



Program

17.06.2025

The list of accepted reports.

Please note that the first author should be the presenter.

If someone did not find themselves in the list, please inform us by email dlcp@sinp.msu.ru [<mailto:dlcp@sinp.msu.ru>]

Section 1. Machine Learning in Fundamental Physics

36. Реконструкция энергии космических лучей ультравысоких энергий, зарегистрированных флуоресцентным телескопом: одного такта времени может быть достаточно

М.Ю. Зотов (1), А.А. Трусов (2)

(1) НИИЯФ МГУ, (2) Физический факультет МГУ

Мы рассматриваем задачу реконструкции энергии космических лучей (КЛ) ультравысоких энергий по данным флуоресцентного телескопа EUSO-TA. Данные были собраны в 2015 г. на сайте эксперимента Telescope Array (TA). EUSO-TA – это небольшой телескоп-рефрактор с диаметром линз 1 м, полем зрения 10x10 градусов и временным разрешением 2.5 мкс, созданный для наземных тестов аппаратуры, создаваемой в рамках научной программы коллаборации JEM-EUSO. Основные трудности реконструкции энергии возникают вследствие маленького поля зрения (в 30 раз меньше, чем у флуоресцентных телескопов TA) и низкого временного разрешения (в 25 раз ниже, чем у TA). Это приводит к тому, что EUSO-TA способен регистрировать лишь часть треков широких атмосферных ливней, и весь сигнал оказывается, как правило, записан в течение всего одного такта времени. Мы покажем, что ансамбль свёрточных нейронных сетей позволяет даже в таких условиях оценить энергию первичных КЛ с приемлемой точностью.

40. Фильтрация ложных максимумов ШАЛ с помощью нейросетевых методов в эксперименте СФЕРА-3

Энтина Е.Л.(1), Подгрудков Д.А.(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.

Проектируемый в настоящее время телескоп СФЕРА-3 предназначен для изучения космических лучей в диапазоне энергий 1–1000 ПэВ методом регистрации излучения Вавилова–Черенкова, отражённого от снежной поверхности. Одной из задач обработки данных является корректная идентификация событий, для которых ось широкого атмосферного ливня (ШАЛ) лежит в поле зрения телескопа. При больших расстояниях между осью ливня и центром поля зрения телескопа за счёт флуктуаций могут формироваться ложные изображения, в которых максимум яркости не соответствует истинному положению оси, что приводит к систематическим ошибкам в методах восстановления параметров первичной частицы ШАЛ, основанных на аппроксимации образа ШАЛ. В представленной работе предложен метод фильтрации событий ШАЛ с ложными максимумами яркости на основе двух подходов машинного обучения. Первый — регрессия расстояния от оси телескопа до оси ШАЛ с использованием сверточной нейронной сети, обученной на смоделированных изображениях. Второй — применение автоэнкодера, обученного реконструировать только истинные изображения, разделяющего события по значению ошибки восстановления. Совмещение этих двух методов позволяет разделять события с корректно и ошибочно определённой осью без зависимости от энергии, массы и угла прихода первичной частицы. Показано, что такой подход обеспечивает высокую точность фильтрации: сохраняется большинство истинных событий при эффективном подавлении ложных. Предложенный метод может быть использован как часть программного триггера или системы отбора событий для последующего физического анализа.

71. Использование нейронного автокодировщика для генерации показаний поверхностных детекторов Telescope Array

Фитагдинов Р.Р. (1,2), Харук И.В.(1,2)

(1) МФТИ, (2) ИЯИ РАН

Исследована возможность применения автоэнкодера (АЕ) для обнаружения аномалий в данных широких атмосферных линий (ШАЛ), смоделированных методом Монте-Карло. АЕ обучался исключительно на событиях с протоном в качестве первичной частицы, достигнув низкой ошибки реконструкции ($\sim 4 \times 10^{-3}$) в латентном пространстве размерности 16. При тестировании на смеси протонных (типичных) и фотонных (аномальных) событий, вопреки ожиданиям, ошибка реконструкции для фотонных событий оказалась в ~ 3 раза ниже, чем для протонных. Анализ латентных представлений АЕ и применение алгоритмов поиска аномалий без учителя (Isolation Forest, One-Class SVM, KDE) подтвердили парадоксальный результат: фотонные события не идентифицируются как аномалии, а скорее соответствуют “наиболее типичным” протонным событиям в латентном пространстве. Это указывает на то, что выбранный формат входных данных ШАЛ (координаты детекторов, интегральный сигнал, времена фронтов) не позволяет АЕ эффективно выделять признаки, существенные для различения типов частиц классификаторами (с точностью > 0.9).

52. Графовая нейронная сеть с механизмом внимания для кластеризации треков частиц по событиям в эксперименте SPD на ускорителе NICA

Омелянчук С.С.(1,2), Ососков Г. А.(2), Талочка Е.Н.(2)

(1) Филиал МГУ, г. Дубна, (2) Объединенный Институт Ядерных Исследований ОИЯИ

Данная работа посвящена разработке методов глубокого обучения для кластеризации треков элементарных частиц по событиям. В данной работе рассматривается архитектура графовой нейронной сети с механизмом внимания (GANN) для классификации треков по событиям в каждом временном срезе на эксперименте SPD. В работе представлен новый подход к сортировке треков, исследование динамики обучения и тестирование модели в разных условиях. Модель реализована и обучена с применением современных инструментов глубокого машинного обучения, предоставляющих возможность параллельных тензорных вычислений.

69. Machine Learning Approach for Lattice Quantum Field Theory Calculations

Vsevolod Chistiakov (1) Dmitry Salnikov (1,2) Artyom Vasiliev (1) Aleksandr Ivanov (1)

(1) MSU. (2) INR RAS

In modern quantum field theory and statistical physics, the expectation values of observables are represented as integrals over function space. In most interesting problems, such integrals can only be computed numerically using lattice approximations, where the functional integral is replaced by a finite-order integral. The resulting multidimensional integrals are computed using Markov chain Monte Carlo methods. Contemporary deep machine learning generative algorithms allow for a significant acceleration of Monte Carlo calculations. The talk will discuss the applications of generative models in quantum scalar field theory.

75. Нейросетевое моделирование оптических солитонов, описываемых обобщённым нелинейным уравнением Шредингера шестого порядка с высокой нелинейностью

Молошников Иван (1), Кувакин Михаил (1,2), Сбоев Александр (1,3)

(1) НИЦ Курчатовский институт, (2) НИУ МФТИ, (3) НИЯУ МИФИ

В работе рассматривается моделирование распространения импульсов в оптической нелинейной среде с использованием обобщенного нелинейного уравнением Шредингера (ОНУШ) шестого порядка производной и с нелинейностью седьмого порядка. Проводится исследование несколько модификаций PINNs (гиперпараметры, топология, методы выбора точек коллокаций, оптимизаторы) на задачах с известным аналитическим решением для одного солитона. Результатом исследования является наиболее оптимальная конфигурация нейронной сети с наилучшими, из рассмотренных, параметрами точности и скорости сходимости. С помощью полученной конфигурации производится моделирование двух и трёх последовательно идущих солитонов. Для такой задачи уже нет известного

аналитического решения и контроль точности модели производится с помощью законов сохранения. В результате показано, что полученная модель PINNs даёт хорошие результаты как для моделирования единичных солитонов так и для много солитонных задач с погрешностью на законах сохранения менее 1%.

65. Temporal difference modulated spiking actor learning

Yunes Tihomirov (1), Roman Rybka (2), Alexey Serenko (2), Alexander Sboev (2)

(1) National Research University Higher School of Economics (HSE), (2) National Research Center Kurchatov Institute

While neuromorphic computing offers substantial energy savings via spiking neural networks (SNNs), developing effective methods suited for hardware deployment for reinforcement learning in SNNs remains a challenge. We present novel spiking neural network architecture for the actor part of the actor-critic framework. The proposed approach incorporates a two-layer network trained using temporal difference modulated spike-timing dependent plasticity (TD-STDP). Evaluated on the classic Acrobot and CartPole control tasks, our SNN-based actor demonstrates competitive performance. Using local plasticity learning rules is important for future implementation on neuromorphic hardware.

41. SBI в задачах анализа динамических изображений многоканального детектора

Сараев Р.Е. (1,2), Шаракин С.А. (1)

(1) НИИЯФ МГУ, (2) Физический Факультет МГУ

Традиционно байесовские модели, формулируемые на языке вероятностного программирования (таких как PyMC или STAN), помимо априорных распределений на параметры содержат функцию правдоподобия. Однако для сложных моделей не всегда удается указать правдоподобие в явном виде, но можно создать симулятор, который генерирует сэмплы данных при каждом заданном значении параметров. Байесовский вывод при таком likelihood-free подходе реализуется в виде последовательной аппроксимации апостериорного распределения (посредством Sequential Monte-Carlo методов) и зависит от ряда управляющих этой аппроксимацией параметров. В докладе на примере задачи эльфокации (восстановления параметров грозового разряда по «ионосферным отпечаткам» - эльфам) показано как можно применять SBI при анализе динамических изображений орбитальных многоканальных детекторов.

76. Simulation of trawl processes using SINN architectures

Belkova Kseniia(1), Mikhailov Mikhail(2)

1. National Research University Higher School of Economics, Saint Petersburg, 2. St. Petersburg State University, St. Petersburg,

This work proposes an approach to simulating trawl processes using Statistics-Informed Neural Networks (SINN) — a stochastic counterpart to Physics-Informed Neural Networks (PINN). Trawl processes are a special case of ambit processes, which are used to model a broad class of spatio-temporal phenomena. These processes are defined via integrals over

Lévy bases on moving sets, allowing for the modeling of various dependency structures in time series. A notable special case is the Gaussian Ornstein–Uhlenbeck process, which has an analytical representation. However, existing modeling methods are limited to a narrow class of trawl processes due to computational complexity, especially when the Lévy basis distribution does not admit a closed-form expression. The main result of this work is a training scheme for SINN based on the characteristic functions of the process’s finite-dimensional distributions. Unlike the original SINN training framework, the proposed approach does not require external simulation of the process during training. The effectiveness of the method is demonstrated on trawl processes, including the Ornstein–Uhlenbeck process, and compared with existing approaches. To validate the method, we used data from the Met Office MIDAS archive of land and marine weather stations: weather conditions for 2013 were modeled based on weather data from 2012.

