

DOI: [10.14515/monitoring.2024.5.2550](https://doi.org/10.14515/monitoring.2024.5.2550)**И. И. Буханский**

**ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ — ЭТО ТВОЕ.
РЕЦ. НА КН.: ХОВАРД ДЖ., ГУГГЕР С. ГЛУБОКОЕ ОБУЧЕНИЕ
С FASTAI И PYTORCH: МИНИМУМ ФОРМУЛ,
МИНИМУМ КОДА, МАКСИМУМ ЭФФЕКТИВНОСТИ.
СПБ.: ПИТЕР, 2022**

Правильная ссылка на статью:

Буханский И. И. Искусственный интеллект — это твое // Мониторинг общественного мнения: экономические и социальные перемены. 2024. № 5. С. 203—217. <https://doi.org/10.14515/monitoring.2024.5.2550>. Рец. на кн.: Ховард Дж., Гуггер С. Глубокое обучение с fastai и PyTorch: минимум формул, минимум кода, максимум эффективности. СПб.: Питер, 2022.

For citation:

Bukhansky I. I. (2024) Artificial Intelligence Is What You Can Do. *Monitoring of Public Opinion: Economic and Social Changes*. No. 5. P. 203–217. <https://doi.org/10.14515/monitoring.2024.5.2550>. Book Review: Howard J., Guggen S. (2022) Deep Learning with fastai and PyEorch: Minimum Formulas, Minimum Code, Maximum Efficiency. St. Petersburg: Piter. (In Russ.)

Получено: 28.12.2023. Принято к публикации: 20.06.2024.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ — ЭТО ТВОЕ РЕЦ. НА КН.: ХОВАРД ДЖ., ГУГГЕР С. ГЛУБОКОЕ ОБУЧЕНИЕ С FASTAI И PYTORCH: МИНИМУМ ФОРМУЛ, МИНИМУМ КОДА, МАКСИМУМ ЭФФЕКТИВНОСТИ. СПБ.: ПИТЕР, 2022

*БУХАНСКИЙ Илья Игоревич — студент 4 курса бакалавриата направления «Социология» Высшей школы современных социальных наук, МГУ им. М. В. Ломоносова, Москва, Россия; исследовательский менеджер методлаборатории, ВЦИОМ, Москва, Россия
E-MAIL: buhanskijii@my.msu.ru
<https://orcid.org/0009-0004-7051-4960>*

Аннотация. Книга разработчика и бизнес-аналитика Джереми Ховарда и инженера-исследователя в области науки о данных Сильвейна Гуггера, вышедшая в 2020 г. и переведенная на русский язык в 2022 г., представляет собой попытку доказать читателю, что он, вне зависимости от образования и профессии, способен самостоятельно обращаться к одной из знаковых технологий XXI века — искусственному интеллекту (ИИ) и создавать продукты с его использованием. В работе проведен тщательный анализ тенденций ИИ, рассмотрен ряд технологий, необходимых для эффективной деятельности в данной сфере; описаны этапы разработки проекта с ИИ от постановки проблемы и сбора данных до создания графического интерфейса и размещения приложения на сервере; раскрыты наиболее острые проблемы, связанные со все более широким применением ИИ в компаниях; актуализирован тезис о необходимости «этического» подхода при работе с ИИ. Основная часть книги посвящена объяснению программ, реализующих тот или иной механизм ИИ (точнее, ту или иную архитектуру нейросетей), этапов и принципов написания кода, посредством которого решается одна из заранее заданных практических про-

ARTIFICIAL INTELLIGENCE IS WHAT YOU CAN DO. BOOK REVIEW: HOWARD J., GUGGER S. (2022) DEEP LEARNING WITH FASTAI AND PYTORCH: MINIMUM FORMULAS, MINIMUM CODE, MAXIMUM EFFICIENCY. ST. PETERSBURG: PITER

*Iliya I. BUKHANSKY^{1,2} — Bachelor Student in Sociology; Research Manager
E-MAIL: buhanskijii@my.msu.ru
<https://orcid.org/0009-0004-7051-4960>*

¹ Moscow State University, Moscow, Russia

² Russian Public Opinion Research Center (VCIOM), Moscow, Russia

Abstract. The book by developer and business analyst Jeremy Howard and data science research engineer Sylvain Gugger, published in 2020 and translated into Russian in 2022, is an attempt to prove to the reader that he or she, regardless of education and profession, is capable of independently accessing one of the iconic technologies of the 21st century — artificial intelligence (AI) — and creating products using it to incorporate the latter into their activities. The work provides a thorough analysis of AI trends, reviews the technology stack necessary for effective work in this field, and describes the stages of developing a project with artificial intelligence, from problem formulation and data collection to creating a graphical user interface and placing the application on the server, reveals the most acute problems associated with the increasing use of AI in companies, and updates the thesis on the need for an “ethical” approach when working with artificial intelligence. The book’s central part is devoted to explaining programs implementing one AI mechanism or another (or rather, one neural network architecture or another), as well as the stages and principles of writing a code using which one of the predetermined practical problems is solved. However, for the reader who does not want to immerse him-

блем. Однако для читателя, не желающего погружаться в программирование, работа предлагает материалы о теории ИИ, о методах получения максимальной полезности от ИИ в случае, если вы заказчик, а не исполнитель проекта, о правилах предупреждения негативных последствий включения ИИ в деятельность фирмы, следование которым требуется не только от разработчиков, но и от всей команды, работающей над продуктом, включая директорат. Таким образом, рецензируемая книга может быть полезна для представителей неограниченного круга профессий, должностей и интересов.

