



THEORY



Сборник аннотаций

[PDF файл.](#)

Секция 1. Машинное обучение в фундаментальной физике

32. Machine Learning in Gamma Astronomy

A.P.Kryukov(1), A.P.Demichev(1)

(1) Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia

The review aims to summarize the most common deep learning methods that are used to analyze the astroparticle data collected with the help of the Imaging Atmospheric Cherenkov Telescopes (IACTs) as well as to provide references to original works.

34. Calibrating for the Future: Enhancing Calorimeter Longevity with Deep Learning

Ali Saraa (1), Bocharnikov Vladimir (1), Derkach Denis (1), Ryzhikov Artem (1)

(1) HSE University, Moscow, Russia

In the realm of high-energy physics, the calibration and longevity of calorimeters are

paramount. Our research introduces a deep learning strategy to refine the calibration process of calorimeters used in particle physics experiments. We develop a Wasserstein Generative Adversarial Network inspired architecture that adeptly calibrates the misalignment in calorimeter data due to aging. Leveraging the Wasserstein distance for loss calculation, this innovative approach requires a significantly lower number of events to achieve high precision, minimizing absolute errors effectively. Our work extends the operational lifespan of calorimeters, thereby ensuring the accuracy and reliability of data in the long term, and is particularly beneficial for experiments where data integrity is crucial for scientific discovery.

57. Методы машинного обучения в задаче предсказания дифференциальных сечений и структурных функций электророждения пиона на протоне в резонансной области.

А.В.Голда

МГУ имени М.В.Ломоносова, Физический факультет, Москва, Россия

От первых теоретических моделей построения искусственных нейронных сетей - до современных, основанных на трансформерах, архитектур нейронных сетей был проделан колоссальный путь в развитии методов искусственного интеллекта. С течением времени подобные методы машинного обучения все глубже проникают в набор повседневных инструментов исследователей в разных областях науки. Физика элементарных частиц не является исключением - на основе алгоритмов искусственного интеллекта уже можно строить генераторы данных, подавлять шумы и фоновые процессы, определять треки частиц и так далее. Исследователи физики элементарных частиц часто работают с большими объемами данных, которые позволяют эффективно строить подобные модели машинного обучения. В данной работе исследуются методы искусственного интеллекта в задаче предсказания дифференциальных сечений реакций в процессах рождения положительно заряженных пионов под действием электронов. При рассмотрении реакции рождения пиона был использован набор данных, на основе которого предсказывалось сечение реакции в различных областях фазового пространства для разных энергий пучка электронов.

В представленной работе решалась задача регрессии и основным алгоритмом, разработанным для решения данной задачи, является глубокая полносвязная нейронная сеть, с архитектурой состоящей из 19 скрытых слоев, которая обучалась с функцией потерь, не включающей в себя каких-либо априорных теоретических знаний о процессе. Для обучения были взяты данные из эксперимента с детектора CLAS, которые хранятся в базе данных «CLAS Physics Database». Поскольку предсказаниями нейронной сети являются дифференциальные сечения, то это позволило нам проверить в фиксированных областях фазового пространства соответствующие сохраняющиеся зависимости, например, в силу ротационной инвариантности амплитуды, мы показываем то, что предсказания сети сохраняют зависимость дифференциального сечения от угла между плоскостью реакции и плоскостью рассеяния. Помимо этого нами были проведены сравнения предсказанных структурных функций, пересчитанных из предсказанных значений дифференциальных сечений, с

экспериментальными данными. На основе данного алгоритма можно интерполировать и экстраполировать как значения сечений, так и значения структурных функций в различных областях фазового пространства.

43. Gamma/hadron separation in the TAIGA experiment with neural network methods

E.O.Gres(1,2), A.P.Kryukov(2), P.A.Volchugov(2), A.P.Demichev(2), J.J.Dubenskaya(2), D.P.Zhurov (1,2), S.P.Polyakov(2), E.B.Postnikov(2), A.A.Vlaskina(2)

1. Applied Physics Institute of Irkutsk State University, 2. Lomonosov MSU, SINP MSU

In this work, the ability of rare VHE gamma ray selection with neural network methods is investigated in the case when cosmic radiation flux strongly prevails (ratio up to 10^4) over the gamma radiation flux from a point source. This difference is valid for the Crab Nebula in the TeV energy range, since the Crab is a well-studied source for calibration and test of various methods and installations in gamma astronomy. TAIGA-IACT, the part of TAIGA experiment, which includes three Imaging Atmospheric Cherenkov Telescopes, observes this gamma-source too. Cherenkov telescopes obtain images of Extensive Air showers (EASs). Hillas parameters can be extracted from images in standard processing method, or images can be processed with convolutional neural networks (CNN). In this work we would like to describe the main steps and results obtained in the gamma/hadron separation task from the Crab Nebula with neural network methods. The results obtained are compared with standard processing method applied in the TAIGA collaboration and using Hillas Parameter cuts. It's demonstrated that a signal was received at the level of higher than 5.5σ in 21 hours of Crab Nebula observations after processing the experimental data with the neural network method.

57. Boosting Novelty Detection Neural Networks with Rational Activations

Zaborenko A.(1,2), Abasov E.(1,2), Boos E.(1), Bunichev V.(1), Volkov P.(1), Vorotnikov G.(1), Dudko L.(1), Iudin E.(1,2), Markina A.(1), Perfilov M.(1)

(1) Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Leninskie gory, GSP-1, Moscow 119991, Russian Federation, (2) Lomonosov Moscow State University, Faculty of Physics, Leninskie gory, GSP-1, Moscow 119991, Russian Federation

Deep Neural Networks (DNNs) have proven effective in unsupervised novelty detection tasks, including model-independent searches in High-Energy Physics. This study investigates the impact of hidden layer activation functions on the accuracy of novelty detection. Our findings highlight that both the accuracy and training stability of the final model are significantly influenced by the choice of activation function. Rational activations, a recent advancement in learnable activation functions, offer the potential for high expressiveness with a minimal number of learnable parameters. By incorporating these activations into one-class DNNs, we aim to enhance their accuracy and stability.

67. Кодирование входных сигналов в терминах ориентированных подграфов в спайковой нейронной сети с локальной пластичностью.

Я.П. Ивина(1), Р.Б. Рыбка(1), А.В. Серенко(1), М.Ю. Христинченко(2), В.А. Ильин(1)

(1) НИЦ Курчатовский институт, (2) МФТИ

Представлены результаты нового способа кодирования входных сигналов во внутренних терминах спайковой нейронной сети - в терминах ориентированных подграфов. В таком подходе ключевым является локальная пластичность спайковой сети - локального динамического механизма STDP - spike timing dependence plasticity. На примере простой модели входных сигналов, состоящей из 16 бинарных сигналов, показано, что каждый из этих сигналов уникально кодируется в виде ориентированного подграфа. Причем множество этих ориентированных подграфов изоморфно множеству входных сигналов. Каждый из этих подграфов содержит свой (уникальный) набор симплексов и гомологий. Этот результат открывает возможности по разработке новых подходов к анализу данных, например, к кластеризации и классификации, на основе алгебро-топологических методов, без этапа обучения нейронной сети.

46. Определение направления ШАЛ по данным TAIGA HiSCORE с помощью полносвязных нейросетей

Ю.Ю. Дубенская (1), А.П. Крюков (1), С.П. Поляков (1), А.А. Власкина (1), П.А. Волчугов (1), Е.О. Гресь (1,2), А.П. Демичев (1), Д.П. Журов (1,2), Е.Б. Постников (1)

(1) НИИЯФ МГУ, (2) НИИПФ ИГУ

Направление прихода широких атмосферных ливней (ШАЛ) позволяет определить источник гамма-излучения и играет важную роль для оценки энергии первичной частицы. Улавливающие черенковское излучение станции HiSCORE в проекте TAIGA распределены на площади около 1 кв. км и регистрируют время прихода фотонов и их количество, что позволяет определять направление оси ШАЛ с высокой точностью. В данной работе мы используем для определения направления ШАЛ полносвязные нейросети. Сети обучались по данным, смоделированным с помощью алгоритма Монте-Карло для гамма-квантов. Показана возможность определения направления со средней ошибкой около 0.3 градуса. В дальнейшем предложенный подход будет использован для мультимодального анализа данных, получаемых с нескольких типов установок, входящих в комплекс TAIGA.

Работа выполнена при поддержке Российского научного фонда (грант № 24-11-00136).

41. Морфологическая классификация джетов активных ядер галактик

Загоруля Д. С. (1), Лисаков М. М. (2)

(1) Московский физико-технический институт, (2) Астрокосмический центр Физического института им. П. Н. Лебедева Российской академии наук

Объем радиоинтерферометрических наблюдений квазаров позволяет использовать методы машинного обучения для классификации объектов по их визуальным особенностям. Ранее Фанаровым и Райли была проведена классификация релятивистских струй (джетов) в радиогалактиках на масштабах килопарсек и была выявлена корреляция между полной светимостью объекта и его морфологией. Морфологическая классификация на масштабах парсек, гораздо ближе к центральной чёрной дыре, позволила бы создать выборки, основанные на визуальных особенностях струй: детальнее исследовать характерные пространственные структуры джетов, а также увеличить точность для систем глобального позиционирования (ГЛОНАСС/GPS) с использованием точечных источников. Для создания подобных выборок нами были выделены четыре морфологических класса: одиночный точечный источник, двойной точечный источник, источник с одно- и двусторонним джетом. Мы разработали синтетические модели джетов активных ядер галактик (АЯГ), которые использовались для генерации обучающей выборки для классификатора. Морфологическим классификатором стала разработанная нами свёрточная нейронная сеть, обученная на синтетических данных. Обученная нейронная сеть использовалась для мультиклассовой классификации 89 тысяч изображений АЯГ из базы Астрогео. Для количественной оценки качества работы классификатора дополнительно был разработан веб-интерфейс для ручной разметки карт АЯГ. Распределение изображений по выделенным классам, предсказанное нейросетью, качественно соотносится с предполагаемым. Результирующая классификация позволит создавать выборки для астрофизических и астрометрических задач. В дальнейшей перспективе планируется адаптировать разработанный классификатор под инструмент для поиска изображений АЯГ со схожей морфологией, а также массово сравнить морфологию джетов на масштабах парсек и килопарсек.

42. Application of Kolmogorov-Arnold Networks in high energy physics

Abasov E.(1,2), Boos E.(1), Bunichev V.(1), Volkov P.(1), Vorotnikov G.(1), Dudko L.(1), Zaborenko A.(1,2), Iudin E.(1,2), Markina A.(1), Perfilov M.(1)

(1) Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, (2) Lomonosov Moscow State University, Faculty of Physics

Kolmogorov-Arnold Networks represent a recent advancement in machine learning, with the potential to outperform traditional perceptron-based neural networks across various domains as well as provide more interpretability with the use of symbolic formulas and pruning. This study explores the application of KANs to specific tasks in high-energy physics. We evaluate the performance of KANs in distinguishing multijet processes in proton-proton collisions and in reconstructing missing transverse momentum in events involving dark matter.

44. Генерация показаний наемных детекторов и поиск аномалий в данных с помощью нейронных сетей

Фитагдинов Роберт Равильевич (1,2), Харук Иван Вячеславович (1,2)

(1) МФТИ, (2) ИЯИ РАН

Основной целью данной работы было разработать генеративную нейронную сеть для реконструкции показаний поверхностных детекторов эксперимента Telescope Array и поиска аномалий в данных сгенерированных с помощью различных моделей основанных на методе Монте-Карло. Для решения этой задачи была выбрана генеративно-состязательная сеть Вастершейна с градиентным штрафом. Для обучения моделей были использованы данные, сэмплированные с помощью метода Монте-Карло. В рамках работы было решено две задачи. Первой задачей стояло написание нейронной сети для генерации показаний и поиска аномалий на наиболее активном детекторе в каждом регистрируемом событии. Написанная модель показала реалистичные результаты и помогла найти детальные различия между реальными данными и данными, сэмплированными с помощью метода Монте-Карло. Также для численного сравнения различных моделей был применен алгоритм поиска аномалий. Алгоритм выявил отличное поведение сэмплированных данных с помощью моделей QGSJET-II-03 и QGSJET-II-04 на протонах с вертикальными углами падения. Второй задачей стояло написание нейронной сети для генерации показаний 36 детекторов. В рамках этой задачи требовалось научить модель генерировать 4 показателя для каждого детектора: интегральный сигнал детектора, время прихода реконструированного плоского фронта, время прихода сигнала, и маска, показывающая сработал ли детектор. Для этой задачи были написаны функции потерь, основанные на физике распространения широкого атмосферного ливня. Модель показала визуально схожие данные и применимость функции потерь, основанной на величине правдоподобия хи-квадрат, для генерации данных физических экспериментов. Модели смогли превзойти скорость генерации данных, в сравнении с генерацией методами, основанных на методе Монте-Карло, на несколько порядков.

xx. Нейросетевое моделирование оптических солитонов, описываемых обобщенными нелинейными уравнениями Шредингера

Завертяев Савелий(1), Молошников Иван(1,2), Сбоев Александр(1,2), Кувакин Михаил(2)

1. НИЯУ МИФИ, 2. НИЦ Курчатовский институт

В работе рассматривается моделирование распространения импульсов в нелинейной среде с использованием двух уравнений в частных производных, а именно уравнения Шредингера второго порядка и обобщенного нелинейного уравнением Шредингера (ОНУШ) четвертого порядка. Показана применимость методов физически мотивированных нейронных сетей (PINN) для решения ОНУШ для анализа физических эффектов с солитонами на примере задач

взаимодействия солитона с уединенной волной. Проведено исследование эффективности методов балансировки для ОНУШ на краевой задаче с нулевыми граничными условиями, а также оценка точности метода PINN с сегментацией путем сопоставления с точным решением для одиночных солитонов ОНУШ 2 и 4 порядка. Экспериментально обосновано использование законов сохранения в качестве средства контроля достоверности решения, в отсутствие возможности сопоставления его с точным решением, в связи с чем они могут использоваться в качестве дополнительной метрики валидации полученных решений.

