный урок тридцатипятилетних исследований ИИ заключается в том, что трудные проблемы легки, а легкие проблемы трудны. Умственные способности четырехлетнего ребенка, которые мы считаем само собой разумеющимися: распознавание лица, поднятие карандаша, хождение по комнате, ответ на вопрос, – на самом деле решают некоторые из самых сложных инженерных задач, когда-либо задуманных... По мере появления нового поколения интеллектуальных устройств именно биржевые аналитики, нефтехимические инженеры и члены совета директоров по условно-досрочному освобождению рискуют быть замененными машинами. Садоводы, администраторы и повара уверены в своей работе на десятилетия вперед» [50].

Литература

1. *McCarthy J., Minsky M.L., Rochester N., Shannon C.E.* A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955 // *AI Mag.* – 2006. – V. 27, No 4. – P. 12–14. – doi: 10.1609/aimag.v27i4.1904.
2. *van Lent M., Fisher W., Mancuso M.* An explainable artificial intelligence system for small-unit tactical behavior // *Proc. 16th Conf. on Innovative Applications of Artificial Intelligence.* – AAAI Press, 2004. – P. 900–907.
3. *Dellermann D., Calma A., Lipusch N., Weber Th., Weigel S., Ebel P.* The future of human-AI collaboration: A taxonomy of design knowledge for hybrid intelligence systems // *Proc. 52nd Hawaii Int. Conf. on System Sciences.* – 2019. – P. 274–283.
4. *Kirsh D.* Foundations of AI: The big issues // *Artif. Intell.* – 1991. – V. 47, No 1–3. – P. 3–30. – doi: 10.1016/0004-3702(91)90048-O.
5. *Monett D., Lewis C.W.P.* Getting clarity by defining artificial intelligence – a survey // Müller V. (Ed.) *Philosophy and Theory of Artificial Intelligence 2017. PT-AI 2017. Studies in Applied Philosophy, Epistemology and Rational Ethics*, V. 44. – 2018. – P. 212–214. – doi: 10.1007/978-3-319-96448-5_21.
6. *Turing A.M.* Computing machinery and intelligence // *Mind. New Ser.* – 1950. – V. 59, No 236. – P. 433–460.
7. *Hayes P., Ford K.* Turing test considered harmful // *IJCAI'95: Proc. 14th Int. Joint Conf. on Artificial intelligence.* – 1995. – V. 1. – P. 972–977.
8. *Marcus G., Rossi F., Veloso M.* Beyond the Turing test // *AI Mag.* – 2016. – V. 37, No 1. – P. 3–4. – doi: 10.1609/aimag.v37i1.2650.
9. *WIPO Technology Trends 2019 – Artificial Intelligence.* – Geneva, Switzerland: WIPO, 2019. – 154 p. – URL: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf/.
10. *Batarseh F.A., Freeman L. Huang Ch.-H.* A survey on artificial intelligence assurance // *J. Big Data.* – 2021. – V. 8. – Art. 60, P. 1–30. – doi: 10.1186/s40537-021-00445-7.
11. *Blagec K., Barbosa-Silva A., Ott S., Samwald M.* A curated, ontology-based, large-scale knowledge graph of artificial intelligence tasks and benchmarks // *Sci. Data.* – 2022. – V. 9, No 1. – Art. 322, P. 1–10. – doi: 10.1038/s41597-022-01435-x.

