

УДК 004.75

ПРОЕКТИРОВАНИЕ РЕПОЗИТОРИЯ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ: АНАЛИЗ ФУНКЦИОНАЛЬНЫХ ТРЕБОВАНИЙ И РАЗРАБОТКА МЕТАЯЗЫКА ДЛЯ ОПИСАНИЯ МОДЕЛЕЙ

¹Ямашкин С.А., ¹Ямашкин А.А., ²Ямашкина Е.О., ³Занозин В.В.

¹ФГБОУ ВО «Национальный исследовательский Мордовский государственный университет им. Н.П. Огарёва», Саранск, e-mail: yamashkinsa@mail.ru;

²ФГБОУ ВО «МИРЭА – Российский технологический университет», Москва;

³ФГБОУ ВО Астраханский государственный университет», Астрахань

В данной работе представлен процесс разработки функциональных требований к репозиторию нейросетевых моделей анализа пространственных данных для обеспечения поддержки процесса принятия управленческих решений в области обеспечения условий устойчивого развития территорий. Дана характеристика ключевых вариантов использования системы. С точки зрения программной реализации каркас репозитория глубоких нейросетевых моделей функционирует на базе паттерна MVC, предполагающего декомпозицию фреймворка проекта на контроллеры, модели и виды. Сделан акцент на разработку адаптивных веб-интерфейсов, позволяющих пользоваться репозиторием при помощи компьютера, подключенного к сети Интернет. Вторая часть исследования посвящена разработке формализованной схемы хранения моделей глубокого машинного анализа пространственных данных в форме метаязыка, описанного на основе теоретико-множественного подхода. Представленное решение позволяет осуществить конвертацию в представления, используемые современными фреймворками машинного обучения в рамках репозитория нейросетевых моделей для анализа пространственных данных. Программный компонент визуального редактирования модели позволяет визуализировать нейросетевые модели в виде граф-схемы, а прикладной программный интерфейс обеспечивает возможность унифицированного взаимодействия с целью обмена данными с репозиторием для импорта и экспорта моделей и получения информации о них. Спроектированная архитектура и ее программная реализация позволяют подойти к решению научной проблемы интеграции нейронных сетей, предварительно обученных моделей с возможностью их последующего использования для решения проектных задач цифровой экономики.

Ключевые слова: глубокое обучение, машинное обучение, репозиторий, функциональные требования, метаязык, теоретико-множественный подход

DEEP NEURAL NETWORK REPOSITORY DESIGNING: ANALYSIS OF FUNCTIONAL REQUIREMENTS AND DEVELOPMENT OF A META-LANGUAGE FOR MODELS DESCRIPTION

¹Yamashkin S.A., ¹Yamashkin A.A., ²Yamashkina E.O., ³Zanozin V.V.

¹Federal State Budgetary Educational Institution of Higher Education «National Research Ogarev Mordovian State University», Saransk, e-mail: yam-ashkinsa@mail.ru;

²Federal State Budgetary Educational Institution of Higher Education
«MIREA – Russian Technological University», Moscow;

³Federal State Budgetary Educational Institution of Higher Education
«Astrakhan State University», Astrakhan

This paper presents the process of developing functional requirements for the repository of neural network models for the analysis of spatial data to support the process of making managerial decisions in the field of ensuring conditions for sustainable development of territories. The characteristics of the key options for using the system are given. From the point of view of software implementation, the framework of the repository of deep neural network models operates on the basis of the MVC pattern, which involves the decomposition of the project framework into controllers, models and views. Emphasis is made on the development of adaptive web interfaces that allow using the repository using a computer connected to the Internet. The second part of the study is devoted to the development of a formalized storage scheme for models of deep machine analysis of spatial data in the form of a meta-language, described on the basis of the set-theoretic approach. The presented solution allows converting into representations used by modern machine learning frameworks within the repository of neural network models for analyzing spatial data. The software component for visual editing of the model allows visualizing neural network models in the form of a graph-diagram, and the application programming interface provides the possibility of unified interaction in order to exchange data with the repository for importing and exporting models and obtaining information about them. The designed architecture and its software implementation allows us to approach the solution of the scientific problem of integrating neural networks, pre-trained models with the possibility of their subsequent use to solve design problems of the digital economy.