Ключевые слова: искусственный интеллект, глубокое обучение, нейронные сети, обработка данных, моделирование, интерпретируемость моделей, этика искусственного интеллекта

self in programming, the work offers materials about the theory of artificial intelligence, the methods of obtaining maximum usefulness from AI in case you are a customer and not a project executor, about the rules of preventing negative consequences of including artificial intelligence in the firm's activity, following which is required not only from developers but also from the whole team working on the product, including the directors. Thus, the reviewed book may be helpful for representatives of an unlimited range of professions, positions, and interests.

Keywords: artificial intelligence, deep learning, neural networks, data processing, modelling, interpretability of models, ethics of artificial intelligence

Искусственный интеллект (ИИ) в последние несколько лет развивается с колоссальной скоростью: областей применения основных его продуктов — моделей глубокого обучения (*deep learning*, *DL*) — становится все больше, качество работы ИИ достигает уровня человеческих способностей (а в ряде случаев и превышает его)¹². Чтобы не отставать от современных разработок, необходимо искать пути интеграции этого инструмента в собственную деятельность. Социология не исключение: в рамках работы социального исследователя существует множество задач, например сбор и обработка данных, предсказательное моделирование, в которых использование ИИ стало устоявшейся практикой либо находится на стадии активного освоения [Котов, 2023]. Ограничением выступает лишь представление о чрезмерной сложности для «гуманитария» научиться взаимодействию с ИИ, программированию, работе с кодом. Существование этого представления признается современными социологами, однако с ним пытаются бороться. Одним из путей противодействия сложившемуся стереотипу выступают императивные указания на ненаучный характер социологического знания в случае, если его авторы посвящают себя построению «литературных теорий», забывая о программировании как необходимом инструменте исследователя в современном инфор-

¹ Раевский Р. Искусственный интеллект научился читать лучше человека // Forbes. 2018. 20 января. URL: <https://www.forbes.ru/tehnologii/355971-iskusstvennyy-intellekt-nauchilsya-chitat-luchshe-cheloveka> (дата обращения: 20.06.2024).

² На что способен искусственный интеллект сегодня и каков его потенциал // РБК Тренды. Индустрия 4.0. 2021. 19 ноября. URL: <https://trends.rbc.ru/trends/industry/cmrm/619766d59a79471862e77e8a> (дата обращения: 20.06.2024).

мационном обществе³⁴. Другой способ менее категоричен: его сторонники, обращаясь к идее о «компьютерном мышлении» как «базовом навыке» исследователя сегодня, особое внимание акцентируют на разных уровнях этого навыка, говоря в том числе о требуемом для взаимодействия с миром технологий минимуме, доступном для освоения каждому⁵.

Развенчание этого представления — миссия рецензируемой книги, в аннотации которой заявлено: не будучи «доктором математических наук» или «ботаном», после прочтения вы сможете разворачивать «самые модные и актуальные приложения» с применением ИИ [Ховард, Гуггер, 2022: 4]. Чтобы вызвать доверие к подобному высказыванию, следует указать на детали биографии его авторов, убеждающие в профессионализме и признанности в научно-техническом сообществе будущих «наставников» читателя. Джереми Ховард — исследователь-основатель компании fast.ai, почетный профессор Школы информационных технологий и электротехники Квинслендского университета в Австралии. Ховард является автором одной из наиболее влиятельных разработок в области нейросетевого моделирования естественного языка — алгоритма ULMFiT для обучения больших языковых моделей (LLM), на котором основаны все современные LLM, включая ChatGPT. Помимо ULMFiT, вклад Ховарда в развитие глубокого обучения включает в себя создание множества инструментов и программных библиотек, нацеленных на ускорение развертывания различного рода цифровых продуктов, среди которых программная библиотека fastai — один из наиболее популярных в мире высокоуровневых фреймворков для глубокого обучения с открытым исходным кодом. О значимости Ховарда в сообществе исследователей нейронных сетей говорит и его карьерный трек: он занимал позицию президента и главного научного сотрудника *Kaggle*, крупнейшего в мире сообщества специалистов по data science, основал *Enlitic* — одну из первых в мире компаний, специализирующихся на глубоком обучении для медицины, в течение двух лет после создания удерживавшую 14-е место в списке «самых умных компаний мира» по версии *MIT Tech Review*. На сегодняшний день Ховард является председателем-основателем подразделения *The AI & Medicine Initiative (AIM)* в Университете Сан-Франциско — организации, действующей в партнерстве с Гарвардом, Стэнфордом и иными академическими медицинскими центрами.

Сильвейн Гуггер — выпускник кафедры фундаментальной и прикладной математики Высшей нормальной школы в Париже, магистр точных наук университета Париж-Дофин, преподаватель математики и информатики. Профессиональная траектория Гуггера как педагога включает в себя преподавание общей математики и введения в программирование на курсе по подготовке к поступлению в инженерные и бизнес-вузы CPGE — дополнительной школы интенсивной подготовки выпускников школ к поступлению в престижные высшие учебные заведения

³ Давыдов А. А. Системная социология: языки программирования // Институт социологии ФНИСЦ РАН. Блоги. Блог А. А. Давыдова. 2009. 22 мая. URL: https://www.isras.ru/index.php?page_id=1075 (дата обращения: 20.06.2024).

⁴ Дополнительно о трансформации социологии, ее методологии и методов вследствие развития искусственного интеллекта см. [Резаев, Трегубова, 2021; Семина, Го, 2022].

⁵ Евдокимова К. Социолог СПбГУ Наталья Трегубова: «Сейчас компьютерное мышление нужно, по сути, всем, а не только программистам или сотрудникам ИТ» // Пять углов. 2023. 5 декабря. URL: https://5uglov.ru/post/19320_sociolog_spbgu_natalya_tregubova_seychas_komputernoe_mishlenie_nuzhno_po_suti_vsem_a_ne_tolko_programmistam_ili_sotrudnikam_it (дата обращения: 20.06.2024).