Ключевые слова: PINN, НУШ, ОНУШ, оптические солитоны.

25. Реконструкция ШАЛ, зарегистрированных флуоресцентными телескопами, с помощью нейронных сетей

Михаил Зотов(1), Павел Захаров(2) для коллаборации JEM-EUSO

(1) НИИЯФ МГУ (2) ВМК МГУ

Флуоресцентные телескопы (ФТ) являются одним из важных инструментов для регистрации свечения широких атмосферных ливней (ШАЛ), порождаемых космическими лучами сверхвысоких энергий. На основе модельных данных для стратосферного телескопа EUSO-SPB2 и наземного телескопа EUSO-TA, разработанных в международной коллаборации JEM-EUSO, мы покажем, каким образом различные методы машинного обучения могут быть использованы для распознавания треков ШАЛ, зарегистрированных с помощью таких приборов. Мы также представим предварительные результаты реконструкции энергии и направления прихода первичных частиц с помощью свёрточных нейронных сетей.

Моделирование и анализ для EUSO-TA поддержаны грантом РФФ 22-62-00010; разработка нейронных сетей для EUSO-SPB2 была выполнена при поддержке гранта РФФ 22-22-0367.

46. Определение направления ШАЛ по данным TAIGA HiSCORE с помощью полносвязных нейросетей

А.П. Крюков (1), С.П. Поляков (1), А.А. Власкина (1), П.А. Волчугов (1), Е.О. Гресь (1,2), А.П. Демичев (1), Ю.Ю. Дубенская (1), Д.П. Журов (1,2), Е.Б. Постников (1)

(1) НИИЯФ МГУ, (2) НИИПФ ИГУ

Направление прихода широких атмосферных ливней (ШАЛ) позволяет определить источник гамма-излучения и играет важную роль для оценки энергии первичной частицы. Улавливающие черенковское излучение станции HiSCORE в проекте TAIGA распределены на площади около 1 кв. км и регистрируют время прихода фотонов и их количество, что позволяет определять направление оси ШАЛ с высокой точностью. В данной работе мы используем для определения направления ШАЛ полносвязные нейросети. Сети обучались по данным, смоделированным с помощью алгоритма Монте-Карло для гамма-квантов. Показана возможность определения направления со

средней ошибкой около 0.3 градуса. В дальнейшем предложенный подход будет использован для мультимодального анализа данных, получаемых с нескольких типов установок, входящих в комплекс TAIGA.

Работа выполнена при поддержке Российского научного фонда (грант № 24-11-00136).

45. Применение нейронных сетей для вычисления континуальных интегралов в квантовой теории

Сальников Д.В.(1,2), Васильев А.В.(1), Иванов А.С.(1), Чистяков В.В.(1)

(1) МГУ имени М.В. Ломоносова, (2) ИЯИ РАН.

В квантовой теории поля средние значения наблюдаемых могут быть представлены в виде функционального интеграла. В общем случае его не удаётся вычислить аналитически. Существуют различные приближённые методы расчётов на решётке, например, метод Монте-Карло. В настоящее время развивается подход к решению этой задачи, использующий нейронные сети. В нашей работе мы вычислили функциональные интегралы в различных моделях квантовой механики с помощью алгоритма нормализующих потоков. Для быстрых вычислений с высокой точностью этот алгоритм применялся совместно с методом генерации при помощи марковских цепей.

40. Исследование механизма генерации гамма-вспышек в блазарах

Сяйфетдинов А.Р., Лисаков М.М.

Московский физико-технический институт, Астрокосмический центр
Физического института им. П. Н.Лебедева РАН

Блазары, разновидность активных галактических ядер, запускающих узконаправленные струи релятивистской плазмы, известны крайней изменчивостью яркости в всем электромагнитном спектре. В гамма диапазоне их яркость может удвоиться всего за несколько часов. Однако механизмы образования таких вспышек до сих пор не понятны. В данной работе мы разработали модель, позволяющую определить диапазон параметров вспышек в гамма диапазоне. Для этого мы создали датасет, выделив 869 вспышек в гамма диапазоне из кривых блеска, полученных с помощью космической обсерватории Fermi в диапазоне энергий 20 МэВ - 300 ГэВ. После создали нейросетевую модель на основе вариационного автоэнкодера (VAE) и рассмотрели геометрический механизм образования вспышек, основанный на спиральной форме струи. Модель VAE позволила получить 2-мерное информативное пространство всех вспышек, а сгенерированные геометрической моделью вспышки определили распределение параметров в этом пространстве. Полученная модель позволяет не только эффективно кластеризовать вспышки по форме, но и быстро генерировать вспышки с заданными параметрами. В дальнейшем планируется сделать более сложное моделирование образования вспышки и расширить количество определяемых физических параметров.

70. Многомерная оптимизация детекторных систем на примере мюонной защиты в эксперименте SHiP

Курбатов Е.О., Ратников Ф.Д.

НИУ ВШЭ

SHiP (Search for Hidden Particles) - новый эксперимент общего назначения на на кольце SPS в CERN, в задачи которого входит поиск скрытых частиц, предложенных многочисленными теориями вне Стандартной Модели. Важным элементом эксперимента является мюонная защита. С одной стороны, она должна обеспечивать хорошее подавление фона, а с другой - не быть слишком тяжелой. В данной работе представлены результаты получения конфигураций мюонной защиты с помощью байесовской оптимизации с применением нескольких типов суррогатов. Это позволило эффективно провести глобальную многомерную оптимизацию в 42-мерном пространстве и снизить поток мюонов в 2.5 раза при сохранении исходной массы защиты.

39. Решение задач математической физики на сетях радиальных базисных функций

Стенькин Дмитрий Александрович, Горбаченко Владимир Иванович

Пензенский государственный университет

Большинство задач математической физики не удается решить аналитически. Решение популярными методами конечных разностей и конечных элементов требует построения сетки и решения систем сеточных уравнений высокой размерности. Построить сетку зачастую довольно сложно, особенно для областей сложной конфигурации. Сеточные уравнения имеют высокую размерность и плохую обусловленность. Решить эти проблемы позволяют физически информированные нейронные сети (physics-informed neural networks). Такие сети аппроксимируют решение краевой задачи, минимизируя невязку уравнения в точках коллокации внутри и на границе области решения. Физически информированные нейронные сети позволяют найти решение в произвольных точках области без построения сетки, позволяют решать прямые и обратные краевые задачи. В физически информированных нейронных сетях обычно применяют полносвязные нейронные сети (многослойные персептроны). В качестве альтернативы авторы предлагают использовать физически информированные сети радиальных базисных функций. Сети радиальных базисных функций содержат только два слоя, что упрощает процесс обучения. Полносвязные сети сложнее обучить из-за большого числа слоёв. При решении на сетях радиальных базисных функций краевых задач математической физики сложной проблемой является выбор параметров радиальных базисных функций. Авторы предлагают в процессе обучения физически информированных сетей радиальных базисных функций настраивать не только линейные параметры сетей — веса, но и нелинейные параметры базисных функций. Для обучения физически информированных сетей радиальных базисных функций нами адаптированы градиентные алгоритмы первого порядка (градиентного спуска, Нестерова, алгоритмы с адаптивной

скоростью обучения) и алгоритм второго порядка — алгоритм Левенберга-Марквардта. Авторами впервые разработаны, реализованы и экспериментально исследованы бессеточные алгоритмы решения на физически информированных сетях радиальных базисных функций прямых краевых задач и обратных коэффициентных краевых задач для кусочно-однородных сред, отличающиеся использованием быстрого метода Левенберга-Марквардта с аналитическим вычислением матрицы Якоби. Для решения обратной задачи была применена параметрическая оптимизация. Авторами были разработаны, реализованы и экспериментально исследованы бессеточные алгоритмы решения на физически информированных сетях радиальных базисных функций задач гидродинамики. В докладе представлены результаты решения ряда модельных задач, демонстрирующие преимущества разработанных алгоритмов.

58. Enhanced Image Clustering with Random-Bond Ising Models Using LDPC Graph Representations and Nishimori Temperature Estimation

Usatyuk Vasiliy, Sapozhnikov Denis, Egorov Sergey

South-West State University, T8 LTD, South-West State University

This article investigates the application of Random-bond Ising Models (RBIMs) for clustering image data while utilizing graph representations derived from Low-Density Parity-Check (LDPC) and Quasi-Cyclic LDPC codes. These codes are constructed using an approach rooted in Energy-Based Models (EBMs), incorporating topology-aware graph models to enhance dynamic analysis within Deep Neural Networks (DNNs). By leveraging the graph structures inherent in tree-like, QC-LDPC codes, our approach aims to capture intricate patterns and latent features present in image datasets. The integration of EBMs facilitates a deeper understanding of the underlying energy landscape, enabling effective representation learning and clustering. Through empirical evaluation, we demonstrate the efficacy of our methodology in uncovering complex data structures and elucidating the interplay between graph representations, energy dynamics, and clustering performance. Our findings underscore the potential of RBIMs and LDPC-based graph models in advancing unsupervised learning tasks and enhancing the interpretability of clustering algorithms in diverse application domains. The article demonstrates the improvement of image classification algorithms based on Erdos-Rényi graph RBIM and EMD Spectrum and Hamming Distance optimized LDPC codes bipartite graphs RBIMs. The proposed approach also refines a numerical method for accurately estimating the Nishimori temperature using the eigenvalues of the Bethe Hessian matrix from the curvature landscape of optimized weighted graphs.

Секция 2. Машинное обучение в науках об окружающей среде

9. Сравнительный анализ методов машинного и глубокого обучения в задаче классификации волновых форм полного

электронного содержания

Тен А.С.(1), Сорокин А.А.(1), Шестаков Н.В.(2,3)

(1) Вычислительный центр ДВО РАН, г. Хабаровск, (2) Институт прикладной математики ДВО РАН, г. Владивосток, (3) Дальневосточный федеральный университет, г. Владивосток

Многие опасные природные явления вызывают ионосферный отклик, который может быть зафиксирован во временных рядах полного электронного содержания (ПЭС), реконструированных из данных ГНСС-наблюдений. Развитие сетей ГНСС и соответственно рост объема инструментальных данных усложняет их обработку, что значительно увеличивает время на поиск и анализ события, при этом растет риск возникновения различных технических ошибок. Цель нашего исследования – сравнить некоторые классические модели машинного обучения и нейронные сети в задаче классификации волновых форм ПЭС и оценить их возможности для создания алгоритма автоматизированного поиска ковулканических ионосферных возмущений (КИВ) по данным ГНСС. В предложенном алгоритме роль классификатора выполняет обученная модель. Модели были обучены для решения задачи бинарной классификации с использованием наборов данных волновых форм ПЭС, содержащих примеры двух классов – шум и КИВ. Волновые формы размечались на примере извержения вулкана Пик Сарычева (Курильские острова), произошедшего 11-16 июня 2009 г. Для обучения классических моделей были предложены специальные признаки данных волновых форм, подобранные с учетом вычислительной эффективности алгоритма. Проведено сравнение моделей по метрикам качества классификации: для модели на основе градиентного бустинга коэффициент корреляции Метьюса на тестовой выборке равен 0,83, а лучший результат среди нейронных сетей показала сеть FCN со значением 0,79. В результате практического применения предложенного алгоритма на тестовых данных алгоритм с классификатором на основе градиентного бустинга показал 96% обнаруженных возмущений, при небольшом количестве ложных срабатываний (48 на 200 тестовых файлов), а алгоритм с нейронной сетью FCN – 95% при 105 ложных срабатываниях. Показано, что при условии нехватки размеченных данных использование классических моделей машинного обучения более оправданно, т.к. нейронные сети имеют признаки переобучения. Полученные нами результаты позволяют утверждать, что модель, обученная на предложенных признаках данных, обеспечивает высокую точность обнаружения КИВ, и также, что предложенный алгоритм можно использовать в системах оперативного мониторинга за вулканической активностью.

72. Application of machine learning methods for analyzing data from semiconductor gas sensors in dynamic temperature mode.