39. ML-Based Optimum Sub-system Size Heuristic for the GPU Implementation of the Tridiagonal Partition Method

Milena Veneva (1)

(1) RIKEN Center for Computational Science, R-CCS, 7-1-26 Minatojima-minami-machi, Chuo-ku, Kobe, Hyogo 650-0047, Japan

The parallel partition algorithm for solving systems of linear algebraic equations (SLAEs) suggested in [1] is an efficient numerical technique for solving SLAEs with tridiagonal coefficient matrices which consists of three stages. It works by splitting the original matrix into smaller sub-matrices and solving these smaller SLAEs in parallel. Originally designed for use with a large number of processors, this algorithm was implemented using MPI (Message Passing Interface) technology in [1]. The development of HPC applications typically consists of two key phases: writing code that functions correctly and then optimizing that code to enhance performance. The nature of the parallel partition method is such that the initial SLAE with N unknowns is partitioned into a number of sub-systems with m unknowns each. The size of the SLAE N that the user solves is usually determined by the size of the problem they need to solve, while the size of the sub-system within the parallel partition method m is a parameter that needs to be tuned. We present one of the optimizations made to our CUDA [2] implementation, namely building a heuristic for finding the optimum sub-system size by using tools frequently used in modern AI-focused approaches. Computational experiments for different SLAE sizes are conducted, and the optimum sub-system size for each of them is found empirically. To estimate a model for the sub-system size, we perform the k -nearest neighbors (k NN) classification method [3]. Statistical analysis of the results is done. By comparing the predicted values with the actual data, the algorithm is deemed to be acceptably good. Next, the heuristic is expanded to work for the recursive parallel partition algorithm as well. An algorithm for determining the optimum sub-system size for each recursive step is formulated. A k NN model for predicting the optimum number of recursive steps for any SLAE size is built.

[1] Austin, T.~M., and Berndt, M., and Moulton, J.~D., A Memory Efficient Parallel Tridiagonal Solver, Preprint LA-VR-03-4149, 13 p. (2004).

[2] NVIDIA, NVIDIA CUDA C++ Programming Guide. <https://docs.nvidia.com/cuda/cuda-c-programming-guide/> [<https://docs.nvidia.com/cuda/cuda-c-programming-guide/>] (2025).

[3] Fix, H., and Joseph, L., Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties, International Statistical Review/Revue Internationale de Statistique, 57 (3) pp. 238–47, doi: 10.2307/1403797 (1989).

44. Natural Image Classification via Quasi-Cyclic Graph Ensembles and Random-Bond Ising Models with Enhanced Nishimori Temperature Estimation

V.S.Usatyuk (1,2), D.A.Sapozhnikov (1), S.I.Egorov(2)

(1) T8 LLC, Moscow, Russia, (2) SWSU University, Kursk, Russia

Recent advances have demonstrated the effectiveness of spectral clustering on the beta-Hessian of Graham matrices constructed from quasi-cyclic graphs in the context of Random-Bond Ising Models (RBIMs). Notably, at paper [1] showed that combining LDPC-inspired graph with VGG16-extracted features from GAN-generated two-class images (e.g., dog vs. cat) significantly outperformed Erdős–Rényi baselines in clustering accuracy, improving overlap from 73.21% to 90.60%—and up to 93.23% when using cosine similarity [2]. In this paper, we extend these insights to natural multi-class datasets, specifically ImageNet-10 and ImageNet-100. We introduce a refined approach to estimating the Nishimori temperature and propose a mixture-of-graphs model built from an ensemble of optimized RBIMs. These models leverage diverse quasi-cyclic graph families—including Spherical graphs and Multi-Edge Type LDPC graphs—to create sparse, expressive interaction structures. Feature embeddings are extracted from a lightweight MobileNetV2-based CNN, compressing 1280-dimensional activations to 32–64 feature maps per image. Using ensembles of 3 to 9 graph models, our approach achieves classification accuracies of up to 98.7% on ImageNet-10 and 82.5% on ImageNet-100 under optimal conditions with 32-dimensional embeddings. We demonstrate that significant parameter reduction in the MLP classification head (from 1280 to 32) improves both computational efficiency and robustness to feature puncturing. Furthermore, this graph-based framework shows promise for enhancing the representation power of knowledge graphs and feed-forward layers in transformer architectures. These results highlight the scalability of quasi-cyclic RBIM spectral embeddings from binary-class GAN-generated images to complex, real-world, multi-class image datasets. Our findings suggest that structural graph design—particularly girth, spectral gap, and ensemble diversity—plays a crucial role in optimizing spectral separability for high-dimensional natural image classification tasks.

[1] Usatyuk, V.S., Sapozhnikov, D.A., & Egorov, S.I. (2024). Enhanced Image Clustering with Random-Bond Ising Models Using LDPC Graph Representations and Nishimori Temperature. *Moscow Univ. Phys.*, 79(Suppl 2), S647–S665.

[2] Dall'Amico, L. et al. (2021). Nishimori meets Bethe: A Spectral Method for Node Classification in Sparse Weighted Graphs. *J. Stat. Mech.*, 093405.

43. Analysis of the TAIGA-HiSCORE Data using the Latent Space of Autoencoders

Yu. Dubenskaya(1), S. Polyakov(2), A. Kryukov(1), A. Demichev(1), P. Volchugov(1), E. Gres(1,3); D. Zhurov(1,3), E. Postnikov(1), A. Razumov(1)

(1) SINP MSU, (2) IIAF NAS RA, (3) IPA IGU

The aim of extensive air shower (EAS) analysis is to reconstruct the physical parameters of the primary particle that initiated the shower. The TAIGA experiment is a hybrid detector system that combines several telescopes and arrays of detector stations to record and analyze EAS data. At present, data from the telescopes and the detector station arrays is analyzed by deriving different sets of auxiliary parameters related to the physical features

of the recording hardware. These sets of parameters are chosen empirically, so there is no certainty that they retain all important information contained in the experimental data and are the best suited for the respective problems. Moreover, because the event parameters recorded by different detector types differ in physical nature, their direct merging is unfeasible, which complicates multimodal analysis. We propose to use autoencoders (AE) for the analysis of TAIGA experimental data and replace the conventionally used auxiliary parameters with the parameters of the AE latent space. The advantage of the AE latent space parameters is that they are not biased by pre-established assumptions and constraints and still contain in a compressed form the physical information obtained directly from the experimental data. A separate artificial neural network is used to reconstruct the parameters of the EAS primary particle from the AE latent space parameters. In this paper, the proposed approach is used to reconstruct the energy of the EAS primary particle based on Monte Carlo simulation data for the TAIGA-HiSCORE detector array. The dependence of the energy determination accuracy on the latent space dimension is analyzed, and these results are also compared with the results obtained by the conventional method. For events recorded by TAIGA-HiSCORE, it is shown that when using the AE latent space, the energy of the primary particle is reconstructed with satisfactory accuracy.

This study was supported by the Russian Science Foundation, grant no. 24-11-00136.

81. Проблема аугментация данных атмосферных черенковских телескопов в стерео режиме на примере установки TAIGA-IACT

Д.Журов(1,3), А.Крюков(1), Ю.Дубенская(1), Е. Грес(1,3); С.Поляков(3), ЕюПостников(1), А.Разумов(1), П.Волчугин(1), А.Демичев(1)
(1) SINP MSU, (2) ИАФ НАС РАН, (3) ИРА ИГУ

Изучение источников гамма-излучения высоких энергий (более 1 ТэВ) во Вселенной возможно только с использованием наземных установок большой площади для регистрации широких атмосферных ливней (ШАЛ). Регистрация ШАЛ осуществляется детекторами заряженных частиц и/или детекторами черенковского света. По данным этих детекторов необходимо определить направление прихода, энергию и тип первичной частицы. Определение типа частицы критически важно для подавления фона заряженных космических лучей и выделения гамма-сигнала. В эксперименте TAIGA, который включает разные типы детекторов, регистрирующие мультимодальные данные, эта задача решается, в том числе, с помощью атмосферных черенковских телескопов (АЧТ) установки TAIGA-IACT. Телескопы регистрируют угловое распределение (изображения) черенковского света от ШАЛ в двух режимах: моно (регистрация одним АЧТ) и стерео (регистрация несколькими АЧТ). На основе полученных изображений решаются задачи классификации и регрессии. Современные тенденции в обработке больших данных в области гамма-астрономии методами машинного обучения показывают, что применение нейросетевых моделей для анализа данных АЧТ позволяют оценить параметры космических лучей с очень хорошей точностью. Для обучения нейросетевых моделей необходимы большие размеченные наборы данных. Поскольку провести разметку экспериментальных данных для обучения нейросетевых моделей практически невозможно, то для обучения используются данные имитационного моделирования. Оно включает моделирование развития самого ШАЛ, а также моделирование оптики телескопа, детектора и регистрирующую электронику.

Наиболее ресурсоемким этапом в моделировании данных АЧТ является моделирование развития ШАЛ в атмосфере Земли. Моделирование одного ливня космических лучей высоких энергий может занимать несколько часов работы современного вычислительного сервера. В связи с этим получение больших выборок, содержащих сотни тысяч и миллионы событий, существенно затруднено, что делает задачу аугментации данных особенно актуальной. Аугментация данных путем простого вращения изображений атмосферных черенковских телескопов (АЧТ) продемонстрировала свою эффективность для задач в монорежиме. В ряде случаев такой метод может быть интерпретирован как наблюдение ШАЛ с другого положения в пространстве. Однако, данный способ не применим в случае обработки данных АЧТ в стерео режиме. В данной работе рассматривается возможность аугментации данных АЧТ путем вращения положений телескопов вокруг оси ШАЛ для обучения нейросетевых моделей при наблюдениях в стереорежиме.

Работа выполнена при финансовой поддержке Российского научного Фонда, грант 24-11-00136.

42. Возможность применения метода нормализующих потоков для извлечения редких гамма событий в эксперименте TAIGA

А.Крюков(1), А.Разумов(1), Д.Журов(1,3), Ю.Дубенская(1), Е. Gres(1,3); С.Поляков(3), Е.Постников(1), П.Волчугов(1), А.Демичев(1)
(1) SINP MSU, (2) ИАФ НАС РАН, (3) IPA IGU

Среди многих методов исследования процессов, происходящих в различных уголках Вселенной, изучение гамма-лучей высоких и сверхвысоких энергий является одним из наиболее перспективных направлений в области космических лучей. В отличие от заряженных космических лучей (протоны и атомные ядра), которые подвержены влиянию галактических и межгалактических магнитных полей, гамма лучи в силу своей электронейтральности сохраняют информацию об источнике своего происхождения. Это предопределило в последние годы бурное развитие экспериментальной гамма-астрономии в мире. Однако отметим, что поток гамма-лучей очень мал по сравнению с общим потоком космических лучей. Так для Крабовидной туманности, одного из ярких гамма источников, доля гамма-лучей в общем потоке составляет около 0.01%. Поэтому крайне важно разрабатывать новые эффективные методы классификации зарегистрированных событий. В настоящее время большую популярность приобретают новые алгоритмы для разделения гамма-лучей и заряженных космических лучей, основанных на глубоком обучении. В этой работе мы рассмотрели новый метод обнаружения редких гамма-квантов на адронном фоне с помощью модели глубокого обучения на основе нормализующих потоков. Эти модели являются генеративными моделями, которые строят обратимое преобразование нормального многомерного распределения случайного вектора в сложное распределение, представленной в виде экспериментальной выборки. Мы рассмотрели две версии метода одноклассовой классификации в двух вариантах. В первом случае в качестве обучающего класса выбраны протоны, а гамма-лучи рассматриваются как аномалии, а в втором случае наоборот. Метод апробировался на модельных Монте-Карло данных гамма-событий для проекта TAIGA.