Keywords: deep learning, machine learning, repository, functional requirements, metalanguage, set-theoretic approach

Развитие инфраструктур пространственных данных (ИПД), направленных на оценку состояния природно-социально-производственных систем (ПСПС) и прогнозирование чрезвычайных процессов и явлений, имеет системообразующее значение

в решении задачи усиления связанности территории стран и регионов. Решение задач классификации, кластеризации, распознавания образов, принятия решений и прогнозирования на основе больших массивов пространственных данных играет большую роль в народном хозяйстве. Существенное место среди моделей и алгоритмов анализа данных занимают искусственные нейронные сети [1].

При решении задач эффективного использования глубоких нейросетевых моделей возникает множество трудностей, важное место среди которых занимает проблема интеграции глубоких нейросетевых моделей в единую систему с целью формирования удобного инструментария специалистов в области анализа данных [2, 3]. В настоящее время общепринятой практикой является консолидация алгоритмов анализа данных в репозиториях – информационных системах, ориентированных на формирование возможности поиска, хранения, развития и эффективного использования накапливаемых проектно-ориентированных решений [4].

Данная публикация посвящена решению научной проблемы формирования архитектуры репозитория глубоких нейросетевых моделей анализа пространственных данных для обеспечения поддержки процесса принятия управленческих решений в области обеспечения условий устойчивого развития территорий. Отдельно поднимается проблема разработки формализованной схемы хранения модели глубокого машинного анализа пространственных данных в форме метаязыка, которая позволила бы осуществить возможность их конвертации в представления, используемые современными фреймворками машинного обучения. Практические результаты данного исследования внедрены в репозиторий нейросетевых моделей для анализа пространственных данных.

Материалы и методы исследования

Разработка репозитория основывается на онтологической модели хранилища, представленной в работе [5], которая определяет принципы систематизации глубоких моделей анализа пространственных данных по классам решаемых задач, природе и размерности анализируемых данных, архитектуре и топологии, свойствам эффективности.

Важной функцией репозитория глубоких нейросетевых моделей является предоставление конечным пользователям, играющим разные роли в системе, адаптивных веб-интерфейсов для быстрого получения систематизированной информации

об оптимальной в использовании глубокой нейросетевой модели, которая должна включать структурированное описание, показатели эффективности, архитектурную и топологическую организацию, а также рекомендации по гибкой настройке гиперпараметров модели, примеры прикладного использования при решении проектно-ориентированных задач [6].

После авторизации в системе пользователь получает доступ к отдельным модулям репозитория исходя из имеющихся прав. С точки зрения программной реализации описываемая программная система функционирует на базе паттерна MVC, предполагающего декомпозицию фреймворка проекта на контроллеры (модули, предназначенные для описания программной бизнес-логики), модели (компоненты для манипулирования данными) и виды (наборы шаблонов для формирования адаптивных веб-интерфейсов). Такая организация системы позволяет повысить связанность отдельных модулей репозитория и ослабить сцепленность между ними.

Результаты исследования и их обсуждение

Стартовый вариант использования репозитория описан прецедентом «L. Авторизация в репозитории», целью которого является обеспечение разграничения прав на чтение, использование и редактирование нейросетевых моделей. Обозначенный функционал реализуется путем предоставления не авторизованному в системе пользователю формы для ввода аутентификационных данных.

Группа вариантов использования «Администрирование» включает два прецедента: «AU. Управление пользователями» (осуществляющего управление распределением ролей, добавлением, редактированием, удалением метаданных о пользователях) и «AL. Управление логами» (формирующего возможность просмотра логов системы с возможностью поиска и фильтрации с целью модерации операционных процессов работы с репозиторием).

Группа прецедентов «Управление нейросетевыми моделями» включает варианты использования, формирующие ядро системы. Интегрирующий прецедент «M. Администрирование базы моделей» декомпируется на функциональный блок «MN. Навигация по каталогу моделей» (посредством системы фильтров), включающий модуль «MR. Рекомендательная система для выбора модели» (позволяющий обеспечивать релевантный поиск, подбор и конфигурирование глубоких нейронных сетей,

тонкую настройку моделей для решения конкретных проектных задач в области анализа пространственных данных). CRUD-подсистема управления данными о моделях включает также вариант использования «МА. Добавление модели» и расширяющий его функционал «МЕ. Модификация модели», позволяющий создавать и редактировать глубокие нейронные сети.