Франции с одной из самых высоких рабочих нагрузок в Европе; по указанному курсу под авторством Гуггера составлены нескольких учебников. Помимо этого, он был ученым-исследователем в fast.ai, где работал над совершенствованием техник ускоренного обучения моделей ИИ в условиях ограниченности ресурсов. На текущий момент Сильвейн занимает позицию инженера-исследователя в HuggingFace — американской компании, ведущей деятельность в области разработки программного обеспечения с использованием машинного обучения; одно из наиболее значимых достижений HuggingFace — собственная платформа, позволяющая разработчикам в сфере ИИ свободно обмениваться обученными моделями; компания также поддерживает программную библиотеку Transformers, обеспечивающую всем желающим доступ к инструментам искусственного интеллекта различных модальностей.

Книга представляет собой практическое пособие, соответствующее запросу социологического сообщества на развитие социогуманитарного знания в рамках цифровизации. Работа состоит из введения и четырех частей, построенных по принципу перехода от высокоуровневого взаимодействия с искусственным интеллектом к более низкоуровневому, и двух приложений. Часть I, «Глубокое обучение на практике», содержит три главы, вводящие читателя в теорию нейросетей, ограничения области применения технологий ИИ, техническую часть создания нейронных сетей. Здесь предлагается руководство по развертыванию наиболее простых инструментов ИИ и обсуждаются проблемы (в том числе этические), возникающие при переходе от моделирования в рамках той или иной задачи до использования полученной модели на практике. Часть II, «Понимание приложений на базе FastAI», состоит из восьми глав, описывающих работу с данными и обучение типовых моделей (т. е. моделей с заранее заданной архитектурой или предопределенной последовательностью слоев (базовых структурных элементов DL-модели), внесение изменений в устройство и порядок которых со стороны разработчика не предполагается) для задач компьютерного зрения, табличного моделирования, обработки естественного языка с помощью готовых структур fastai. Часть III, «Основы глубокого обучения», представляет собой пять глав, посвященных самостоятельному созданию архитектур моделей искусственного интеллекта в сферах компьютерного зрения, табличного моделирования и обработки естественного языка при помощи примитивных слоев фреймворка PyTorch, одного из наиболее популярных в сфере DL, а также написанию цикла обучения для полученных моделей (иначе говоря, в данной части читатель учится подбирать и упорядочивать слои нейронной сети для создания уникальных архитектур, в отличие от части II, где структурные элементы модели уже отобраны и расположены тем или иным образом, вследствие чего разработчик может лишь обучать и использовать предоставленные ему готовые архитектуры). Часть IV, «Глубокое обучение с чистого листа», содержит четыре главы, посвященные повествованию и рассказывающие о создании с нуля простейших слоев моделей искусственного интеллекта и основных структур, используемых в частях I, II и III.

Введение формирует образ ИИ как одной из лидирующих технологий XXI века, динамика развития и возможности которой все возрастают благодаря ее вклю-

чению в неограниченный круг экспертных областей⁶. Для продуктивного применения DL в своем предметном поле есть лишь два обязательных требования: наличие здравого смысла и упорства. Для скорейшего освоения материала желательно иметь базовые навыки программирования и знания математики уровня средней школы, однако книга построена так, что их можно приобрести (или освежить) по мере чтения, обращаясь при необходимости к дополнительным материалам [там же: 22—23].

Авторы обещают научить читателя работать с современными моделями DL, способными выполнять «распознавание, локализацию изображений и обнаружение; классификацию документов и языковое моделирование; обработку табличных данных с категориальными, непрерывными и смешанными данными, включая временной ряд; совместную фильтрацию (рекомендации)» [там же: 23]. Знание о применении ИИ не исчерпывается техническим аспектом: Ховард и Гуггер затронут в книге множество не менее важных, чем написание кода, теоретических и философских сторон DL [там же: 23—24].

В главе 1 авторы предпринимают попытку доказать два ключевых тезиса книги: 1) работа с DL доступна для всех — не требуется ученой степени по математике, десятилетнего опыта в написании кода, огромных массивов данных и излишне дорогого оборудования [там же: 28]; 2) продуктивное освоение DL возможно лишь на практике, поэтому структура книги подчинена правилу демонстрации решения конкретной проблемы и его последующего подробного рассмотрения, чтобы читатель смог осознать ход мысли и реализовать его уже в своем коде [там же: 35—36]. Первое положение подтверждается историей авторов: будучи специалистами в сфере машинного обучения (*machine learning*, ML), Джереми Ховард (несмотря на прочную связь с техническими науками) — бакалавр философии, специалист по данным, но без формального технического образования, то есть «не математик», а Сильвейн Гуггер — академик в области точных наук, автор нескольких учебников по математике, однако «не программист» [там же: 33—35]. Второй тезис звучит в книге рефреном: совершенство в DL достигается лишь через практический опыт, поэтому написание кода и решение задач (и, что особенно интересно и полезно, задач из собственной жизни) должны стать основой для включения читателя в мир DL. К какой-либо специальной теории можно обратиться и после, и во время практики, если это будет необходимо [там же: 36—38], а минимум общей теории лаконично изложен в перечне основных терминов, главный из которых, конечно, само «глубокое обучение», определяемое как «компьютерная методика для извлечения и преобразования данных... [реализуемая] через несколько слоев нейронных сетей, каждый из которых получает вводные от предыдущего и постепенно их уточняет; слои обучаются через алгоритмы, которые минимизируют ошибки и одновременно повышают точность; таким образом сеть учится выполнять конкретные задачи» [там же: 28—29]. В дополнение к основным терминам авторы, находя это важным для понимания принципов глубокого обучения, расширяют теоретический минимум кратким экскурсом в историю ис-

⁶ О современных тенденциях в области ИИ в России см.: Яков и партнеры, Яндекс. Искусственный интеллект в России — 2023: тренды и перспективы. М., 2023. URL: https://yakov.partners/upload/iblock/c5e/c8t1wrkdne5y9a4nqlicderalwny7xh4/20231218_AI_future.pdf (дата обращения: 20.06.2023).

