

Надежный и логически прозрачный искусственный интеллект:

технология, верификация и применение
при социально-значимых и инфекционных
заболеваниях

Крупный научный проект

Надежный и логически прозрачный искусственный интеллект: технология, верификация и применение при социально-значимых и инфекционных заболеваниях



Научный руководитель:
А.Н. Горбань
(Университет Лейстера, ННГУ)

Координатор: М.В. Иванченко,
проректор по научной работе ННГУ

Консорциум:
ННГУ, ИТМО, ИПМ РАН, ИСП РАН, ПИМУ

Факультеты и институты ННГУ:
ИИТММ, РФ, ХФ, ИББМ

204 исполнителя
(24 доктора наук, 65 кандидатов наук,
24 аспиранта, 130 участников – до 39 лет)

Международное сотрудничество:

- Potsdam Institute for Climate Research,
- Imperial College London,
- Leicester University,
- Harvard Medical School,
- University of Bologna,
- Institut Curie (Paris),
- Oslo Metropolitan University,
- Gent University,
- Paris Sciences et Lettres University,
- Universidad Complutense de Madrid,
-

План доклада

- Проблемы и задачи
- Работы по разработке технологии робастного и логически прозрачного ИИ
- Сбор и анализ биомедицинских данных
- Апробация в задачах лазерной физики
- Апробация в задачах квантовой динамики
- Исследование опасных природных явлений
- Анализ климатических данных и предсказание экстремальных событий
- Методы ИИ для нахождения материалов с заранее заданными свойствами
- Оценка качества машинного обучения в анализе временных рядов
- Разработка и использование моделей спайковых нейронных сетей
- Опережающая диагностики стрессовых состояний с/х растений
- Генеративный дизайн математических моделей с элементами ИИ
- Заключение

Наш закон – календарный план

1 год.

- Создание математических основ для оценки вероятностей ошибок и сертифицирования робастности и качества обучения.
- Формулировка и доказательство теоремы стохастического разделения для смесей лог-вогнутых с явными оценками вероятности разделения, в том числе для зависимых выборок. Будут использоваться логарифмические неравенства Соболева и марковские
- деревья.
- Новые теоремы существования состязательных возмущений, выявляющих негрубость и уязвимость систем.
- Тестирование технологий корректоров в сочетании с алгоритмами состязательного обучения.

1 год (продолжение).

- Создание нового семейства алгоритмов получения явного вывода для действий систем ИИ на основе нейронных сетей.
- Теория и алгоритмы взаимообучения и коррекции в больших сообществах МО и ИИ систем.
- Разработка алгоритмов многокритериальной оптимизации, основанных на сочетании МО с оригинальными алгоритмами глобального поиска.
- Развитие методов МО и ИИ для генеративного дизайна математических моделей в форме ДУЧП и ИДУЧП на основе данных наблюдений и априорных знаний в форме эталонных моделей математической физики.

2 год.

- Разработка, имплементация и тестирование алгоритмов для сертификации систем МО и ИИ на синтетических и реальных стандартных бенчмарках. Анализ новых типов скрытых атак на ИИ, меняющих структуру модели незаметным для тестирования образом. Имплементация и тестирование алгоритмов взаимобучения в больших сообществах МО и ИИ систем.
- Суперкомпьютерная имплементация и тестирование новых алгоритмов оптимизации.
- Обеспечение биомедицинских и других стратегически важных задач реальной сложности алгоритмами надежного и логически прозрачного ИИ.
- Разработка алгоритмов кодизайна структуры математической модели и параллельного алгоритма ее реализации, обеспечивающей заданные характеристики параллельной масштабируемости на конкретной суперкомпьютерной архитектуре.

3 год.

- Доработка разработанных технологий и их валидация на разнородных реальных данных, подготовленных при решении биомедицинских и других стратегически важных задач реальной сложности.
- Получение конкретных верифицируемых результатов в заявленных практических приложениях.
- Подготовка пользовательского программного обеспечения.

Б. Задачи анализа биомедицинских данных

1 год: Сбор новых локальных данных и подготовка имеющихся данных к анализу (систематизация, анализ недостающих компонент, импутация и различные методы предобработки). Разработка алгоритмов для генерации синтетических данных и тестирование методов репрезентации с помощью графов, сетевого анализа, ИИ и МО. Адаптация имеющейся методологии к анализу многомерных серийных данных.

2 год: Применение разработанных методов для анализа молекулярных и фундаментальных механизмов различных заболеваний, таких как синдром Дауна и связанное с ними ускоренное старение, нейродегенеративных заболеваний и различных видов коронавирусов, включая COVID-19. Применение методов графового анализа, ИИ и МО для решения задач ранней диагностики и связанной с ней уменьшения смертности.

3 год: Адаптация методов объяснимого и логически прозрачного ИИ в разработанные алгоритмы и программы для изучения механизмов болезней, ранней диагностики и идентификации возможных молекулярных и генетических мишеней для возможного лечения.