Отдельно приведем описание функционального компонента «MV. Визуальное редактирование модели» как составного блока описываемой CRUD-подсистемы, позволяющего визуализировать нейросетевые модели глубокого обучения в виде граф-схемы, с возможностью интерактивного онлайн-редактирования топологии и архитектуры модели посредством тонкого клиента (веб-браузера). Для реализации функционала визуального редактирования модели использован язык программирования JavaScript, компонент апробирован при формировании моделей нейросетевых архитектур, представленных в открытой нейросетевой библиотеке Keras.

Согласно разработанной ранее онтологической модели, домен глубоких нейросетевых моделей должен быть ассоциирован с доменами проектных задач и анализируемых данных. Соответствующие обозначенным предметным областям прецеденты также реализованы в репозитории в виде CRUD-подсистем «Администрирование базы задач» (основана на вариантах использования TN. Навигация по реестру задач, TA. Добавление информации о задаче, TE. Редактирование информации о задаче) и «Управление данными» (сформирована прецедентами DN. Навигация по базе данных, DA. Добавление данных, DE. Редактирование данных).

С целью оптимизации процессов интеграции и практического использования глубоких нейросетевых моделей была разработана формализованная схема хранения моделей в форме метаязыка, которая делает возможной их конвертацию в представления, используемые современными фреймворками машинного обучения. Представим описание спроектированного метаязыка с позиции теоретико-множественного подхода и примем в качестве универсального множества набор моделей репозитория $MODELS$, мощность которого определяется числом моделей репозитория:

$$MODELS = \{MODEL_i \mid 1 \leq i \leq N \wedge i \in \mathbb{Z}\}.$$

Топология модели конкретной глубокой нейросетевой модели $MODEL_i$ может быть представлена в виде граф-схемы $GRAPH_i$ и структурированного метаописания $META_i$.

В свою очередь, метаописание модели формируется на основе массива дескрипторов $DESCRIPTORS_i$ (включающего название модели, аннотированное описание, метрики эффективности и т.д.) и параметров сборки $COMPILATION_i$. Граф-схема модели $GRAPH_i$ представляет собой ориентированный граф, множество вершин $LAYERS_i$ которого определяет набор слоев глубокой нейросетевой модели, а множество дуг $LINKS_i$ определяет структуру нейронной сети, устанавливает направленные связи между слоями.

$$MODEL_i = \langle META_i, GRAPH_i \rangle = \\ = \langle \langle DESCRIPTORS_i, COMPILATION_i \rangle, \\ \langle LAYERS_i, LINKS_i \rangle \rangle.$$

Кортеж параметров сборки модели $COMPILATION_i$ включает объект $optimizer_i$, описывающий методы и алгоритмы оптимизации нейросетевой модели (в том числе стохастический градиентный спуск (SGD), адаптивную оценку модели (Adam) и др.), а также параметры их функционирования. Функция потерь $loss_i$ определяет параметры, которые модель должна стремиться минимизировать в процессе обучения для решения задач регрессии и классификации. Наконец, объект $metrics_i$ определяет функцию, используемую для оценки эффективности модели, при этом рассчитываемые значения метрик не используются в процессе обучения модели. Различные метрики могут быть использованы при формировании моделей для решения задач мультиклассовой и бинарной классификации, регрессии, сегментации. Описание кортежа параметров сборки модели с позиции теоретико-множественного подхода выглядит следующим образом:

$$COMPILATION = \langle optimizer_i, loss_i, metrics_i \rangle.$$

Важным компонентом графовой модели глубокой нейросетевой модели i выступает слой $LAYER_{ij}$, который с точки зрения формализованного описания модели может быть представлен в виде набора объектов, определяющих тип и архитектуру слоя $TYPE_{layer_{ij}}$, набор интерфейсов слоя $INTERFACES_{ij}$, аргументов $PROPERTIES_{ij}$, определяющих особенности функционирования слоя, дескрипторы визуализации $VISUAL_{ij}$, определяющие особенности визуализации модели в рамках адаптивных веб-интерфейсов:

$$LAYER_{ij} = \langle TYPE_{layer_{ij}}, INTERFACES_{ij}, \\ PROPERTIES_{ij}, VISUAL_{ij} \rangle.$$