Isaev I.V. (1,2,3), Chernov K.N. (4), Shchurov N.O.(1,4), Dolenko S.A. (1), Krivetskiy V.V. (3,5)

(1) D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, (2) Kotelnikov Institute of Radio Engineering and Electronics, Russian

Academy of Sciences, (3) Scientific-Manufacturing Complex Technological Centre, (4) Physics Department, M.V. Lomonosov Moscow State University, (5) Chemistry Department, M.V. Lomonosov Moscow State University

This work considered the problem of environmental monitoring of air in cities and industrial areas, which consists in detecting gases and volatile organic compounds using semiconductor gas sensors. To provide selectivity in the detection of certain gases, several semiconductor sensors with different doping components were used. Also, to ensure selectivity of gas determination, as well as high temporal resolution of the sensors, nonlinear temperature operating conditions were used - the so-called heating dynamics. Due to the high complexity of the model describing the processes of interaction between gases and sensors, machine learning methods based on the use of physical experiment data were used to process the sensor response. In this study, we considered such machine learning problem statements as the regression problem, the binary and the multi-class classification problem. The regression task was to determine the concentration of a specific gas using data from a single physical experiment. Also, within the framework of one physical experiment, the presence/absence of a specific gas was determined in a binary classification problem. Using a merged data set including several physical experiments, a multi-class classification problem was solved, which consisted of determining the presence of one of the gases under consideration. Each individual computational experiment used response data from a single sensor and within a single heating dynamic. Based on the results of the work, conclusions were drawn regarding the selection of optimal sensors and heating dynamics for a specific gas/all gases.

The study was carried out at the expense of the grant No. 22-19-00703 from the Russian Science Foundation, <https://rscf.ru/en/project/22-19-00703/> [<https://rscf.ru/en/project/22-19-00703/>].

64. ПРОГНОЗИРОВАНИЕ ХАРАКТЕРИСТИК ВОЗВРАТНОЙ МИГРАЦИИ НЕРКИ В УСТЬЕ РЕКИ ФРЕЙЗЕР С ПРИМЕНЕНИЕМ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ

М. А. Борисов (1,2) М. А. Креницкий(1,2)

(1)Московский физико-технический институт, (2)Институт океанологии имени П.П. Ширшова РАН

В регионе устья реки Фрейзер в провинции Британская Колумбия, Канада, ведется промысловый вылов нерки предприятиями из Канады и США. Этот регион характеризуется сложной системой проливов между островом Ванкувер и материковой частью Канады. Промысловые сезоны и соотношение вылова нерки между США и Канадой определяются характеристиками возвратной миграции нерки в этой системе проливов. В данном исследовании рассматриваются следующие ключевые параметры возвратной миграции нерки (*Oncorhynchus nerka*) в качестве целевых переменных: медианная дата возвратной миграции для подвидов *Chilko* и *Early Stuart* и доля северного отклонения. Медианная дата возвратной миграции характеризует момент, когда 50% возвращающихся нерки подвида *Chilko* и *Early Stuart* из северо-восточной части Тихого океана прибывают в устье пролива Хуан-де-Фука между Канадой и США. Доля нерки с северным маршрутом миграции (NDR - North

Diversion Rate), возвращающейся в реку Фрейзер через пролив Джонстон, в сравнении с возвращающимися через пролив Хуан-де-Фука, отражает процентный коэффициент ухода рыбы на северный маршрут в течение года. В настоящем исследовании предполагается связь поведения нерки с химическим составом морской воды, температурой поверхности моря и динамикой течений в верхних слоях океана. В качестве обучающей выборки используются различные переменные из данных открытых реанализов, в частности, данные Glorgys12, такие как локальные течения, скорости течений у поверхности моря, соленость и температура поверхности моря, а также переменные, связанные с взаимодействием воздуха и моря, такие как скорость ветра, характеристики морских волн, давление на уровне моря, температура воздуха, влажность и осадки. Подход Momentum Contrast (MoCo) применяется для извлечения признакового описания из данных реанализа и уменьшения размерности данных, далее в качестве моделей используются случайный лес, градиентный бустинг и гребневая регрессия. В результате получены модели, позволяющие прогнозировать характеристики возвратной миграции нерки.

56. Обзор поточечных и комплексных мер качества в исследованиях атмосферы и океана

Резвов В.Ю. (1,2), Креницкий М.А. (1,2)

(1) Московский физико-технический институт, (2) Институт океанологии им. П.П. Ширшова РАН

В науках об океане и атмосфере для описания качества моделирования и масштабирования используются обобщенные количественные показатели, называемые метриками. Метрики, или меры, качества дают представление о точности воспроизведения моделью физических процессов, и позволяют оценить неопределенности результатов моделей. В настоящей работе мы представляем классификацию наиболее часто встречающихся в научной литературе метрик качества. Для каждой группы мер качества приведены примеры их использования в научных задачах. Помимо оценки достоинств и недостатков традиционных поточечных метрик исследуется актуальность применения комплексных методов, рассматривающих различные аспекты результатов моделирования. Среди таких специфических метрик выделяются меры с акцентом на пространственной структуре и неоднородностях прогнозируемых полей и вероятностные методы проверки ансамблевых прогнозов. Отдельное внимание в данной работе также посвящено росту популярности феноменологических метрик, включая метрики, основанные на редких и экстремальных явлениях. Все рассматриваемые меры дополняются существующими в научной литературе оценками их применения к научным задачам исследования атмосферы и океана.

36. Статистическое прогнозирование уровня загрязнения воздуха частицами PM10 и PM2.5 в городских агломерациях в условиях сложного рельефа с применением методов машинного обучения на примере г. Гренобль

Суслов Александр Иванович (1), Креницкий Михаил Алексеевич (1,2), Staquet

Chantal (3), Le Bouëdec Enzo (3)

(1) Институт океанологии им. П. П. Ширшова РАН, (2) Московский физико-технический институт, Институтский переулок, (3) Université Grenoble Alpes, France

В настоящем исследовании мы предлагаем несколько методов, основанных на подходах машинного обучения, для прогнозирования уровня загрязнения воздуха в городах, расположенных в горных долинах. В качестве примера рассматривается Гренобль (Франция). Задача прогнозирования уровня загрязнения решается как в постановке регрессии, так и в постановке классификации факта превышения пороговых уровней. Для прогнозирования мы применяем подход, основанный на данных, привлекая различные модели машинного обучения. На основе исторических данных за период с 2012 по 2018 год, собранных на нескольких метеорологических станциях, расположенных в долине Гренобля, мы обучили несколько моделей машинного обучения для прогнозирования среднесуточных значений концентраций мелкодисперсных частиц PM₁₀ и PM_{2.5} на три дня вперед. Наибольший интерес в нашем исследовании представляют дни с высоким уровнем концентрации PM, превышающим пороговые значения определенные Всемирной Организацией Здравоохранения (ВОЗ). Было обнаружено, что наличие локальных метеорологических условий приводит к формированию температурной инверсии, которая имеет статистическую взаимосвязь с уровнем загрязнения воздуха в этом регионе. Несмотря на то, что уровень загрязнения в значительной степени обусловлен локальными метеороусловиями, модели машинного обучения, рассмотренные в нашем исследовании, могут быть адаптированы и для других городов, находящихся в низинах. Для этого их необходимо обучить на релевантных данных. В настоящем исследовании использовались методы ансамблирования, основанные на деревьях решений, такие как Catboost и Balanced random forest. Данные методы комбинируют прогнозы нескольких моделей для повышения точности и устойчивости окончательного решения. Наилучшие результаты в прогнозировании уровня загрязнения воздуха как в постановке регрессии, так и классификации показал алгоритм Catboost. В задачах классификации также хорошо зарекомендовал себя алгоритм Balanced random forest, предназначенный для работы с несбалансированными данными. В рамках задачи регрессии Catboost эффективно аппроксимировал пики загрязнения. Значения RMSE составили 7.35 ± 0.58 мкг/м³ для прогноза на три дня вперед. В подходе классификации Catboost достиг F1-меры 0.65, а BRF — 0.64 за аналогичный период.

49. Аппроксимация пространственно-временной изменчивости городского острова тепла Москвы методами машинного обучения

Варенцов Михаил Иванович (1,2,3,4), Криницкий Михаил Алексеевич (4,5,1), Степаненко Виктор Михайлович (1,2)

(1) МГУ имени М.В. Ломоносова, НИВЦ, (2) МГУ имени М.В. Ломоносова, Географический факультет, (3) Гидрометеорологический научно-исследовательский центр Российской Федерации, (4) Институт физики атмосферы имени А.М. Обухова РАН, (5) Институт океанологии им. П.П. Ширшова РАН, (6) Московский физико-технический институт,

Одним из перспективных направлений применения методов машинного обучения (МО) в метеорологии является задача статистического масштабирования или даунскейлинга, то есть повышение разрешения метеорологических данных. В рамках этой задачи методы МО часто рассматриваются как альтернатива вычислительно-затратному методу динамического даунскейлинга с использованием региональных гидродинамических моделей атмосферы [1]. Задача даунскейлинга особо актуальна для урбанизированных территорий в силу их пространственной неоднородности и наличие локальных погодноклиматических аномалий, в первую очередь эффекта городского острова тепла – положительной аномалии температуры, оказывающей значимое влияние на термический комфорт и здоровье населения [2]. В докладе рассмотрен опыт применения методов МО для статистической аппроксимации пространственно-временной изменчивости острова тепла (городской аномалии температуры воздуха) на основе предикторов, характеризующих крупномасштабные атмосферные процессы, на примере Московского мегаполиса. На первом этапе исследования решена задача «даунскейлинга в точку» – аппроксимации наблюдаемой разности температуры «город-фон» для центра Москвы [Varentsov et al., 2023a]. В такой постановке выполнено сравнение нескольких моделей МО типа регрессии, включая линейную модель, методы случайных лесов, градиентного бустинга, опорных векторов, а также полносвязную нейросеть (Multilayer Perceptron). Показано, что модели МО могут успешно воспроизводить временную динамику разности «город-фон» на суточном, синоптическом и сезонном масштабах. При этом наилучшее качество было достигнуто моделью градиентного бустинга CatBoost, она же показала максимальное быстроедействие. На втором этапе исследования разработанная модель на основе метода градиентного бустинга в реализации CatBoost была адаптирована для аппроксимации городской аномалии температуры уже как трёхмерной (пространственно-временной) величины. В качестве дополнительных предикторов использованы детализированные данные о свойствах подстилающей поверхности и городской среды. В качестве данных для оптимизации моделей МО использованы результаты суперкомпьютерного моделирования метеорологического режима Московского региона с региональной гидродинамической моделью атмосферы COSMO с шагом сетки 1 км, дополненной специальной городской параметризацией TERRA_URB [3]. Для описания пространственной структуры острова тепла предложен оригинальный «квазилокальный» подход к применению методов МО. Он заключается в задании признакового описания в форме таблицы с данными для множества формально не связанных друг с другом точек (узлов расчётной сетки), и порождении дополнительных признаков для учета нелокальных пространственных взаимодействий между предикторами. Для этого к исходным двумерным полям характеристик поверхности применяется набор преобразований, в том числе самонастраивающиеся направленные фильтры, формируемые с учетом крупномасштабных данных о скорости и направлении ветра. Показано, что такой подход позволяет достаточно реалистично воспроизводить пространственную структуру острова тепла в различных условиях, в том числе воспроизводить нелокальные эффекты, связанные с перемешиванием и адвекцией, например тепловые шлейфы. В рамках продолжения исследований

представляется перспективной задача сравнения предложенного «квазилокального» подхода с популярным в задаче даунскейлинга «нелокального» подхода на основе свёрточных ИНС, подразумевающего задание признакового описания в виде двумерных пространственных полей [1]. Работа выполнена при поддержке некоммерческого фонда развития науки и образования «Интеллект». Анализ результатов мезомасштабного моделирования выполнен при частичной поддержке Российского научного фонда, проект №24-17-00155.

Ссылки: 1. Sun Y. et al. Deep learning in statistical downscaling for deriving high spatial resolution gridded meteorological data: A systematic review. ISPRS Journal of Photogrammetry and Remote Sensing. 2024. Vol. 208. P. 14–38. 2. Oke T.R. et al. Urban Climates. Cambridge: Cambridge University Press. 2017. 509 p. 3. Varentsov M. et al. Megacity-Induced Mesoclimatic Effects in the Lower Atmosphere: A Modeling Study for Multiple Summers over Moscow, Russia. Atmosphere. 2018. Vol. 9, № 2. P. 50. 4. Varentsov M., Krinitskiy M., Stepanenko V. Machine Learning for Simulation of Urban Heat Island Dynamics Based on Large-Scale Meteorological Conditions. Climate. 2023. Vol. 11, № 10. P. 200.