12. *Pearson K.* LIII. On lines and planes of closest fit to systems of points in space // London, Edinburgh, Dublin Philos. Mag. J. Sci. Ser. 6. – 1901. – V. 2, No 11. – P. 559–572. – doi: 10.1080/14786440109462720.
13. *LeCun Y., Bengio Y., Hinton G.* Deep learning // Nature. – 2015. – V. 521, No 7553. – P. 436–444. – doi: 10.1038/nature14539.
14. *Russell S.J., Norvig P.* Artificial Intelligence: A Modern Approach. – Prentice Hall, 2010. – xviii, 1132 p.
15. *Shannon C.* XXII. Programming a computer for playing chess // Philos. Mag., Ser. 7. – 1950. – V. 41, No 314. – P. 1–18.
16. *Brown T.B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D.M., Wu J., Winter C., Hesse Ch., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner Ch., McCandlish S., Radford A., Sutskever I., Amodei D.* Language models are few-shot learners: arXiv:2005.14165v4. – 2020. – doi: 10.48550/arXiv.2005.14165.
17. *Littman M.L., Ajunwa I., Berger G., Bouilrier C., Currie M., Doshi-Velez F., Hadfield G., Horowitz M.C., Isbell Ch., Kitano H., Levy K., Lyons T., Mitchell M., Shah J., Sloman St., Vallor Sh., Walsh T.* Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report. – Stanford, CA: Stanford Univ., 2021. – 82 p. – URL: <https://ai100.stanford.edu/gathering-strength-gathering-storms-one-hundred-year-study-artificial-intelligence-ai100-2021-study/>.
18. *Sarker I.H.* Machine learning: Algorithms, real-world applications and research directions // SN Comput. Sci. – 2021. – V. 2. – Art. 160, P. 1–21. – doi: 10.1007/s42979-021-00592-x.
19. *Pearl J., Mackenzie D.* The Book of Why: The New Science of Cause and Effect. – N. Y.: Basic Books, 2018. – 432 p.
20. *Nilsson N.* The Quest for Artificial Intelligence: A History of Ideas and Achievements. – Cambridge Univ. Press, 2010. – 562 p. – doi: 10.1017/CBO9780511819346.
21. *Levesque H.J., Davis E., Morgenstern L.* The Winograd schema challenge // Proc. 13th Int. Conf. on the Principles of Knowledge Representation and Reasoning. – Inst. Electr. Electron. Eng. Inc., 2012. – P. 552–561.
22. *Лекун Я.* Как учится машина: Революция в области нейронных сетей и глубокого обучения. – М.: Альпина PRO, 2021. – 335 с.
23. *Barricelli N.* Esempi numerici di processi di evoluzione // Methodos. – 1954. – V. 6. – P. 45–68.
24. *Rajkumar R., Ganapathy V.* Bio-inspiring learning style Chatbot inventory using brain computing interface to increase the efficiency of E-learning // IEEE Access. – 2020. – V. 8. – P. 67377–67395. – doi: 10.1109/ACCESS.2020.2984591.
25. *Ashby R.W.* An Introduction to Cybernetics. – London: Chapman & Hall, 1956. – ix, 295 p.
26. *Boisot M., McKelvey B.* Complexity and organization-environment relations: Revisiting Ashby's law of requisite variety // Allen P., Maguire St., McKelvey B. (Eds.) The Sage Handbook of Complexity and Management. – London: Sage Publ., 2011. – P. 279–298. – doi: 10.4135/9781446201084.
27. *Шеллинг Т.* Динамические модели сегрегации // Журн. матем. социологии. – 1971. – № 1–2. – С. 143–186.
28. *Розенблатт Ф.* Принципы нейродинамики. Перцептроны и теория механизмов мозга. – М.: Мир, 1965. – 480 с.