В. Апробация на задачах реальной сложности из разных областей наук

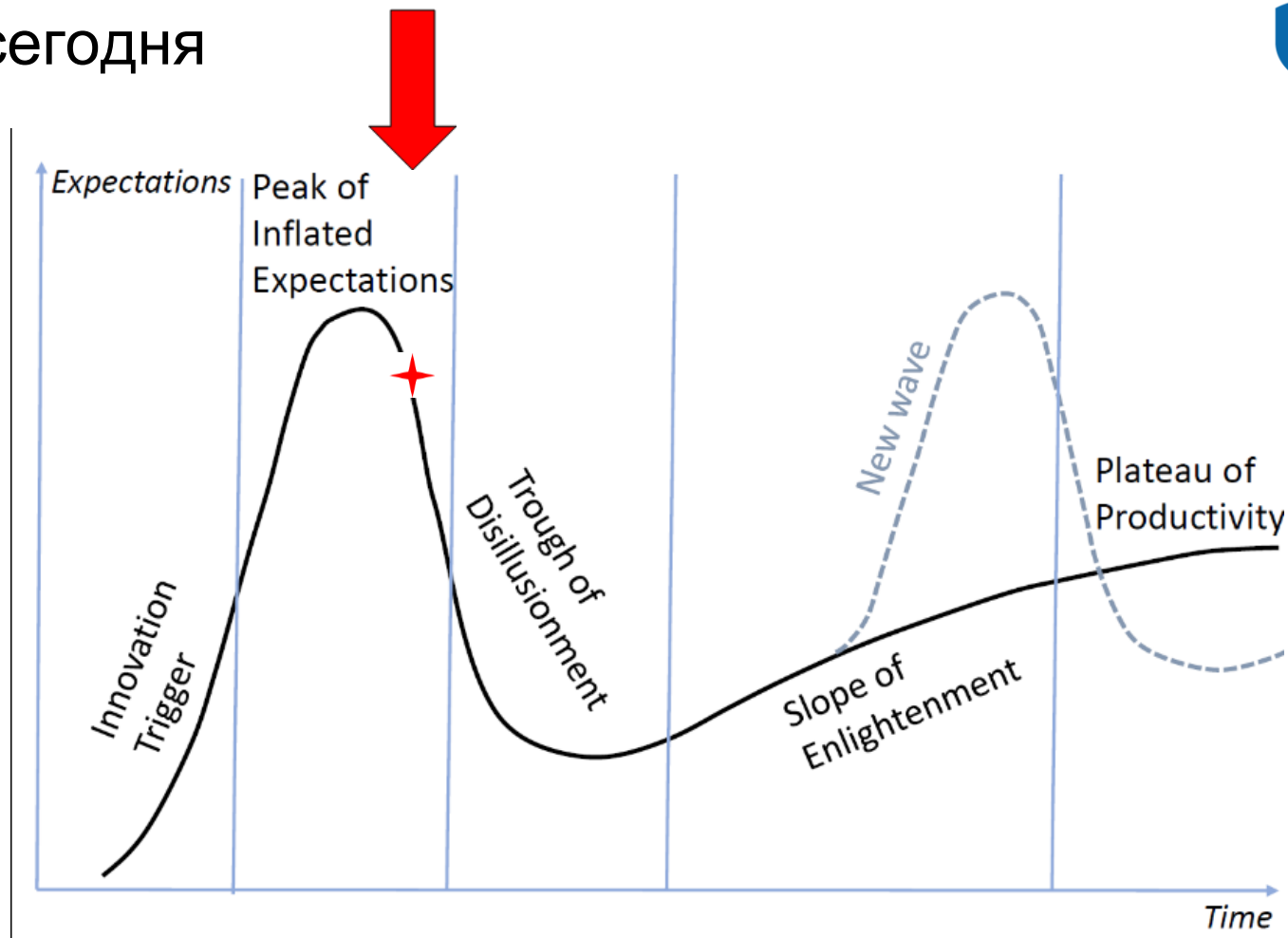
- 1 год. Сбор экспериментальных данных и подготовка вычислимых данных для систем МО и ИИ.
- Разработка нескольких вариантов систем для тестирования.
- 2 год. Подготовка программного обеспечения для обучения. Обучение систем МО и ИИ. Диагностика ошибок. Модификация систем. Продолжение сбора экспериментальных данных и результатов вычислительных экспериментов.
- 3 год. Имплементация методов коррекции ошибок и получение надежных систем ИИ с объясняемыми решениями. Тестирование систем на новых данных. Анализ и приложения результатов.

Проблемы ИИ, основанного на данных, на решение которых нацелен проект

- Ошибки ИИ: как с ними работать и их исправлять
 - быстро (неитерационно),
 - надежно (избегая повторения),
 - без разрушения имеющихся навыков.
- Необъяснимость решений ИИ:
 - как перевести неявную логику ИИ в явную логически прозрачную форму?

Без решения этих проблем нас ждет новая «Зима ИИ» (как минимум, в тех областях, в которых от решений зависит безопасность и жизнь людей).

ИИ сегодня



Использование приложений

Наше кредо

- С самого начала - **реальные приложения**.
- **Актуальность и значимость** приложений.
- **Разнообразие** приложений для **универсальной** технологии.
- Особое место занимают **медико-биологические** приложения.
- Кроме того, мы работаем с динамикой климата, предсказанием экстремальных событий, разработкой новых материалов, квантовых и оптических технологий.
- **Ориентация на лидеров** – в каждой области исследования ведутся лидирующими группами (в национальном и международном масштабе).

Цели и задачи

Создание систем искусственного интеллекта (ИИ) и машинного обучения (МО) нового поколения, способных к выявлению и быстрой коррекции ошибок, в том числе, благодаря имплементации логически объяснимых решений.

Новая технология надежного и объяснимого нейросетевого ИИ будет реализована в стратегически важных задачах реальной сложности, где требование надежности и логической объяснимости является критичным:

- анализ больших биомедицинских данных и выявление сверхранних предикторов заболеваний
- анализ климатических данных и предсказание экстремальных событий
- инженерия новых материалов, квантовых и оптических технологий
- разработка нейросетей, реализующих информационно-вычислительные (интеллектуальные) функции мозга.