3. Опыт использования методов машинного обучения в задачах диагностирования геоиндуцированных токов в высокоширотных энергосистемах

Воробьев Андрей Владимирович (1), Воробьева Гульнара Равиловна (2)
(1) Геофизический центр РАН, Москва, (2) Уфимский университет науки и технологий

Как известно, геоиндуцированные токи (ГИТ) возникают в пространственно-распределенных токопроводящих технических системах (магистральные трубопроводы, линии электропередачи (ЛЭП) и телеграфной связи, объекты железнодорожной инфраструктуры и пр.) вследствие геомагнитных вариаций (ГМВ), скорость изменения которых в высокоширотных регионах часто составляет несколько сотен нТл/мин. Протекая по заземленным обмоткам силовых трансформаторов системообразующих электрических цепей, экстремальные ГИТ способны перевести их магнитные системы в режим насыщения, что, в свою очередь, может вызвать сбой или отказ соответствующих электротехнических систем. Однако по причине малой изученности механизмов возникновения и развития ГИТ, а также фрагментарности и неоднородности имеющихся эмпирических данных, задача их прогнозирования и диагностирования на сегодняшний день сопряжена со множеством неопределенностей и остается практически нерешенной. В работе на основе методов машинного обучения рассматриваются подходы к диагностированию уровня ГИТ в магистральной электрической сети «Северный транзит». При этом в качестве входных параметров выступают как геомагнитные данные, регистрируемые магнитными станциями в заданном субрегионе (Кольский п-ов, Россия), так и естественные (видимые) индикаторы экстремальной геомагнитной активности. На примере годовой выборки, включающей более 35 000 записей было показано, что подход к диагностированию ГИТ, на базе множественной линейной регрессии, обеспечивает среднеквадратическую ошибку (СКО) $\sim 0.122 \text{ A}^2$.

Использование искусственной нейронной сети с ReLu функцией активации способно несколько улучшить точность диагностирования (СКО $\sim 0.119 \text{ A}^2$), однако при этом существенно снижается интерпретируемость и теоретическая значимость модели. Применение в свою очередь Байесовского классификатора к данным оптических наблюдений полярных сияний показало, что апостериорная вероятность того, что во время сияний на севере уровень ГИТ на станции «Выходной» превысит 2 А составляет 5.78 %, в то время как вероятность превышения данного значения при сияниях в зените и на юге составляет 10.04 % и 14.93 % соответственно. В отсутствие же сияний данная модель указывает на то, что вероятность достижения ГИТ аналогичного уровня не превышает 0.26 %, а вероятность превышения 3 А практически равна нулю.

72. Application of machine learning methods for analyzing data from semiconductor gas sensors in dynamic temperature mode

Isaev I.V. (1,2,3), Chernov K.N. (4), Shchurov N.O.(1,4), Dolenko S.A. (1), Krivetskiy V.V. (3,5)

(1) D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, (2) Kotelnikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences, (3) Scientific-Manufacturing Complex Technological Centre, (4) Physics Department, M.V. Lomonosov Moscow State University, (5) Chemistry Department, M.V. Lomonosov Moscow State University

This work considered the problem of environmental monitoring of air in cities and industrial areas, which consists in detecting gases and volatile organic compounds using semiconductor gas sensors. To provide selectivity in the detection of certain gases, several semiconductor sensors with different doping components were used. Also, to ensure selectivity of gas determination, as well as high temporal resolution of the sensors, nonlinear temperature operating conditions were used - the so-called heating dynamics. Due to the high complexity of the model describing the processes of interaction between gases and sensors, machine learning methods based on the use of physical experiment data were used to process the sensor response. In this study, we considered such machine learning problem statements as the regression problem, the binary and the multi-class classification problem. The regression task was to determine the concentration of a specific gas using data from a single physical experiment. Also, within the framework of one physical experiment, the presence/absence of a specific gas was determined in a binary classification problem. Using a merged data set including several physical experiments, a multi-class classification problem was solved, which consisted of determining the presence of one of the gases under consideration. Each individual computational experiment used response data from a single sensor and within a single heating dynamic. Based on the results of the work, conclusions were drawn regarding the selection of optimal sensors and heating dynamics for a specific gas/all gases. The study was carried out at the expense of the grant No. 22-19-00703 from the Russian Science Foundation, <https://rscf.ru/en/project/22-19-00703/> [<https://rscf.ru/en/project/22-19-00703/>].

49. Машинное обучение и глубокое обучение в науках об атмосфере, океане и климате: достижения и перспективы в

2024 году

Креницкий М.А. (1,2)

(1) Институт океанологии П.П. Ширшова Российской академии наук, (2) Московский физико-технический институт

Методы машинного обучения (ML) и глубокого обучения (DL) активно применяются в различных сферах изучения океана, атмосферы и климата, таких как океанография, метеорология и климатология. Эти подходы позволяют эффективно обрабатывать большие объемы данных, выявлять скрытые закономерности, уменьшать или оценивать неопределенность в климатических и погодных прогнозах, автоматизировать мониторинг и ускорять аналитические исследования. Среди успешных примеров можно упомянуть анализ данных дистанционного зондирования, моделирование геофизических процессов, аппроксимацию неизвестных физических параметров и решение задач статистического прогнозирования погоды и климата. Тем не менее, существуют определенные сложности, такие как необходимость в больших объемах данных, вычислительные и технические особенности методов наук о данных, а также обеспечение физической правдоподобности результатов. В перспективе ожидается развитие гибридных моделей, которые объединяют физические и статистические методы, а также улучшение интерпретируемости моделей ML и DL. В докладе мы рассмотрим текущие достижения в применении ML и DL в изучении океана, атмосферы и климата, а также обсудим проблемы и перспективы их дальнейшего развития.

71. Towards the Statistical Correction of High-Resolution Weather Forecasts from the WRF Model employing Deep Learning

Голиков Виктор Артёмович (1, 2), Креницкий Михаил Алексеевич (2, 3), Вановский Владимир Валерьевич (1, 2)/

(1) Сколковский институт науки и технологий, (2) Московский физико-технический институт, (3) Институт океанологии им. П.П. Ширшова РАН

Numerical weather prediction models, like the WRF, are crucial for simulating atmospheric conditions and forecasting significant geophysical parameters. However, due to limitations such as coarse spatial resolution and imperfect parameterizations, these models often exhibit systematic biases when compared to observed data. To mitigate these discrepancies, numerous statistical correction techniques are employed. In this study, we introduce a neural network-based method for statistical correction of WRF model forecasts. Framed as a time series prediction problem, the model inputs are hourly WRF forecasts, with corrected WRF fields as the target outputs. The correction algorithm features a U-Net inspired architecture augmented with a transformer mechanism for processing latent representations. In our study, we are focused on correcting three environmental variables denoted by T2, U10 and V10: temperature and two wind components on heights two and ten meters above the surface, respectively. The model is trained, validated, and tested using ERA5 reanalysis data, meteorological station observations, and satellite scatterometer data from 2019 to 2023. Only the WRF forecast data is required for practical application (inference mode). In order to effectively merge data from multiple sources while retaining the inherent characteristics of the WRF outputs, an

additional term is included in the loss function. We compared a few neural network architectures, demonstrating that the proposed method significantly reduces forecasting errors and retains the resolution of mesoscale processes. Experiments on the corrected environmental fields show improved representation of mesoscale phenomena.

74. The Ob-Yenisei plume movement analysis in the Kara Sea using artificial neural networks

A.S. Savin (1,2), M.A. Krinitskiy (1,2), A.A. Osadchiev (1,2)

(1) Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia,

(2) Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, Russia

The movement of the Ob-Yenisei plume is a key factor determining the dynamics of the surface layer in the Kara Sea during the ice-free period. Traditionally, this movement is described by the action of wind shear stress on the sea surface, which leads to the formation of a drift current in accordance with Ekman's theory. In this study, the movement of the Ob-Yenisei plume under the action of wind stress is investigated using Artificial Neural Networks. The evolution of the plume from its initial position to its final position over a period of 2-3 weeks is considered, taking into account the action of the wind field from ERA-5 reanalysis data. Wind data are processed using a Convolutional LSTM Neural Network, which also allows for the consideration of the evolution of the wind field and its influence on the plume's position throughout the entire period examined. The importance of input features for the trained Neural Network model is analyzed. This approach allows for conclusions about the most significant wind conditions that determine the movement of the Ob-Yenisei plume in the Kara Sea.

72. Прогнозирование состояния магнитосферы Земли с помощью специального алгоритма для работы с многомерными временными рядами

Владимиров Р.Д. (1), Широкий В.Р. (1), Баринов О.Г. (1), Доленко С.А. (1), Мяжкова И.Н. (1)

(1) НИИ ядерной физики имени Д.В. Скобельцына МГУ имени М.В. Ломоносова

Данное исследование посвящено адаптации и применению специального 4-ступенчатого алгоритма на основе методов машинного обучения, разработанного для анализа многомерных временных рядов при решении задач прогнозирования определенных событий и выявления их предвестников - явлений, представленных неизвестной комбинацией значений параметров, описывающих объект [1]. Помимо прогнозирования событий, алгоритм можно использовать для прогнозирования значений непрерывных величин. При этом используемый алгоритм позволяет адаптивно отбирать как физические входные признаки, так и конкретные значения задержки при учете истории каждого физического признака в рамках погружения (топологического вложения) соответствующего ему временного ряда. Такой отбор может обеспечить лучшее понимание процессов, происходящих в изучаемом объекте. Общая схема 4-ступенчатого алгоритма включает в себя следующие этапы: 1. Выбор наиболее

значимых физических признаков (переменных) среди тех, которые, по мнению исследователя, могут повлиять на прогнозируемое целевое значение. В текущей реализации для этой цели используется итеративный подход. В рамках этого подхода система оценивает корреляцию (Пирсона/Спирмена) между входным признаком и его задержками и прогнозируемой переменной. Затем на основе заданного порога выбирается часть переменных, которые будут использоваться на следующих этапах. 2. Выбор диапазона используемых задержек.

Рассматриваемые наборы входных признаков создавались следующим образом. Набор 0 включал выбранные входные переменные в текущий момент. В набор 1 вошли все входные переменные из набора 0, а также все эти переменные с задержкой в 1 временной шаг. В набор 2 вошли все входные переменные из набора 1, а также все эти переменные с задержкой в 2 временных шага и так далее до определенного исследователем предела. Модель машинного обучения обучается на каждом наборе данных, созданном в рамках этого цикла, и ее качество оценивается на предварительно подготовленном наборе тестовых данных. Цикл останавливается, когда по заданному критерию увеличение диапазона задержки больше не приводит к значительному повышению точности прогнозирования. 3. Выбор наиболее важных входных признаков из полученного пространства признаков, ограниченного на первых двух этапах. Для этого этапа могут быть использованы стандартные подходы к оценке существенности входных признаков. 4. Настройка гиперпараметров модели. В данной работе алгоритм применялся для прогнозирования значений потока релятивистских электронов с $E > 2$ МэВ на геостационарной орбите, а также значений геомагнитных индексов Dst и Kp. Результаты сравнивались между собой по значению относительной ошибки прогнозирования и по наборам наиболее значимых входных признаков, выбранных алгоритмом.

Рассматривались временные ряды с шагом в один час. В качестве входных физических признаков использовались значения следующих величин: 1. Параметры солнечного ветра (СВ) в точке Лагранжа L1 между Землей и Солнцем: скорость СВ v (км/с), температура СВ T (К), плотность протонов в СВ n (см^{-3}). 2. Параметры вектора межпланетного магнитного поля (ММП) в той же точке L1 в системе GSM: компоненты ММП V_x, V_y, V_z , модуль ММП V_{mag} (нТл). 3. Геомагнитные индексы: экваториальный геомагнитный индекс Dst (нТл), планетарный геомагнитный индекс Kp (безразмерный). 4. Специально добавленные функции случайного шума, чтобы проверить, распознает ли система их как нерелевантные для прогнозирования целевой переменной. Набор также включал четыре гармонические переменные (две с суточным периодом и две с годовым периодом) для учета вращения Земли вокруг своей оси и вокруг Солнца (для этих переменных погружение временных рядов не использовалось). Показано, что применение 4-ступенчатого алгоритма к рассматриваемым задачам прогнозирования позволяет повысить качество прогнозирования по сравнению с прогнозированием на полном наборе признаков без использования отбора. При этом отбираемые алгоритмом входные признаки соответствуют существующим физическим представлениям о физических величинах, существенных для прогнозирования рассмотренных целевых переменных. Исследование выполнено за счёт гранта Российского научного фонда № 23-21-00237, <https://rscf.ru/project/23-21-00237/> [<https://rscf.ru/project/23-21-00237/>]. 1. Dolenko, S. et al. LNCS 5769, 295-304. https://doi.org/10.1007/978-3-642-04277-5_30 [https://doi.org/10.1007/978-3-642-04277-5_30]

36. Методы вероятностного программирования при реконструкции событий многоканального изображающего детектора: ЭЛЬФЫ и ТРЕКИ

Шаракин Сергей Александрович (1) Сараев Роман Евгеньевич (2)

(1) МГУ им. М.В. Ломоносова, НИИЯФ им. Д.В. Скобельцына (2) МГУ им. М.В. Ломоносова, аспирантура Физического факультета

