

Кафедра интеллектуальных систем

Стипендия им. К.В. Рудакова

Осень 2024

Дедлайн 2 сентября 2024

Пожалуйста, скомпилируйте этот файл

Об утверждении Порядка назначения и выплат стипендии

В соответствии с Положением о стипендиальном обеспечении и других формах материальной поддержки студентов, аспирантов, слушателей подготовительного отделения МФТИ, утвержденным приказом ректора № 1053-1а от 29.06.2019 ПРИКАЗЫВАЮ:

1. Утвердить Порядок назначения и выплат Научной академической стипендии.
2. Заведующему канцелярией административного отдела М. А. Гусевой довести данный приказ до сведения начальников подразделений в течение 7 дней с даты регистрации.
2. Контроль за исполнением настоящего приказа возложить на директора Физтех-школы прикладной математики и информатики А. М. Райгородского.

Ректор МФТИ

Порядок назначения и выплат Научной академической стипендии Физтех-школы ПМИ обучающимся 2-4 курсов бакалавриата и 1-2 курсов магистратуры

1. Настоящее Положение регламентирует условия и порядок назначения и выплат *Научной академической стипендии им. К.В. Рудакова* (далее – стипендии).
2. Стипендия вводится с целью дополнительного поощрения обучающихся 2–4 курса бакалавриата и 1–2 курса магистратуры Кафедры интеллектуального анализа данных Физтех-школы прикладной математики и информатики (далее – ФПМИ), осуществляющих научно-исследовательскую деятельность в области прикладной математики: машинного обучения, анализа данных, и других разделов математики.
3. Сбор анкет-заявок и подтверждающие документы для участия в конкурсе на стипендию осуществляется специалистами Учебного офиса ФПМИ после выхода соответствующего объявления на информационных ресурсах ФПМИ.
4. Отбор стипендиатов производит комиссия, в которую входят представители Кафедры интеллектуального анализа данных и дирекции ФПМИ (далее – Комиссия), на основании представленных достижений в области прикладной математики один раз в семестр.
5. Состав Комиссии определяется жертвователем. Председателем комиссии является директор ФПМИ.
6. Критериями для отбора стипендиатов являются:
 - a. Вероятность своевременного выполнения научной работы высокого качества.
 - b. Ожидаемый вклад в исследуемую проблему или решение задачи.
 - c. Качество представленной заявки и научного задела.
7. Решение Комиссии о предоставлении стипендий оформляется протоколом заседания Комиссии.
8. Стипендия назначается каждый семестр сроком на один семестр не позднее 1 октября в осеннем семестре и 1 апреля в весеннем семестре.
9. Размер стипендии составляет 40 000 рублей.
10. Стипендии выплачиваются ежемесячно.
11. Выплаты стипендии осуществляются за счет средств, полученных по договору пожертвования между компанией ООО “Форексис” и МФТИ.
12. В случае если на текущий семестр средства на выплату стипендий отсутствуют, то стипендия не выплачивается.
13. Выплата стипендии прекращается при наступлении хотя бы одного из следующих обстоятельств:
 - a. отчисление обучающегося из МФТИ;
 - b. уход в отпуск (академический, по беременности и родам, по уходу за ребенком до достижения им возраста трех лет).

Семестр 5 бакалавры / Semester 5, bachelors

Шаблон свободный

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Дорогова София Юрьевна, 3 курс, ФПМИ ПМФ

Средний балл в ЛК/ GPA: 8.24

Научный руководитель (указать ученую степень) / Scientific adviser: Корчагин Сергей Андреевич, бакалавр, младший научный сотрудник ИППО РАН

Тема: / Title Создание биологических фантомов для анализа оптических свойств кожи человека на основе гиперспектральной визуализации

Аннотация: / Annotation: В статье рассматривается метод неинвазивной оценки состояния кожи с использованием гиперспектральной визуализации, позволяющей анализировать концентрации хромофоров для диагностики кожных заболеваний. Основное внимание уделено разработке моделей для расчета концентраций гемоглобина, меланина и билирубина, что актуально для оценки возрастных изменений гематом, гемангиом и меланом, а также для диагностики рака. Исследование включает изготовление биомиметических образцов кожи с известными концентрациями хромофоров. Сравнительный анализ моделей Бугера–Ламберта–Бера и Кубелки–Мунка показал высокую точность предсказания относительных и абсолютных концентраций хромофоров, что может повысить качество медицинской диагностики.

Ожидаемый результат: / Expected result: 48-я школа-конференция «Информационные технологии и системы 2024», 16-20 сентября, (ВГУ) г. Воронеж

Ранее полученные результаты: / Delivered results: Статья уже подана на участие в конференции «Информационные технологии и системы 2024»

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development:

Дополнительные ссылки: / Additional links:

Конкурсная заявка на Научную академическую стипендию им. К.В. Рудакова

Student

Nikolai Rekut, 5th semester

Scientific advisor

Alexander Beznosikov, PhD in Computer Science

Title

Efficient transformer-based model for chemical compounds

Abstract

1. I am working on this project with two students-teammates under supervision of Beznosikov A. N.

2. Our project is devoted to applying transformer-based architecture for molecular compounds. Despite that there are few works in this sphere, we present new bimodal architecture. Also models in this sphere are usually trained on representations of molecules, which not suitable for machine learning tasks.

3. We have compared our model with the existing approaches and confirmed our assumptions about architecture improvements and data format changes.

4. Intermediate results of our work has been presented at 66th All-Russian Scientific Conference of MIPT ("Fine-tuning as an effective alternative to regression for transformers of molecular structures").

Expected results

1. We plan to finish the writing an article for this project by end of September 2024. After that we will send the paper and the abstract to International Conference on Learning Representations (ICLR).
2. I plan to use this transformer-based model for prediction of molecule structure by it's NMR (its my project by supervision of Valentin V. Novikov in MIPT, in collaboration with Grimmlab INEOS RAS). We would like to send paper about this project to JCIM journal.

References on my articles

Oral presentation "Fine-tuning as an effective alternative to regression for transformers of molecular structures" during 66th All-Russian Scientific Conference of MIPT

P.S.