29. *Middleton F.A., Strick P.L.* The cerebellum: An overview // *Trends Neurosci.* – 1998. – V. 21, No 9. – P. 367–936. – doi: 10.1016/s0166-2236(98)01330-7.
30. *Herculano-Houzel S.* Coordinated scaling of cortical and cerebellar numbers of neurons // *Front. Neuroanat.* – 2019. – V. 4. – Art. 12, P. 1–8. – doi: 10.3389/fnana.2019.00012.
31. *Herculano-Houzel S., Avelino-de-Souza K., Neves K., Porfírio J., Messeder D., Mattos Feijó L., Maldonado J., Manger P.R.* The elephant brain in numbers // *Front. Neuroanat.* – 2014. – V. 8. – Art. 46, P. 1–8. – doi: 10.3389/fnana.2014.00046.
32. *Kawato M.* Internal models for motor control and trajectory planning // *Curr. Opin. Neurobiol.* – 1999. – V. 9, No 6. – P. 718–727. – doi: 10.1016/s0959-4388(99)00028-8.
33. *Broucke M.E.* Adaptive internal models in neuroscience // *Found. Trends® Syst. Control.* – 2022. – V. 9, No 4. – P. 365–550. – doi: 10.1561/26000000027.
34. *Welniarz Q., Worbe Y., Gallea C.* The Forward Model: A Unifying Theory for the Role of the Cerebellum in Motor Control and Sense of Agency // *Front. Syst. Neurosci.* – 2021. – V. 15. – Art. 644059, P. 1–14. – doi: 10.3389/fnsys.2021.644059.
35. *Wolpert D.M., Miall R.C.* Forward models for physiological motor control // *Neural Networks.* – 1996. – V. 9, No 8. – P. 1265–1279. – doi: 10.1016/s0893-6080(96)00035-4.
36. *Green A.M., Hirata Y., Galiana H.L., Highstein S.M.* Localizing sites for plasticity in the vestibular system // Highstein S.M., Fay R.R., Popper A.N. (Eds.) *The Vestibular System. Springer Handbook of Auditory Research*, V. 19. – N. Y.: Springer, 2004. – P. 423–495. – doi: 10.1007/0-387-21567-0_10.
37. *Kawato M., Gomi H.* A computational model of four regions of the cerebellum based on feedback-error learning // *Biol. Cybern.* – 1992. – V. 68. – P. 95–103. – doi: 10.1007/BF00201431.
38. *Wolpert D.M., Ghahramani Z., Jordan M.I.* An internal model for sensorimotor integration // *Science.* – 1995. – V. 269, No 5232. – P. 1880–1882. – doi: 10.1126/science.7569931.
39. *Albus J.* A new approach to manipulator control: The cerebellar model articulation controller (CMAC) // *J. Dyn. Syst., Meas., Control.* – 1975. – V. 97, No 3. – P. 220–227. – doi: 10.1115/1.3426922.
40. *Rosenblatt F.* *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms.* – Washington, DC: Spartan Books, 1962. – 616 p.
41. *Albus J.S.* A theory of cerebellar function // *Math. Biosci.* – 1971. – V. 10, No 1–2. – P. 25–61. – doi: 10.1016/0025-5564(71)90051-4.
42. *Мунский М., Пе́йперт С.* Перцептроны. – М.: Мир, 1971. – 264 с.
43. *Gonzalez-Serrano F.J., Figueiras-Vidal A.R., Artes-Rodriguez A.* Generalizing CMAC architecture and training // *IEEE Trans. Neural Networks.* – 1998. – V. 9, No 6. – P. 1509–1514. – doi: 10.1109/72.728400.
44. *Tsa Y., Chu H.-C., Fang S.-H., Lee J., Lin C.-M.* Adaptive noise cancellation using deep cerebellar model articulation controller // *IEEE Access.* – 2018. – V. 6. – P. 37395–37402. – doi: 10.1109/ACCESS.2018.2827699.
45. *Huynh T.-T., Lin Ch.-M., Le T.-L., Cho H.-Y., Pham Th.-Th.T., Le N.-Q.-K., Chao F.* A new self-organizing fuzzy cerebellar model articulation controller for uncertain nonlinear systems using overlapped Gaussian membership functions // *IEEE Trans. Ind. Electron.* – 2020. – V. 67, No 11. – P. 9671–9682. – doi: 10.1109/TIE.2019.2952790.
46. *Fan R., Li Y.* An adaptive fuzzy trajectory tracking control via improved cerebellar model articulation controller for electro-hydraulic shovel // *IEEE/ASME Trans. Mechatron.* – 2021. – V. 26, No 6. – P. 2870–2880. – doi: 10.1109/TMECH.2021.3094284.