95. Гамма-астрономия ультравысоких энергий и проект TAIGA-100

L.Kuzmichev, SINP MSU

14

Section 2. Machine Learning for Environmental Sciences

15. Восстановление приповерхностной влажности атмосферы над океаном с применением методов машинного обучения

С. А. Вострикова (1), М. А. Криницкий (1,2), С. К. Гулёв (2), М. П. Александрова (2)
(1) Московский физико-технический институт, (2) Институт океанологии им. П. П. Ширшова РАН

Влажность воздуха в приповерхностном слое атмосферы над океаном является ключевым климатическим параметром, влияющим на процессы переноса влаги и тепла между океаном и атмосферой, а также на динамику атмосферных процессов в целом. Международный массив данных о характеристиках океана и атмосферы (ICOADS) указывает на недостаточную плотность измерений влажности в начале XX века по сравнению с более поздними периодами, что создает сложности для адекватного анализа климатических тенденций относительной влажности. Представленные в литературе методы восстановления временных рядов влажности зачастую демонстрируют ограниченную точность, основываясь преимущественно на статистических и эвристических подходах. Наша работа направлена на повышение качества решения этой задачи за счёт применения методов машинного обучения. В настоящей работе решена задача в формулировке аппроксимации моментальных значений относительной влажности по данным сопутствующих измерений атмосферного давления, температуры воздуха, скорости и направления ветра, температуры поверхности океана, а также наблюдений количества и типов облачности на трёх ярусах. Кроме этого, в составе сопутствующих переменных используется код погоды по стандарту ВМО и расчетная высота солнца. В исследовании использованы модели машинного обучения следующих типов: линейная регрессия, случайный лес, градиентный бустинг (CatBoost) и полносвязная искусственная нейронная сеть. Для повышения территориальной и временной специфичности разрабатываемых моделей мы провели исследование для каждой ячейки размером 5 градусов по широте и долготе и каждого сезона по отдельности. На основе полученных результатов были построены карты пространственного распределения ошибок моделей, которые позволили выявить регионы с высокой и низкой точностью аппроксимации влажности. Исследование подтвердило эффективность методов машинного обучения для восстановления климатических рядов, определило наиболее подходящие модели для этой задачи и обозначило перспективные направления для дальнейшей работы.

16. Сравнение моделей машинного обучения в задаче

идентификации аномалий в данных визуальной съемки поверхности моря

Белоусова О.(1,2), Криницкий М.А.(1,2)

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

Обнаружение морского мусора является важной задачей для защиты экосистем океана, поскольку загрязнение угрожает биологическому разнообразию и здоровью водных ресурсов. Однако сложность задачи заключается в большом разнообразии объектов, их малых размерах, частичном погружении под воду и слиянии с окружающей средой, что делает обнаружение затруднительным. В настоящей работе мы проводим сравнительный анализ современных методов нахождения морского мусора на примере двух подходов. В первом из них проведена детекция объектов на цельных фотоснимках с помощью нейронной сети YOLO. Во втором подходе осуществлены поиск аномалий на отдельных фрагментах фотоснимков с помощью свёрточной нейронной сети ResNet50, обученной по методу Momentum Contrast (MoCo), а затем их классификация с помощью градиентного бустинга CatBoost. Особенности предлагаемого во втором подходе алгоритма обучения стали разные способы сэмплирования фрагментов изображений поверхности моря, необходимых для контрастного самоконтролируемого обучения, и использование различных функций потерь. Качество обучения оценивалось с помощью метрики F1-Score в задаче распознавания объектов трёх типов: плавающий морской мусор, птицы и блики на камере.

45. Application of Convolutional Neural Networks for Upper Ionosphere Remote Sensing Using All-Sky Camera Data

Andrei Vorobev (1), Gulnara Vorobeva (2)

(1) The Geophysical Center of the Russian Academy of Sciences, (2) Ufa University of Science and Technology

This study proposes an original approach to the automatic classification of the upper ionosphere state through machine identification of images captured by sky cameras, also known as all-sky imagers. Based on 10 years of sky observations within the auroral oval (Kola Peninsula, Russia), represented by 163,899 images with a 10-minute sampling interval, an intelligent information system was developed using convolutional neural networks. This system identifies whether an input image belongs to one of seven predefined classes and subsequently interprets the result. The analysis of performance metrics for the system, built on the ResNet50 neural network architecture, demonstrated a classification accuracy of 96%, a level practically unattainable through manual processing of datasets of this scale. This approach holds the highest practical significance in Russia's polar regions, where reliable and accurate geomagnetic data coverage is sparse (Taymyr Peninsula, Gydan Peninsula, northern areas of Yakutia, etc.). In these regions, auroras serve as the only widely accessible indicator of space weather conditions and the state of the upper ionosphere.

54. Foundation models of ocean and atmosphere in 2025: milestones and perspectives.

Krinitkiy M.A. (1,2)

Over the past two years, large-scale deep-learning foundation models have evolved from atmospheric-only emulators into first-generation, coupled ocean–atmosphere surrogates capable of delivering global forecasts in seconds. In this talk, we will trace the key milestones that enabled that leap: kilometre-resolution training corpora, self-supervised spatio-temporal transformers, cross-fluid conservation tokens, and mixed-precision inference on modern GPU/TPU hardware. Benchmarks such as GraphCast, Pangu Weather, NVIDIA Earth-2 FourCastNet NIM now meet or exceed state-of-the-art numerical weather prediction at 1–10-day horizons, while emerging oceanic counterparts reproduce mesoscale eddy energetics and seasonal heat content with orders-of-magnitude lower cost. Building on our earlier neural super-resolution and statistical correction modules for surface winds and temperature, we outline a roadmap toward a Russian kilometre-scale foundation models. The presentation will (i) review the present performance frontier, (ii) discuss outstanding challenges—coupling stability, representation of rare extremes, uncertainty quantification, and open benchmarking and (iii) propose community actions, including BRICS-wide validation campaigns and resource-aware fine-tuning strategies. By compressing decades of observations and petabytes of high-resolution simulations into a reusable latent representation, foundation models promise to democratise high-impact forecasts, accelerate climate-risk adaptation, and shift the focus of geophysical research from parameter tuning to physics-aware learning.

60. Сравнение моделей машинного обучения в задаче идентификации аномалий в данных визуальной съемки поверхности моря

Белоусова О.(1,2), Криницкий М. А.(1,2)

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

Обнаружение морского мусора является важной задачей для защиты экосистем океана, поскольку загрязнение угрожает биологическому разнообразию и здоровью водных ресурсов. Однако сложность задачи заключается в большом разнообразии объектов, их малых размерах, частичном погружении под воду и слиянии с окружающей средой, что делает обнаружение затруднительным. В настоящей работе мы проводим сравнительный анализ современных методов нахождения морского мусора на примере двух подходов. В первом из них проведена детекция объектов на цельных фотоснимках с помощью нейронной сети YOLO. Во втором подходе осуществлены поиск аномалий на отдельных фрагментах фотоснимков с помощью свёрточной нейронной сети ResNet50, обученной по методу Momentum Contrast (MoCo), а затем их классификация с помощью градиентного бустинга CatBoost. Особенности предлагаемого во втором подходе алгоритма обучения стали разные способы сэмплирования фрагментов изображений поверхности моря, необходимых для контрастного самоконтролируемого обучения, и использование различных функций потерь. Качество обучения оценивалось с помощью метрики F1-Score в задаче распознавания объектов трёх типов: плавающий морской мусор, птицы и блики на камере.

56. Detection of Irminger Rings in high resolution ocean hydrodynamic modeling data using artificial neural networks

M.Kalinin(1), M.Krinitzkiy(1,2), P.Verezemskaya(1)

(1) Shirshov Institute of Oceanology, Russian Academy of Sciences, (2) Moscow Institute of Physics and Technology

Deep convection in the Labrador Sea is a key component in the formation of the lower branch of the Atlantic Meridional Overturning Circulation (AMOC). It is known that mesoscale eddy activity in the Labrador Sea, represented by Irminger Rings (IR), influences the convection process. In order to analyze the impact of IRs on the spatial-temporal variability of the mixed layer depth, it is necessary to create a trajectory database of eddy motion, which poses the problem of IRs detection and tracking with high accuracy. In this study, we propose the novel technique for detection of IRs in high-resolution ocean numerical simulation. The research is based on the regional model of the Subpolar North Atlantic NNATL12. There are known automated eddy identification methods that are widely used as a tool for studying eddy activity in statistically significant samples. The most commonly used local extrema search method depends strongly on a number of parameters chosen by an expert exploiting this approach. In order to alleviate the subjectivity issue, we first implemented the automatic identification scheme for IRs based on the local extrema search. We optimized the scheme employing Bayesian optimization framework resulting in optimal values of the hyperparameters of this eddy identification algorithm. While the optimization significantly improved the quality of the identification, we found that there is a room for further improvement of IRs detection. As a promising alternative to the heuristic local extrema search algorithm, we propose using artificial neural networks. In this study, we employed a convolutional neural networks similar to U-Net which we trained to segment the eddies. We first pretrained it on the results of heuristic IR detection algorithm. We then further trained it on the expert-labeled IRs. The resulting IR detection quality is high enough to further implement tracking algorithms. The application of artificial neural networks, specifically convolutional neural networks akin to U-Net, has demonstrated considerable potential in enhancing the detection of Irminger Rings in high-resolution oceanic simulations. By leveraging a two-stage training process, initially utilizing heuristic algorithm results followed by expert-labeled IRs, we achieved a detection accuracy that surpasses traditional methods, thus providing a robust foundation for subsequent eddy tracking endeavors. The integration of machine learning techniques with traditional oceanographic methodologies holds significant promise for advancing the precision and reliability of IR detection and tracking. This approach not only mitigates the subjectivity inherent in parameter selection for heuristic methods but also capitalizes on the adaptability and learning capabilities of neural networks. As such, this method presents a substantial improvement over existing techniques and contributes to a more nuanced understanding of mesoscale eddy dynamics and their influence on deep convection processes in the Labrador Sea.