кусственного интеллекта [там же: 30—33] (дополнительно о терминологии и истории ИИ см. [Deng, Yu, 2014]).

Переход к непосредственной практике предваряется замечаниями о технической стороне машинного обучения (ML) вообще и глубокого обучения (DL) в частности: работа с ИИ невозможна без взаимодействия с программным обеспечением, представляющим собой в данной сфере довольно широкую палитру из фреймворков и приложений. При этом авторы считают, что принципиальной разницы, какое ПО изучать, нет: связанные с нейросетями новые технологии замещают старые решения крайне быстро, поэтому учиться следует скорее не конкретному инструменту, а умению адаптироваться к новшествам [там же: 39]. Однако начать с чего-то одного все же нужно, и для этого книга предлагает зарекомендовавший себя набор из языка программирования Python, среды программирования *Jupyter Notebook*, фреймворка для нейросетевого моделирования PyTorch и авторской разработки *fastai* (*fast artificial intelligence*) — библиотеки для высокоуровневого (на самом деле и средне-, и низкоуровневого) обращения к нейронным сетям [там же: 38—39].

Практика начинается с разработки классификатора изображений кошек и собак. Сначала Гуггер и Ховард дают рекомендации по настройке рабочей станции для развертывания нейросети, предлагая в качестве рабочего места облачные интернет-сервисы, обеспечивающие доступ к необходимому оборудованию бесплатно [там же: 40—41]. Затем приводится сам код классификатора с пояснениями к каждой строке: всего в шести строках программы помещаются все этапы моделирования, от загрузки набора данных для обучения до отладки нейросети. Результат работы кода — таблица, описывающая процесс обучения: какая по счету итерация обучения происходила, каково значение потерь, то есть ошибок работы модели, оценивая и минимизируя которые, нейросеть учится выполнять данную ей задачу лучше, на тренировочном массиве. Затем полученная нейросеть тестируется на случайном фото кота: машина действительно определяет картинку как фото именно кота [там же: 43—46].

На этом примере читатель уясняет ряд важных положений:

1) Чтобы получить обученную модель, потребовались только входные и выходные данные, то есть массив фотоизображений и массив указаний на классы изображенных на нем объектов. На основе этой информации нейросеть выработала в себе ряд алгоритмов для определения, кто изображен на картинке: кот или собака. Таким образом, если для выполнения задачи требуется лишь входная и выходная информация, то нейросети в принципе применимы к любой задаче, для которой есть исходные данные и определена цель, что с ними требуется сделать (подробно процесс обучения сети см. [там же: 46—53]).

2) Полученный классификатор позволяет рассмотреть онтологические ограничения ML, в том числе DL: для успешного обучения нейросети нужны не просто «сырые», но размеченные данные: метки, имеющие содержательную связь с поставленной целью⁷, а это, в свою очередь, означает, что в данных должны присутствовать регулярности (шаблоны, паттерны), доступные для обучения, поскольку в хаосе нейросети бесполезны [там же: 52].

⁷ Подробнее о разметке на примере задач обработки изображений см. [Макаров, Чертков, Лукьянчиков, 2023].

3) Есть у ML и моральные ограничения. К наиболее значимым относят петлю обратной связи. Это ситуация, когда результаты нейросети могут воздействовать на мир, на данных о котором обучаются следующие поколения этой сети. Например, полицейские использовали нейросеть для предсказания района с наивысшим уровнем преступности, повышенный полицейский контроль фиксирует в предсказанном нейросетью районе больше правонарушений, чем обычно, и эти данные подаются в нейросеть для дальнейшего дообучения. Так машина перестает выявлять объективные закономерности и начинает генерировать предсказания «из себя», а не основываясь на закономерностях реальности. Опасность от петли обратной связи подчеркивает необходимость здравого смысла и постоянной методологической рефлексии при проверке правильности полученных результатов [там же: 53—54].

4) Входные данные всегда следует делить на обучающую, валидирующую (отложенную) и контрольную выборки. Долгое обучение модели на одних и тех же данных приводит к переобученности, что снижает эффективность нейросети в работе с новыми данными. Контроль за переобученностью ведется через использование валидирующей части датасета. Однако хорошая метрика на отложенной выборке может быть следствием подгонки настроек сети именно под валидационные данные. Чтобы избежать «привыкания», вводят третью выборку, на которой работа машины оценивается после обучения и настройки [там же: 56—58, 77—82]⁸.

5) Полученный классификатор обучался не с нуля, а был дообучен. В этом раскрывается суть *переноса обучения (transfer learning)* — одной из базовых технологий концепции *fastai*: под свою задачу вы берете нейросеть, показавшую хороший результат на аналогичной задаче, и ее *донастраиваете* и *дообучаете*⁹. Например, если нужно создать классификатор, распределяющий фото на пять классов, *fastai* предлагает взять зарекомендовавшую себя в классификации изображений модель (допустим, ResNet), задать нужное количество классов и обучить на ваших данных¹⁰.