I have some projects at the intersection of machine learning and natural sciences, especially in chemistry, medicine and neurobiology and would like to further develop in this sphere.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Карпеев Глеб Андреевич, 4 курс, ФПМИ ПМИ

Средний балл в ЛК/ GPA: 7.52

Научный руководитель (указать ученую степень) / Scientific adviser: к. ф.-м. н., Стрижов Вадим Викторович.

Тема: / Title Порождающие модели для прогнозирования (наборов временных рядов) в метрическом вероятностном пространстве

Аннотация: / Annotation: Решается задача прогнозирования наборов временных рядов с высокой ковариацией и высокой дисперсией. Для решения данной задачи предлагается построить пространство парных расстояний, представляющего метрическую конфигурацию временных рядов. Прогноз осуществляется в данном пространстве, а затем результат возвращается в исходное пространство с использованием метода многомерного шкалирования. В данной работе предлагаются порождающие модели для прогнозирования наборов временных рядов в метрическом вероятностном пространстве. Новизна работы заключается в применении римановых моделей для прогнозирования и использовании римановых генеративных диффузных моделей.

Ожидаемый результат: / Expected result: Публикация в Journal of Industrial and Management Optimization.

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development: Ассистент на международных курсах Neuromatch Academy <https://neuromatch.io/>.

Дополнительные ссылки: / Additional links: <https://github.com/intsystems/2024-Project-152> - ссылка на работу.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Дивильковский Максим Михайлович, 4 курс, ФПМИ
ПМИ

Средний балл в ЛК/ GPA: 7.14

Научный руководитель (указать ученую степень) / Scientific adviser: к.
ф.-м. н., Стрижов Вадим Викторович.

Тема: / Title Восстановление значения многомерного временного ряда в точке по
спрогнозированной матрице попарных корреляций

Аннотация: / Annotation: Решается задача поточечного прогнозирования набора
временных рядов с высокой ковариацией и высокой дисперсией. Для решения данной
задачи предлагается построить пространство парных расстояний. В этом простран-
стве прогнозируется матрица попарных расстояний, а затем по известной матрице
восстанавливаются значения временных рядов. В данной статье изучается способ вос-
становления прогноза в пространстве временных рядов по известной матрице попар-
ных расстояний. Показывается существование нескольких значений временного ряда,
удовлетворяющих одной матрице попарных расстояний. Предлагается несколько ал-
горитмов, основанных на использовании матриц, построенных по различным времен-
ным интервалам с использованием попарной корреляции. Так же, в статье выводится
явный вид восстановленных значений через матрицу попарных корреляций. Помимо
этого, приводится оценка качества восстановления при добавлении шума в матри-
цы попарных расстояний. Новизна метода заключается в том, что прогнозирование
делается не в исходном пространстве, а в пространстве попарных расстояний.

Ожидаемый результат: / Expected result: Публикация в Journal of Industrial
and Management Optimization.

Дополнительные ссылки: / Additional links: [https://github.com/intsystems/2024-
Project-152a](https://github.com/intsystems/2024-Project-152a) - ссылка на работу.

Семестр 9 магистры / Semester 9, masters

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Киселев Никита Сергеевич

Средний балл в ЛК/ GPA: 8.99

Научный руководитель (указать ученую степень) / Scientific adviser: Грабовой Андрей Валериевич, канд. физ.-мат. наук

Тема: / Title Leveraging Loss Landscape Convergence for Effective Sample Size Estimation

Аннотация: / Annotation: Sample size determination is a crucial aspect of training deep neural networks, as it directly impacts model performance and efficiency. While there are existing results on the properties of loss landscapes in deep neural networks, the relationship between sufficient sample size and loss landscape convergence remains under-explored. In this paper, we develop a method for estimating the optimal sample size required to achieve good model performance. Our approach leverages the convergence properties of the loss landscape as the sample size increases, providing a novel perspective on sample size determination. We demonstrate the effectiveness of our method through extensive experiments on various datasets, showcasing its practical applicability and potential to enhance the efficiency of deep learning models.

Ожидаемый результат: / Expected result:

1. Текущее исследование было начато не так давно и находится на стадии проведения вычислительных экспериментов. Это связано с тем, что работа 1 была завершена только 20 августа, непосредственно перед подачей на конференцию AI Journey 2024.
2. Ожидается, что предлагаемый метод может быть эффективно использован при дообучении нейронных сетей, в частности больших языковых моделей. Для этого проводится изучение точности и обобщаемости модели, обученной на выбранном с помощью метода наборе данных.
3. Работа активно идет, планируется завершить ее в течение 2024 года. Ожидается, что полученные результаты можно будет представить в одном из журналов уровня Q1-Q2 или на одной из конференций.

Ранее полученные результаты: / Delivered results: По теме текущего исследования были подготовлены 2 статьи, а также 1 выступление на конференции:

1. N. Kiselev, A. Grabovoy. Unraveling the Hessian: A Key to Smooth Convergence in Loss Function Landscapes. // Submitted to the AI Journey 2024 conference
Исследование поведения ландшафта функции потерь в глубокой нейронной сети при изменении размера обучающей выборки. Показывается, что в полносвязной нейронной сети, используемой для задачи многоклассовой классификации, в окрестности точки экстремума поверхность функции потерь перестает значительно изменяться, начиная с определенного числа обучающих объектов. Получены оценки на скорость сходимости разности значений функции потерь на подвыборках, отличающихся на один объект. Полученные результаты используются в текущем исследовании для создания метода определения достаточного размера выборки.
2. N. Kiselev, A. Grabovoy. Sample Size Determination: Posterior Distributions Proximity. // Submitted to the Computational Management Science journal
Определение достаточного размера выборки по близости апостериорных распределений параметров модели. Рассматриваются две подвыборки, отличающиеся на один объект. Если в среднем по таким выборкам апостериорные распределения близки, то размер выборки считается достаточным. Исследуются только линейные модели.
3. Киселев Н.С., Грабовой А.В. Определение достаточного размера выборки по апостериорному распределению параметров модели. // 66-я Всероссийская научная конференция МФТИ. – 2024.