73. Моделирование турбулентного переноса примесей в планетарном пограничном слое с применением методов крупных вихрей и методов машинного обучения

И. А. Герасимов (1), М. А. Креницкий (1, 2), Е. В. Мортиков (3,4)

(1) Московский физико-технический институт (национальный исследовательский

университет), (2) Институт океанологии им. П. П. Ширшова РАН, (3) Научно-исследовательский вычислительный центр Московского государственного университета имени М.В. Ломоносова, (4) Московский центр фундаментальной и прикладной математики, Россия

Планетарный пограничный слой атмосферы (ППС) играет ключевую роль в контексте переноса загрязняющих веществ, что непосредственно влияет на экологическое состояние атмосферы и подстилающей поверхности [1]. Поэтому моделирование распространения примесей в ППС является фундаментальной задачей, поскольку точное прогнозирование необходимо для оценки качества воздуха на различных масштабах. Традиционно считается, что распространение газовых и аэрозольных примесей в ППС хорошо аппроксимируется с использованием функции плотности нормального распределения [2], как показано на рис. 1, однако существующие методы расчета ограничены вычислительными ресурсами из-за сложности моделирования турбулентных процессов, возникающих при взаимодействии воздушных масс с подстилающей поверхностью. Полуэмпирические подходы к прогнозированию переноса загрязнений, демонстрируют невысокую точность, особенно в условиях сложной орографии и неоднородной подстилающей поверхности [3]. Целью настоящего исследования является разработка методологии аппроксимации первых моментов пространственного распределения примесей в турбулентном ППС с применением методов машинного обучения (МО). Для получения референсной коллекции данных мы используем метод крупных вихрей для численного моделирования атмосферы с высоким разрешением. В рамках исследования были проанализированы первые два центральных момента распределения примеси по вертикальной и горизонтальной осям, осредненные по времени, в условиях стационарного ППС. Для формирования коллекции обучающих данных было проведено численное моделирование распространения примесей методом крупных вихрей [4]. В качестве варьируемых входных параметров модели задавались внешние условия, такие как шероховатость подстилающей поверхности, градиент температуры над верхней границей ППС и т.д., одинаковые для всего домена. Также, для каждого источника задавались координаты и его мощность. Эти же параметры использовались для формирования признаков описания событий (вариантов распространения примесей в атмосфере при заданных внешних параметрах) при решении задачи в подходе машинного обучения. Полученные результаты численного моделирования сравнивались с аппроксимациями, рассчитанными при помощи трех различных подходов: традиционной теории на основе Гауссова распределения [3], классических алгоритмов МО и искусственных нейронных сетей (ANN). В качестве классических алгоритмов МО были использованы: линейная модель, модель случайных лесов и модель градиентного бустинга Catboost. Сравнительный анализ результатов продемонстрировал превосходство методов машинного обучения в сравнении с традиционным полуэмпирическим подходом. В перспективе планируется расширение методологии для учета большего количества атмосферных параметров и более разнообразных условий подстилающей поверхности. Также рассматривается возможность применения методов глубокого обучения для повышения точности прогнозирования на больших временных масштабах.

[1] Hendrik Tennekes; The atmospheric boundary layer. Physics Today 1 January 1974; 27 (1): 52-63. <https://doi.org/10.1063/1.3128397> [<https://doi.org/10.1063/1.3128397>] [2] Sutton O. G., "A theory of eddy diffusion in the atmosphere", Proc. Roy. Soc. London, A, 135, 1932,

pp.143-165. [3] Ražnjević, A., van Heerwaarden, C., and Krol, M.: Evaluation of two common source estimation measurement strategies using large-eddy simulation of plume dispersion under neutral atmospheric conditions, *Atmos. Meas. Tech.*, 15, 3611–3628, <https://doi.org/10.5194/amt-15-3611-2022>, 2022 [<https://doi.org/10.5194/amt-15-3611-2022>, 2022]. [4] Tkachenko, E.V., Debolskiy, A.V. & Mortikov, E.V. Intercomparison of Subgrid Scale Models in Large-Eddy Simulation of Sunset Atmospheric Boundary Layer Turbulence: Computational Aspects. *Lobachevskii J Math* 42, 1580–1595 (2021) <https://doi.org/10.1134/S1995080221070234> [<https://doi.org/10.1134/S1995080221070234>]

51. Нейросетевое пространственное масштабирование полей приповерхностного ветра над Баренцевым и Карским морями

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

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

В настоящем исследовании изучается возможность применения искусственных нейронных сетей к задаче масштабирования приповерхностного ветра над Баренцевым и Карским морями. Используются различные конфигурации модели глубокого обучения с пропускными соединениями, способной к выявлению сложных нелинейных соотношений, соответствующих различным пространственным масштабам. В нашем исследовании в качестве входных данных низкого разрешения для такой модели применяются результаты глобального атмосферного реанализа ERA5 с пространственным разрешением $0,25^\circ$, а опорные данные в высоком разрешении представлены результатами численного моделирования с применением модели Weather Research and Forecasting (WRF) с разрешением 6 км за период с 2015 по 2023 гг. Начальными и граничными условиями для модели WRF послужили данные оперативного анализа GFS (Global Forecast System, NOAA, USA). Бикубическая интерполяция данных реанализа является опорным решением и сравнивается с нейросетевым моделированием с точки зрения среднеквадратичной ошибки абсолютной скорости ветра, пикового отношения сигнала к шуму и индекса структурной схожести. Для оценки качества масштабирования мы также демонстрируем валидацию результатов нейросетевого моделирования с точки зрения воспроизведения атмосферных явлений на различных масштабах.

82. Сравнение методов машинного обучения для учета связей с запаздыванием при моделировании городского острова тепла.

Назмутдинов К.Ф. (1, 2), Варенцов М.И. (1, 3)

(1) РГГМУ, (2) ААНИИ, (3) НИВЦ МГУ

В работе исследуются методы машинного обучения для аппроксимации разницы температур между городской и сельской местностью (интенсивности городского острова тепла) на примере Москвы и Санкт-Петербурга. В роли предикторов выступают долгосрочные осредненные по региону данные наблюдений с загородных станций и данные глобального реанализа ERA5 с шагом сетки 0.25° с 2012 по 2023 гг. Особенностью метеорологических данных является наличие сильной автокорреляции (связей с запаздыванием). Для учета этих зависимостей исследуются два подхода: (1) явное порождение признаков, характеризующих связи

с запаздыванием (тенденции, скользящие средние) для классической модели CatBoostRegression, показавшей ранее наилучшую точность, и (2) применение архитектурно учитывающей последовательности рекуррентной искусственной нейронной сети LSTM. Результаты экспериментов показали, что подход с LSTM не превзошел по точности CatBoost с временными признаками, порожденными экспертным образом. В планах – эксперименты с моделью GRU и обучение нейросетей на данных, включающих связи с запаздыванием.

72. Deep Learning-Based Estimation of wind induced waves parameters from X-Band Radar Imagery

Alexander Suslov (1), Mikhail Krinitskiy (1, 2), Alexander Gavrikov (1), Mikhail Borisov (2), Natalia Tilinina (1)

(1) Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, 117997 Russia

(2) Moscow Institute of Physics and Technology, Dolgoprudny, Moscow oblast, 141701 Russia

Shipborne navigation radars are essential for safe maritime navigation, detecting vessels and obstacles. Reflections from the sea surface—known as Bragg scattering—produce sea clutter, typically filtered out as noise. When the sea surface is rough enough (wind speed > 3 m/s, significant wave height (swh) > 0.5 m), this clutter becomes visible in unfiltered radar images, allowing for the retrieval of wind-induced ocean wave parameters.

Traditional wave parameter estimation relies on three-dimensional Fourier analysis and linear dispersion relationships, which require modulation transfer functions, signal-to-noise ratios, and radar-specific calibrations, resulting in high computational costs and limited accuracy gains. Deep learning techniques, particularly convolutional neural networks (CNNs), offer robust image processing capabilities, handling noisy data without the need for Fourier transforms or long radar image series. In this study, we introduce a CNN-based SeaVision package which estimates wave characteristics from shipborne radar data. We trained various CNN architectures (ResNet52, ResNet152, etc.) to infer significant wave height and period, using Spotter buoy measurements as ground truth. Our approach outperforms classical methods by requiring only a single radar snapshot to estimate wave characteristics, whereas traditional 3D Fourier-based methods require over 12 minutes of radar data.

83. USING MACHINE LEARNING METHODS FOR JOINT PROCESSING OF DATA FROM MULTIPLE SEMICONDUCTOR GAS SENSORS

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

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

Scientific-Manufacturing Complex Technological Centre, (3) MIREA – Russian Technological

University, (4) Physics Department, M.V. Lomonosov Moscow State University (5) Chemistry

Department, M.V. Lomonosov Moscow State University

This study addresses 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, as well as high temporal resolution of the sensors, nonlinear temperature operating conditions were used - the so-called heating dynamics. Due to high complexity of physical and chemical models 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. To provide additional selectivity in the detection of specific gases, this study considers simultaneous use of data from multiple semiconductor sensors with various doping components with building machine learning models capable of providing joint processing. Based on the results of the study, conclusions were made regarding the selection of optimal combinations of 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-P from the Russian Science Foundation.

67. Доменная адаптация нейронных сетей в задаче диагностики природных вод по спектрам комбинационного рассеяния света

Л.С.Утегенова (1), К.А.Бузанов (1), А.А.Гуськов (1,2), Т.А.Доленко (1,2), С.А.Доленко (2) (1) Московский государственный университет имени М.В. Ломоносова, физический факультет, (2) Научно-исследовательский институт ядерной физики имени Д. В. Скобельцына Московского государственного университета имени М. В. Ломоносова

Современная экологическая ситуация характеризуется возрастающим антропогенным воздействием на природные жидкие среды, в частности, из-за выбросов в окружающую среду ионов тяжелых металлов. Тяжелые металлы, обладая высокой токсичностью и способностью к биоаккумуляции, представляют серьезную угрозу для водных экосистем и здоровья человека. В связи с этим разработка эффективных методов мониторинга концентраций тяжелых металлов в природных водах является одной из актуальных задач современной экологии. В настоящей работе разрабатывается оптический метод диагностики сложного состава природных вод, основанный на спектроскопии комбинационного рассеяния (КР) света, поскольку данный подход позволяет проводить дистанционный и экспрессный анализ, в отличие от применяющихся на сегодняшний день методов аналитической химии. Для решения многопараметрических обратных задач спектроскопии КР активно применяются искусственные нейронные сети (ИНС), которые за последние десятилетия зарекомендовали себя как мощный инструмент для решения некорректно поставленных обратных задач с выраженной нелинейностью. Однако применение ИНС для решения указанных задач требует получения представительного набора данных реальных сред (тысячи спектров), что существенно ограничивает возможность трансфера разработанных технологий из области научных исследований в область практического применения. В данной работе предлагаются подходы, способные решить проблему нехватки данных. Во-первых, рассматривается обучение ИНС на спектрах КР модельных растворов, полученных в лабораторных условиях. Во-вторых, для преодоления проблемы различия в распределении спектральных данных модельных и речных растворов исследовались и применялись алгоритмы доменной адаптации. В рамках исследования был проведен сравнительный анализ альтернативных подходов: переноса обучения и доменно-состязательного обучения ИНС при переходе от большой выборки спектров модельных растворов тяжелых металлов в дистиллированной воде (исходный домен) к малой выборке спектров растворов, полученных на основе реальных речных вод Москвы-реки, Яузы, Битцы и Сетуни (целевой домен). Было получено уменьшение ошибки для всех исследуемых ионов (Zn^{2+} , Cu^{2+} , Li^{+} , Fe^{3+} , Ni^{2+} , NH_4^{+} , SO_4^{2-} , NO_3^{-}) в среднем на 50% по сравнению с применением ИНС, обученных только на малой базе спектров КР растворов на

основе реальных речных вод.