В докладе предложены новые способы анализа динамических изображений, регистрируемых многоканальными высокочувствительными детекторами с высоким временным разрешением. Орбитальный детектор “УФ атмосфера” наблюдает за атмосферой Земли с борта МКС и регистрирует ряд транзиентных оптических явлений, связанных с грозовой активностью, в том числе и события типа ELVES (“эльфы”). Вероятностная модель эльфа в качестве параметров включает координаты и ориентацию породившего свечение грозового разряда, а также высоту ионизированного слоя, в котором регистрируется свечение. Байесовский вывод (реализованный средствами библиотеки PyMC) позволяет вычислить апостериорные распределения на эти параметры на основании времен пиков сигналов в отдельных каналах детектора. Система наземных многоканальных детектор PAIPS кроме изучения разных типов полярных сияний служит также своеобразным полигоном отработки алгоритмов вероятностной реконструкции. Для этого используется широкий класс так называемых трековых событий – метеоров, пролетов спутников и самолетов, движение звезд по небу. Байесовская модель включает в себя как параметры самого трекового события, так и особенности его регистрации (аппаратную функцию детектора). Кроме того, при таком подходе удастся включить в модель априорную информацию о событии (горизонтальность движения спутника, характерную скорость и направления движения потокового метеора и т.п.). Данные методы могут быть обобщены на стерео-события (регистрацию треков двумя детекторами с пересекающимися полями зрения) или применяться при реконструкции КЛ ПВЭ в детекторах флуоресцентного свечения ШАЛ. Работа выполнена при финансовой поддержке гранта РФФИ № 22-62-00010

71. Применение методов машинного обучения для идентификации полярных мезомасштабных циклонов в данных численного моделирования атмосферы

Левковская Ю. А. (1), Криницкий М. А. (2, 3), Веземская П. С. (2)

(1) Национальный исследовательский университет “Высшая школа экономики” (2) Институт океанологии им. П. П. Ширшова РАН (3) Московский физико-технический институт

Полярные мезомасштабные циклоны (ПМЦ) — это интенсивные атмосферные вихри небольшого размера, формирующиеся над океаном на границе арктической и умеренной атмосферных воздушных масс. ПМЦ характерны сильными ветрами, поэтому представляют угрозу морской инфраструктуре и судоходству, а также оказывают влияние на глубокую конвекцию океана. На сегодня существует два основных метода получения информации о мезомасштабных циклонах - ручное дешифрирование спутниковых снимков и

идентификация ПМЦ по различным эмпирическим критериям в данных численного моделирования атмосферы. У каждого из способов есть серьезные недостатки. Из-за короткого времени жизни ПМЦ данные дистанционного зондирования, характерные разреженностью во времени, могут недостаточно полно и детально отражать динамику этих вихрей. Обнаружение мезомасштабных циклонов в модельных данных затруднено эмпирической природой критериев идентификации, разрабатываемых исследователями. В настоящей работе мы применили подход интерпретации моделей машинного обучения, натренированных идентифицировать ПМЦ в данных моделирования атмосферы, для составления списка наиболее значимых критериев. Для обучения моделей используются характеристики мезомасштабных циклонов, размеченных в ретроспективном анализе RAS NAAD по данным календарей ПМЦ (Rojo et al., 2019), (Golubkin et al., 2021), (Noer et al., 2011). Для оценки значимости признаков использовались встроенные методы моделей, а именно: значения коэффициентов модели для логистической регрессии и метода опорных векторов с учетом стандартизации признаков; подход оценки среднего интегрального нормализованного снижения функции потерь для каждого из признаков в случае моделей случайных лесов и градиентного бустинга. По результатам проведенного исследования самыми значимыми признаками являются: индекс холодных вторжений на 500 гПа, разница между температурой на высоте 700 гПа и на уровне моря, завихренность на 850 гПа и давление на уровне моря. Эти признаки были выбраны как наиболее значимые при анализе большинства моделей МО, и их состав хорошо согласуется с физическим аспектом идентификации мезомасштабных циклонов. Результаты настоящего исследования в отношении отбора наиболее значимых признаков могут быть в дальнейшем использованы для построения надежных объективных систем идентификации и построения траекторий ПМЦ, использующих подходы, основанные на данных, а именно методы машинного обучения или глубокого обучения.

Секция 3. Машинное обучение в естественных науках

47. Предсказание структуры дефектов MoS₂ по заданным свойствам

Г. Карлинский (1), М. Лазарев (1)

(1) Национальный исследовательский университет «Высшая школа экономики»

Дизайн материала с заданными свойствами является принципиально важной задачей, в свою очередь, свойства материала сильно зависят от типа и концентрации дефектов. В данной работе мы создали алгоритм генерации дефектов, по заданным свойствам, в двумерных материалах на примере MoS₂. Предлагаемый алгоритм работает на порядки быстрее и качественнее алгоритмов на основе нейросетевых генеративных моделей. За основу был взят подход определения свойств кристалла, энергии формирования кристалла и разницы LUMO-HOMO, с помощью метода символьной регрессии SEGVAE. Формулы взаимодействия каждого типа дефектов между собой в явном виде

были найдены в приближении попарного взаимодействия. Были изучены характеристик кристаллов с разнообразными конфигурациями дефектов. Разработанный метод является интерпретируемым и может быть использован для генерации конфигураций дефектов с заданными свойствами и в других двумерных материалах.

61. Estimation of signals in white noise using neural networks modeling

Kurdoshev Z.M.(1), Pchelintsev E. A.(1)

(1) Tomsk State University

We consider the problem of statistical signal processing using neural networks. Let the observed process obey the following stochastic differential equation [1] $dy_t = S_t dt + \varepsilon dw_t$, $0 \leq t \leq 1$, (1) Here S_t is the unknown signal, (y_t) is the observations, $\varepsilon > 0$ is the intensity noise, and the noise process (w_t) is the standard Brownian motion (Wiener) process. This work explores the use of artificial intelligence (AI) in signal processing, its applications and results. AI is used to reconstruct noisy signals. As a result, signal processing time and efficiency increased. Methodology: Different types of signals are pre-modeled to solve this problem. White noise of varying intensity is added to the modeled signals. The generated signal is processed according to traditional methods, that is, it is cleaned of noise and it is evaluated with the original signal, and in the next step, a neural network model is built and this neural network is trained. . Previously created noise functions and conventionally evaluated signals are used as a database for training. LSTM and Danse models were used to construct the neural network model. Adam's optimizer was used for the neural network and mean squared error as the loss function. Main results: Signal processing through neural networks improves accuracy up to 10^{-4} and reduces time by more than 20 times. Conclusion: The use of neural networks in this field, as in other fields, leads to effective results. Signal processing tasks have been simplified and accuracy has been increased. Based on the obtained results, experiments can be carried out in practice.

37. Архитектура "трансформер" для анализа рисков группового влияния нутриентов пищи

Баландина А.Н.¹, Груздев Б.В.¹, Савельев Н.А.², Будамян Я.С.¹, Кисиль С.И.³, Грачев Е.А.¹

(1) Физический факультет МГУ имени М.В.Ломоносова; (2) ПАО "Вымпелком" (в период работы над публикацией) (3) Институт русского языка и культуры МГУ имени М.В.Ломоносова.

В медицине крайне важно учитывать контекст при выставлении диагноза пациента. Один и тот же показатель в различном контексте может означать очень разные вещи. Нейросетевые модели, основанные на архитектуре трансформера, не применялись ранее в анализе диетологических данных анамнезов пациентов с подозрением на наличие у них метаболического синдрома. Вместе с тем, основные плюсы этих архитектур, такие как учет глобального контекста, методы интерпретации весов внимания, получения информативных векторов элементов входной последовательности, которые

могут быть затем эффективно использованы в прикладных задачах — крайне важны как раз в анализе биомедицинских данных. Способность механизма attention выявлять многофакторные связи дает возможность экономить время врача и акцентировать его внимание на определенных, выявленных нейронной сетью, паттернах. В данной работе Encoder Transformer был адаптирован под работу с табличными данными и применен на данных анамнезов пациентов и их диетологической истории для классификации наличия и источников метаболических нарушений. Исследования по применению архитектуры трансформера на табличных данных в общем и диетологических данных в частности имеют большой потенциал. В данной работе был получен результат, скоррелированный с известными медицинскими результатами, и были разработаны новые методы применения attention на табличных данных, а также новые методы векторизации этих данных.

68. Метод сегментации очагов опухолевого поражения глаза ретинобластомой на основе глубокого обучения

Волков Егор Николаевич(1), Аверкин Алексей Николаевич(1), Нечаевский Андрей Васильевич(1)

(1) Университет «Дубна»

Использование искусственных нейронных сетей (ИНС) в задачах диагностики заболеваний на основе анализа медицинских изображений различных доменов не только автоматизирует рутинные задачи клиницистов за счёт уменьшения времени обработки изображений, но и является одной из отличительных черт парадигмы Здравоохранение 5.0. Автоматический анализ снимков с помощью ИНС внедрён в подавляющем большинстве областей медицины (снимки лучевой диагностики, снимки кожных покровов и т.д.). Одной из областей медицины, где внедрение продуктов на основе ИНС для диагностики заболеваний остаётся крайне низким является офтальмология. Ретинобластома является редкой злокачественной опухолью, возникающей из клеток сетчатки. В подавляющем большинстве случаев возникает в раннем детском возрасте. Из-за своей редкой встречаемости опухоль сложно диагностируема особенно в условиях учреждений первичного звена здравоохранения. Основным методов диагностики является офтальмоскопия и анализ ретинальных снимков. Использование глубокого обучения может помочь в диагностике и стадировании заболевания, особенно на ранних стадиях, когда частично сохранить зрение ещё представляется возможным. В работе представлен метод сегментации участков опухолевого поражения глаза на основе использования ИНС U-Net-подобной архитектуры. Приводится полный процесс разработки пайплайна от создания собственного набора данных до проверки работы сети на тестовых данных. Оценка качества сегментации производилась с использованием значений метрик IoU и DSC. Корректность сегментации оценена с помощью метода объяснительного искусственного интеллекта Grad-CAM. Рассматриваются возможности совершенствования метода в условиях низкой доступности обучающих данных. Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № FEEM-2024-0014 Применение объяснительного искусственного интеллекта для интерпретации алгоритмов машинного обучения).

73. Nonlinear relevance estimation of multicollinear features for reducing the input dimensionality of optical spectroscopy inverse problem

Shchurov N.O.(1,2), Isaev I.V.(1,3), Burikov S.A.(1,2), Laptinskiy K.A.(1,2), Sarmanova O.E.(1,2), Dolenko T.A.(1,2), Dolenko S.A.(1)

(1) D.V.Skobel'tsyn Institute of Nuclear Physics, M.V.Lomonosov Moscow State University, (2) Physics Department, M.V. Lomonosov Moscow State University, (3) Kotelnikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences

In this work we considered inverse problems of optical spectroscopy, which consist in determining the ingredient ions concentrations of multicomponent water solutions by their spectra. Due to the nonlinearity of the problem under consideration, as well as to the lack of an adequate mathematical model describing the spectra of multicomponent solutions, machine learning methods based on the use of physical experiment data were used to solve this problem. At the same time, inverse spectroscopy problems are characterized by high input dimensionality with a large number of features, both relevant and irrelevant. In turn, some of the corresponding relevant features are redundant due to their multicollinearity, caused by the fact that the characteristic lines of the solution components have a certain width and cover several spectrum channels at once. This leads to a deterioration in the quality of machine learning solution of the problem, and therefore there is a need for a feature selection procedure that takes into account their relevance and redundancy, as well as the nonlinear relationship with the determined parameters. In this study, we considered a feature selection procedure based on the iterative selection of features with the highest relevance to the target variable and on the elimination of features with a high relationship with each other. In this selection procedure, the method of weight analysis of a trained neural network was used as a nonlinear measure of relevance, and the Pearson correlation coefficient was used as a measure of the relationship of features with each other. Finally, the quality of a neural network solution of the problem of determining the concentrations of solution components from spectroscopic data was compared on the full set of input features and on its subsets compiled using the selection procedure under consideration, as well as using traditional methods for selecting significant input features. The study was carried out at the expense of the grant No. 24-11-00266 from the Russian Science Foundation, <https://rscf.ru/en/project/24-11-00266/> [<https://rscf.ru/en/project/24-11-00266/>].

69. Use of Neural Network Approximation of the Parameters-Property Relationship in Synthesis of Carbon Dots

Dolenko S.A. (1), Laptinskiy K.A. (1), Korepanova A.A. (2), Burikov S.A. (1,2), Dolenko T.A. (1,2)

(1) D.V. Skobel'tsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, (2) Faculty of Physics, M.V. Lomonosov Moscow State University

The unique optical properties and simplicity of the methods for obtaining carbon dots (CD) open up wide prospects for their application in optoelectronics, biomedicine, sensors, analytical chemistry, etc. The physico-chemical properties of

CD are determined by their synthesis method and by the precursors used. One of the most common methods for producing CD is solvothermal/hydrothermal synthesis, based on heating solutions of carbonaceous materials in a closed volume. Synthesis conditions, sets of precursors, and the ratio of their concentrations significantly affect the photoluminescent properties of CD – the wavelength of photoluminescence radiation, its quantum yield, and photostability. However, in the modern world, for specific applications, it is necessary to understand the interrelationships of Synthesis – Structure – Properties and the ability to synthesize CD with specified properties. That is, it is extremely important to determine precursors and optimal synthesis parameters for obtaining nanoparticles with predetermined properties, e.g. quantum yield of luminescence (QYL). To solve the specified inverse problem of the type “Synthesis Parameters - Properties” it is advisable to use machine learning (ML) methods, which have great potential in identifying hidden correlations in multiparametric systems. In this study, ML was used to determine the optimal conditions for the hydrothermal synthesis of CD from citric acid (CA) and ethylenediamine (EDA) with a wide range of precursor ratios, temperature and reaction time to obtain nanoparticles with a given QYL. As is known, it is from this pair of precursors that CD with the largest range of QYL changes are synthesized – from 0 to 100%, which significantly complicates the solution of the task. In physical experiment, we synthesized 343 CD samples from CA and EDA. CA concentration was fixed at 0.1 M. EDA concentration varied from 0.01 to 2 M, the time of the hydrothermal reaction from 30 to 360 minutes, the temperature of synthesis from 80°C to 200°C. For each CD sample, optical absorption spectrum and 2D fluorescence spectrum were obtained, and QYL value was calculated using the method of the reference dye. The purpose of the study was to determine the areas of change in the parameters of hydrothermal synthesis of CD, in which the nanoparticles are likely to have the highest QYL. The described experimental data was used to build a neural network approximation of the dependence of QYL on the three synthesis parameters (EDA concentration, time and temperature of synthesis). The optimal architecture turned out to be a multi-layer perceptron with 8 neurons in the single hidden layer. According to the results of the approximation, the highest QYL of the CD was estimated to be 99.15%, obtained at EDA concentration 0.35 M, synthesis temperature 145 °C, and synthesis duration 240 sec. The experimental value of the QYL obtained at the specified values of the parameters, was 98.9%. The study was carried out at the expense of the grant No. 23-12-00138 from the Russian Science Foundation, <https://rscf.ru/en/project/23-12-00138/> [<https://rscf.ru/en/project/23-12-00138/>].