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development:

1. В этом семестре проведу 2 лекции + 2 семинара по базовым подходам к инициализации, оптимизации и регуляризации глубоких нейронных сетей в рамках курса «Методы глубокого обучения»
2. Являюсь научным консультантом в работе по определению размера выборки со студентом 4-го курса нашей кафедры Владиславом Мешковым.

Дополнительные ссылки: / Additional links: Помимо представленных выше результатов по размеру выборки, моя научно-исследовательская деятельность также включает следующее:

1. D. Yarmoshik, D. Kovalev, A. Rogozin, N. Kiselev, D. Dorin, A. Gasnikov. Decentralized Optimization with Coupled Constraints. // Submitted to the **NeurIPS 2024** conference with average rating 5.5
Работа посвящена децентрализованной оптимизации с афинными ограничениями. Предлагается новый ускоренный метод первого порядка, который при сравнительном анализе оказывается лучше своих конкурентов. В данном исследовании я отвечал за реализацию большей части вычислительных экспериментов,

а также за написание соответствующей части статьи. Работа прошла стадию rebuttle, в скором времени появится окончательное решение от рецензентов.

2. D. Dorin, N. Kiselev, A. Grabovoy. Forecasting fMRI Images From Video Sequences: Linear Model Analysis. // Submitted to the Health Information Science and Systems journal

Настоящее исследование посвящено анализу зависимости между видеорядом, просматриваемым человеком, и снимками проводимого в этот момент обследования фМРТ. Совместно с Даниилом Дориным, мы предложили метод для анализа исследуемой зависимости, заключающийся в предсказании попарных разностей между снимками в предположении марковости этой последовательности. Работа была выполнена в рамках курса «Моя первая научная статья», а затем подана в журнал на публикацию.

Более подробно про мою научно-исследовательскую деятельность можно посмотреть в моем резюме.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Веприков Андрей Сергеевич

Средний балл в ЛК/ GPA: 8.96

Научный руководитель (указать ученую степень) / Scientific adviser: Хри-
танков Антон Сергеевич, канд. физ.-мат. наук

Тема: / Title Задачи стохастического управления в системах машинного обучения
со скрытыми петлями обратной связи

Аннотация: / Annotation: В системах машинного обучения часто возникает ситу-
ация, когда на результаты работы прогностической модели влияют на решения поль-
зователя, данные о которых используются для улучшения модели. При этом возника-
ет вынужденное смещение входных данных. Такие эффекты называются скрытыми
петлями обратной связи. В магистерской работе предлагается продолжить тематику
бакалаврского диплома, и рассмотреть для интеллектуальной системы задачу управ-
ления эффектом обратной связи. Для этого планируется решить задачу стохастиче-
ского управления для достижения заданного предельного распределения входных
данных. Полученные ранее результаты позволят определить условия существования
петель обратной связи и вид предельного множества системы. Также планируется
провести вычислительный эксперимент для подтверждения теоретических результа-
тов, полученных в работе.

Ожидаемый результат: / Expected result:

1. Использовать результаты из методов стохастического управления в постановке
доверенного искусственного интеллекта и многократного машинного обучения
2. Ожидается, что полученный результат будет подан в виде статьи в Q1-Q2 жур-
нал, а также будет использован в качестве магистерского диплома
3. Также ожидается, что результаты из предыдущих исследований по теме много-
кратного машинного обучения удастся успешно применить в исследуемой задаче

Ранее полученные результаты: / Delivered results: По теме текущего иссле-
дования были подготовлены три работы:

1. Veprikov A., Afanasiev A., Khritankov A. A Mathematical Model of the Hidden
Feedback Loop Effect in Machine Learning Systems //arXiv preprint arXiv:2405.02726.
– 2024. Submitted to the JLMR journal (Q1)

2. Веприков А., Хританков А. Преобразования признаков пространства в модели процесса многократного машинного обучения //66 конференция МФТИ. – 2024.
3. Веприков А., Хританков А., Афанасьев А. Математическая модель эффекта обратной связи в системах искусственного интеллекта //Математические методы распознавания образов (ММО-21). – 2023.
4. Веприков А., Хританков А. Преобразования признаков пространства в модели процесса многократного машинного обучения // Подано в журнал Искусственный интеллект и принятие решений (ИИиПР).

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development:

1. Планируется ассистировать по курсу Программная инженерия для анализа данных.

Дополнительные ссылки: / Additional links: Помимо представленных выше результатов, моя научно-исследовательская деятельность также включает работу в лаборатории Математических методов оптимизации МФТИ. Вместе с коллегами из этой лаборатории мною было написано несколько работ:

1. Veprikov A. et al. Markov Compression: Looking to the Past Helps Accelerate the Future // Submitted to NeurIPS 2024 conference with average rating 6.0
2. Lobanov A., Veprikov A. et al. Non-smooth setting of stochastic decentralized convex optimization problem over time-varying graphs //Computational Management Science. – 2023. – Т. 20. – №. 1. – С. 48.
3. Veprikov A. et al. New Aspects of Black Box Conditional Gradient: Variance Reduction and One Point Feedback // Submitted to the Chaos, Solitons and Fractals journal (Q1)
4. Solodkin V., Veprikov A., Beznosikov A. Methods for Optimization Problems with Markovian Stochasticity and Non-Euclidean Geometry //arXiv preprint arXiv:2408.01848. – 2024. Submitted to the AAAI 2024 conference
5. Chernyavskiy A., Veprikov A. An Out-of-Shelf Multi-Level Monte Carlo Approach for Average-Reward Reinforcement Learning // Submitted to the ICOMP 2024 conference with average rating 7.0
6. Веприков А. и др. Алгоритм Нулевого Порядка для Решения Задач Децентрализованной Оптимизации // Подано на конференцию AIJ 2024 от Сбербанка

Более подробно про мою научную деятельность и мои проекты можно посмотреть в моем резюме.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Семёнов Андрей

Средний балл в ЛК/ GPA: 8.92

Научный руководитель (указать ученую степень) / Scientific adviser: Безносиков Александр Николаевич, канд. физ.-мат. наук

Тема: / Title Just a Simple Transformation is Enough for Data Protection in Vertical Federated Learning

Аннотация: / Annotation: Vertical Federated Learning (VFL) enables collaborative training of deep learning models while maintaining privacy protection. However, the VFL procedure still has components that are vulnerable to attacks by malicious parties. In our work, we consider feature reconstruction attacks – a common risk targeting input data compromise. We theoretically claim that feature reconstruction attacks cannot succeed without knowledge of the prior distribution on data. Consequently, we demonstrate that even simple model architecture transformations can significantly impact the protection of input data during VFL. Confirming these findings with experimental results, we show that MLP-based models are resistant to state-of-the-art feature reconstruction attacks.