Исследование выполнено за счёт гранта Российского научного фонда № 24-11-00266. Работа Л.Утегеновой поддержана фондом теоретической физики и математики «Базис» (договор № 24-2-1-72-1).

89. Восстановление высоты зданий с использованием машинного обучения и цифровой модели поверхности ArcticDEM

Окунева Влада Викторовна(1), Самсонов Тимофей Евгеньевич(1,2), Варенцов Михаил Иванович (1,2)

(1) Московский государственный университет имени М.В. Ломоносова, (2) ФГБУ «Гидрометцентр России»

Моделирование погоды и климата для урбанизированных территорий требует детального описания городской подстилающей поверхности. Стремительное преобразование городской среды требует регулярного обновления необходимых для расчета данных. Несмотря на появление все большего количества моделей городского климата в разных масштабах, на данный момент все еще отсутствует полный, глобальный и согласованный набор пространственных данных, характеризующих релевантные свойства городской застройки. Существующие сейчас продукты, показывают, что разнообразие параметров и неоднородность городской среды в разных частях мира осложняют создание такого набора данных. Именно поэтому важна интеграция существующих наборов данных и усовершенствование уже существующих баз, путем дополнения необходимых атрибутивных характеристик и заполнения недостающих данных. Целью исследования является восстановление высотности зданий, одного из ключевых геометрических параметров городской среды, на основе разнообразных источников с использованием машинного обучения. Работа выполнена на примере города Санкт-Петербурга. В работе используется новая векторная база пространственных данных Overture Maps, уникальностью которой является интеграция нескольких источников данных (OpenStreetMap, Microsoft Buildings, Google Open Buildings), регулярный выпуск обновлений и стандартизированная структура хранения данных. Несмотря на существенные достоинства, проблемой остается отсутствие данных о высоте и этажности для значительной доли зданий, особенно для территории России. На решение этой проблемы направлено наше исследование. В качестве предикторов для восстановления этажности использован набор морфометрических признаков соответствующих зданиям полигонов: его площадь, его компактность, пропорции минимального по площади ограничивающего прямоугольника и отношение площади здания к площади этого прямоугольника. Кроме того, были использованы характеристики окружающей территории, а именно локальные климатические зоны (ЛКЗ) из базы данных WUDAPT, которые являются уникальными регионами с однородными по структуре, покрову и материалом поверхности. В качестве дополнительных характеристик также были использованы назначение здания и его класс из базы данных Overture Maps. Модель восстановления этажности реализована методом регрессионных деревьев с применением градиентного бустинга. Техническая реализация выполнена с помощью библиотеки CatBoost. Кроме того, исследована возможность использования в качестве дополнительного предиктора оценки высоты здания цифровая модель поверхности (ЦМП) высокого разрешения. В качестве последней использована ЦМП ArcticDEM с разрешением в 2 м и покрывающая территорию севернее 60 градусов с.ш. На основе

информации о дорожной сети, хранящейся в Overture Maps, была создана цифровая модель рельефа (ЦМР), характеризующая высоту поверхности без учета зданий и деревьев. Путем нахождения разности исходного набора данных и рассчитанной ЦМР были получены оценки высоты зданий, которые использовались в качестве одного из признаков в модели машинного обучения. Обучение модели проводилось как на уровне всей территории Санкт-Петербурга, так и отдельно для разных типов ЛКЗ. Такой подход позволил оценить влияние морфологических особенностей территории на качество предсказаний и значимость признаков. В ходе обучения на тренировочных выборках модель демонстрировала высокую точность и достигала значений 0,94, что указывает на хорошее согласование модели с исходными данными. На валидационных выборках точность была ниже и варьировалась от 0,68 до 0,79. Анализ важности признаков в построении модели показал, что их значимость существенно меняется в зависимости от типа ЛКЗ. Это подтверждает наличие выраженных морфологических особенностей в пределах конкретной зоны и подчеркивает необходимость индивидуального подхода к построению модели для различных территорий. Так, для ЛКЗ 10, соответствующей промышленными территориям, ключевым оказался высота, полученная по ArcticDEM, а для открытой малоэтажной застройки (ЛКЗ 6) – класс здания. Сравнение моделей, обученных на разных масштабах (в пределах городских районов и всего города), также выявило различия в распределении признаков. Это указывает на влияние масштаба анализа на итоговый результат и точность модели. Таким образом, модель градиентного бустинга на основе данных из базы данных Overture Maps продемонстрировала высокую эффективность при восстановлении высот зданий, имеющих различные характеристики. В дальнейшем повышение точности модели предполагается путем расширения используемых признаков и внедрения новых источников данных.

92. Intercomparison of machine learning approaches for identifying hail from basic weather parameters

***Blinov P.D.** (1), Chernokulsky A.V. (2), Krinitsky M.A. (3, 4), Bugrimov S.A. (5)
(1) National Research University - Higher School of Economics, Moscow, Russia, (2) A.M. Obukhov Institute of Atmospheric Physics RAS, Moscow, Russia, (3) P.P. Shirshov Institute of Oceanology RAS, Moscow, Russia (4) Moscow Institute of Physics and Technology, Russia (5) Lomonosov Moscow State University, Moscow, Russia*

This work presents an integrated approach to hail diagnosis using ERA5 reanalysis data and Russian ground observations. We investigate the efficacy of three distinct methodologies: a Convolutional Neural Network (CNN), a Gradient Boosting on Trees (CatBoost) model, and a traditional threshold approach based on the composite WMAXSHEAR index. Interpretability analysis was conducted using SHAP (SHapley Additive Explanations) and reparameterization techniques. A comparative study of the models' performance was carried out. The practical applicability of the proposed methods is further illustrated through a real-case example.

55. Enhancing the Quality of Kp Index Machine Learning Forecasting Using Higher-Frequency Data and Feature Transformations

***I.M.Gadzhiev** (1), I.N.Myagkova (2), O.G.Barinov (2), S.A.Dolenko (2)
(1) Physical Department, M.V.Lomonosov Moscow State University, Moscow, Russia (2) D.V.Skobel'tsyn Institute of Nuclear Physics, M.V.Lomonosov Moscow State University,*

Moscow, Russia

In this study, we investigate the problem of increasing the quality of existing models for Kp index forecasting up to 24 hours ahead with hourly step. We show that one way to do so is to incorporate 5-minute frequency data on the parameters of the interplanetary magnetic field and solar wind into the training data. We also estimate the effect of adding feature transformations such as adding time-series differences. Forecasting the Kp index is of great practical importance, since strong geomagnetic disturbances lead to undesirable effects, such as the occurrence of geomagnetically induced currents (the strength of which during magnetic storms can reach tens of amperes) in long conductors with low resistance - communication and power lines, pipelines, railways; failures in radio communication systems and satellite navigation systems. Given the 3-hour frequency of the Kp-index, the task is formulated as forecasting next 8 values of the Kp index every hour. We use gradient boosting and perceptron type neural networks, which showed best performance in this task in our previous studies. Previously [1] we used only hourly frequency data available from the ACE Science Center [2] and from other sources (e.g. Dst index from [3]). This is a common approach, because it is relatively simple, it looks consistent, and it does not require a lot of computational resources. However, our analysis, as well as the underlying physics of geomagnetic processes, suggested that higher-frequency data (especially for the Bz component) could serve as a good predictor for geomagnetic disturbances. The ACE Science Center provides historical data on the IMF and SW parameters with a 5-minute frequency (updated daily) [ссылка], and the National Oceanic and Atmospheric Administration provides real-time 1-minute frequency stream for some of the parameters [2], so it is possible to use 5-minute frequency data for real-time forecasting. We incorporate the 5-min frequency data on the IMF and SW into the training data and show that this improves the quality of the forecasting. We also consider and evaluate feature transformation techniques like differencing times series. Additionally, we provide evaluation results on the real-time data from NOAA. Practical conclusions are drawn on which combination works best for the perceptron and gradient boosting. We carry out model interpretation and compare the results with our previous findings. This study has been performed within the framework of the state assignment of M.V.Lomonosov Moscow State University.

[1] Gadzhiev I.M., Barinov O.G., Dolenko S.A., Myagkova I.N., Comparative Analysis of the Procedures to Forecast the Kp Geomagnetic Index by Machine Learning, Moscow University Physics Bulletin 79(2), P. 854-865, <http://dx.doi.org/10.3103/S002713492470231X> [<http://dx.doi.org/10.3103/S002713492470231X>]

[2] ACE Science Center <https://izw1.caltech.edu/ACE/ASC/> [<https://izw1.caltech.edu/ACE/ASC/>]

[3] World Data Center for Geomagnetism, Kyoto <https://wdc.kugi.kyoto-u.ac.jp/> [<https://wdc.kugi.kyoto-u.ac.jp/>]

[4] National Oceanic and Atmospheric Administration, ACE Real Time Solar Wind, <https://www.swpc.noaa.gov/products/ace-real-time-solar-wind> [<https://www.swpc.noaa.gov/products/ace-real-time-solar-wind>]

96. Машинное обучение для статистической детализации характеристик пространственного распределения осадков в Московском регионе

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

(1) Московский государственный университет имени М.В. Ломоносова, Научно-

исследовательский вычислительный центр, (2) Институт физики атмосферы имени А.М. Обухова РАН, (3) Гидрометеорологический научно-исследовательский центр Российской Федерации, (4) Институт океанологии им. П.П. Ширшова РАН, Нахимовский проспект, (5) Московский физико-технический институт

В связи с наблюдаемыми изменениями климата учащающиеся экстремальные осадки оказывают влияние на различные регионы, включая Северную Евразию, и особенно разрушительны в крупных городах. Глобальные климатические изменения обычно оцениваются путем уменьшения масштаба крупномасштабных климатических характеристик, которые лучше разрешаются в климатических моделях, до мелкомасштабных переменных, которые не могут быть явно разрешены на сетке климатической модели. В предыдущих исследованиях изучались подходы машинного обучения для уменьшения масштаба осадков в нескольких регионах, но территория Московской агломерации, крупнейшей в России и Европе, осталась нетронутой, а ежегодный риск наводнений из-за экстремальных осадков в этом регионе остается очень высоким. В данной работе методы машинного обучения в статистической детализации используются для получения характеристик пространственного распределения осадков (максимального значения, моментов распределения, квантилей) из полей гидродинамического моделирования с низким разрешением (реанализа ERA5). Для обучения статистических моделей (гребневой регрессии, градиентного бустинга, случайного леса) используются данные за 33 года наблюдений на метеостанциях Московского региона. В качестве признакового описания выступают физически обоснованные крупномасштабные предикторы осадков, рассчитанные с использованием данных реанализа ERA5. В их число входят как простые величины в толще атмосферы (температура, влажность и др.), так и сложные комплексные характеристики (CAPE, лапласиан приземного давления и др.). Наилучший результат по выбранным метрикам качества (RMSE, R2) достигается с использованием модели градиентного бустинга, при следующей конфигурации набора предикторов: осреднённые за сутки, стандартизованные, включая модельный предиктор осадков (среднее значение по площади). В работе показан рейтинг важности признаков крупномасштабных атмосферных предикторов для территории Московской области для различных конфигураций экспериментов и применённых моделей.