Укрупненный план-график проекта

№	Наименование работ	2020	2021	2022
1	Технология надежного и логически прозрачного ИИ			
1.1	Разработка теоретических основ и базовых алгоритмов			
1.2	Имплементации и тестирование, разработка сертификации			
1.3	Разработка пользовательского программного обеспечения			
2	Задачи анализа биомедицинских данных			
2.1	Сбор и подготовка данных для анализа			
2.2	Применение ИИ для анализа механизмов заболеваний			
2.3	Разработка прикладных надежных и логически прозрачных ИИ			
3	Апробация на различных важных задачах			
3.1	Сбор и подготовка данных для обучения ИИ			
3.2	Обучение систем ИИ, диагностика ошибок, модификация систем			
3.3	Разработка прикладных надежных и логически прозрачных ИИ			

Укрупненный план-график первого года

№	Наименование работ
1.1	Разработка теоретических основ и базовых алгоритмов
1.1.1	Разработка технологии неитеративной коррекции ошибок и взаимообучения ИИ
1.1.2	Создание алгоритмов получения явных знаний из данных для нейросетевого ИИ
1.1.3	Разработка алгоритмов многокритериальной оптимизации с использованием МО
1.1.4	Развитие методов ИИ для дизайна математических моделей
1.2	Имплементации и тестирование, разработка сертификации
1.2.1	Имплементация и тестирование новых алгоритмов
1.2.2	Анализ новых типов скрытых атак на ИИ
2.1	Сбор и подготовка данных для анализа
3.1	Сбор и подготовка данных для обучения ИИ

Основная проблема ИИ 2020: ошибки ИИ и их обработка. Ошибки - неизбежные спутники ИИ, управляемого данными.

А.Н. Горбань

«Message 1.» Кейнот пленарный доклад
по нейронным сетям на Всемирном Конгрессе
Вычислительного Интеллекта
(IEEE WCCI2020)

Задел создан во время выполнения мегагранта Горбаня в
ННГУ: «Масштабируемые сети систем искусственного
интеллекта для анализа данных растущей размерности»

Как справиться с ошибками в мире многомерных больших данных?

Нужно внешнее
устройство -
корректор ошибок
ИИ



Технические условия

Корректор должен:

- быть простым;
- не менять навыков унаследованной системы;
- допускать быстрое неитеративное обучение;
- исправлять новые ошибки без разрушения предыдущих исправлений.

Конкуренты:

Donoho – Stanford

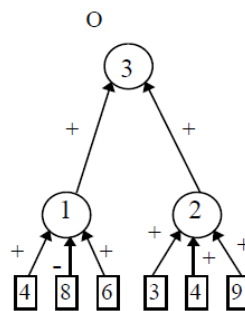
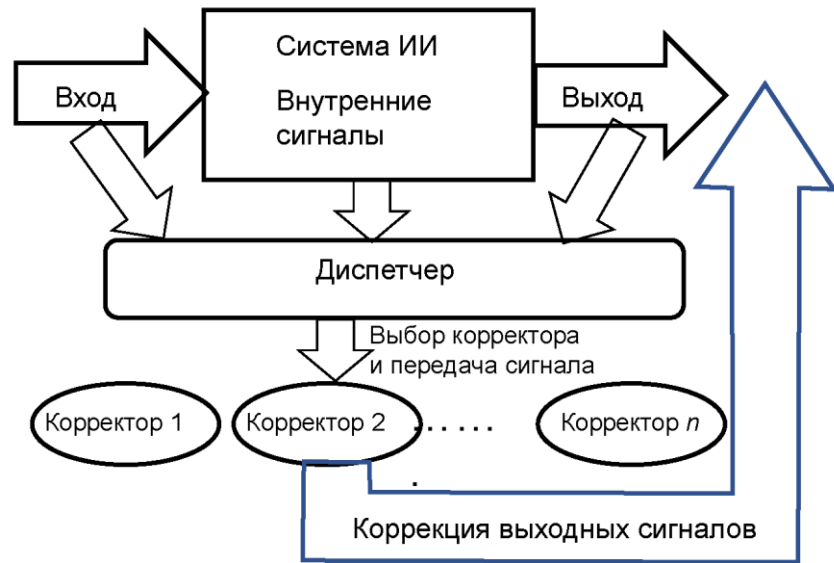
Tanner – Oxford

Vershinin - University of California, Irvine

Технологии МО и ИИ

Рук. проф. А.Н.Горбань

- Создан новый подход к проблеме коррекции и анализу робастности систем ИИ.
- Получены теоремы о стохастической отделимости семейств компактных множеств в больших размерностях — основа построения корректоров.
- Выявлены новые типы уязвимости систем ИИ к внешнему вмешательству.
- Созданы новые методы получения явных знаний из нейронных сетей.
(ХАИ)