Ожидаемый результат: / Expected result: Пройти на ICLR 2025.

Ранее полученные результаты: / Delivered results: Статья была подана на NeurIPS 2024.

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development: Планирую выступить с докладом на научном семинаре группы.

Дополнительные ссылки: / Additional links: Моя научно-исследовательская деятельность также включает несколько других статей:

1. Semenov A. et al. "Sparse Concept Bottleneck Models: Gumbel Tricks in Contrastive Learning" // Submitted to the NeurIPS 2024 conference
2. Beznosikov A., Dvinskikh D., Semenov A. "Bregman Proximal Method for Efficient Communications under Similarity" // Submitted to the ICOMP 2024 conference with average rating 8.67

3. Chezhegov S., Semenov A. et al. "Gradient Clipping Improves AdaGrad when the Noise Is Heavy-Tailed" // Submitted to the NeurIPS 2024 conference with average rating 5.0

4. "Mixed Newton Method for Optimization in Complex Spaces" // Submitted to the NeurIPS 2024 conference.

Полный список моих работ обновляется здесь.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Дорин Даниил Дмитриевич

Средний балл в ЛК/ GPA: 8.83

Научный руководитель (указать ученую степень) / Scientific adviser: Грабовой Андрей Валериевич, канд. физ.-мат. наук

Тема: / Title Spatial-temporal characteristics in a simultaneous EEG and fMRI data decoding.

Аннотация: / Annotation: There is a significant relationship between EEG signals and fMRI. This relationship is expressed after stimulus onset. The event-related potential, observed after the event, manifests a change in the fMRI image with a small delay. EEG is a superposition of several signals generated in different parts of a human brain. The examined relationship is the signal generator changes the blood oxygen level in the neighborhood. The EEG signal analysis requires a robust decomposition algorithm. The EEG-fMRI relationship provides prior information about the signal generators and their features. It solves the inverse EEG problem and boosts the quality of the EEG signal analysis which is mandatory for real-time physical motion prediction. To test the relationship the casualty inference goes in a low-dimensional latent space using the Canonical Correlation Analysis Principle. To select an adequate model, both spaces, source EEG and target fMRI, are decomposed separately and then fine-tuned with fixed model hyperparameters and pre-tuned parameters.

Ожидаемый результат: / Expected result:

1. Исследование является продолжением моей научной деятельности, связанной с анализом данных фМРТ. Работа начата недавно, так как предыдущее исследование подано на конференцию AI Journey 2024 только 20 августа.
2. На данный момент найден подходящий для исследования датасет. Начаты первичные эксперименты. Ожидается, что предлагаемый метод эффективно себя покажет в задаче декодирования, в частности в задаче классификации стимула.
3. Планируется завершить работу в течение 2024 года. Подать результаты в журнал уровня Q1-Q2 или выступить с результатами работы на одной из конференций.

Ранее полученные результаты: / Delivered results: По теме текущего исследования были подготовлены 2 работы и имеется выступление на конференции:

1. D. Dorin, N. Kiselev, A. Grabovoy. Forecasting fMRI Images From Video Sequences: Linear Model Analysis. // Submitted to the **Health Information Science and Systems** journal.

В исследовании анализируется взаимосвязь между показаниями датчиков фМРТ и просмотренным человеком видеорядом. Предлагается метод аппроксимации показаний фМРТ с использованием видеоряда. Метод предполагает неизменяемую во времени гемодинамическую реакцию и построение линейной модели для каждого вокселя фМРТ. Эффективность модели оценивается с помощью вычислительных экспериментов на большом наборе данных.

2. Дорин Д.Д., Грабовой А.В., Стрижов В.В. Пространственно-временные характеристики в задаче декодирования данных фМРТ. // Submitted to the **AI Journey 2024** conference.

В работе исследуются пространственно-временные характеристики в задаче декодирования данных функциональной фМРТ. Предложен метод снижения пространственной размерности временных рядов фМРТ, основанный на взвешивании стимулированных областей мозга с использованием кросс-корреляционной функции. На основе этого метода предложена модель классификации сегментов временных рядов фМРТ, использующая кодировщик с применением римановой геометрии для извлечения пространственно-временных характеристик.

3. Дорин Д.Д., Киселев Н.С., Грабовой А.В. Пространственно-временные методы анализа временных рядов. // 66-я Всероссийская научная конференция МФТИ. – 2024.

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development:

1. В предыдущем семестре являлся научным консультантом на курсе mlp. В дальнейшем планирую продолжать работу научным консультантом на кафедре.
2. В этом семестре проведу 2 лекции в рамках курса «Методы глубокого обучения» по теме сегментации и детекции.

Дополнительные ссылки: / Additional links: Помимо представленных выше результатов, моя научно-исследовательская деятельность также включает следующее:

1. A. Filatov*, D. Dorin*, N. Barinov, U. Izmesteva, I. Ignashin, I. Stepanov, V. Vasilev, M. Kurkin, D. Yudin, A. Alanov, S. Zagoruyko, D. Dimitrov, A. Kuznetsov. Garage: Generative Augmentation Framework for Transforming Object Representations in Images. // Submitted to the **EMNLP 2024 System Demonstration Track**

conference.

В исследовании предлагается новый подход к генеративным аугментациям изображений с изменением аннотаций. Реализована библиотека для автоматического аугментирования датасета и демо интерфейс. В данном исследовании я отвечал за идейную часть, а также за большую часть реализации библиотеки. Работа находится в процессе рецензирования.