Объект $TYPE_{layer}$ определяет архитектуру слоя и в решающей степени задает особенности обработки данных и настройки гиперпараметров при обучении модели. Базовые слои модели (группа BASE) представлены следующими архитектурами: входной слой (Input), полносвязный слой (Dense), а также пользовательский программируемый слой (Custom Layer), архитектурная организация которого может быть задана пользователем. Слои для построения сверточных моделей (группа CONVOLUTION) включают сверточные слои, обрабатывающие данные различной размерности (Convolution), слои разделимой свертки (Depthwise Separable Convolution), а также слои субдескрипции (Pooling). Рекуррентные слои представлены архитектурами полносвязного рекуррентного слоя

(RNN), слоя долгой кратковременной памяти (LSTM), управляемых рекуррентных блоков (GRU), сверточным слоем долгой кратковременной памяти (Convolutional LSTM). Наконец, метаязык описания глубоких нейросетевых моделей поддерживает слои изменения размерности (Reshape), объединения и слияния (Fusion), поэлементного слияния на основе математических операций (Add, Average, Maximum, Minimum, Multiply, Subtract). Отдельную группу составляют слои регуляризации и прореживания (Regularization).

Параметры интерфейса $INTERFACES_{ij}$ слоя $LAYER_{ij}$ (входов $INPUTS_{ij}$ и выходов $OUTPUTS_{ij}$) определяются кортежем параметров $\langle name, type \rangle$, где объект $name$ задает имя интерфейса, $type$ – тип и размерность данных.

$$INTERFACES_{ij} = \langle INPUTS_{ij}, OUTPUTS_{ij} \rangle = \\ = \left\langle \left\{ INPUTS_{ij\lambda} \mid 1 \leq \lambda \leq \Lambda \wedge \lambda \in \mathbb{Z} \right\}, \left\{ OUTPUTS_{ij\mu} \mid 1 \leq \mu \leq M \wedge \mu \in \mathbb{Z} \right\} \right\rangle.$$

Набор аргументов $PROPERTIES_{ij}$ задает гиперпараметры, определяющие модель функционирования слоя $LAYER_{ij}$. Данный набор именованных аргументов может отличаться для слоев разной архитектуры и может включать, например, указание методов и алгоритмов инициализации весов слоя и регуляризации, функций активации.

Наконец, с целью обеспечения наглядной визуализации глубоких нейросетевых моделей в рамках адаптивных веб-интерфейсов метаязык предполагает хранение параметров группы $VISUAL_{ij}$, не отвечающих за функциональные особенности нейронной сети, но определяющих аспекты ее наглядного графического отображения. Объект группы $META_{ij}$ характеризуется заголовком $title_{ij}$ и схемой визуализации $scheme_{ij}$, определяющей цветовое оформление слоя в рамках веб-интерфейсов и другие стили. Элементы кортежа $POSITION_{ij}$ определяют координаты top_{ij} и $left_{ij}$ расположения слоя на полотне визуализации, а кортежа $SIZE_{ij}$ – его размеры $width_{ij}$ и $height_{ij}$.

$$VISUAL_{ij} = \langle META_{ij}, POSITION_{ij}, SIZE_{ij} \rangle = \\ = \langle \langle title_{ij}, scheme_{ij} \rangle, \langle top_{ij}, left_{ij} \rangle, \\ \langle width_{ij}, height_{ij} \rangle \rangle.$$

Для объединения слоев $LAYERS_i$ модели i в единую модель вводится множество связей $LINKS_i$, в котором каждый объект характеризуется набором параметров, определяющих слой-источник ($LAYER_{origin}$) и его

выходной интерфейс ($OUTPUT_{origin}$), а также слой-приемник ($LAYER_{target}$) и его входной интерфейс ($INPUT_{target}$).

$$LINK_{ij} = \langle \langle LAYER_{origin}, OUTPUT_{origin} \rangle, \\ \langle LAYER_{target}, INPUT_{target} \rangle \rangle.$$