47. *Ji D., Shin D., Park J.* An error compensation technique for low-voltage DNN accelerators // IEEE Trans. Very Large Scale Integr (VLSI) Syst. – 2021. – V. 29, No 2. – P. 397–408. – doi: 10.1109/TVLSI.2020.3041517.
48. *Agrawal K.* To study the phenomenon of the Moravec’s paradox: arXiv:1012.3148. – 2010. – doi: 10.48550/arXiv.1012.3148.
49. *Moravec H.* Mind Children: The Future of Robot and Human Intelligence. – Cambridge, Mass.: Harvard Univ. Press, 1988. – 214 p.
50. *Pinker S.* The Language Instinct: How the Mind Creates Language. – N. Y.: William Morrow, 1994. – 494 p.

Поступила в редакцию
16.05.2022

Нурутдинов Айрат Рафкатович, генеральный директор

ПАО «Таттелеком»

ул. Н. Ершова, д. 57, г. Казань, 420061, Россия

E-mail: *ayrat.nurutdinov@gmail.com*

Латыпов Рустам Хафизович, доктор технических наук, заведующий кафедрой системного анализа и информационных технологий

Казанский (Приволжский) федеральный университет

ул. Кремлевская, д. 18, г. Казань, 420008, Россия

E-mail: *roustam.latypov@kpfu.ru*

ISSN 2541-7746 (Print)

ISSN 2500-2198 (Online)

**UCHENYE ZAPISKI KAZANSKOGO UNIVERSITETA.
SERIYA FIZIKO-MATEMATICHESKIE NAUKI
(Proceedings of Kazan University. Physics and Mathematics Series)**

2022, vol. 164, no. 2–3, pp. 244–265

REVIEW ARTICLE

doi: 10.26907/2541-7746.2022.2-3.244-265

**Potentials of the Bio-Inspired Approach
in the Development of Artificial Intelligence
Systems (Trends Review)**

A.R. Nurutdinov^{a}, R.Kh. Latypov^{b**}*

^a*Tattelecom Company, Kazan, 420061 Russia*

^b*Kazan Federal University, Kazan, 420008 Russia*

E-mail: **ayrat.nurutdinov@gmail.com, **roustam.latypov@kpfu.ru*

Received May 16, 2022

Abstract

Artificial intelligence (AI) efficiently builds predictive models in engineering, politics, economics, and science, as well as provides optimal strategies for solving various problems. However, modern AIs are often far from being as accurate as one might have expected a few decades ago. As a result, a number of problems linked to the widespread use of AI hinder the positive

effects of the tasks it solves. This article focuses on the difficulties and limitations in using AI systems that have arisen to date and possible ways to overcome them.

Keywords: artificial intelligence, machine learning, bio-inspired approach, cerebellum model

Figure Captions

Fig. 1. Divergence of the reference function and the function constructed by neural network outside the training set.

Fig. 2. Fragment of the algorithm.

Fig. 3. Splitting of the function at the branch point.

Fig. 4. Changes in the dimensionality of the function with an increase in the number of parameters.

Fig. 5. Distribution of publications in five largest areas of areas machine learning.

Fig. 6. Selection of the coefficients of a polynomial (model) using gradient descent.

Fig. 7. Schelling segregation model.

Fig. 8. Basic CMAC scheme.