2. D. Yarmoshik, D. Kovalev, A. Rogozin, N. Kiselev, D. Dorin, A. Gasnikov. Decentralized Optimization with Coupled Constraints. // Submitted to the **NeurIPS 2024** conference. Исследование посвящено децентрализованной оптимизации с аффинными ограничениями. Разработан новый метод первого порядка, который превосходит существующие аналоги. В рамках исследования я отвечал за проведение вычислительных экспериментов и написание соответствующей части статьи. Работа прошла стадию рецензирования с средним рейтингом 5.5, в ближайшее время ожидаются окончательные результаты.

Подробнее про мою исследовательскую деятельность и мои достижения можно посмотреть в резюме.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Игнашин Игорь Николаевич

Средний балл в ЛК/ GPA: 7.80

Научный консультант Ярмошик Демьян Валерьевич.

Тема 1 Edge-based methods for solving the min-cost problem in barrier cost function models.

Аннотация: / Annotation: In this paper, we investigate the min-cost problem, particularly focusing on its relevance to telecommunication networks. The motivation for this study stems from the need to design robust networks with high throughput and low latency, which are critical in telecommunications. Achieving this requires fast and efficient methods for solving convex minimum cost multicommodity flow problems, especially when dealing with barrier cost functions, as modeled by Kleinrock/Davidson. To address these challenges, we implement and compare various edge-based methods. On the primal side, we explore modifications of the Frank-Wolfe algorithm, including the self-concordant Frank-Wolfe method proposed by Dvurechensky, to ensure efficient and accurate convergence. On the dual side, we implement adaptive composite gradient methods and the cutting plane method proposed by Fukushima. Our aim is to derive the most effective algorithms across these different approaches, supported by theoretical complexity estimates. Additionally, we plan to implement the Universal Similar Triangles Method (USTM) introduced by Kubentayeva and the dual ascent method based on the cutting plane method proposed by Hearn. These methods are expected to further improve performance in solving the min-cost problem, particularly in terms of convergence speed and solution accuracy.

Ожидаемый результат: / Expected result:

1. Данная работа была начата этим летом. На данный момент есть датасеты с графами телекоммуникаций, предоставленные Huawei. Решается задача Min-Cost с ограничениями на потоки. По сути мы оптимизируем количество передаваемой информации, изменяя загруженность графа в соответствии с заданными корреспонденциями.
2. Реализованы алгоритмы Dual Subgradient и модификации FW, описанные в статье. А также алгоритмы протестированы на графах телекоммуникаций.
3. Планируется реализовать алгоритмы USTM и Dual ascent для модели с барьерным типом функций затрат, а также провести теоретический анализ данных методов.

4. Планируется завершить данную работу к концу 2024 года и подать ее на конференцию IEEE Q1 в области сетевых технологий.

Ранее полученные результаты: / Delivered results:

1. Публикация статьи в журнале «Компьютерные исследования и моделирование». В этой работе сравниваются существующие модификации Frank-Wolfe метода и предлагаются новые модификации: NFW, WFFW.
2. Научная статья «Conjugate Frank-Wolfe in Machine Learning». Подана на конференцию по оптимизации OPTIMA. Находится на рецензировании. Проводится в октябре 2024. В данной работе рассматривается задача выпуклой оптимизации с ограничениями. То есть для задачи бинарной классификации используется лог-лосс с l_1 ограничениями на веса в качестве регуляризатора. Сравняется качество решения данной задачи методом Frank-Wolfe и его модификацией Conjugate Frank-Wolfe и более общей модификацией N Conjugate Frank-Wolfe. Также получена оценка на сходимость алгоритма CFW, предложенного в статье.
3. Выступление на 66-ой Всероссийской научной конференции МФТИ 2024 с темой доклада «О задаче поиска равновесного распределения потоков».

Научный руководитель (указать ученую степень) / Scientific adviser: Грабовой Андрей Валериевич, канд. физ.-мат. наук

Тема 2 Bayesian distillation of transformer models with decoder layer pruning

Аннотация: / Annotation: This paper explores various approaches to simplifying approximating models through Bayesian distillation of deep neural networks, with a particular focus on transformer-based models. The motivation for this study stems from the need to simplify models in order to achieve more efficient performance in terms of both quality and computational speed. The study utilizes the concepts of "teacher" and "student" models, investigating methods to derive prior distributions for the student model's parameters, which are fewer than those of the teacher model, based on the posterior distribution of the teacher model's parameters.

Building on previous work on Bayesian distillation of transformers, which provided theoretical insights into the posterior distribution when pruning the decoder layer, this paper plans to experimentally evaluate the removal of the decoder layer using Bayesian distillation, with the expectation of improved performance compared to existing distillation methods. Additionally, the research examines the distillation of convolutional neural networks (CNNs). The primary model used in this study is a seq2seq transformer for translating text from Russian to English. The distilled models are expected to demonstrate better convergence and superior quality compared to models initialized with parameters from arbitrary distributions under the same training conditions.

Ожидаемый результат: / Expected result:

1. Ранее были получены теоретические выкладки для байесовской дистилляции трансформера и модели RNN с вниманием: априорные распределения для параметров модели ученика, где модель ученика – это модель учителя с удаленными слоями декодера/энкодера (в случае трансформеров), либо модель RNN с удаленным блоком attention. А также был проведен эксперимент для случая RNN.
2. По сравнению с предыдущей работой, в данной работе основной акцент поставлен на экспериментальных результатах. Планируется провести эксперимент с байесовской дистилляцией модели seq2seq трансформера с удалением слоя энкодера, продолжающий предыдущую работу – бакалаврский диплом. Также в данной работе планируется получить теоретические выкладки для случая модели сверточных нейронных сетей.
3. Ожидаются следующие результаты экспериментов: улучшение сходимости и качества дистиллированной модели с помощью байесовского подхода в сравнении с обычными методами дистилляции.
4. Планируется завершить данную работу к концу 2024 года – началу 2025 года и податься на зимнюю/весеннюю конференцию уровня Q1-Q2.