Работа выполнена при поддержке Некоммерческого фонда содействия развитию науки и образования «ИНТЕЛЛЕКТ».

16

Section 3. Machine Learning in Natural Sciences

37. Neutron spectrum unfolding with deep learning models for tabular data

Chizhov Konstantin Alekseevich (1,2), Bely Artyom Alekseevich (2)

(1) Joint Institute for Nuclear Research, Laboratory of Information Technologies named after M.G. Meshcheryakov, (2) University "Dubna"

Estimation of the effective dose and unfolding the spectrum of neutron radiation at

nuclear power facilities and charged particle accelerators is complicated by the lack of direct methods for detecting neutrons and the need to register secondary particles. The main difficulties are related to the wide energy range of neutrons from 1 meV to several hundred MeV, complex dependence of the neutron interaction cross section on energy. One of the main devices used for neutron spectrometry is the Bonner multi-sphere spectrometer (BSS). The measurement results and the desired spectrum, discretized on the energy grid (or decomposed into basis functions) are tabular data. However, due to the limited set of moderator spheres and correlations in its response functions, the number of input features is limited. In this paper, it is proposed to transform the original scalar continuous features into vectors. And then unfold the spectra for the transformed features using deep learning models included in the Mambular framework: a sequential model from Mamba architecture blocks based on autoregressive state-space models; a model using transform encoders (FT-Transformer), a multilayer perceptron (MLP) model, and batch ensembling for MLP (TabM). The results are compared with the spectra unfolded using the automated machine learning (AutoML) framework implementing an ensemble of linear and decision-tree based (LightGBM, CatBoost, and Random Forest) regression models. For training and validation of the models, a set of 5×10^5 synthetic spectra was generated, modeled as a superposition of four weighted components describing the spectra of thermal, epithermal, fast, and high-energy neutrons. A comparison was made with calculated and measured spectra from the IAEA compendium database, 251 spectra. The uncertainty of the spectra unfolding was estimated using the Monte Carlo method, in which random perturbations were introduced into the input data. The model was trained on the JINR Multifunctional Information and Computing Complex.

The research was carried out within the framework of the state assignment of the Ministry of Science and Higher Education of the Russian Federation (topic No. 124112200072-2).

49. Применение концепции переноса обучения для градиентного бустинга при решении обратных задач разведочной геофизики

М.К.Шалеев (1), И.В.Исаев (2), И.Е.Оборнев (2), С.А.Доленко (2)

(1) Физический факультет Московского государственного университета имени М.В. Ломоносова, Москва, (2) Научно-исследовательский институт ядерной физики им. Д.В.Скобельцына, Москва

Данное исследование посвящено изучению эффективности методов переноса обучения, применяемых к моделям градиентного бустинга, для решения обратных задач разведочной геофизики. Основная проблема, решаемая в работе - необходимость наличия большого объема обучающих данных при адаптации моделей к новым геологическим структурам или наборам данных с различной степенью сложности (например, при переходе между разрезами с фиксированными и переменными свойствами слоев). В отличие от предыдущих работ, использовавших нейронные сети, в данной работе рассматриваются алгоритмы градиентного бустинга. Исследуются специфические стратегии переноса обучения, адаптированные для ансамблевых методов: 1. Стандартное дообучение: модель, предварительно обученная на исходном наборе данных, продолжает обучение на целевом наборе. 2. Перенос с обогащением признаков: модель обучается "с нуля" на целевом наборе, но входные признаки дополняются ещё одним, содержащим ответы модели, предварительно обученной на исходном наборе данных. 3. Перенос с усечением модели: после предобучения на исходных данных определяется

итерация с минимальной ошибкой на валидационном поднаборе целевого набора. Далее ансамбль усекается до этой итерации, и полученная модель используется для инициализации дообучения. 4. Перенос с адаптивным ансамблированием: на целевом наборе данных обучается новая модель “с нуля”. Итоговое предсказание формируется как взвешенная сумма предсказаний модели-донора и целевой модели. Весовой коэффициент адаптивно подбирается с помощью алгоритма оптимизации. Эксперименты проводятся в рамках геофизической постановки, включающей задачи гравиметрии, магнитометрии и магнитотеллурического зондирования. Используются три синтетических набора данных (“Лёгкий”, “Средний”, “Сложный”), различающихся характером изменения параметров среды. Производится сравнение эффективности стратегий переноса обучения и традиционного обучения. Оценивается точность восстановления глубин границ геологических слоев и вычислительная эффективность при различных объемах обучающей выборки. Результаты показывают, что применение методов переноса обучения для градиентного бустинга позволяет существенно сократить необходимый размер обучающей выборки, сохраняя или даже улучшая точность решения по сравнению с моделями бустинга без переноса. Это подтверждает, что перенос обучения позволяет существенно повысить эффективность и расширить возможности применения градиентного бустинга при решении обратных задач разведочной геофизики в условиях ограниченных данных. Исследование выполнено за счёт гранта Российского научного фонда № 24-11-00266.

48. Анализ стратегий обучения FBPINNs

*Алимов Павел Геннадьевич (1), Гориховский Вячеслав Игоревич (1)
(1) Санкт-Петербургский государственный университет*

Физически-информированные нейронные сети (PINNs) активно применяются для решения задач механики, но сталкиваются с трудностями в сложных областях. Метод Finite-Basis PINNs (FBPINNs), использующий разбиение области на подобласти с локальными сетями, предлагается как развитие PINN для повышения точности и сходимости. Однако FBPINNs на данный момент является незрелым подходом: отсутствуют надежные реализации и глубокое понимание факторов их устойчивости. Целью данной работы является исследование влияния различных стратегий обучения на стабильность и точность FBPINNs. В работе систематически изучено влияние различных стратегий взвешивания компонент функции потерь и методов обучения локальных сетей. На основе этого анализа разработана стабильная реализация FBPINNs, обеспечивающая лёгкое внедрение новых уравнений для постановки экспериментов. Работа завершается формулировкой практических рекомендаций по эффективному конфигурированию и применению FBPINNs для задач механики, способствуя развитию этого перспективного подхода.

62. The creation of reasonable robot control behavior in the form of executable code

*Skorokhodov Maksim (1,2), Latalin Vladislav (1), Rybka Roman (1), Sboev Alexander (1,2)
(1) National Research Center “Kurchatov Institute”, Moscow, Russia (2) National Research Nuclear University MEPHI (Moscow Engineering Physics Institute), Moscow, Russia*

A system for robot control behavior based on Large Language Models using Prompt

Engineering methods and answer generation in the form of executable program blocks is presented. The method employed is to convert human instructions into code snippets taking into account information from the robot's world. The resulting system was evaluated using the set of 60 hand-written instructions. This set is divided into three parts: "direct" instructions relate to objects actually present in the environment, "abstract" instructions describe general tasks without specific details, and "out-of-space" instructions refer to actions beyond the robot's capabilities or mention objects not shown in its environment. Modern Large Language Models, such as Deepseek, Qwen, Mistral were tested with the different types of training and numbers of parameters. The results show these models are capable of interpreting robot control instructions as code snippets and delegating complex interaction logic within these fragments. Generating control instructions directly as executable code increases the flexibility of performing robot control tasks and minimizes the risk of hallucinations, for example in carrying out precise arithmetic operations.

63. Применение сетей Колмогорова-Арнольда для решения обратной задачи спектроскопии при создании мультимодального наносенсора ионов металлов на основе углеродных точек

Г.А.Куприянов(1, 2), И.В.Исаев(2), К.А.Лаптинский(1, 2), С.А.Доленко(1, 2), Т.А.Доленко(1)

(1) Физический факультет Московского государственного университета имени М.В. Ломоносова. (2) Научно-исследовательский институт ядерной физики им. Д.В.Скобельцына

Разработка мультимодального флуоресцентного наносенсора на основе углеродных точек (УТ) направлена на создание перспективного инструмента для анализа жидких сред и количественной оценки примесей металлов [1]. Одним из ключевых этапов является создание модели, оценивающей концентрацию катионов металлов по спектрам флуоресценции УТ, введенных в исследуемый образец. В качестве таких моделей могут использоваться перцептроны и сверточные нейронные сети. Благодаря универсальности, нелинейности и наличию эффективных алгоритмов обучения, нейросетевые модели приобрели широкую популярность в исследованиях для выявления сложных закономерностей в больших наборах данных. Нейронные сети Колмогорова-Арнольда (СКА) представляют собой подход, альтернативный стандартным нейросетевым моделям. В отличие от поиска матричных коэффициентов (весов), обучение СКА заключается в поиске непрерывных функций активации [2]. Это позволяет СКА проявлять большую нелинейность по сравнению со стандартными нейросетевыми подходами, что в ряде случаев может обеспечить повышенную точность аппроксимации [2, 3]. Архитектурно СКА основана на суммах и суперпозиции скалярных функций одной переменной. Такая структура позволяет осуществлять наглядную визуализацию модели, что упрощает интерпретацию ее работы [2]. В данном исследовании на основе экспериментальных данных [1] были построены модели СКА для одновременного определения концентраций шести катионов металлов и одного аниона по двумерным картам флуоресценции растворов с УТ. СКА продемонстрировала точность определения концентрации ионов, сопоставимую с референсными методами (многослойный перцептрон, случайный лес, градиентный бустинг, метод группового учета аргументов). Был предложен и применен усовершенствованный метод визуальной интерпретации СКА, учитывающий

существенные взаимосвязи между входными признаками. Полученные результаты позволили выявить ключевые закономерности, обнаруженные и используемые СКА при определении концентраций ионов.