References

1. McCarthy J., Minsky M.L., Rochester N., Shannon C.E. A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Mag.*, 2006, vol. 27, no. 4, pp. 12–14. doi: 10.1609/aimag.v27i4.1904.
2. van Lent M., Fisher W., Mancuso M. An explainable artificial intelligence system for small-unit tactical behavior. *Proc. 16th Conf. on Innovative Applications of Artificial Intelligence*. AAAI Press, 2004, pp. 900–907.
3. Dellermann D., Calma A., Lipusch N., Weber Th., Weigel S., Ebel P. The future of human-AI collaboration: A taxonomy of design knowledge for hybrid intelligence systems. *Proc. 52nd Hawaii Int. Conf. on System Sciences*, 2019, pp. 274–283.
4. Kirsh D. Foundations of AI: The big issues. *Artif. Intell.*, 1991, vol. 47, nos. 1–3, pp. 3–30. doi: 10.1016/0004-3702(91)90048-O.
5. Monett D., Lewis C.W.P. Getting clarity by defining artificial intelligence – a survey. In: Müller V. (Ed.) *Philosophy and Theory of Artificial Intelligence 2017. PT-AI 2017. Studies in Applied Philosophy, Epistemology and Rational Ethics*, 2018, vol. 44, pp. 212–214. doi: 10.1007/978-3-319-96448-5_21.
6. Turing A.M. Computing machinery and intelligence. *Mind. New Ser.*, 1950, vol. 59, no. 236, pp. 433–460.
7. Hayes P., Ford K. Turing test considered harmful. *IJCAI'95: Proc. 14th Int. Joint Conf. on Artificial intelligence*, 1995, vol. 1, pp. 972–977.
8. Marcus G., Rossi F., Veloso M. Beyond the Turing test. *AI Mag.*, 2016, vol. 37, no. 1, pp. 3–4. doi: 10.1609/aimag.v37i1.2650.
9. *WIPO Technology Trends 2019 – Artificial Intelligence*. Geneva, Switzerland, WIPO, 2019. 154 p. Available at: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf/.
10. Batarseh F.A., Freeman L. Huang Ch.-H. A survey on artificial intelligence assurance. *J. Big Data.*, 2021, vol. 8, art. 60, pp. 1–30. doi: 10.1186/s40537-021-00445-7.
11. Blagec K., Barbosa-Silva A., Ott S., Samwald M. A curated, ontology-based, large-scale knowledge graph of artificial intelligence tasks and benchmarks. *Sci. Data*, 2022, vol. 9, no. 1, art. 322, pp. 1–10. doi: 10.1038/s41597-022-01435-x.