Ранее полученные результаты: / Delivered results:

1. Игнашин И.Н. Байесовская дистилляция моделей на базе трансформеров. На данный момент выбирается конференция для подачи данной статьи. В данной работе исследуется байесовская дистилляция моделей с вниманием. Получены теоретические выкладки для априорного распределения параметров модели ученика, с удаленным слоем внимания. Проведен эксперимент для случая модели seq2seq RNN с блоком внимания.

Дополнительные ссылки: / Additional links: Помимо представленных выше работ, я работал над следующими статьями:

1. A. Filatov*, D. Dorin*, N. Barinov, U. Izmesteva, I. Ignashin, I. Stepanov, V. Vasilev, M. Kurkin, D. Yudin, A. Alanov, S. Zagoruyko, D. Dimitrov, A. Kuznetsov. Garage: Generative Augmentation Framework for Transforming Object Representations in Images. Работа подана на конференцию EMNLP 2024 System Demonstration Track. В работе предлагается новый подход к генеративным аугментациям изображений с изменением аннотаций. Реализована библиотека для автоматического аугментирования датасета и демо интерфейс. В данной работе я отвечал за обучение моделей Faster-CNN и DETR на аугментированных датасетах. Работа находится в процессе рецензирования.

Конкурсная заявка на Научную академическую стипендию им. К.В. Рудакова

Student

Nikolai Rekut, 5th semester

Scientific advisor

Alexander Beznosikov, PhD in Computer Science

Title

Efficient transformer-based model for chemical compounds

Abstract

1. I am working on this project with two students-teammates under supervision of Beznosikov A. N.

2. Our project is devoted to applying transformer-based architecture for molecular compounds. Despite that there are few works in this sphere, we present new bimodal architecture. Also models in this sphere are usually trained on representations of molecules, which not suitable for machine learning tasks.

3. We have compared our model with the existing approaches and confirmed our assumptions about architecture improvements and data format changes.

4. Intermediate results of our work has been presented at 66th All-Russian Scientific Conference of MIPT ("Fine-tuning as an effective alternative to regression for transformers of molecular structures").

Expected results

1. We plan to finish the writing an article for this project by end of September 2024. After that we will send the paper and the abstract to International Conference on Learning Representations (ICLR).
2. I plan to use this transformer-based model for prediction of molecule structure by it's NMR (its my project by supervision of Valentin V. Novikov in MIPT, in collaboration with Grimmlab INEOS RAS). We would like to send paper about this project to JCIM journal.

References on my articles

Oral presentation "Fine-tuning as an effective alternative to regression for transformers of molecular structures" during 66th All-Russian Scientific Conference of MIPT

P.S.

I have some projects at the intersection of machine learning and natural sciences, especially in chemistry, medicine and neurobiology and would like to further develop in this sphere.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Сапронов Юрий Феликсович

Средний балл в ЛК/ GPA: 7.60 (был, сейчас отсутствует)

Научный руководитель (указать ученую степень) / Scientific adviser: Гасников Александр Владимирович, доктор физ.-мат. наук

Тема: / Title Optimal Approximation of Average Reward Markov Decision Processes

Аннотация: / Annotation: We continue to develop the concept of studying the ε -optimal policy for Average Reward Markov Decision Processes (AMDP) by reducing it to Discounted Markov Decision Processes (DMDP). Existing research often stipulates that the discount factor must not fall below a certain threshold. Typically, this threshold is close to one, and as is well-known, iterative methods used to find the optimal policy for DMDP become less effective as the discount factor approaches this value.

Our work distinguishes itself from existing studies by allowing for inaccuracies in solving the empirical Bellman equation. Despite this, we have managed to maintain the sample complexity that aligns with the latest results. We have succeeded in separating the contributions from the inaccuracy of approximating the transition matrix and the residuals in solving the Bellman equation in the upper estimate so that our findings enable us to determine the total complexity of the epsilon-optimal policy analysis for DMDP across any method with a theoretical foundation in iterative complexity.

Текущий результат: / Current result:

1. Данная работа получила accept на конференции ICOMP 2024

Ожидаемый результат: / Expected result:

1. Ожидается, что она будет добавлена в журнал Q1-Q2.
2. Планируется продолжить работу в данном направлении, но уже с бесконечными средами, а не табулярными.

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development:

1. Хотел бы поучаствовать в каком-то из преподавательских процессов на кафедре.

Семестр 11 магистры / Semester 11, masters

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Яцулевич Владимир Владимирович, 6 курс, ФПМИ ИВТ

Средний балл в ЛК/ GPA: 9.55

Научный руководитель (указать ученую степень) / Scientific adviser: Трушечкин Антон Сергеевич, доктор физико-математических наук, старший научный сотрудник МИАН, профессор МФТИ, профессор НИТУ МИСиС

Тема: / Title Протокол квантовой криптографии BB84 при отличающихся эффективностях детекторов в случае пассивного выбора базиса

Аннотация: / Annotation: Научно-исследовательская работа посвящена протоколу квантовой криптографии BB84. В ней будет рассмотрено состояние исследований по этому протоколу на сегодняшний день. А именно случай, когда принимающие детекторы сигнала работают неидеально. В частности, будет представлена формула скорости генерации секретного ключа для протокола квантовой криптографии BB84 при отличающихся эффективностях детектора при активном выборе базиса, а также некоторые результаты в случае пассивного выбора базиса. Также будут рассмотрены результаты реального применения данного протокола.

Один из самых распространённых протоколов квантового распределения ключей – это протокол BB84. В этом протоколе информация кодируется поляризацией или фазой сильно когерентного импульса, описывающее состояние одиночного фотона. Практическая реализация данного протокола использует детекторы, которые реагируют на поступающие сигналы. В идеальном случае детектор должен срабатывать только один раз, если прилетел ровно один фотон. В реальном же случае детектор будет срабатывать только с некоторой вероятностью $0 < \eta < 1$, которая называется эффективностью детектора.

Ожидаемый результат: / Expected result: Полученные результаты будут представлены в магистерской работе на кафедре Методов современной математики МИАН

Ранее полученные результаты: / Delivered results:

1. Постерный доклад на 4-ой международной школе «Сверхпроводниковые технологии для обработки квантовой информации», 2024 г.