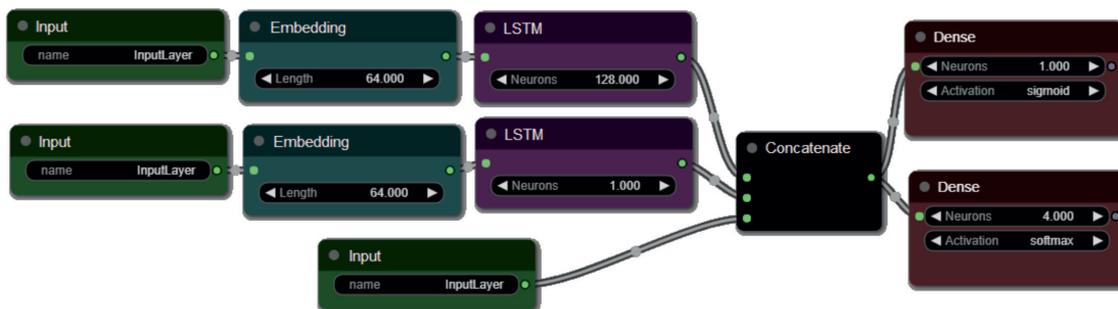
Модуль визуального редактирования модели позволяет визуализировать нейросетевые модели в виде граф-схемы с возможностью интерактивного онлайн-редактирования топологии и архитектуры модели (рисунок).

Прикладной программный интерфейс репозитория обеспечивает возможность унифицированного взаимодействия с целью обмена данными с репозиторием для импорта и экспорта нейросетевых моделей и получения информации о них.

Выводы

В статье дана характеристика проекту, посвященному решению научной проблемы накопления и систематизации глубоких нейросетевых моделей посредством проектирования и разработки репозитория обучаемых алгоритмов для анализа и прогнозирования развития пространственных процессов.

С точки зрения программной реализации каркас системы функционирует на базе паттерна MVC, предполагающего декомпозицию фреймворка проекта на контроллеры, модели и виды. Сделан акцент на разработку адаптивных веб-интерфейсов, позволяющих пользоваться репозиторием при помощи компьютера, подключенного к сети Интернет.



Визуализация нейросетевой модели в виде граф-схемы: модель с тремя входами и двумя выходами

Разработка формализованной схемы хранения моделей глубокого машинного анализа пространственных данных в форме метаязыка позволяет осуществить возможность их конвертации в представления, используемые современными фреймворками машинного обучения.

Спроектированная архитектура репозитория глубоких нейросетевых моделей и ее программная реализация позволяют подойти к решению научной проблемы интеграции нейронных сетей, предварительно обученных моделей с возможностью их последующего использования для решения проектных задач цифровой экономики.

Работа выполнена при финансовой поддержке гранта Президента Российской Федерации (грант № МК-199.2021.1.6).

Список литературы

1. Schowengerdt R.A. Remote sensing: models and methods for image processing. 3rd ed. Orlando, FL, USA: Academic Press. 2006. P. 387-456.
2. LeCun Y., Bengio Y., Hinton G. Deep learning. Nature. 2015. Vol. 521. № 7553. P. 436.
3. Zhu X.X., Tuia D., Mou L., Xia G. S., Zhang L., Xu F., Fraundorfer F. Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine. 2017. Vol. 5. № 4. P. 8-36.
4. Xiu M., Jiang Z.M.J., Adams B. An Exploratory Study on Machine-Learning Model Stores. arXiv preprint. 1905.2020. 10677.
5. Ямашкин С.А., Ямашкин А.А., Занозин В.В. Формирование репозитория глубоких нейронных сетей в системе цифровой инфраструктуры пространственных данных // Потенциал интеллектуально одарённой молодежи – развитию науки и образования: материалы IX Международного научно-форума молодых ученых, инноваторов, студентов и школьников / Под общ. ред. Т.В. Золиной. 2020. С. 370–375.
6. Скворцов М.А., Большакова М.В., Ямашкин С.А., Ямашкин А.А. Сравнительный анализ подходов к управлению базами данных для организации хранилища репозитория нейросетевых моделей // Современные наукоемкие технологии. 2021. № 6 (1). С. 108–113.