12. Pearson K. LIII. On lines and planes of closest fit to systems of points in space. *London, Edinburgh, Dublin Philos. Mag. J. Sci. Ser. 6.*, 1901, vol. 2, no. 11, pp. 559–572. doi: 10.1080/14786440109462720.
13. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*, 2015, vol. 521, no. 7553, pp. 436–444. doi: 10.1038/nature14539.
14. Russell S.J., Norvig P. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2010. xviii, 1132 p.
15. Shannon C. XXII. Programming a computer for playing chess. *Philos. Mag., Ser. 7*, 1950, vol. 41, no. 314, pp. 1–18.
16. Brown T.B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D.M., Wu J., Winter C., Hesse Ch., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner Ch., McCandlish S., Radford A., Sutskever I., Amodei D. Language models are few-shot learners. *arXiv:2005.14165v4*, 2020. doi: 10.48550/arXiv.2005.14165.
17. Littman M.L., Ajunwa I., Berger G., Boutilier C., Currie M., Doshi-Velez F., Hadfield G., Horowitz M.C., Isbell Ch., Kitano H., Levy K., Lyons T., Mitchell M., Shah J., Sloman St., Vallor Sh., Walsh T. *Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report*. Stanford, CA, Stanford Univ., 2021. 82 p. Available at: <https://ai100.stanford.edu/gathering-strength-gathering-storms-one-hundred-year-study-artificial-intelligence-ai100-2021-study/>.
18. Sarker I.H. Machine learning: Algorithms, real-world applications and research directions. *SN Comput. Sci.*, 2021, vol. 2, art. 160, pp. 1–21. doi: 10.1007/s42979-021-00592-x.
19. Pearl J., Mackenzie D. *The Book of Why: The New Science of Cause and Effect*. New York, Basic Books, 2018. 432 p.
20. Nilsson N. *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge Univ. Press, 2010. 562 p. doi: 10.1017/CBO9780511819346.
21. Levesque H.J., Davis E., Morgenstern L. The winograd schema challenge. *Proc. 13th Int. Conf. on the Principles of Knowledge Representation and Reasoning*. Inst. Electr. Electron. Eng. Inc., 2012, pp. 552–561.
22. LeCun Y. *Kak uchitsya mashina: Revolyutsiya v oblasti neironnykh setei i glubokogo obucheniya* [How the Machine Learns: The Revolution in Neural Networks and Deep Learning]. Moscow, Al'pina PRO, 2021. 335 p. (In Russian)
23. Barricelli N. Esempi numerici di processi di evoluzione. *Methodos*, 1954, vol. 6, pp. 45–68. (In Italian)
24. Rajkumar R., Ganapathy V. Bio-inspiring learning style Chatbot inventory using brain computing interface to increase the efficiency of E-learning. *IEEE Access*, 2020, vol. 8, pp. 67377–67395. doi: 10.1109/ACCESS.2020.2984591.
25. Ashby R.W. *An Introduction to Cybernetics*. London, Chapman & Hall, 1956. ix, 295 p.
26. Boisot M., McKelvey B. Complexity and organization-environment relations: Revisiting Ashby's law of requisite variety. In: Allen P., Maguire St., McKelvey B. (Eds.) *The Sage Handbook of Complexity and Management*. London, Sage Publ., 2011, pp. 279–298. doi: 10.4135/9781446201084.
27. Schelling C. Dynamic models of segregation. *J. Math. Soc.*, 1971, vol. 1, no. 2, pp. 143–186. doi: 10.1080/0022250X.1971.9989794.
28. Rosenblatt F. *Printsipy neirodinamiki. Pertseptrony i teoriya mekhanizmov mozga* [Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms]. Moscow, Mir, 1965. 480 p. (In Russian)