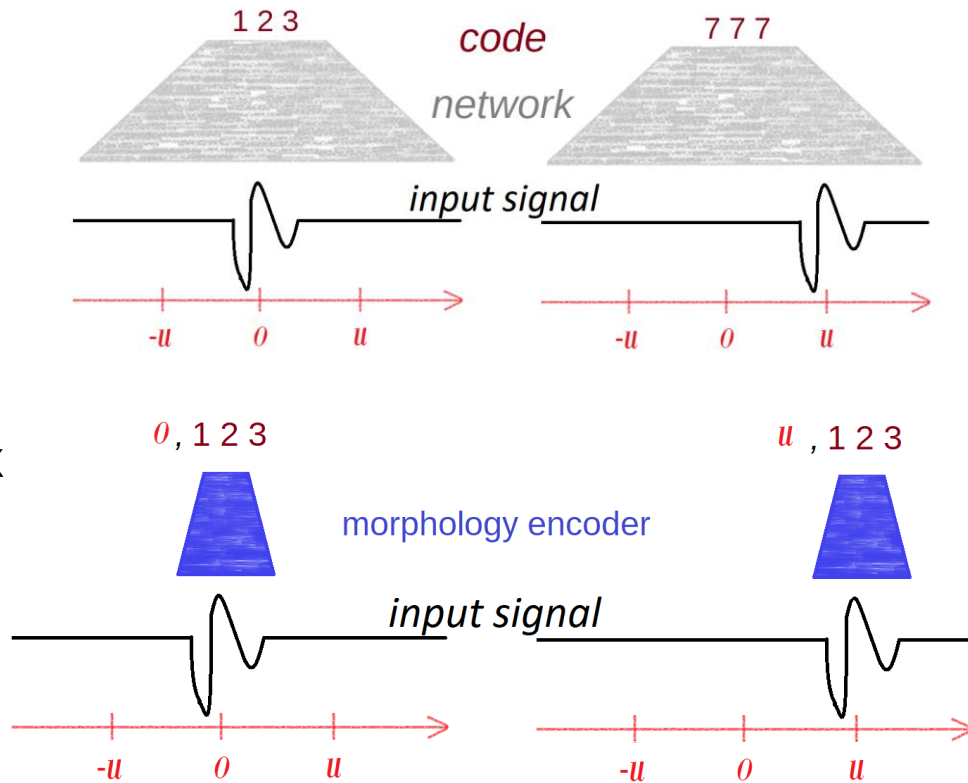


Что нужно для победы оппозиции в США? Конфликт в президентской партии, или Сочетание активности 3-й партии с социальными волнениями, или пассивность президента на фоне экономического спада

Интерпретируемое машинное обучение для анализа ЭКГ

Рук. проф. Г.В. Осипов

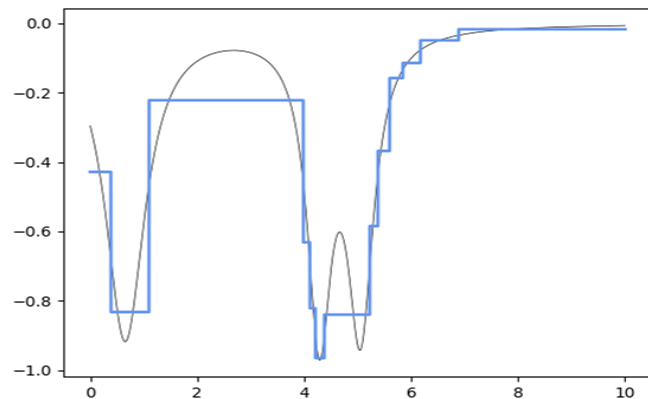
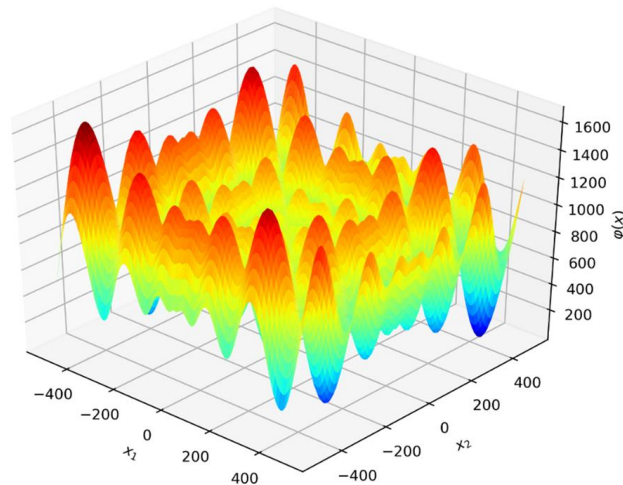
- Исследована возможность использования технологии саккадо-подобного зрения для решения проблемы интерпретируемости глубоких НС при обработке ЭКГ



Интеллектуальный глобальный поиск...

Рук. проф. В.П. Гергель

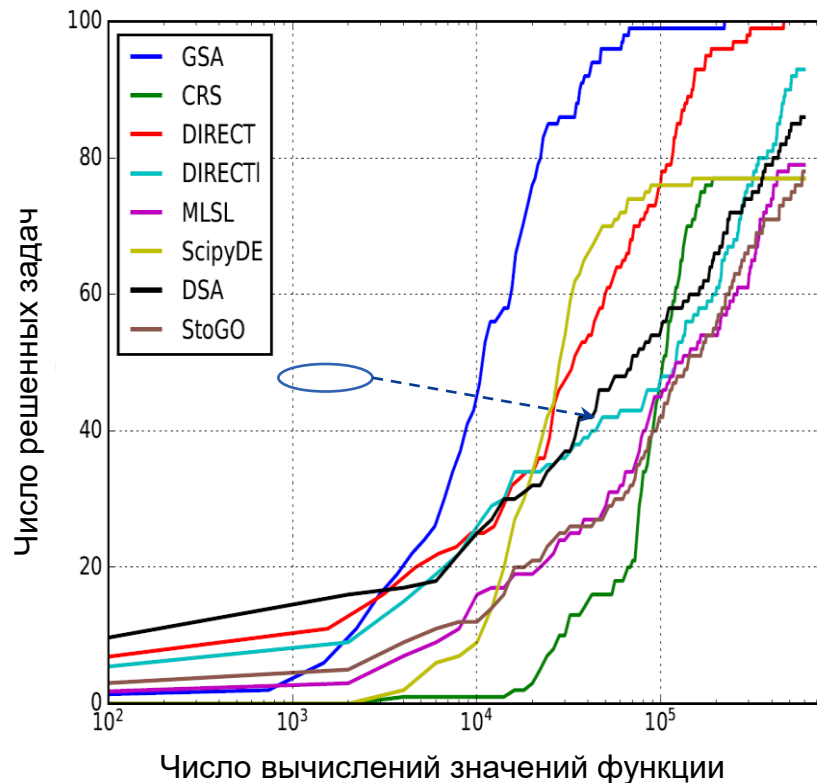
- Разработан новый подход к решению задач многокритериальной глобальной оптимизации.
- Получены теоретические оценки сокращения числа поисковых испытаний, требующихся для решения задачи с заданной точностью.
- Проведены исследования результативности применения методов машинного обучения для повышения эффективности решения задач глобальной оптимизации.