62. Improving Representativity of Spectroscopic Data using Variational Autoencoders: Approaches and Problems

Mushchina A.S. (1,2), Isaev I.V. (1), Sarmanova O.E. (1,2), Dolenko T.A. (1,2), Dolenko S.A. (1)

(1) D.V.Skobel'tsyn Institute of Nuclear Physics, M.V.Lomonosov Moscow State University, (2) Faculty of Physics, M.V.Lomonosov Moscow State University

This study constitutes an investigation into the potential improvement of neural network methods through the augmentation of the training dataset using a variational autoencoder (VAE). We consider the inverse problem of spectroscopy of

multi-component water solutions, aimed at determining the concentrations of various ions in the solutions based on their spectral data (Raman, IR or optical absorption spectroscopy). While the shape of the spectra is sensitive to the concentrations of ions, the dependence of spectral intensities on ion concentrations in multi-component solutions is complex and non-linear, thus requiring analysis of many spectral channels at once. Such analysis may be performed using machine learning methods, e.g. neural networks. However, to train a neural network, a large dataset is required. Adequate modeling of spectra of multi-component solutions is yet far beyond reasonable computational capabilities. The required dataset may be obtained through laboratory measurements, but such experiment is very laborious and expensive. Furthermore, the experimental data may also exhibit a significant imbalance in the space of target values that can interfere with the training of a regression neural network model. Thus, the issue of adequately expanding the training dataset is highly pertinent. This study examines the possibility of augmenting the training dataset by generating extra spectra using a VAE. The aim is to provide reduction of the error of solving the inverse problem. There are several possible approaches to such generation. It seems that the simplest way is to use a conditioned VAE (cVAE) trained on experimental data. However, we demonstrate that the quality of such generation is insufficient – generated spectra differ too much from experimental spectra with the same ion concentrations. Expanding the experimental dataset with such generated spectra even increases the error of determination of ion concentrations. Two other approaches use VAE. In this case, we require a way to determine the target ion concentrations for each of the generated spectra. This may be done by some ML regression model trained on experimental data – either in the feature space of the spectra or in the latent space of VAE. Subsequently generated spectra can be used in various ways along with experimental ones during the training of regression neural networks solving the inverse problem. In this study, we compare these two approaches and discuss their merits and shortcomings. The impact of the level of noise in the experimental data, and of the level of noise added to experimental spectra during training of the VAE, on the generated spectra is also explored. Possible ways of improving data distribution during spectra generation are also discussed. The study was carried out at the expense of the grant No. 24-11-00266 from the Russian Science Foundation, <https://rscf.ru/en/project/24-11-00266/> [<https://rscf.ru/en/project/24-11-00266/>].

68. Астрофизика частиц и анализ экспериментальных данных

Л.Кузьмичев (1)

(1) НИИЯФ МГУ

65. Comparative Analysis of the Procedures to Forecast the Kp Geomagnetic Index by Machine Learning

Gadzhiev I.M. (1,2), Dolenko S.A. (1), Barinov O.G. (1), Myagkova I.N. (1)

(1) D.V.Skobel'tsyn Institute of Nuclear Physics, M.V.Lomonosov Moscow State University, (2) Faculty of Physics, M.V.Lomonosov Moscow State University

Geomagnetic disturbances are one of the most important factors in space weather,

the role of which will increase with the development of the space industry and the global digital industry, both on Earth and in near-Earth space. Geomagnetic activity is usually characterized by special indices. One of the most common geomagnetic indices is the Kp index, first introduced by Julius Bartels in 1939. In this study, we explore the possibility of predicting the following Kp index values during the next day (24 hours) using machine learning (ML) models based on the hourly values of the parameters of solar wind and interplanetary magnetic field, and of the hourly Dst geomagnetic index. We use such ML models as linear regression, gradient boosting and multilayer perceptrons. We test to what extent the use of time series delay embedding improves the performance of ML models. Due to the specifics of the Kp index (it is measured once every three hours, different from other geomagnetic indices), several ML models are calculated separately, depending on the number of hours remaining until the next Kp index value is calculated, to comprise a composite model making hourly predictions. A special procedure using a synthetic KpH index with an hourly frequency as a target variable for ML models training to improve the quality of the forecast is also being tested. The quality of all the models is assessed on an hourly basis on various testing data depending on Kp index range, corresponding to various levels of disturbance of the magnetosphere. Finally, conclusions were drawn about the optimal procedure of creating and applying of a machine learning model to solve the Kp index forecasting problem. The best results by most of the quality metrics were demonstrated by the composite model. The significance of the features that are used by the model for prediction was also studied. The most significant input features detected were preceding values of the Kp index itself, phase of the Earth's daily rotation, solar wind velocity, modulus and z-component of the interplanetary magnetic field. This list well matches the existing physical notions on feature significance based on the results of physical experiment. The study was carried out at the expense of the grant No. 23-21-00237 from the Russian Science Foundation, <https://rscf.ru/en/project/23-21-00237/> [<https://rscf.ru/en/project/23-21-00237/>].

50. Сети глубокого обучения для построения виртуальных датчиков технологических процессов нефтепереработки

Лазухин Иван Сергеевич(1), Петровский Михаил Игоревич(1), Машечкин Игорь Валерьевич(1)

(1) МГУ имени Ломоносова

В настоящее время, в связи с развитием систем автоматического и автоматизированного управления, предприятия накапливают большие объемы данных о состоянии технологических процессов. Множества физических датчиков фиксируют состояние системы в каждый момент времени, тем самым отвечая за управление системой и поддержку ее параметров в допустимых пределах. Одновременно с этим, значительную часть качественных характеристик таких процессов, в особенности, связанных с нефтехимической промышленностью, составляют лабораторные исследования. Такие исследования проводятся с нерегулярной частотой — существующие на практике средства их интерполяции основаны на линейных или же кусочно-постоянных интерпретациях. Математические модели, обобщающие лабораторные исследования на частоту, соответствующую физическим

датчикам будем называть виртуальными датчиками. Таким образом, исследуемая в работе задача прогнозирования виртуальных датчиков, то есть получения химических показателей производственного процесса в реальном времени на основе значений физических датчиков является актуальной. Авторы исследуют набор данных, соответствующий нефтеперерабатывающей установке, содержащий около 10,000 показаний 300 физических датчиков, собираемых с частотой 1 час, однако соответствующие лабораторные исследования составляют набор всего в 300 показаний, собираемых с частотой около суток. Соответственно, возникает несколько задач, связанных с высокой размерностью входов, разреженностью и в целом небольшим числом показаний целевой переменной, а сама задача прогнозирования лабораторных исследований содержит подзадачу аппроксимации. Авторами были предложены несколько подходов, объединяющих задачи прогноза и заполнения пропусков для моделирования лабораторных исследований. Были рассмотрены методы отбора признаков на основе PLS регрессии (иерархическая кластеризация), градиентного бустинга, деревьев Байеса, графовой нейронной сети, базовые методы в виде Lasso и корреляций. Каждый из предложенных методов отбора был адаптирован для возможности учитывать экспертное мнение специалистов области. Были предложены методы прогнозирования лабораторных исследований на основе графовых нейросетей, генеративных подходов, полносвязных и рекуррентных сетей, рассмотрены интерполяции отклика с использованием Loess и сплайнов, а также вариант регуляризации, учитывающей корреляцию с откликом. Полученные авторами экспериментальные результаты показывают преимущество использования для данной задачи простых рекуррентных сетей, графовых нейросетей с предварительной интерполяцией, отбора признаков на основе деревьев Байеса. Отдельно стоит отметить неоднозначность оценки качества полученных моделей; авторами был предложен комбинированный подход, учитывающий адекватность модели, корреляцию ее с истинными значениями лаборатории и стандартные ошибки.

37. Нейронные операторы для гидродинамического моделирования подземных хранилищ газа (ПХГ)

Д.Д. Сирота (1), К.А. Гуцин (1), С.А. Хан (1), С.Л. Костиков (1), К.А. Бутов (1)

(1) ПАО “Газпром”

Значительная часть исследований в области глубокого обучения сосредоточена на изучении отображений между конечномерными пространствами. Вместе с тем гидродинамические процессы фильтрации газа в ПХГ, описываемые дифференциальными уравнениями в частных производных (ДУЧП), требуют изучения отображений между функциональными пространствами бесконечной размерности, что отличает данную задачу от традиционных. В настоящее время подходы и методы глубокого обучения активно развиваются и все чаще достигают высокой точности в различных задачах численного моделирования. Одним из перспективных подходов является построение нейронных операторов – обобщение нейронных сетей для аппроксимации отображений между функциональными пространствами. В данной работе был построен и обучен оператор Фурье с добавлением слоев U-Net (U-FNO) для гидродинамического

моделирования процессов фильтрации газа в ПХГ. Впервые показано, что данный метод может быть применен не только для задач моделирования фильтрации газа в цилиндрической системе координат с одной скважиной, но и для задач трехмерной фильтрации газа в декартовой системе координат с множеством скважин. При этом нейронному оператору не требуется большой размер обучающей выборки для достижения требуемого качества аппроксимации ДУЧП, что демонстрируется не только на тестовой выборке, но и на искусственно сгенерированных сценариях с внесением существенных изменений в структуру моделируемого объекта.

46. Модель машинного обучения на гауссовских процессах для предсказания энерговыделения в топливных ячейках активной зоны исследовательского ядерного реактора

Н.В.Смольников (1), А.Г.Наймушин (1), И.И.Лебедев (1), М.Н.Аникин (1)

(1) Национальный исследовательский Томский политехнический университет

Благодаря компактным размерам активной зоны и большому количеству экспериментальных устройств исследовательские ядерные установки (ИЯУ) нашли широкое применение для проведения фундаментальных и прикладных исследований физики твердого тела, нейтронного рассеяния, производства изотопов, материаловедения и медицины. Малые размер активной зоны и частичная перегрузка топлива, при которой происходит замена только наиболее выгоревших тепловыделяющих сборок (ТВС), являются причинами формирования участков с высокой неоднородностью распределения излучения. Как результат, происходит перераспределение энерговыделения по топливным ячейкам, что оказывает влияние на эффективность использования топлива и условия проведения научных и производственных работ. Более того, в условиях, когда тепловыделяющие элементы работают при небольшом запасе до кризиса теплообмена, высокая неоднородность излучения может стать причиной ограничения эксплуатации ТВС. Выравнивание неравномерности распределения энерговыделения в активной зоне исследовательского реактора, как правило, осуществляется за счет перестановок «свежих» ТВС от центра к периферии. Однако, подготовка к процедуре профилирования требует определения характеристик активной зоны для каждого шага при перестановках, что осуществляется на основе ресурсозатратного расчетного моделирования процессов переноса излучения и баланса нейтронов. В настоящей работе предложен подход к использованию модели машинного обучения на гауссовских процессах для прогнозирования энерговыделения в гетерогенной структуре ядерного реактора. Это позволит оценивать изменения параметров энерговыделения в топливных ячейках без проведения моделирования. Модель обучалась на основе характеристик активной зоны для различных топливных кампаний реактора с применением кросс-валидации. По результатам обучения и настройки общая точность работы модели по коэффициенту детерминации составила 0.99.

47. An effective algorithm for predicting human fatigue using a portable Brain-Computer Interface

Ажогин Константин Эдуардович (1), Шевченко Михаил Геннадьевич (1), Щербань Игорь Васильевич (1)

(1) Южный федеральный университет, г. Ростов-на-Дону, Россия.