Исследование выполнено за счет гранта Российского научного фонда № 22-12-00138, <https://rscf.ru/project/22-12-00138/> [<https://rscf.ru/project/22-12-00138/>]. Г.А.Куприянов является стипендиатом Фонда развития теоретической физики и математики «Базис» и Некоммерческого Фонда развития науки и образования «Интеллект».

[1] Sarmanova, O.E., Laptinskiy, K.A., Burikov, S.A., Chugreeva, G.N., Dolenko, T.A.: Implementing neural network approach to create carbon-based optical nanosensor of heavy metal ions in liquid media. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 286, 122003 (2023).

[2] Liu, Z. et al.: KAN: Kolmogorov-Arnold Networks. arXiv:2404.19756v4 (2024).

[3] Wang, Y. et al.: Kolmogorov–Arnold-Informed neural network: A physics-informed deep learning framework for solving PDEs based on Kolmogorov–Arnold Networks. arXiv:2406.11045v1 (2024).

64. Optimization of IRT-T research reactor fuel loading pattern by genetic algorithm

*N.V. Smolnikov (1), D.V. Pasko (1), M.N. Anikin (1), I.I. Lebedev (1), A.G. Naimushin (1)
National Research Tomsk Polytechnic University*

Research Nuclear Reactors (RNR) are powerful sources of neutron and gamma radiation with large number of beams and experimental channels that can be used to produce technical and medical isotopes, conduct researches in solid physics, neutron scattering and other fields. Most of RNRs operate in Partial Refueling Mode, where only burnt fuel assemblies are replaced during refueling. This leads to power density increase in a localized sectors (fuel assemblies), changes in neutron intensity in experimental channels and variations in the fuel cycle length. In this research we do propose using genetic algorithm to optimize the power density distribution and fuel cycle length via fuel loading pattern optimization fuel loading pattern.

59. Камни: Коллективная игра между агентами разнообразных типов, разработанная для изучения взаимодействия человека и искусственного интеллекта в многоагентной среде

*Чернов К.Н.(1), Исаев И.В.(2), Гуськов А.А.(1,2), Куприянов Г.А.(1,2), Макаров А.С.(1), Мушина А.С.(1,2), Самсонович А.В.(3), Доленко С.А.(2)
(1) Физический факультет, Московский государственный университет имени М.В. Ломоносова; (2) Научно-исследовательский институт ядерной физики имени Д.В. Скобельцына, Московский государственный университет имени М.В.Ломоносова; (3) Национальный исследовательский ядерный университет «МИФИ»*

В работе представлена агент-ориентированная игровая система “Камни”, предназначенная для моделирования и анализа взаимодействий между игроками-людьми, алгоритмическими агентами, биологически вдохновленными агентами когнитивной архитектуры (ВІСА) и нейро-агентами в кооперативной игре. Система включает в себя несколько типов агентов: агент-человек, случайные агенты, «двойки» (агенты, взаимодействующие в паре), «тройки» (агенты, работающие по

трое), ВІСА-агенты с адаптивной когнитивной архитектурой и нейро-агенты, которые обучаются стратегии игры по данным уже сыгранных игр. В игре агенты делают ход от одного из нескольких “камней” к другому; камень убирается, если после хода ровно два агента (не меньше и не больше) находятся у этого камня. Кооперативная цель агентов – убрать все камни, кроме двух, за минимальное количество ходов. Игрок-человек может взаимодействовать со средой через Telegram-бота, что позволяет участвовать в игре в режиме реального времени. Такая архитектура системы обеспечивает динамичный игровой процесс, в котором агенты с различными стратегиями соревнуются или сотрудничают, в то время как на процессы принятия решений агентами ВІСА влияют действия человека и других агентов. Разработанная платформа является исследовательским инструментом для сбора и анализа поведенческих данных различных типов агентов в смешанных средах с использованием искусственного интеллекта. Сравнивая эффективность и стратегии адаптации ВІСА-агентов, мы стремимся выявить закономерности, возникающие в сценариях сотрудничества и конкуренции. Интеграция с Telegram обеспечивает доступность, а протоколируемые взаимодействия дают представление о том, как модели ВІСА реагируют на непредсказуемость человека. Кроме того, мы анализируем протокольные записи (логи) игрового процесса, чтобы совершенствовать ВІСА-агентов с помощью генетических алгоритмов, оптимизируя их когнитивные и поведенческие параметры в разных поколениях. Этот подход способствует разработке более надежных и совместимых с человеком систем искусственного интеллекта.

70. Построение нейродифференциальных уравнений с применением методов обратных задач динамики

Шорохов С.Г.

Российский университет дружбы народов им. П.Лумумбы

В обратных задачах динамики (inverse problems of dynamics) необходимо определить неизвестные силы или параметры динамической системы по известным характеристикам ее движения, в том числе по известной траектории движения. Математически, решение обратных задач динамики часто сводится к построению правых частей системы обыкновенных дифференциальных уравнений по заданному интегральному многообразию (метод Еругина). Задача построения системы дифференциальных уравнений может иметь множество решений, поэтому в методе Еругина строится вся совокупность систем дифференциальных уравнений, которые удовлетворяют заданным условиям, и за счет выбора произвольных функций, входящих в решение, можно, в частности, добиться того, чтобы движение по интегральному многообразию было устойчиво по Ляпунову. В нейродифференциальных уравнениях (neural ODEs) правые части системы обыкновенных дифференциальных уравнений аппроксимируются нейронной сетью, а в качестве обучающего набора данных может выступать набор наблюдений за движением динамической системы в различные моменты времени. В докладе обсуждается применение методов обратных задач динамики (метода Еругина) при построении и обучении нейродифференциальных уравнений. В качестве примера рассматривается плоское движение по заданной траектории – овалу Кассини.

74. СОЗДАНИЕ ДИНАМИЧЕСКОГО КОГНОВИЗОРА – РАСПОЗНАВАНИЕ

КОГНИТИВНЫХ СОСТОЯНИЙ С ПОМОЩЬЮ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ

Макаров Александр Сергеевич (1), Гаджиев Исмаил Маратович (2), Доленко Сергей Анатольевич (2)

(1) Московский государственный университет имени М.В.Ломоносова (МГУ), (2) Научно-исследовательский институт ядерной физики имени Д.В. Скобельцына МГУ

Данная работа является продолжением исследования, в котором рассматривалась возможность разработки «Когновизора» с помощью методов машинного обучения и анализа главных компонент. Когновизор – это устройство, которое на основе анализа нейрофизиологических данных о мозговой активности распознает и визуализирует когнитивные состояния и переходы между ними. В данной работе в качестве методов машинного обучения рассматривались свёрточные нейронные сети, а в качестве когнитивных состояний были выбраны 6 типов мышления с постепенным изменением свойств пространственно/образного и вербально/логического типа и состояние покоя головного мозга (как базовый уровень работы процессов сознания). Для получения нейрофизиологических данных использовался метод функциональной магнитно-резонансной томографии (фМРТ). Было рассмотрено три типа задач классификации: 1) Бинарная классификация – состояние покоя и любое активное состояние; 2) Классификация по трем классам – состояние покоя, пространственный и вербальный типы мышления; 3) Многоклассовая классификация – состояние покоя, 2 пространственных и 4 вербальных типа мышления. Поскольку данные для этой задачи имеют слишком большую размерность (4D-объемы $\sim 91 \times 109 \times 91 \times 3620$), то на первом этапе стояла задача подобрать оптимальную для имеющихся вычислительных ресурсов архитектуру сверток в сверточной нейронной сети. Было рассмотрено три вида сверток: 1) Модель свертки, которая принимает на вход несколько срезов 3D изображения и интерпретирует их как срезы 2D изображения, после чего к каждому из них применяется 2D свертка; 2) Разделяемая 3D свертка; 3) Полная 3D свертка. По результатам работы были сделаны выводы о способе выбора оптимального типа свертки и о том, какую свертку использовать в дальнейших исследованиях.

78. Analysis of PINN Training Strategies for Heat Conduction Problems

Tarasov A.A.(1), Stepanova M.M.(1), Orlov S.E.(1)

(1) Saint-Petersburg State University

In recent years, physics-informed neural networks (PINNs) have been increasingly employed to solve applied problems in mathematical physics. PINNs incorporate equations and boundary conditions directly into the model architecture through automatic differentiation. This approach allows solving differential equations without explicit spatiotemporal discretization, proving especially effective for multidimensional and nonstationary problems. However, despite their versatility, conventional PINNs have several limitations that complicate its training. This study analyzes common issues and explores methods to overcome them using heat conduction problems as examples. In particular, it addresses the suppression of spectral bias in neural network solutions via the Random Fourier Features (RFF) method, dynamic balancing of loss function weights, the issue of violating temporal causality in nonstationary problems, various sampling

strategies for generating training data, and optimal selection of neural network architecture and hyperparameters. It is demonstrated that implementing these strategies with careful consideration of problem-specific features significantly accelerates convergence of the numerical solution toward the exact solution in various cases. Practical recommendations on selecting and combining strategies are provided for researchers dealing with heat conduction and related differential models.

66. Comparison of Data Generation Methods for Spectral Analysis Using Variational Autoencoders

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) Physical Department, M.V.Lomonosov Moscow State University

This study explores application of data representativity enhancement using variational autoencoders (VAEs) to the inverse problem of Raman spectroscopy of multicomponent aqueous solutions of inorganic salts. By extending our earlier work on optical absorption spectroscopy to Raman scattering, we assess the transferability of VAE-based dataset expansion methods used to solve inverse problems in spectroscopy across spectroscopic techniques. The objective of the considered spectroscopic studies is to determine the concentrations of various ions in multi-component aqueous solutions based on spectral information. Unlike other spectroscopic techniques such as infrared or optical absorption spectroscopy, Raman spectroscopy provides more detailed information about the vibrational states of molecules, making it particularly sensitive to changes in the ionic composition of a solution. Raman spectra are high-dimensional, correlated, and nonlinearly dependent on the sample composition, what complicates their interpretation. To address this challenge, machine learning methods, particularly regression-based neural networks, can be employed. A critical factor influencing model is the representativity of the training dataset. We attempt to expand the training dataset by generating synthetic spectra using VAEs and investigate the potential of this approach to improve the representativity of the training data, which may in turn lead to a reduction in concentration determination errors when solving the inverse problem. While conditioned VAEs (cVAEs) offer a direct way to incorporate target analyte concentrations, we also examine alternative strategies using standard VAEs paired with auxiliary regression models to assign target concentrations to generated spectra. While further validation is needed, the considered approaches may provide a basis for developing synthetic data generation methods that could potentially better capture the physical characteristics of Raman spectra.

This study has been performed at the expense of the grant of the Russian Science Foundation no. 24-11-00266.