29. Middleton F.A., Strick P.L. The cerebellum: An overview. *Trends Neurosci.*, 1998, vol. 21, no. 9, pp. 367–936. doi: 10.1016/s0166-2236(98)01330-7.
30. Herculano-Houzel S. Coordinated scaling of cortical and cerebellar numbers of neurons. *Front. Neuroanat.*, 2019, vol. 4, art. 12, pp. 1–8. doi: 10.3389/fnana.2010.00012.
31. Herculano-Houzel S., Avelino-de-Souza K., Neves K., Porfírio J., Messeder D., Mattos Feijó L., Maldonado J., Manger P.R. The elephant brain in numbers. *Front. Neuroanat.*, 2014, vol. 8, art. 46, pp. 1–8. doi: 10.3389/fnana.2014.00046.
32. Kawato M. Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.*, 1999, vol. 9, no. 6, pp. 718–727. doi: 10.1016/s0959-4388(99)00028-8.
33. Broucke M.E. Adaptive internal models in neuroscience. *Found. Trends® Syst. Control*, 2022, vol. 9, no. 4, pp. 365–550. doi: 10.1561/26000000027.
34. Welniarz Q., Worbe Y., Gallea C. The forward model: A unifying theory for the role of the cerebellum in motor control and sense of agency. *Front. Syst. Neurosci.*, 2021, vol. 15, art. 644059, pp. 1–14. doi: 10.3389/fnsys.2021.644059.
35. Wolpert D.M., Miall R.C. Forward models for physiological motor control. *Neural Networks*, 1996, vol. 9, no. 8, pp. 1265–1279. doi: 10.1016/s0893-6080(96)00035-4.
36. Green A.M., Hirata Y., Galiana H.L., Highstein S.M. Localizing sites for plasticity in the vestibular system. In: Highstein S.M., Fay R.R., Popper A.N. (Eds.) *The Vestibular System. Springer Handbook of Auditory Research*. Vol. 19. New York, Springer, 2004, pp. 423–495. doi: 10.1007/0-387-21567-0-10.
37. Kawato M., Gomi H. A computational model of four regions of the cerebellum based on feedback-error learning. *Biol. Cybern.*, 1992, vol. 68, pp. 95–103. doi: 10.1007/BF00201431.
38. Wolpert D.M., Ghahramani Z., Jordan M.I. An internal model for sensorimotor integration. *Science*, 1995, vol. 269, no. 5232, pp. 1880–1882. doi: 10.1126/science.7569931.
39. Albus J. A new approach to manipulator control: The cerebellar model articulation controller (CMAC). *J. Dyn. Syst., Meas., Control*, 1975, vol. 97, no. 3, pp. 220–227. doi: 10.1115/1.3426922.
40. Rosenblatt F. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Washington, DC, Spartan Books, 1962. 616 p.
41. Albus J.S. A theory of cerebellar function. *Math. Biosci.*, 1971, vol. 10, nos. 1–2, pp. 25–61. doi: 10.1016/0025-5564(71)90051-4.
42. Minsky M., Papert S. *Perseptrony [Perceptrons]*. Moscow, Mir, 1971. 264 p. (In Russian)
43. Gonzalez-Serrano F.J., Figueiras-Vidal A.R., Artes-Rodriguez A. Generalizing CMAC architecture and training. *IEEE Trans. Neural Networks*, 1998, vol. 9, no. 6, pp. 1509–1514. doi: 10.1109/72.728400.
44. Tsa Y., Chu H.-C., Fang S.-H., Lee J., Lin C.-M. Adaptive noise cancellation using deep cerebellar model articulation controller. *IEEE Access*, 2018, vol. 6, pp. 37395–37402. doi: 10.1109/ACCESS.2018.2827699.
45. Huynh T.-T., Lin Ch.-M., Le T.-L., Cho H.-Y., Pham Th.-Th.T., Le N.-Q.-K., Chao F. A new self-organizing fuzzy cerebellar model articulation controller for uncertain nonlinear systems using overlapped Gaussian membership functions. *IEEE Trans. Ind. Electron.*, 2020, vol. 67, no. 11, pp. 9671–9682. doi: 10.1109/TIE.2019.2952790.
46. Fan R., Li Y. An adaptive fuzzy trajectory tracking control via improved cerebellar model articulation controller for electro-hydraulic shovel. *IEEE/ASME Trans. Mechatron.*, 2021, vol. 26, no. 6, pp. 2870–2880. doi: 10.1109/TMECH.2021.3094284.

47. Ji D., Shin D., Park J. An error compensation technique for low-voltage DNN accelerators. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, 2021, vol. 29, no. 2, pp. 397–408. doi: 10.1109/TVLSI.2020.3041517.
48. Agrawal K. To study the phenomenon of the Moravec's paradox. *arXiv:1012.3148*, 2010. doi: 10.48550/arXiv.1012.3148.
49. Moravec H. *Mind Children: The Future of Robot and Human Intelligence*. Cambridge, Mass., Harvard Univ. Press, 1988. 214 p.
50. Pinker S. *The Language Instinct: How the Mind Creates Language*. New York, William Morrow, 1994. 494 p.

Для цитирования: Нурутдинов А.Р., Латыпов Р.Х. Перспективы биоинспирированного подхода в разработке систем искусственного интеллекта (обзор тенденций) // Учен. зап. Казан. ун-та. Сер. Физ.-матем. науки. – 2022. – Т. 164, кн. 2–3. – С. 244–265. – doi: 10.26907/2541-7746.2022.2-3.244-265.

For citation: Nurutdinov A.R., Latypov R.Kh. Potentials of the bio-inspired approach in the development of artificial intelligence systems (trends review). *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki*, 2022, vol. 164, no. 2–3, pp. 244–265. doi: 10.26907/2541-7746.2022.2-3.244-265. (In Russian)