Разбирая шесть строк кода, авторы демонстрируют принципы работы DL-архитектур с загруженными данными. Благодаря этому для читателя DL перестает быть «черным ящиком», что, в свою очередь, позволяет разработчику объяснять, что происходит внутри используемых приложений, и тем самым рассеивать естественное для всяких нововведений недоверие со стороны как заказчиков, так и коллег-социологов [там же: 61—63].

Таким образом, в главе 1 наглядно представлен теоретический минимум о DL, требованиях и ограничениях в сфере ML, принципах построения и обучения моделей, а также советы, как выбирать программное обеспечение.

Помимо непосредственно разработки нейросети, создание DL-продукта включает в себя подготовку данных до моделирования, а после построения модели — операции по интеграции построенной нейросети в приложение и тестированию ее работы. Об этом говорится в главе 2.

⁸ О четвертом типе выборки — контролирующем, применяемом периодически после внедрения ИИ в работу, — см. [Четвериков и др., 2023].

⁹ Подробнее о *transfer learning* см.: [Pan, Yang, 2009].

¹⁰ Подробно о *переносе обучения* с помощью «*fastai*» см. с. 59—61 и в документации к фреймворку.

Прежде всего необходимо определить, какой вид нужно придать данным для решения поставленной задачи и как их разметить. Можно использовать массив фотографий, размеченный на кошек и собак, а можно создать таблицу, в которой звери описаны по отличительным признакам, и обучать модель на табличных данных. Оба способа хороши, но для разных задач. Так, известен метод классификации звука, в котором аудиоданные трансформируются в спектрограммы, то есть в картинки, к которым применяется классификатор изображений. Этот подход — неочевидный, прийти к нему можно лишь через перебор разных методик, однако именно он показывает SOTA-результаты (наилучшие из достигнутых на данный момент) в задаче аудиоклассификации. Помимо этого, в новых задачах часто хороши «старые» подходы, подобранные по аналогии данных: так, при работе с белковыми цепями, похожими по своей структуре на текст, текстуальные нейросети оказались крайне продуктивными [там же: 90, 93]. Вывод следующий: чтобы плодотворно сегментировать операции и подбирать под них инструменты, нужно не просто анализировать данные, но и быть готовым трансформировать их, причем так, чтобы не подменить стоящую задачу какой-то иной задачей, имеющей более очевидное решение. Трансформации подлежат «речь» задачи, которую надо адекватно перевести с языка непригодного для решения метода на язык пригодного метода.

За определением вида данных следует их сбор [там же: 95—100], после которого нужно снова решить, какие преобразования доступных для сбора данных будут полезны. Например, при классификации фото различные манипуляции с картинками, такие как изменение размера путем растягивания или сжатия изображения, вырезания из него случайного фрагмента, позволяют добиться единообразия подаваемого в нейросеть материала, что необходимо для ее работы. Помимо этого, такие искажения, как поворот фото на случайный градус, добавление в него случайного шума и другие, называемые аугментацией, улучшают качество сети, поскольку делают каждую итерацию (эпоху) ее обучения как бы новой (к тому же синтетически увеличивая объем датасета, что особенно полезно в случае, когда исходных данных мало): нейросети «не понимают» одинаковости измененных объектов, поэтому учатся искать на фото не конкретные сущности, а общие признаки, позволяющие решить поставленную задачу [там же: 103—105]. Ясно, что подобные трансформации применимы не только к картинкам, но и к аудио, и к табличным данным, и т. д.¹¹ Таким образом, умение ML-специалиста работать с подаваемой в сеть информацией позволяет, во-первых, создать набор данных, отвечающий требованиям к обучению DL-инструмента, и, во-вторых, улучшить качество работы модели.

Обозначенные процессы преобразования данных и их загрузки в нейросеть обернуты разработчиками `fastai` в удобный и простой инструмент — `DataLoader` в совокупности с используемым для его тонкой настройки `DataBlock` [там же: 100—104]. После загрузки данных и обучения нейросети следует еще один важный этап — анализ ошибок. Если точность модели не устраивает, следует посмотреть, где она ошибается чаще всего: возможно, ошибка проявляется в особой части данных и для исправления положения нужно увеличить их долю в обучаю-

¹¹ Об аугментации текстовых данных и ее особенностях см. [Матвеева, Махныткина, 2023].

щей выборке. Когда этот шаг не помогает, следует переходить к перестройке нейросети. Если же точность модели устраивает, анализ ошибок позволяет установить, а не достигается ли хороший результат тем, что данные не сбалансированы, и поэтому сеть просто относит, например, всю информацию к модальному классу, представляющему абсолютное большинство в массиве. Таким образом, рассмотрение ошибок действительно полезно для корректировки и проверки перформанса сети. Для его осуществления в `fastai` предусмотрено множество удобных методов [там же: 105—109]. Авторы неоднократно отмечают: работа с данными, их чистка и подготовка для загрузки в модель — очень важный этап развертывания DL-решения, занимающий, как правило, 90 % времени всей разработки [там же: 108].

После подготовки данных, обучения и отладки модели обычно возникает потребность интегрировать полученную нейросеть в приложение, которое обеспечит удобство взаимодействия с продуктом пользователям, не обладающим навыками программирования, коими и являются рядовые заказчики. Для решения этой задачи авторы предлагают использовать уже знакомый читателю `Jupyter` с последующей загрузкой программы на сервер; они описывают каждый шаг превращения кода в веб-приложение [там же: 109—117].