Мониторинг состояния операторов ответственных технологических процессов играет ключевую роль в обеспечении безопасности и эффективной эксплуатации сложных технических систем в целом. Своевременное обнаружение усталости, стресса или монотонии человека-оператора помогает предотвратить возможные происшествия и обеспечивает гарантированное выполнение важных задач. Для решения проблемы мониторинга перечисленных физиологических процессов, а также прогноза функционального состояния оператора целесообразным является использование портативных интерфейсов мозг-компьютер (ИМК). Гарнитуры таких ИМК имеют только лишь четыре-восемь электродов для регистрации электроэнцефалограммы (ЭЭГ) мозга человека, оснащены средствами беспроводной связи с миниатюрными вычислителями, эргономичны и не мешают оператору выполнять трудовые функции. Тем не менее, использование сокращенного числа отведений ЭЭГ не позволяет обеспечить полноценный анализ биоэлектрической активности мозга, а регистрируемые сигналы зашумлены вследствие двигательной и идеомоторной активности человека-оператора при выполнении им текущих функциональных обязанностей, его мускульной активности или глазных движений. Подобные проблемы не позволяют обеспечить эффективное использование технологии портативных ИМК для мониторинга состояния человека в настоящее время. Разработан алгоритм прогноза усталости человека на основе регистрации сигналов ЭЭГ от четырех отведений портативного ИМК. Алгоритм работоспособен в условиях деградации системы наблюдения биоэлектрической активности мозга, а также имеет свойства робастности к артефактам двигательной и идеомоторной активности, морганий и глазных движений. За счет применения непрерывного вейвлет-преобразования (НВП) к регистрируемым сигналам ЭЭГ выполняется переход в частотно-временной континуум. Избыточность НВП позволяет оценивать локализованные энергетические характеристики временного ряда и, таким образом, полноценно выявлять структуру сигнала и детектировать шумовые паттерны. Так как энергетический вклад паттерна шумового артефакта определенной природы концентрируется вокруг некоторого уровня НВП-разложения сигнала, то спектральные составляющие распределяются по частотам неравномерно, и вейвлет-энтропия в частотном диапазоне паттерна принимает низкое значение. За счет исключения из рассмотрения тех частотных диапазонов НВП сигнала, которые зашумлены паттернами шумовых артефактов, обеспечивается выше названное свойство робастности алгоритма. Мониторинг усталости человека выполняется на основе анализа поведения функции глобального спектра энергии для выбранного частотного диапазона сигнала ЭЭГ на интервале наблюдений. Снижение этого функционального параметра свидетельствует о накоплении усталости. Для решения задачи прогноза усталости используется модель нейронной иерархической интерполяции N-NiTS. В отличие от близкой модели N-BEATS, выбранная модель обеспечивает эффективный прогноз на вариативных временных масштабах. В качестве дискриминативных признаков используются значения скалограмм в тех частотных диапазонах, где вейвлет-энтропия на интервале

времени минимальна.

55. Алгоритм классификации премоторных потенциалов по сигналу электроэнцефалограммы для нейрореабилитации с помощью интерфейса «мозг-компьютер» замкнутого цикла

Саевский А.И. (1), Шепелев И.Е. (1), Щербань И.В. (1)

(1) Южный федеральный университет

В течение последних десятилетий активно развиваются так называемые интерфейсы «мозг-компьютер» (ИМК). ИМК представляет из себя систему, в которой сигналы активности мозга считываются методами электрофизиологии, после чего подвергаются автоматизированной обработке для конвертации в четкую команду управления. Последнее звено таких систем является наиболее проблемной частью в связи с нестационарностью считываемых биологических сигналов, которая усложняет устойчивое и точное декодирование. В связи с этим разработка устойчивых методов обработки и классификации сигналов, чаще всего энцефалограммы (ЭЭГ), для извлечения команд управления ИМК является крайне актуальной задачей. Внутри этой задачи также существует более узкая, связанная с долей информативного сигнала в обрабатываемых сегментах: она может быть мала, а в случае малой выраженности целевых паттернов они могут просто затеряться в фоновом сигнале из-за грубой разметки выборки. В данной работе предлагается подход для решения вышеуказанных задач на примере классификации премоторных потенциалов, основанный на применении такой модели машинного обучения как линейный дискриминантный анализ (ЛДА) к последовательному выделению информативных составляющих в частотной и временных областях. Первый этап заключается в поиске для каждого целевого движения ЛДА-преобразования выборки на основе спектральных мощностей классических частотных диапазонов многоканального ЭЭГ сигнала, составленной из примеров этого движения и фонового сигнала. В результате получается столько же ЛДА-компонент, сколько было целевых движений, в нашем случае два – движения правой и левой руками, то есть компактно представлена информация из частотной области сигнала за счет понижения размерности с помощью ЛДА. Для дальнейшего выделения информации в частотной области предлагается нарезать сигнал каждого движения и аналогичного по длине фонового сигнала на пересекающиеся короткие сегменты от 250 до 750 мс. На каждом коротком сегменте рассчитываются признаки, после чего создается выборка из разностей признаков фона и движения (класс 1) и разных сегментов фона между собой (класс 0). Столбцы данной выборки соответствуют идентичным временным сегментам, поэтому на данной выборке можно обучить ЛДА (в нашем случае 2-компоненту) для понижения размерности, что в результате сжато будет представлять информацию из временной области. Таким образом, данный шаг, по сути, представляет из себя популярную в области машинного обучения аугментацию данных с последующим понижением размерности с помощью ЛДА. Для окончательной классификации по такому сжатому признаковому пространству применена такая модель машинного обучения как логистическая регрессия. Приведены результаты сравнения качества работы алгоритма в нескольких вариациях на основе признаков разной природы (спектральные

мощности, параметры Юрта, межканальные корреляции) с традиционным подходом на основе общих пространственных паттернов и линейного классификатора, часто используемым в мировых исследованиях, а также между собой. Показано, что уже первый этап с преобразованиями в частотной области дает улучшение точности с 63.9 в традиционном подходе до 77.5% на выборке из 16 экспериментов на разных испытуемых. С добавлением второго этапа с преобразованием во временной области точность повышается до 98.8% при использовании либо признаков корреляций на сегментах длины 750 мс, либо параметров Юрта с сегментами по 500 мс. В среднем по разным конфигурациям моделей длина сегмента 500 мс является наиболее оптимальной. Таким образом, помимо полученного повышения метрик точности, предлагаемый алгоритм на основе классических методов машинного обучения имеет свойство устойчивости к преходящим шуму и артефактам за счет разбиения сигнала на короткие сегменты, а также решает актуальную задачу выделения информативной части сигнала в ситуациях, когда экспертная разметка сигнала может быть лишь примерной. Более того, показано, что после всех преобразования данные от разных испытуемых остаются почти полностью линейно разделимы и могут быть качественно классифицированы общей моделью, что если и встречается в мировых, то с рядом допущений, громоздкими моделями и уступающими показателями точности. Предложенный алгоритм может быть использован для нейрореабилитации – восстановление утраченных функции ЦНС или ПНС, в частности, моторных. Для этого чаще всего используется ИМК замкнутого цикла, в котором пациент получает обратную связь, сигнализирующую об успешном выполнении целевого движения. В случае утраченных моторных функций выполнением целевого стимула будет уже даже генерация в ЭЭГ пространственно-временного паттерна намерения движения какой-либо конечностью (премоторного паттерна), что может быть успешно детектировано предложенным алгоритмом. Что важно, за счет использования простых моделей машинного обучения предложенный алгоритм может использоваться в режиме реального времени без высоких требований к производительности ПК.

75. Spiking neural network actor-critic reinforcement learning with temporal coding and reward modulated plasticity

Vlasov D. S. (1), Rybka R.B. (1), Serenko A.V.(1), Sboev A.G. (1, 2)

(1) NRC “Kurchatov Institute, (2) National Research Nuclear University MEPhI (Moscow Engineering Physics Institute)

The article presents an algorithm for adjusting the weights of the spike neural network of the Actor-Critic architecture. A feature of the algorithm is the use of time coding of input data. The critic neuron is used to calculate the change in the expected value of the action performed based on the difference in spike times received by the critic when processing the previous and current states. The change in the weights of the synaptic connections of the Actor and Critic neurons is carried out under the influence of local plasticity (Spike-timing-dependent plasticity), in which the change in weight depends on the received value of the expected reward. The proposed learning algorithm was tested to solve the problem of holding a cart pole (cart-pole), in which it demonstrated its effectiveness. The proposed algorithm

is an important step towards the implementation of reinforcement learning algorithms for spiking neural networks on neuromorphic computing devices.

48. Сверточные нейронные сети для создания углеродного фотолюминесцентного наносенсора ионов металлов

Г. Чугреева (1), К. Лаптинский (1,2), О. Сарманова (1), Т. Доленко (1)

(1) Физический факультет, МГУ им. М. В. Ломоносова, Москва, РФ (2) НИИЯФ им. Д. В. Скобельцына, МГУ им. М. В. Ломоносова, Москва, РФ

Углеродные точки (УТ) – класс ноль-размерных наночастиц, обладающих стабильной интенсивной фотолюминесценцией (ФЛ), которая зависит от условий синтеза и крайне чувствительна к изменению параметров окружения – температуры, pH, наличия ионов и биомолекул [1]. Такие свойства углеродных точек открывают широкие возможности их использования в качестве наносенсоров параметров среды [2]. Так, актуальной задачей является контроль ионов тяжелых металлов в технологических средах, природных и сточных водах. Такой контроль подразумевает мониторинг одновременно концентраций нескольких катионов металлов в среде. Однако в настоящее время нам известны работы, в которых УТ используются лишь как наносенсоры, максимум, 2-х катионов металлов [3]. В таких работах один из параметров фиксируемый, а концентрация второго определяется с помощью калибровочных кривых. Очевидно, что для создания мультимодального наносенсора необходимо решить многопараметрическую обратную задачу люминесцентной спектроскопии, что успешно достигается с помощью искусственных нейронных сетей (ИНС). В данной работе представлены результаты разработки оптического углеродного наносенсора, способного одновременно определить тип и концентрацию 6 ионов в водной среде по спектрам ФЛ УТ. Для решения указанной 6-параметрической обратной задачи применялись различные архитектуры ИНС: перцептроны с различным количеством скрытых слоев и нейронов них, 1D и 2D сверточные сети. С помощью двумерных сверточных нейронных сетей удалось решить обратную задачу фотолюминесцентной спектроскопии по определению типа и концентрации катионов тяжелых металлов Ni²⁺, Cu²⁺, Co²⁺, Al³⁺, Cr³⁺ и аниона NO₃⁻ со средней абсолютной ошибкой 1.22±0.13 мМ, 0.77±0.15 мМ, 0.79±0.10 мМ, 0.39±0.11 мМ, 0.40±0.10 мМ, 2.26±0.35 мМ, соответственно. Полученная точность решения обратной задачи удовлетворяет потребностям мониторинга водных технологических сред [4].

Исследование выполнено за счёт гранта Российского научного фонда № 22-12-00138, <https://rscf.ru/project/22-12-00138/> [<https://rscf.ru/project/22-12-00138/>].

[1] A.M. Vervald, K.A. Laptinskiy, G.N. Chugreeva, S.A. Burikov, T.A. Dolenko. Quenching of Photoluminescence of Carbon Dots by Metal Cations in Water: Estimation of Contributions of Different Mechanisms. *J. Phys. Chem. C*, vol. 127, pp. 21617-21628, 2023. [2] Wibrianto, A., Khairunisa, S. Q., Sakti, S. C. W., Ni'mah, Y. L., Purwanto, B., & Fahmi, M. Z. (2021). Comparison of the effects of synthesis methods of B, N, S, and P-doped carbon dots with high photoluminescence properties on HeLa tumor cells. *RSC Advances* (Vol. 11, Issue 2, pp. 1098–1108). Royal Society of

Chemistry (RSC). <https://doi.org/10.1039/d0ra09403j> [<https://doi.org/10.1039/d0ra09403j>] [3] Li, P., & Li, S. F. Y. (2020). Recent advances in fluorescence probes based on carbon dots for sensing and speciation of heavy metals. *Nanophotonics* (Vol. 10, Issue 2, pp. 877–908). Walter de Gruyter GmbH. <https://doi.org/10.1515/nanoph-2020-0507> [<https://doi.org/10.1515/nanoph-2020-0507>] [4] F. Akbal, S. Camci, Treatment of metal plating wastewater by electrocoagulation, *Environ. Prog. Sustain*, vol. 31, pp. 340–350, 2012.

73. Kolmogorov-Arnold Networks vs Multi-Layer Perceptron: Solution of an Inverse Problem of Exploration Geophysics

Kupriyanov G.A. (1,2), Isaev I.V. (1), Dolenko S.A. (1)

(1) D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, (2) Faculty of Physics, M.V. Lomonosov Moscow State University

Kolmogorov-Arnold neural networks (KAN) are an alternative approach to standard neural network models. In contrast to the search for matrix coefficients, the KAN searches for continuous activation functions [1]. This allows KAN to show greater non-linearity, and makes the result of its prediction easier to interpret. In this study, we test KAN against usual multilayer perceptron type neural networks on solving the inverse problem of exploration geophysics. Exploration geophysics requires solving specific inverse problems—reconstructing the spatial distribution of the medium properties in the thickness of the earth from the geophysical fields measured on its surface [2]. Dependence of the target variables on a large number of interconnected input features and non-linear interaction among the latter make neural networks one of efficient method of solving such problems. Here we demonstrate that for parameterization schemes with a relatively small number of input features KAN outperform multi-layer perceptron in respect to regression error, at the expense of higher computational cost.