Интеллектуальный глобальный поиск

Рук. проф. В.П. Гергель

- Проблема поиска значений параметров для наиболее эффективного применения методов машинного обучения сформулирована как **задача глобальной оптимизации**.
- Разработанный в рамках выполнения проекта **метод глобального поиска (GSA)** превосходит широко известные алгоритмы (как детерминированные, так и биоинспирированные) по эффективности.

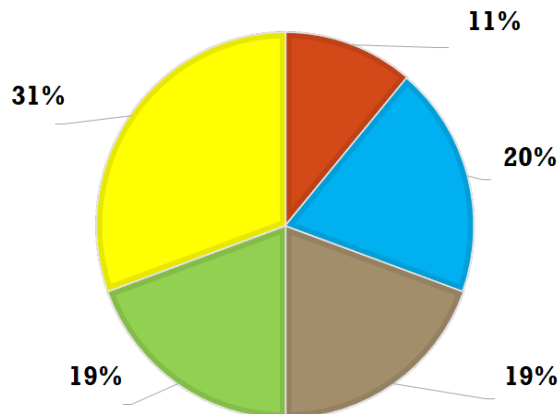


Сбор и анализ биомедицинских данных



Мария Ведунова

тяжелое течение COVID19



- Собрана библиотека биологических образцов плазмы и сыворотки больных COVID-19 в динамике реконвалесценции.
- Проведена систематизация и предобработка клинических данных и характеристика групп больных.
- Проведен первичный анализ транскриптомных данных при экспериментальной неоплазии.
- Разработаны основные подходы к анализу транскриптомных данных клеток иммунной системы.
- Создан протокол сбора библиотеки Биологических образцов для исследования транскриптома онкобольных

Главная проблема: Определение траектории развития патологии («динамического фенотипа») и идентификация точек бифуркации («кризисов»)



Задачи ИИ: определение точек бифуркации и пограничных состояний человека, используя различные методы анализа сложных многомерных клинических данных

Биомедицинские задачи проекта:

1. Анализ и прогноз развития заболеваний ассоциированных с метаболическим синдромом

2. Прогноз развития рецидива и эффективности терапии при обнаружении первичного онкологического заболевания

3. Поиск эпигенетических маркеров развития возраст-ассоциированных заболеваний

4. Анализ функциональной активности нейронных сетей головного мозга

5. Поиск сверххранних маркеров развития осложнений и манифестации хронических заболеваний у больных с COVID-19

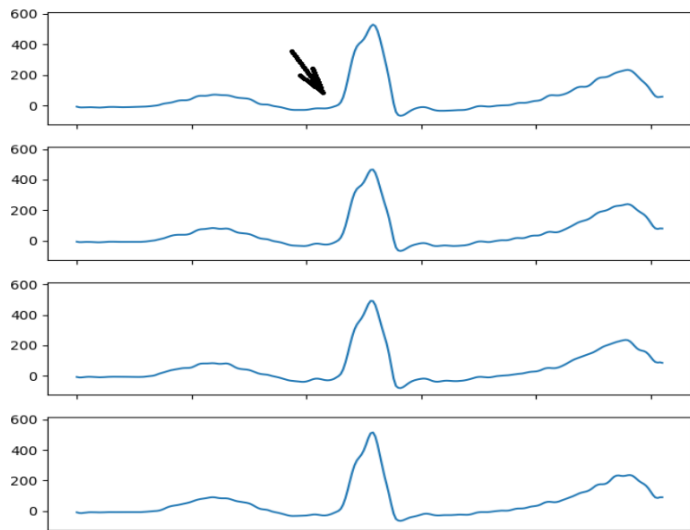
6. Прогнозирование развития эпидемиологического процесса

Анализ клинических данных большой размерности

Григорий
Осипов

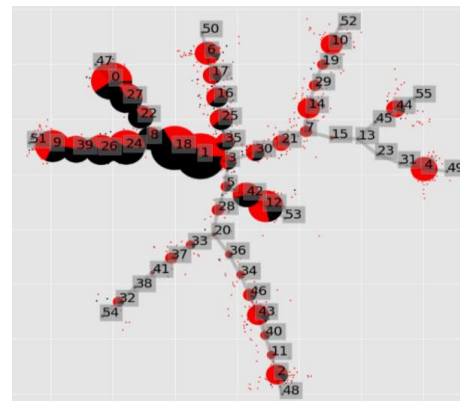


Автоматизированная
диагностика
сердечно-сосудистых
заболеваний



Андрей
Зиновьев

PSL 
UNIVERSITÉ PARIS



Моделирование больших клинических данных как «букета» ветвящихся клинических траекторий. Модели диабета и осложнений инфаркта миокарда.