Наконец, перед передачей нейросети заказчику нужно протестировать поведение модели; этот этап гораздо труднее, чем в большинстве других цифровых решений. Поведение сети задается не заранее созданным алгоритмом, но «стремлением» модели соответствовать обучающим данным, что зачастую порождает две трудности: проблема данных, не соответствующих области, и проблема сдвига области. Первая трудность состоит в том, что нейросети могут попадаться данные, существенно отличающиеся от обучающей выборки. Вторая — в том, что любая модель устаревает в стремительно изменяющемся мире¹². Ни то, ни другое затруднение невозможно полностью устранить заранее. Остается тщательное, кропотливое тестирование нейросети, которое непременно включает в себя на ранних этапах полностью ручное выполнение возложенной на нейросеть функции на разнородных массивах данных [там же: 117—120].

Авторы повторяют мысль: оставлять за сетью лишь вспомогательную функцию, сохраняя ведущую роль за человеком. Прежде чем внедрять системы ML, полезно ответить на вопросы: что будет, если система заработает действительно очень хорошо? Как верные прогнозы могут повлиять на принятие решений? Какими могут быть последствия и как их можно будет фиксировать? Помимо размыслений требуется разработка систем постоянного мониторинга (человеческого надзора) и своевременного реагирования на их сигналы [там же: 117—121]. Только так, по мнению авторов, непредсказуемость DL-разработки не выйдет за рамки исправимой ошибки.

В целом в главе 2 рассмотрены процедуры, следующие до и после непосредственно моделирования, а именно: 1) выбор типа данных, их сбор, проведение необходимых преобразований, загрузка в модель; 2) анализ ошибок, внедрение обученной на данных сети в приложение, тестирование полученного приложе-

¹² Подробнее об указанных проблемах см. [Намиот, Ильюшин, 2022].

ния во избежание проблем, связанных с применением результатов работы нейросети в жизни людей.

Сюжеты главы 2 плавно переходят в размышления о потенциальном вреде и этике использования нейросетей, обсуждаемые в главе 3. Уже есть примеры, когда использование нейросетей приносило людям страдание: ошибочное распределение DL-алгоритмами врачебной помощи, сокращающее количество часов, предназначенных для больных диабетом или церебральным параличом, и др. [там же: 126—129]. Становится очевидно: эти проблемы многогранны, над ними нужно серьезно работать. Причем наиболее эффективен подход, когда над этическим вопросом думает команда, поскольку так достигается триангуляция в понимании текущей ситуации. Разные люди (по полу, возрасту профессии и т. д.), обладая разными точками зрения, позволяют охватить как можно больший объем возможных последствий и, соответственно, предусмотреть как можно больше потенциальных проблем [там же: 126, 155—157].

Почему команда должна думать об этичности своей работы, почему она не может следовать формуле «надо — делаю», оставляя ответственность за негативные аффекты итогового продукта на директорате? Потому что именно аналитики и разработчики имеют прямое отношение к опыту взаимодействия с моделью, именно они могут влиять на то, каким образом и в каких ситуациях сетью будут пользоваться (и будут ли ей пользоваться вообще), и поэтому они должны учитывать, как модель может применяться и как это применение сделать лишь положительным [там же: 129—133].

Проверка своей модели на этичность, продумывание последствий введения нейросети в жизнь — попытка не допустить возникновения ошибок. Однако нужно понимать, что в подавляющем большинстве случаев ошибки все равно рано или поздно возникнут, поэтому помимо предотвращения проблем следует учиться и их исправлению. Разные ошибки требуют разных способов их решения.

Уже рассмотренные нами петли обратной связи, конечно, нужно предугадывать: например, статистика показывает, что на встречи, посвященные технике и технологиям, более склонны ходить мужчины, и система, подбирающая потенциальных участников, при введении пола как переменной начнет воспроизводить указанное неравенство, поэтому исследователю нужно сознательно исключить признак половой принадлежности из обучения модели, что сделает сеть менее точной, но в то же время лишит ее работу дискриминационного характера. Однако часто заметить петлю до запуска сети в производство сложно. Совет один: увидев смещение в результатах модели, необходимо оперативно проанализировать его причины и переобучить модель, исключив проблемные признаки, следовательно, «здесь и сейчас» «развязать» петлю обратной связи [там же: 135—138].

Помимо петель нужно помнить о возможной необъективности ML-продукта. Необъективность — порождение так называемых шести смещений, которые подробно рассмотрены на с. 138—147. Анализ смещений наводит на мысль о «смещенности» самого человека и производимых им данных. Есть ли тогда смысл требовать объективности от алгоритмов, работающих с информацией о человеке и от человека? Да! Технологии сильны, мы доверяем им сложные вычисления, мы склонны думать, что они либо не допускают ошибок вообще, либо делают их в крайне

редких случаях, мы внедряем их в нашу жизнь со все большим размахом, вследствие чего влияние моделей на мир растет. Однако технологии и люди должны сойтись в одном: самые влиятельные персоны всегда вызывают полемику вокруг своих идей, суть которой — установить, точно ли идея, продвигаемая этой личностью, адекватна, не смещена, нацелена на благо. Так должно быть и с моделями. Нам нужно преодолеть эйфорию, вызванную последними достижениями в областях IT, DL, ведущую к ощущению неограниченности способностей машин и нейросетей, и понять, что все эти машины работают с нашими данными, в которых мы вполне можем найти смещения и, что самое главное, их устранить. Так, наши «способные на все» модели, обученные на чистых, адекватных данных, позволят нам радоваться не только их сложности и скорости, но и их объективности и этичности. Раз модели так сильно влияют на наш мир, давайте направим это влияние на позитивные изменения общества [там же: 148—149].