The study was carried out at the expense of the grant No. 24-11-00266 from the Russian Science Foundation, <https://rscf.ru/en/project/24-11-00266/> [<https://rscf.ru/en/project/24-11-00266/>].

1. Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljagic, M., Hou, T.Y., & Tegmark, M. (2024). KAN: Kolmogorov-Arnold Networks. ArXiv, abs/2404.19756. 2. Isaev, I.V., Osborne, I.E., Osborne, E.A., Rodionov, E.A., Shimelevich, M.I., Dolenko, S.A. (2023). Study of the Integration of Physical Methods in Neural Network Solution of the Inverse Problem of Exploration Geophysics with Variable Physical Properties of the Medium. *Moscow University Physics Bulletin*, 2023, Vol. 78, Suppl. 1, pp. S122–S127.

67. Модификация визуального трансформера методом b-cos для повышения интерпретируемости в задаче классификации изображений дерматоскопии

Лукьянов Андрей Николаевич (1), Волков Егор Николаевич (1), Ярушев Сергей Александрович (1), Аверкин Алексей Николаевич (1)

(1) Российский экономический университет им. Г.В. Плеханова

Появление архитектуры «трансформер» совершило революцию в задачах анализа текстов и повысило качество анализа изображений. Доля использования нейросетей архитектуры трансформер растёт с каждым годом. Однако, из-за крайне слабой интерпретируемости и потребности в больших объёмах обучающих данных, использование стандартных архитектур визуального трансформера затруднено. В 2023 году группой исследователей (Manuel Tran, Amal Lahiani и др.) был предложен метод на основе b-cos выравнивания для улучшения показателей работы визуальных трансформеров в задачах анализа изображений. По задумке авторов, замена всех линейных преобразований преобразованием b-cos для обеспечения выравнивания входных данных способно значительно улучшить объяснимость нейросети и показатели отдельных метрик качества. В представленной работе приводится пример практического применения указанного метода путём модификации архитектуры стандартного визуального трансформера. В ходе экспериментов модификация показала хорошие результаты в задаче многоклассовой классификации изображений дерматоскопии (меланома, меланоцитарные невусы и т.д.). Приведён вывод визуального объяснения в виде тепловой карты, отображающей области изображений, внёсшие наибольший вклад в итоговое предсказание. Обсуждаются перспективные методы и возможные пути улучшения объяснимости визуальных трансформеров.

Исследование выполнено за счет Российского научного фонда (грант № 22-71-10112), <https://rscf.ru/project/22-71-10112/> [<https://rscf.ru/project/22-71-10112/>].

73. Machine Learning for NICA SPD Aerogel Reconstruction

Foma Shipilov (1,2), Alexander Barnyakov (3,4), Artem Ivanov (5), Fedor Ratnikov (1)

(1) HSE University, (2) Skoltech, (3) Budker Institute of Nuclear Physics of Siberian Branch Russian Academy of Sciences, (4) Novosibirsk State Technical University (5) Joint Institute for Nuclear Research

In the end-cap region of the SPD detector complex, particle identification will be provided by a Focusing Aerogel RICH detector (FARICH). The FARICH's primary function is to separate pions and kaons in final open charmonia states (momentum range below 5 GeV/c). The optimization of detector parameters, as well as a free-running (triggerless) data acquisition pipeline to be employed in the SPD necessitate a fast and robust method of event reconstruction. In this work, we employ a Convolutional Neural Network (CNN) for particle identification in the aerogel. The CNN model achieves a superior separation between pions and kaons compared with traditional approaches. Unlike algorithmic methods, an end-to-end CNN model is able to process a full 2-dimensional detector response and skip the intermediate step of computing particle velocity, solving the particle classification task directly.

52. Improving Physics-Informed Neural Networks via Quasi-classical Loss Functionals

Sergey Shorokhov

Peoples' Friendship University of Russia named after Patrice Lumumba, Moscow, Russia

As is well-known, the loss functionals of physics-informed neural networks (PINNs) contain the residuals loss functional for the governing partial differential equation (PDE) and the residuals loss functionals for each of the boundary (or initial) conditions of the boundary value problem under consideration. We study the problem of constructing loss functionals for PINN training using methods of the theory of variational principles for nonpotential operators, developed by V.M. Filippov. Generally, a quasi-classical variational functional contains derivatives of lower order compared to the order of derivatives in the PDE with some boundary conditions integrated into the functional, which results in lower computational costs when estimating the loss functional via Monte Carlo integration. We demonstrate advantages of quasi-classical loss functionals over the residuals loss functional for some boundary value problems for hyperbolic PDEs. The first quasi-classical loss functional contains first-order derivatives of the unknown function, requires evaluation of two integrals, and is based on the variational formulation introduced by V.M. Shalov. Another obtained loss functional does not contain derivatives of the unknown function at all, requires evaluation of only one integral, and is based on the variational formulation established by V.M. Filippov.

42. Применение свёрточных нейронных сетей для выделения событий ШАЛ в эксперименте СФЕРА-3

Энтина Е.Л.(1), Подгрудков Д.А.(1), Азра К.Ж.(1), Бонвеч Е.А.(1), Галкин В.И.(1), Зива М.Д.(1), Иванов В.А.(1), Колодкин Т.А.(1), Овчаренко Н.О.(1), Роганова Т.М.(1), Черкесова О.В.(1), Чернов Д.В.(1)

(1) Московский государственный университет имени М. В. Ломоносова, Россия, Москва.

Телескоп СФЕРА-3 разрабатывается для изучения массового состава и энергетического спектра космических лучей в диапазоне энергий 5-1000 ПэВ. Используемый метод детектирования отраженного от снежной поверхности излучения Вавилова-Черенкова от широких атмосферных ливней (ШАЛ) предполагает непрерывную регистрацию света, поэтому требуется разработать триггерный алгоритм для выделения фрагментов, соответствующих пришедшим событиям ШАЛ, в непрерывном потоке фотонов звёздного фона и собственного свечения атмосферы. В работе представлены результаты применения свёрточных нейронных сетей в задаче классификации изображений, полученных с помощью Монте-Карло моделирования черенковского света ШАЛ. К изображениям применяются необходимые модификации, включая оптический фон, учёт геометрии оптической схемы и функции отклика электроники. Дана оценка качества нейронной сети в сравнении с показателями триггерного алгоритма, реализованного в установке СФЕРА-2.

40. Дескрипторы для определения локальной структуры атомистических моделей аморфных льдов

П. А. Яшин (1), Е. Р. Баганцова (1), А. А. Шуплецова (1), В. В. Стегайлов(1)

(1) Московский физико-технический институт (национальный исследовательский университет)

Часто при исследовании воды и её свойств с применением молекулярной динамики требуется распознавать различные фазы льдов. Кроме того, всё большую популярность и значимость получают аморфные льды, среди которых можно выделить аморфные льды низкой плотности (LDA), высокой плотности (HDA), а по недавним исследованиям – и средней плотности (MDA) [1]. Для таких льдов применяются простые методы распознавания локальной структуры (принадлежности отдельной молекулы конкретной фазе), но они обладают малой точностью и в зависимости от задачи могут формулироваться иначе, иметь другие внутренние параметры или вовсе быть неприменимы [2]. Поэтому особый интерес представляет разработка модели, способной корректно обрабатывать в том числе и аморфные структуры. В данной работе мы рассматриваем применение различных дескрипторов для описания кристаллических и аморфных льдов, а также жидкой фазы воды для решения задачи о распознавании этих структур. Сравнение производится посредством обучения классической для данного типа задач нейронной сети, входной вектор которой – структура, описанная одним из дескрипторов. На основе проводимого сравнительного анализа нами разрабатывается оптимальный способ описания локальных атомистических образований, который может быть применён для решения задачи о распознавании льдов, в том числе аморфных, и воды.

1. Rosu-Finsen A. [et al.]. Medium-density amorphous ice. *Science* 2023. V. 379. P. 474-478. DOI: 10.1126/science.abq2105. 2. Garkul A., Stegailov V. Molecular dynamics analysis of elastic properties and new phase formation during amorphous ices transformations. *Sci. Rep.* 2022. V. 12:13325. P. 2045-2322. DOI: 10.1038/s41598-022-17666-2.

71. Deep learning-based eye anatomical structures synthetic images generating: a comparative analysis of techniques and models

Averkin Alexey(1), Matveev Ivan(1), Volkov Egor(1)

(1) FRC CSC RAS

Applications of artificial neural networks (ANN) in the diagnosis of various diseases based on deep learning image analysis is one of the most important technologies developing within the Healthcare 5.0 paradigm. However, the major limitation associated with the applicability of this technology is the availability of data for training ANNs. The labeling of real medical images is performed by clinicians and is therefore extremely costly. In addition, not for all pathologies, not enough images can be obtained for quality ANN training. The problem of data scarcity for ANN training is particularly acute in ophthalmology, where publicly available datasets exist only for the most common pathologies. The solution to the problem of data limitations is the generation of synthetic images of the necessary anatomical structures of the eye (retina, ocular nerve, iris) with the presence of pathologies on them. Deep learning methods and various generative ANN models (GAN, VAE,

Diffusion models, etc.) are also actively used for this purpose. The use of each of the models is associated with different quality of the generated images. In this paper a comparative analysis of techniques and models for generating synthetic images of anatomical structures of the eye is carried out. The principles of realistic image generation and the dependence of image quality on the groups of model architectures used are considered. Conclusions are made about current possibilities and prospects of generation of synthetic images of anatomical structures of the eye using ANN.

This study was supported by the Russian Science Foundation (grant no. 24-21-00330), <https://rscf.ru/en/project/24-21-00330/> [<https://rscf.ru/en/project/24-21-00330/>].

47. Engineering Point Defects in Transition Metal Dichalcogenides for Tailored Material Properties

Abdalaziz Al-Maeeni (1) Tigran Ramazyan (1) Denis Derkach (1) Andrey Ustyuzhanin (2,3)

1) HSE University, Myasnitskaya Ulitsa, 20, Moscow 101000, Russia 2) Constructor University, Campus Ring 1, Bremen 28759, Germany 3) Institute for Functional Intelligent Materials, National University of Singapore, 4 Science Drive 2, Singapore 117544, Singapore

The tunability of physical properties in transition metal dichalcogenides (TMDCs) through point defect engineering offers significant potential for the development of next-generation optoelectronic and high-tech applications. Building upon prior work on machine learning-driven material design, this study focuses on the systematic introduction and manipulation of point defects in TMDCs to tailor their electronic, optical, and mechanical properties. Leveraging a comprehensive dataset generated via density functional theory (DFT) calculations, we explore the effects of various defect types and concentrations on the material characteristics of TMDCs. Our methodology integrates the use of pre-trained large language models to generate defect configurations, enabling efficient predictions of defect-induced property modifications. This research differs from traditional methods of material generation and discovery by utilizing the latest advances in transformer models, which have proven to be efficient and accurate discrete predictors. In contrast to high-throughput methods where configurations are generated randomly and then screened based on their physical properties, our approach not only enhances the understanding of defect-property relationships in TMDCs but also provides a robust framework for designing materials with bespoke properties. This facilitates the advancement of materials science and technology.

- Part of my thesis is funded by the Ministry of Education and some additional funding was obtained through AI Center and RScF.

51. Машинное обучение и 3D анализ сцен в оценке сложноизмеряемых геометрических характеристик тела человека для задач биомедицины

Авторы представляют новый подход, объединяющий технологии обработки 3D изображений и машинного анализа данных, что позволяет автоматизировать процесс определения антропометрических характеристик и функционалов Минковского тела человека, в частности, площади его поверхности и параметров формы. Геометрические характеристики, такие как объем (и вес), площадь поверхности, форма тела (соматотип) пациента, играют в медицинской практике ключевую роль в первичной дифференциальной диагностике многих заболеваний и прогнозе их течения. Те же характеристики используются при назначении объёма врачебного (лекарственного, хирургического и т.д.) вмешательства во многих областях медицины, включая химиотерапию рака, трансплантологию, лечение ожогов, кожных заболеваний, в токсикологии. Традиционные измерительные и расчетные методики анализа указанных параметров часто дают настолько грубые их оценки, что в некоторых случаях они могут значительно влиять даже на смертность пациентов, особенно для детей и людей с неидеальным телосложением: именно они и нуждаются в особенно точных и индивидуализированных измерениях. Для более качественного измерения столь важных геометрических характеристик авторы работы предлагают использование методов машинного обучения. В работе решается задача регрессии оценки геометрических характеристик тела человека по его линейным обмерам (обхваты груди, талии, длины рук, ног и т.д.). Лучшим методом показала себя линейная регрессия, которая позволила улучшить относительную погрешность измерения объема тела с 15% (через оценку плотности тела) до 2%, площади поверхности тела человека в сравнении с популярной формулой Дюбуа с 1.6% до 1.1%. При этом модель обладает большей обобщающей способностью, позволяющей точно вычислять объем и площадь поверхности тела в независимости от его комплекции, что может быть критически важно в случаях, касающихся жизни человека. Результаты исследования показывают высокую точность и эффективность разработанного метода, открывая перспективы для создания инновационных автоматизированных решений в ряде областей биомедицины и здравоохранения, повышения качества диагностики и лечения.