Динамическое фенотипирование – революция в классификации патологических состояний

Ранние диагностические критерии развития онкопатологии на основе анализа транскриптомных данных

Критерии опухоли-ориентированной активации иммунной системы связанные с долгосрочной ремиссией и оценкой эффективности терапии

**Татьяна
Мищенко**



Экспериментальные
модели онкологии и
активации иммунной
системы

**Ольга
Вершинина**



Анализ
транскриптомных
данных



Сбор
биомедицинских
данных

**Дмитрий
Крысько**



Молекулярная
диагностика и
транскриптомика
при развитии
онкологии

Анализ сверххранних маркеров развития осложнений при COVID-19

- Анализ сверххранних маркеров развития осложнений при COVID-19
- Оценка риска манифестации и прогрессирования хронических заболеваний после перенесенной инфекции COVID-19

Мария
Ведунова



Анализ специфического
иммунологического и
молекулярного профиля
больных COVID-19.
Лонгитюдные исследования



Сбор
биомедицинских
данных

Алексей
Заикин



Новые методы машинного
обучения для поиска
прогностических маркеров
течения заболевания COVID-19

Методы ИИ, машинного обучения и сетей для анализа разнородных данных

Рук. А.А. Заикин

международная
коллаборация:
61 автор, 28 организаций

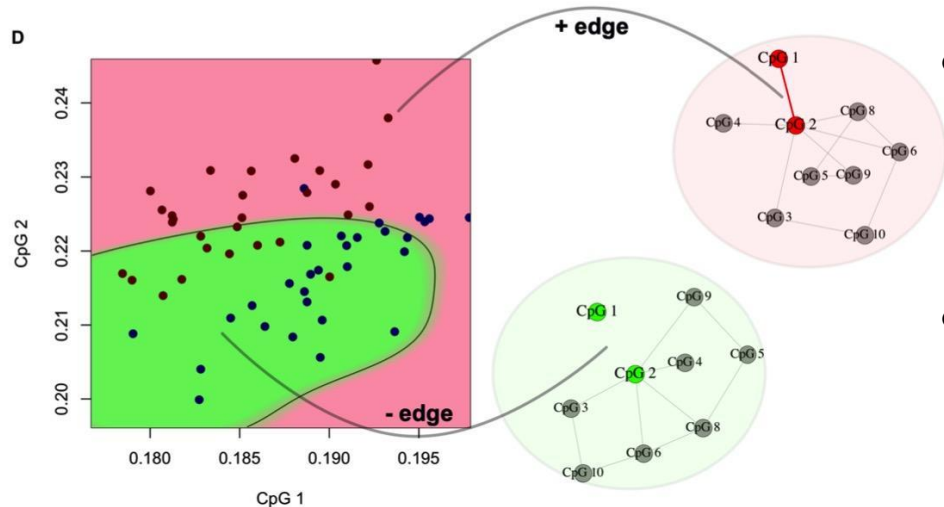


Схема построения
паренклитической сети

- Логически прозрачный интеллект был реализован в виде обучения классификатора по топологии паренклитических графов
- Результат: После перекрестной проверки модель показала отличную точность ($AUC = 0,81$), несмотря на то, что среднее время от выборки до результата составляло 39 дней.
- Особенно удивительные результаты ($AUC = 1$) на полностью независимом наборе данных - 24 пациента с Covid-19, поступивших в больницу Инсбрука в Австрии.

AUC – площадь под ROC кривой

7 Dec 2020

Coronavirus

Covid blood test can predict patient survival chances

Protein analysis provides digital picture of immune response and mortality risk, say scientists

- [Coronavirus - latest updates](#)
- [See all our coronavirus coverage](#)

“We can predict which patients will need oxygen support and ventilator support quite accurately, and we also have markers for patients who are not that severely ill initially, but are at high risk of getting worse,” said Ralser, whose research is published as a [preprint](#) but has not yet been peer reviewed.



▲ It is unlikely that a blood test alone would ever be used to dictate which patients are allocated ICU beds.
Photograph: Murtaja Lateef/EPA

A blood test has been developed that can predict whether Covid patients will need intensive care – or are even likely to survive – shortly after they develop symptoms.



УНИВЕРСИТЕТ
ЛОБАЧЕВСКОГО

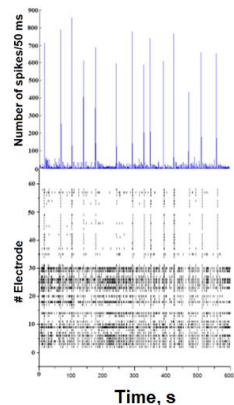
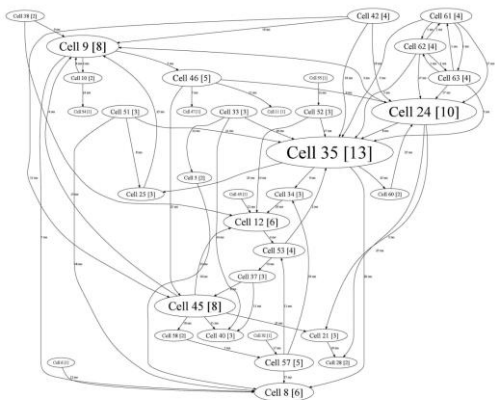
<https://www.theguardian.com/world/2020/dec/07/covid-blood-test-predict-patient-survival-chances>

Реорганизация нейронных сетей головного мозга при ишемических повреждениях и нейродегенеративных заболеваниях