Кроме проверки на «этичность» непосредственно моделей и их работы, необходимо тестировать и данные: в их отношении всегда нужно думать, как информация, которая собирается и хранится, которую модель принимает на входе и генерирует на выходе, может быть использована во вред обществу, например в качестве информационного оружия. Этот тезис возвращает нас к вопросу об ответственности лиц, включенных в ML-проект: в случае, если негативные для мира последствия использования данных возможны, необходимо крайне серьезно подойти к их защите. Если же вы не уверены, что сможете защитить данные от вторжения посторонних лиц, а использование информации не по назначению способно принести значительный вред обществу, нужно решить, стоит ли продолжать проект, сохраняя тем самым возможность возникновения серьезной опасности для окружения [там же: 153].

Ясно, что приведенные рекомендации помогут командам, чувствующим ответственность за свою продукцию, достичь большей «этичности» DL для общества, однако велика вероятность, что останутся «другие», которые все же будут эксплуатировать свои алгоритмы, несмотря на их дискриминационный, смещенный характер [там же: 159]. В случае, когда речь идет о таких проблемах, лежащих в природе целых структур и областей общественной деятельности, необходимо прибегать к помощи государства. Авторское предложение — изменить на государственном уровне систему стимулов для получения прибыли, поскольку именно она ведет бизнес к нарушению общечеловеческих правил ради больших заработков [там же: 159—160]. Иной аргумент для обращения к политике ради защиты обществу — факт, что проблемы неверного применения IT-продуктов затрагивают общественные блага, такие как информационная среда, а точнее ее «чистота», или общественная конфиденциальность, или человеческие права в целом, а общественные блага должны быть защищены государством [там же: 161].

Проблема урегулирования сферы ML сродни обустройству безопасности дорожного движения: ясно, что компаниям не хочется тратить деньги на проведение краш-тестов, на внедрение стандартов безопасности и экологических стандартов, однако без государственного принуждения к таким действиям нормальная, «сохранная» жизнь человека в условиях автомобилизации становится невозможной [там же: 161—162].

Заинтересованный читатель найдет в книге множество других размышлений об этической проблеме при применении искусственного интеллекта, мы же постарались описать авторское видение данного аспекта ИИ в целом, акцентируя внимание на наиболее важных, с нашей точки зрения, положениях¹³.

Вывод из главы 3 таков: этические вопросы в сфере ML в целом и DL в частности действительно сложны, готовых верных ответов на них нет. Поэтому все исследовательское сообщество должно участвовать в поиске удовлетворительных формулировок и требований [там же: 162—163].

В трех первых главах заключена базовая информация для плодотворного знакомства с дальнейшими главами о нейросетях и их архитектурах, включающая в себя рассмотрение процессов сбора и обработки данных и обучения на них нейронных сетей, внедрения полученных DL-разработок в приложения, принципов оценки «этичности» собственных продуктов, шагов по минимизации вероятности возникновения связанных с деятельностью в сфере DL проблем. Остальные три части посвящены непосредственно практике и похожи на курс, осваивать который нужно самостоятельно, поэтому останавливаться на них в рецензии нецелесообразно.

По прочтении книги нам захотелось испробовать инструменты DL, предлагаемые авторами в рамках библиотеки *fastai*, на практике, чтобы понять, действительно ли принципы относительно простого и быстрого развертывания качественных нейросетей, делающие книгу такой привлекательной в глазах читателей, реализованы в программном обеспечении составителей текста.

Для тестов мы выбрали три соревновательные задачи с платформ *Kaggle* и *Zindi*:

а) выявление из массива постов из соцсетей тех записей, в которых говорилось о какой-либо катастрофе;

б) классификация изображений почвы, сделанных телефонными камерами, на классы в зависимости от вида повреждения земли (с целью принятия своевременных решений по сохранению плодородия почв, которым нанесен тот или иной ущерб);

в) классификация новостей, написанных на языке суахили, по темам статей.

Помимо общедоступных задач, нами была предпринята попытка использовать *fastai* для классификации аудиозаписей неоконченных телефонных интервью длительностью до 20 секунд, то есть для выполнения задачи, над которой мы уже достаточно длительное время работаем и в рамках которой у нас скопился обширный инструментарий, продемонстрировавший хороший и воспроизводимый результат, следовательно, имеется возможность сравнить перформанс моделей из новой библиотеки с зарекомендовавшими себя методами.

Проведя ряд экспериментов, мы выяснили: библиотека действительно позволяет получить качественное решение, прилагая минимум усилий к непосредственно построению архитектур моделей и к их обучению или дообучению. Однако важнейших этап подготовки данных, балансировки массивов, их аугментации, их наилучшего способа представления в памяти компьютера и, следовательно, в качестве посылок на вход нейросетей — задача, все же требующая непосредственного участия разработчика и, более того, нуждающаяся в творческом, «ищущем» подходе.

Таким образом, мы пришли к выводу, что *fastai* — заслуживающая своей популярности библиотека, применение которой способно ускорить и облегчить про-

¹³ Подробнее о данной стороне ИИ с позиции социально-философской проблематики см. [Шляпников, 2023; Ивлев, 2023].

цесс разработки нейросетевых решений во многих областях человеческой деятельности. К тому же она проста в освоении, поэтому поможет новичкам в сфере глубокого обучения совершить быстрый и стремительный старт. Тем не менее это ПО — не заменитель человека, но лишь его помощник, который позволяет специалисту ставить больше экспериментов и, следовательно, искать лучшие пути решения стоящих перед ним задач. Fastai хороша как полноценный инструмент для решения относительно простых задач, в деле же более сложных проблем она полезна как способ быстрой разработки прототипа, «прощупывания почвы», осознания, насколько способны нейросетевые модели к решению этих заданий, но все же ведущая роль в таких ситуациях должна оставаться за человеком.