2. Устный доклад на II Всероссийской школе-семинаре НЦФМ «Центр исследования архитектур суперкомпьютеров», 2024 г.
3. Печать в сборнике тезисов на II Всероссийской школе-семинаре НЦФМ «Центр исследования архитектур суперкомпьютеров», 2024 г.

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development:

1. Учебный ассистент кафедры Дискретной математики, осенний семестр 2024/2025 учебного года, ведение семинаров по дисциплине «Математическая логика».
2. Учебный ассистент кафедры Дискретной математики, весенний семестр 2023/2024 учебный год, принятие устных заданий по дисциплине «Математическая логика».
3. Учебный ассистент кафедры Алгоритмов и технологий программирования, весенний семестр 2023/2024 учебного года, ведение семинаров по дисциплине «Математическая логика».
4. Педагог дополнительного образования Технопарка Физтех-лицея, 2024/2025 учебный год, дисциплины «Дискретная математика», «Python и инструменты машинного обучения».
5. Преподаватель ЗФТШ, 2023/2024 учебный год, 2024/2025 учебный год, дисциплины Математика, Информатика и Физика.
6. Разработчик курса «Современные методы квантовых вычислений», кафедра Вычислительной физики, весенний семестр 2023/2024 учебный год.

Дополнительные ссылки: / Additional links:

1. SQH 2024. 4-ая международная школе «Сверхпроводниковые технологии для обработки квантовой информации», 2024 г. Постерный доклад на тему «Формулы скорости генерации секретного ключа для протокола квантовой криптографии BB84 при отличающихся эффективности детекторов при пассивном и активном выборе базиса».
2. Сборник тезисов, Саров 2024. II Всероссийская школа-семинар НЦФМ «Центр исследования архитектур суперкомпьютеров».
3. Современные методы квантовых вычислений.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Боева Галина Леонидовна

Средний балл в ЛК/ GPA: 9.55

Научный руководитель (указать ученую степень) / Scientific adviser: Зайцев Алексей Алексеевич, канд. физ.-мат. наук

Тема: / Title Label Attention Network for Temporal Sets Prediction: You Were Looking at a Wrong Self-Attention.

Аннотация: / Annotation: Most user-related data can be represented as a sequence of events associated with a timestamp and a collection of categorical labels. For example, the purchased basket of goods and the time of buying fully characterize the event of the store visit. Anticipation of the label set for the future event called the problem of temporal sets prediction, holds significant value, especially in such high-stakes industries as finance and e-commerce. A fundamental challenge of this task is the joint consideration of the temporal nature of events and label relations within sets. The existing models fail to capture complex time and label dependencies due to ineffective representation of historical information initially. We aim to address this shortcoming by presenting the framework with a specific way to aggregate the observed information into time- and set structure-aware views prior to transferring it into main architecture blocks. Our strong emphasis on input arrangement facilitates the subsequent efficient learning of label interactions. The proposed model is called Label-Attention NETwork, or LANET. We conducted experiments on four different datasets and made a comparison with four established models, including SOTA, in this area. The experimental results suggest that LANET provides significantly better quality than any other model, achieving an improvement up to 65% in terms of weighted F1 metric compared to the closest competitor. Moreover, we contemplate causal relationships between labels in our work, as well as a thorough study of LANET components' influence on performance. We provide an implementation of LANET to encourage its wider usage.

Ожидаемый результат: / Expected result:

1. Исследование было доведено до этапа публикации и результат принят на конференцию ранга А ECAI 2024.
2. Планируется усовершенствовать данную работу, создав математическое обоснование преимуществ агрегирования по меткам классов.
3. Также в данном исследовании планируется провести ряд работ по усовершенствованию эффективности и ускорения метода.

Ранее полученные результаты: / Delivered results: По теме данного исследования имеется публикация:

1. Elizaveta Kovtun, Galina Boeva, Artem Zabolotnyi, Evgeny Burnaev, Martin Spindler, Alexey Zaytsev Label Attention Network for sequential multi-label classification: you were looking at a wrong self-attention.
2. Список принятый работ на конференцию ECAI. Доклад на конференции планируется 19-24 октября.

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development:

1. Занимаюсь преподаванием и организацией курсов на совместной программе Сколтеха со Сбербанком.
2. Проводила курс по повышению квалификации в Сколтехе по теме "Модели последовательностей событий".
3. Сейчас также провожу курс базового машинного обучения и в сентябре также будет дополнительный курс, связанный с темой последовательностей событий и процессами Хокса.

Дополнительные ссылки: / Additional links: Помимо представленных выше результатов, моя научно-исследовательская деятельность также включает следующее:

1. Vladislav Zhuzhel, Vsevolod Grabar, Galina Boeva, Artem Zabolotnyi, Alexander Stepikin, Vladimir Zholobov, Maria Ivanova, Mikhail Orlov, Ivan Kireev, Evgeny Burnaev, Rodrigo Rivera-Castro, Alexey Zaytsev Continuous-time convolutions model of event sequences. Статья подана в журнал Information Sciences(Q1). Мы предлагаем метод COTIC, основанный на непрерывной сверточной нейронной сети, подходящей для учета неравномерного прохождения событий во времени. В COTIC расширенная и многослойная архитектура эффективно справляется с зависимостями между событиями. Кроме того, модель обеспечивает общую динамику интенсивности в непрерывном режиме времени, включая процессы Хокса. Модель COTIC превосходит существующие подходы к большинству рассмотренных наборов данных, создавая представления для последовательности событий, которые могут быть использованы для решения последующих задач - например, для прогнозирования следующего типа события и времени возврата.
2. Ilya Kuleshov, Galina Boeva, Vladislav Zhuzhel, Evgenia Romanenkova, Evgeni Vorsin, Alexey Zaytsev COTODE: COntinuous Trajectory neural Ordinary Differential Equations for modelling event sequences. Данная статья была подана на KDD 2024. Большинство современных методов, как правило, моделируют такие процессы

по большей части кусочно-непрерывными траекториями. Чтобы решить эту проблему, мы рассматриваем события не как самостоятельные явления, а как наблюдения за гауссовым процессом, который, в свою очередь, управляет динамикой actor. Мы предлагаем интегрировать полученную динамику, что приведет к модификации широко успешной нейронной модели ODE с непрерывной траекторией. С помощью теории гауссовых процессов мы смогли оценить неопределенность в представлении actor, которая возникает из-за отсутствия наблюдения за ними в промежутках между событиями. Эта оценка привела нас к разработке нового, теоретически обоснованного механизма отрицательной обратной связи.