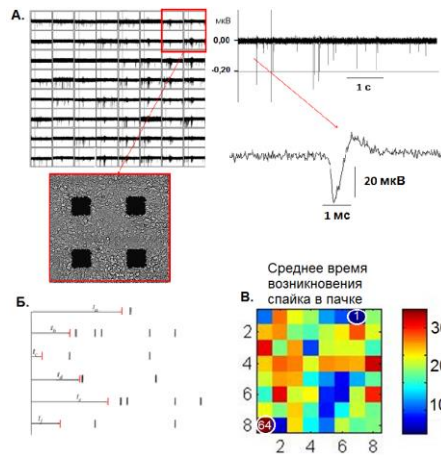
Елена Митрошина

Сбор биологических данных,
разработка новых экспериментальных
моделей патологий головного мозга,
поиск эффективных нейропротекторов



Алена Калякулина

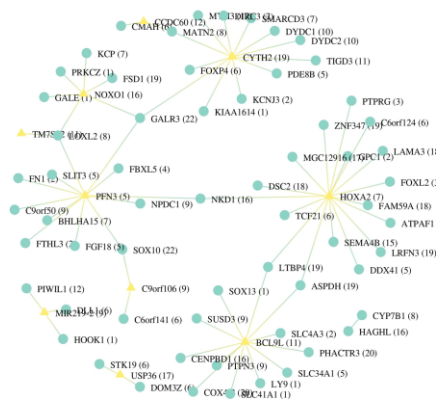
Новые методы оценки
спонтанной биоэлектрической
активности и оценки
функциональной реорганизации
нейронных сетей



УНИВЕРСИТЕТ
ЛОБАЧЕВСКОГО



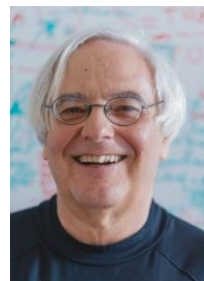
**Михаил
Иванченко**



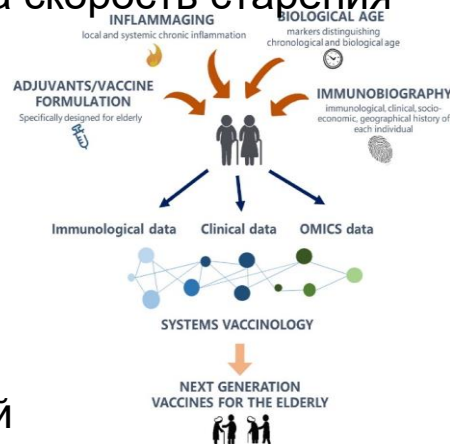
- Построение нейросетей для выявления ранних признаков заболеваний
- Методы прозрачного искусственного интеллекта для выявления маркеров механизмов заболевания и терапевтических мишеней.



Италия
Клаудио
Франчески



Ведущий геронтолог
Европы. Выявление
факторов, влияющих
на скорость старения



Анализ эпигенетических данных. Разработка новых моделей определения биологического возраста на основе многомерных неструктурированных данных



**Елена
Кондакова**

Сбор биобанка, проведение биологических исследований

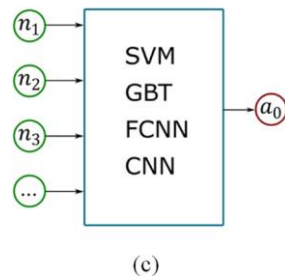
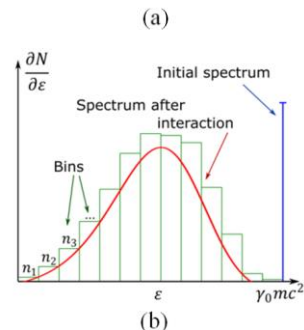
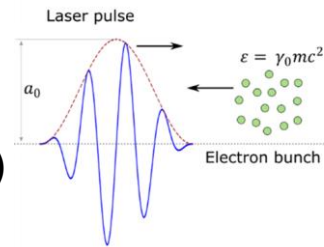
Основные результаты 1-го этапа:

- Разработаны подходы для анализа биологических данных большой размерности, в том числе неструктурированных клинических данных
- Разработана методика сбора биологических данных (малые выборки) с целью верификации разрабатываемых моделей
- Собрана библиотека биологических образцов больных COVID-19 и онкобольных с первично выявленной патологией
- Проведены дополнительные исследования значимых иммунологических и молекулярных маркеров в крови больных