Список литературы (References)

1. Ивлев Д. В. Искусственный интеллект и проблемы этики // Право и практика. 2023. № 4. С. 263—267.
Ivlev D. V. (2023) Artificial Intelligence and Ethical Issues. *Law and Practice*. No. 4. P. 263—267. (In Russ.)
2. Котов Д. А. Алгоритмы искусственного интеллекта в прикладных социологических исследованиях // СоциоДиггер. 2023. Т. 4. № 7—8. URL: <https://sociodigger.ru/articles/articles-page/algorithmy-iskusstvennogo-intellekta-v-prikladnykh-sociologicheskikh-issledovaniyakh> (дата обращения: 12.10.2024).
Kotov D. A. (2023) Artificial Intelligence Algorithms in Applied Sociological Research. *SocioDigger*. Vol. 4. No. 7—8. URL: <https://sociodigger.ru/articles/articles-page/algorithmy-iskusstvennogo-intellekta-v-prikladnykh-sociologicheskikh-issledovaniyakh> (accessed: 12.10.2024). (In Russ.)
3. Макаров К. С., Чертков С. Н., Лукьянчиков К. В. Концептуальная модель системы разметки графических данных для решения задач компьютерного зрения с использованием технологий машинного обучения // Auditorium. Электронный научный журнал Курского государственного университета. 2023. № 2. URL: https://api-mag.kursksu.ru/api/v1/get_pdf/4939/ (дата обращения: 20.06.2024).
Makarov K. S., Chertkov S. N., Lukyanchikov K. V. (2023) Conceptual Model of a Graphic Data Markup System for Solving Computer Vision Problems Using Machine Learning Technologies. *Auditorium. Electronic Scientific Journal of Kursk State University*. No. 2. URL: https://api-mag.kursksu.ru/api/v1/get_pdf/4939/ (accessed: 20.06.2024). (In Russ.)
4. Матвеева А. А., Махныткина О. В. Метод аугментации текстовых данных с сохранением стиля речи и лексики персоны // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23. № 4. С. 743—749. <https://doi.org/10.17586/2226-1494-2023-23-4-743-749>.
Matveeva A. A., Makhnytkina O. V. (2023) Text Augmentation Preserving Persona Speech Style and Vocabulary. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. Vol. 23. No. 4. P. 743—749. <https://doi.org/10.17586/2226-1494-2023-23-4-743-749>. (In Russ.)

5. Намиот Д. Е., Ильющин Е. А. Мониторинг сдвига данных в моделях машинного обучения // *International Journal of Open Information Technologies*. 2022. Т. 10. № 12. С. 84—93.
Namiot D. E., Ilyushin E. A. (2022) Data Shift Monitoring in Machine Learning Models. *International Journal of Open Information Technologies*. Vol. 10. No. 12. P. 84—93. (In Russ.)
6. Резаев А. В., Трегубова Н. Д. От социологии к новой социальной аналитике: кризис социологии и проблема искусственного интеллекта // *Социологическое обозрение*. 2021. Т. 20. № 3. С. 280—301.
Rezaev A. V., Tregubova N. D. (2021) Sociology on the Way to New Social Analytics: The Crisis in Sociology and the Problem of Artificial Intelligence. *Russian Sociological Review*. Vol. 20. No. 3. P. 280—301. (In Russ.)
7. Семина Т. В., Го В. Воздействие технологий искусственного интеллекта на социальные отношения // *Социология*. 2022. № 3. С. 173—178.
Semina T. V., Guo W. (2022) The Impact of Artificial Intelligence Technologies on Social Relations. *Sociology*. No. 3. P. 173—178. (In Russ.)
8. Ховард Дж., Гуггер С. Глубокое обучение с fastai и PyTorch: минимум формул, минимум кода, максимум эффективности. СПб.: Питер, 2022.
Howard J., Guggler S. (2022) Deep Learning for Coders with fastai and PyTorch: AI Applications Without a Ph D. St. Petersburg: Piter. (In Russ.)
9. Четвериков С. Ф., Арзамасов К. М., Андрейченко А. Е., Новик В. П., Бобровская Т. М., Владимирский А. В. Подходы к формированию выборки для контроля качества работы систем искусственного интеллекта в медико-биологических исследованиях // *Современные технологии в медицине*. 2023. Т. 15. № 2. С. 19—27. <https://doi.org/10.17691/stm2023.15.2.02>.
Chetverikov S. F., Arzamasov K. M., Andreichenko A. E., Novik V. P., Bobrovskaya T. M., Vladimirovsky A. V. (2023) Approaches to Sampling for Quality Control of Artificial Intelligence in Biomedical Research. *Sovremennye tehnologii v medicine*. Vol. 15. No. 2. P. 19—27. <https://doi.org/10.17691/stm2023.15.2.02>.
10. Шляпников В. В. Некоторые проблемы этики искусственного интеллекта // *Идеи и идеалы*. 2023. Т. 15. № 2—2. С. 365—376. <https://doi.org/10.17212/2075-0862-2023-15.2.2-365-376>.
Shlyapnikov V. V. (2023) Some Problems with Artificial Intelligence Ethics. *Ideas and Ideals*. Vol. 15. No. 2—2. P. 365—376. (In Russ.) <https://doi.org/10.17212/2075-0862-2023-15.2.2-365-376>.
11. Deng L., Yu D. (2014) Deep learning: Methods and Applications. *Foundations and Trends® in Signal Processing*. Vol. 7. No. 3—4. P. 197—387. <https://doi.org/10.1561/20000000039>.
12. Pan S. J., Yang Q. (2010) A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*. Vol. 22. No. 10. P. 1345—1359. <https://doi.org/10.1109/TKDE.2009.191>.