Подробнее про мою исследовательскую деятельность и мои достижения можно посмотреть в резюме.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Корнилов Никита Максимович

Средний балл в ЛК/ GPA: 9.31

Научный руководитель (указать ученую степень) / Scientific adviser: Гасников Александр Владимирович, доктор физ.-мат. наук

Тема: / Title Optimal Flow Matching: Learning Straight Trajectories just in One Step

Аннотация: / Annotation: Over the several recent years, there has been a boom in development of Flow Matching (FM) methods for generative modeling. One intriguing property pursued by the community is the ability to learn flows with straight trajectories which realize the Optimal Transport (OT) displacements. Straightness is crucial for the fast integration (inference) of the learned flow's paths. Unfortunately, most existing flow straightening methods are based on non-trivial iterative FM procedures which accumulate the error during training or exploit heuristics based on minibatch OT. To address these issues, we develop and theoretically justify the novel **Optimal Flow Matching** (OFM) approach which allows recovering the straight OT displacement for the quadratic transport in just one FM step. The main idea of our approach is the employment of vector field for FM which are parameterized by convex functions. Our OFM can efficiently solve both OT and FM problems. Thus, we pave a novel theoretical bridge between Optimal Transport and Flow Matching. Such direct connection may result in adopting of the methods' strong sides to each other and deepening the understanding of them.

Ранее полученные результаты: / Delivered results: По теме данного исследования имеется публикация, поданная и высоко оцененная на ревью NeuralPS 2024.

Ожидаемые новые результаты: / Expected result:

1. Найти способ исправить все технические Limitations, описанные в работе. Запустить больше экспериментов.
2. Применить полученные результаты в уже имеющихся работах, основанных на Flow Matching
3. Теоретически посчитать сокращение скорости сходимости метода.

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development:

1. Веду семинары по Методам Оптимизации для 3 курса ФИВТ МФТИ.

Дополнительные ссылки: / Additional links: Помимо представленных выше результатов, моя научно-исследовательская деятельность также включает следующее:

1. Пул работ по безградиентной оптимизации. Крайняя работа повествует о медленном клиппинге, который работает в теории и на практике лучше при очень тяжелом шуме, чем обычный клиппинг. Работа подана и высоко оценена на NeuralPS 2024.

Сейчас в работе проект с Александром Безносиковым по доказательству сходимости Adam подобных алгоритмов с высокой вероятностью. После эти результаты могут быть использованы для безградиентной оптимизации с human feedback.

**Конкурсная заявка на Научную академическую стипендию
им. К.В. Рудакова**

Студент / Student's name: Жарова Мария Александровна, 6 курс, ФПМИ ИВТ

Средний балл в ЛК/ GPA: 7.80

Научный руководитель (указать ученую степень) / Scientific adviser: Цурков Владимир Иванович, доктор физико-математических наук, профессор МФТИ, академик РАН

Тема: / Title Применение бустинга в рекомендательных системах

Аннотация: / Annotation: В данном исследовании предлагается построить рекомендательную систему в условиях ограниченного числа объектов, некоторые из которых имеют высокую корреляцию между собой. При таких данных методы коллаборативной фильтрации становятся менее эффективными из-за уменьшения разреженности матриц и появления большого количества дубликатов векторов пользователей, а контентные не подходят вследствие проблемы корреляции.

Для решения описанной проблемы в работе реализована система, состоящая из независимых классификационных моделей для каждого объекта, обучение каждой из которых базируется на метаданных и агрегированных коллаборативных факторах. Далее при помощи калибровки предсказаний получаются сравнимые вероятности заинтересованности пользователей в различных объектах, из которых формируется итоговый список рекомендаций.

Ожидаемый результат: / Expected result: Материалы работы будут представлены в магистерском дипломе на кафедре Интеллектуальных систем. Также по имеющимся результатам планируется публикация статьи в журнале «Известия РАН. Теория и системы управления», 2024, № 6 (пройден этап проверки научным редактором, в данный момент находится на этапе рассмотрения литературным редактором); и презентация доклада на 15-й Международной конференции «Интеллектуализация обработки информации», 27.09.2024 г.

В настоящий момент проводятся эксперименты по повышению точности системы путём учёта отрицательных действий пользователей.

Ранее полученные результаты (по текущей работе): / Delivered results:

1. Участие в 66-й Всероссийской научной конференции МФТИ и публикация в сборнике тезисов материалов доклада «Комплексная рекомендательная система на основе бустинга», 2024 г.

Участие в преподавании и в развитии Кафедры ИС: / Teaching and IS chair development:

1. Преподаватель программ онлайн-магистратур ДПО МФТИ, осенний и весенний семестры 2023/2024 учебного года, осенний семестр 2024/2025 года: чтение лекций по дисциплинам «Модели ML в production», «Инжиниринг данных», «Анализ естественного языка»; сопровождение и оценивание работ студентов в рамках проектного практикума; проведение экзаменов и проверка домашних работ по дисциплинам «Глубокое обучение», «Модели ML в production», «Проектный практикум» .
2. Преподаватель программ онлайн-магистратур УрФУ и ТГУ, осенний и весенний семестры 2023/2024 учебного года: сопровождение и оценивание работ студентов в рамках проектного практикума.
3. Чтение лекций по введению в машинное обучение для учащихся Лицея НИУ ВШЭ в рамках программы «Alfa campus».

Дополнительные ссылки: / Additional links: Результаты моей предыдущей научно-исследовательской деятельности:

1. Публикация статьи «Нейросетевые подходы для рекомендательных систем» в журнале «Известия РАН. Теория и системы управления», 2023, № 6, с. 150–165.
2. Публикация статьи «Neural Network Approaches for Recommender Systems» в журнале «Journal of Computer and Systems Sciences International», 2024, Т. 62, с. 1048–1062.