87. Применение переноса обучения сверточной нейронной сети для повышения точности решения обратной задачи фотолюминесцентной наносенсорике

Г. Чугреева(1), К. Лаптинский(1,2), Т. Доленко(1)
1- Физический факультет, МГУ им. М. В. Ломоносова, 2- НИИЯФ им.Д.В.Скобелевца, МГУ им. М. В. Ломоносова

Углеродные точки (УТ) – класс углеродных наночастиц, обладающих стабильной

интенсивной фотолюминесценцией (ФЛ), зависящей от условий синтеза и крайне чувствительной к изменению параметров окружения [1]. Такие свойства УТ открывают широкие возможности их использования в качестве наносенсоров параметров среды [2]. В настоящее время на основе УТ активно разрабатываются наносенсоры для одновременного определения концентрации ряда ионов, растворенных в воде и тушащих ФЛ углеродных точек [1]. Нейронные сети (НС), будучи универсальными аппроксиматорами, способны успешно решать многопараметрические обратные задачи фотолюминесцентной спектроскопии, в частности, с хорошей точностью одновременно определять концентрацию ионов в водной среде по спектрам ФЛ УТ [3]. Однако, применение НС для решения указанной задачи требует получения больших баз спектральных данных, т.е. использования дорогостоящей аппаратуры, привлечения высококвалифицированных специалистов, длительного времени сбора примеров и др. Одним из способов преодоления приведенных сложностей является применение переноса обучения НС. В данном исследовании показано, как с помощью переноса обучения сверточной нейронной сети (СНС) удалось не только повысить точность определения концентрации ионов в среде по спектрам ФЛ УТ, но и уменьшить временные затраты, необходимые для обучения модели. Экспериментально были получены две базы спектров ФЛ УТ в водных растворах с 6 ионами (Ni^{2+} , Cu^{2+} , Co^{2+} , Al^{3+} , Cr^{3+} , NO_3^-) и в водных растворах с 7 ионами (те же ионы, и добавлены ионы Pb^{2+}). На первом этапе были обучены СНС(6) и СНС(7) для определения типа и концентрации каждого из 6 ионов и 7 ионов по спектрам ФЛ УТ в водной среде, соответственно. На следующем этапе СНС(6), обученная на базе данных, содержащей спектры ФЛ водных растворов УТ с 6 ионами, была дообучена на базе спектров ФЛ УТ в растворах с 7 ионами. Точность определения концентрации исследованных ионов дообученной СНС оказалась выше в 2-3 раза по сравнению с точностью определения концентрации ионов моделью СНС(7), обученной с нуля на базе спектров растворов УТ с 7 ионами.

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/>]. Работа Г. Чугреевой поддержана фондом теоретической физики и математики «Базис» (договор № 23-2-2-22-1).

[1] Vervald A.M., Laptinskiy K.A., Chugreeva G.N., Burikov S.A., Dolenko T.A.. (2023) 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).

[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).

[3] Chugreeva, G. N., Laptinskiy, K. A., Plastinin, I. V., Sarmanova, O. E., & Dolenko, T. A. (2024). Development of a Multimodal Photoluminescent Carbon Nanosensor for Metal Ions in Water Using Artificial Neural Networks. *Moscow University Physics Bulletin*, 79(S2), S844–S853.

86. Probabilistic Spiking Neural Network with Correlation-based Memristive Synaptic Updates

Dmitry Kunitsyn (1,2), Alexander Sboev (1,2), Yury Davydov (1), Danila Vlasov (1), Alexey Serenko (1), and Roman Rybka (1,2)

(1) National Research Centre “Kurchatov Institute”, Moscow, Russia (2) National Research

Spiking Neural Networks (SNNs) are a biologically inspired class of neural models that encode information as discrete temporal impulses (spikes). These networks exhibit low latency and reduced power consumption, particularly when implemented on neuromorphic memristive hardware. However, efficient training of SNNs remains an active research area, as architectures trainable without error backpropagation could yield significant energy savings during both inference and training. Local plasticity rules offer a promising alternative, since they can be implemented in memristive devices. In this work, we introduce a lightweight, single-layer, correlation-based spiking neural network with probabilistic neurons. Owing to its compact and simple architecture, our model is potentially well suited for hardware realization. We evaluate its performance under both the classical spike-timing-dependent plasticity (STDP) and an experimentally-approximated nanocomposite memristive plasticity. Using the proposed correlation-based direct reinforcement training method, our network achieves high F1-scores across several tabular and image classification benchmarks for all considered plasticity models. In particular, the proposed approach yields 96% on the Fisher Iris dataset, 94% on the Wisconsin Breast Cancer dataset, and 89% on the Scikit-Learn Digits dataset when trained with regular STDP, and achieves comparable results on the nanocomposite plasticity. This suggests that the correlation-based SNN with probabilistic neurons is suitable for implementation in memristive hardware.

90. Модель машинного обучения для прогнозирования вентиляторных порогов

Минкин А.С. (1)

(1)Институт прикладной математики им. М.В. Келдыша РАН, 125047, Москва

Одной из задач, связанных с определением состояния легочной и сердечно-сосудистой систем, является определение индивидуальных значений вентиляционных порогов (ВП), которые могут быть найдены по результатам анализа данных газообмена при выполнении кардиореспираторного нагрузочного тестирования. Из-за высокого уровня шума в регистрируемых данных газообмена прогнозирование ВП часто основывается на визуальном осмотре графиков. Результаты этих оценок, полученных с использованием популярного метода V-slope [1], активно применяемого для прогнозирования вентиляторно-анаэробного порога, существенно зависят от используемого эргоспирометра и методики предварительной обработки данных. Поэтому процедуру оценки ВП можно отнести к слабо формализованным задачам, для решения которых обычно используются методы машинного обучения (ML). Цель исследования — разработка ML-модели, позволяющей сделать экспресс оценку ВП по данным эргоспирометрии при выполнении кардиореспираторного нагрузочного тестирования. Для этого используется набор данных тренировок на беговой дорожке здоровых добровольцев (мужчин и женщин) — участников эксперимента «Сириус» с длительной изоляцией, выполняющих тест со ступенчато-возрастающей физической нагрузкой до предела переносимости. Для повышения эффективности оценки ВП сделан предварительный подбор признаков по данным газообмена с точки зрения значимости их влияния на результат классификации. Рассматриваются такие данные газообмена как потребление O₂, выделение CO₂, минутная вентиляция и т. д. в сочетании с данными о частоте сердечных

сокращений. Для оценки значимости признаков с точки зрения точности классификации используется обучение деревьев решений, а для получения итогового прогноза ВП с использованием выделенных признаков — случайный лес. Указанные модели являются интерпретируемыми [2], так как дают заключения о влиянии различных признаков, используемых в качестве сигнала достижения ВП. Таким образом, использование предварительно обученных ML-моделей для предсказания ВП эффективнее и быстрее, чем визуальный осмотр исходных данных экспертом, так как такой подход требует меньшего числа операций предобработки данных, а также позволяет объяснить результат классификации на основе анализа значимых признаков, отсортированных по их важности.

[1] Beaver WL, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. *J Appl Physiol*, Vol. 60, No. 6, 1985, P. 2020-2027. DOI: 10.1152/jappl.1986.60.6.2020.

[2] Mishra, Pradeepa. *Practical Explainable AI Using Python: Artificial Intelligence Model Explanations Using Python-based Libraries, Extensions, and Frameworks*. Apress Berkeley, CA, 2022. DOI: 10.1007/978-1-4842-7158-2.

85. Finding optimal carbon dots synthesis parameters for quantitative analysis of components in multi-component aqueous solutions using machine learning

Guskov A.A. (1, 2), Isaev I.V. (2), Laptinskiy K.A. (2), Dolenko T.A. (1, 2), Dolenko S.A. (2) (1) Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia, (2) D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia

Optical nanosensors based on carbon dots (CD) introduced into the object of the study are widely used for analyzing the content of multicomponent liquid media. Their applicability stems from the high sensitivity of CD photoluminescence to changes in medium parameters, such as pH and solution temperature, type and concentration of dissolved substances. In addition, such sensors offer stable optical properties, biocompatibility, low production costs, and enable both rapid and remote diagnostics of the investigated objects. At the same time, the physicochemical properties of CD strongly depend on the synthesis method and the precursors used. Therefore, CD synthesized under different conditions can show significantly different photoluminescence responses even when placed in identical solutions. This study focuses on finding the optimal CD synthesis parameters that ensure high-precision determination of the concentration of specific heavy metal salts in aqueous solutions. Creation of nanoparticles with such features represents an urgent scientific and practical challenge. In the physical experiment, 74 types of CD were prepared by hydrothermal synthesis from citric acid and ethylenediamine under a wide range of precursor ratios, temperatures and reaction times. Then each type of CD was placed in aqueous solutions of $\text{Co}(\text{NO}_3)_2$ and $\text{Cu}(\text{NO}_3)_2$ with salt concentrations ranging from 0 to 6 mM in increments of 0.67 mM. As the result, 100 samples were prepared for each type of CD. The excitation-emission spectra of carbon dots fluorescence were registered for all obtained aqueous solutions of CD and salts. To solve the stated task, it is reasonable to apply machine learning methods that are capable of revealing hidden relationships in multiparametric systems. In this study, an artificial neural network based on a multilayer perceptron architecture was used to approximate the dependence of the target variable – the error in determining the concentrations of the studied heavy metal salts – on the CD synthesis parameters. The target variable values for training the approximating model

were obtained as a result of solving the inverse problem of determining the concentrations of heavy metal salts from the excitation-emission spectra of fluorescence of CD by machine learning algorithms. This two-step approach may allow one to find carbon dot synthesis parameters that yield the desired accuracy in determining the concentrations of various components in multicomponent solutions.

This study has been performed at the expense of the grant of the Russian Science Foundation no. 22-12-00138-P.

91. Classifying Russian speech commands with a hardware-deployable spiking neural network transferred from an artificial neural network

Roman Rybka (1, 2), **Alexey Serenko** (1), **Alexander Naumov** (1), **Alexander Sboev** (1, 2)
(1) National research centre Kurchatov Institute, Moscow, Russia (2) National research nuclear university MEPhI, Moscow, Russia

We present a baseline accuracy of classifying audio recordings of command words in Russian from a recent dataset RuSC using a 7-layer convolutional spiking neural network of Integrate-and-Fire neurons. The network is obtained by transferring weights from a trained network of ReLU neurons of same topology, and then by adjusting neuron thresholds using a same-topology network with the ClipFloor activation function. In order to make the network prospectively deployable to neuromorphic processors, its synaptic weights are quantized to 8-bit integer. When the duration of presenting one input sample is 200 time steps of spiking network, the resulting performance is the f1-micro of 98%, which is just 1% lower than originally reported on that dataset with artificial neural networks. This result might be a starting point against which further spiking network solutions for keyword spotting in Russian could be compared.