Апробация в задачах лазерной физики

Рук. проф. А.А. Гоносков, доц. И.Б. Мееров

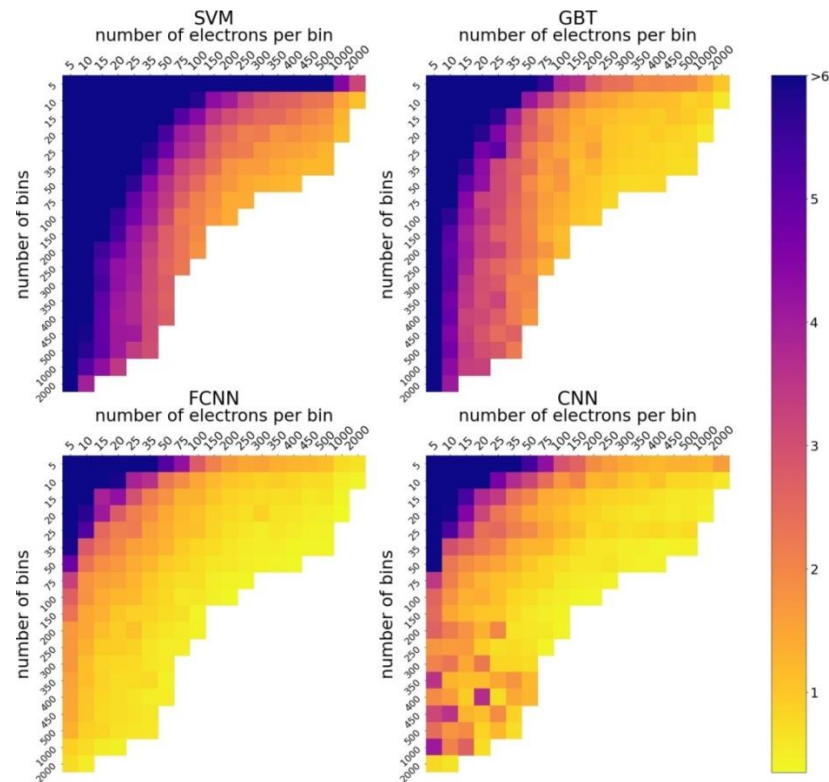
- Изучено, с какой точностью можно восстанавливать пиковую амплитуду лазерного импульса методами Машинного Обучения (МО) (GBT, SVM, FCNN, CNN).
- Сформирован набор данных путем численных моделирований пакетом программ hi-Chi.
- Сконструирована архитектура ИНС, эмпирически настроены параметры методов МО, натренированы модели, исследовано влияние стратегии распределения спектров по бинам на точность восстановления пиковой амплитуды импульса.
- Установлено, что все рассмотренные методы способны решить задачу с высокой точностью, однако полносвязные НС стабильно показывают лучший результат, устойчивый относительно выбора гиперпараметров метода, а также числа электронов и бинов.



Апробация в задачах лазерной физики

Рук. проф. А.А. Гоносков, доц. И.Б. Мееров

- *Перспектива использования результатов:*
управление параметрами лазерных установок
- *Потребность в ХАИ:* необходимость понимания построенных моделей (достоверность и надежность результатов моделирования в дорогостоящих экспериментах)
- *Направления работы:*
 - а) построение новых моделей в задачах, приближенных к эксперименту;
 - б) сокращение размеров построенных сетей (PCA, Principal Graphs...)



Апробация в задачах квантовой динамики

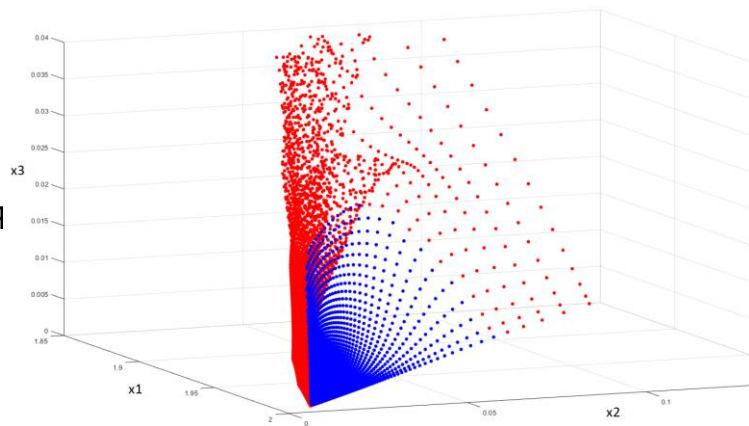
Рук. проф. С.В. Денисов, доц. И.Б. Мееров

- Рассматривается задача о существовании независимых от времени генераторов марковской эволюции, которые дают стробоскопическую динамику, идентичную динамике исходной периодически модулированной квантовой системы.
- Задача сводится к NP-трудной задаче.
- Ведется поиск методов, позволяющих решать задачу методами МО и ИНС.
- Выдвинута гипотеза о разделимости ответов «да»/«нет» при решении вопроса о существовании Флоке-Линдбладана (ФЛ) в пространстве спектров матриц Чои.
- Сгенерированы данные для тренировки и тестирования моделей методами МО.
- Выполнена тренировка моделей и получены первичные результаты их тестирования.

Апробация в задачах квантовой динамики

Рук. проф. С.В. Денисов, доц. И.Б. Мееров

- *Перспектива использования результатов:*
продвижение в понимании процессов в кв. системах.
- *Потребность в ХАИ:* необходимость понимания моделей – выявление сути может дать инструмент для решения задачи и ключ к пониманию фундаментальных физических процессов.
- *Направления работы:*
 - а) построение и упрощение моделей МО;
 - б) изучение полученных методами МО функциональных зависимостей, попытка формирования обобщенных решающих правил;
 - в) выявление связи построенных моделей с физикой квантовых систем с использованием методов ХАИ.



Исследование опасных природных явлений

Состав научной группы:

5 к.ф.-м. н., 3 аспиранта, 2 студента. Из 10 человек 7 моложе 39 лет. Руководитель группы: Стрижак Сергей Владимирович

Актуальность исследований:

Возникновение опасных природных явлений на территории РФ, приводящих к гибели людей и техногенным катастрофам

Вызовы:

Отсутствие достоверных геофизических данных, разработка новых математических моделей, использование искусственного интеллекта

Направление исследований:

Построение математических моделей машинного обучения для задач, связанных с изучением опасных природных явлений
(обледенение модельных тел, потоки на склоне, сдвиговое перемешивание в стратифицированной жидкости).

Содержание работ:

Построение алгоритма обнаружения предвестника опасного природного явления и алгоритма обучения нейронной сети, разработка архитектур нейронной сети (CNN, RNN, FNN) и методики оценки вероятности ошибки, построение классов состояний



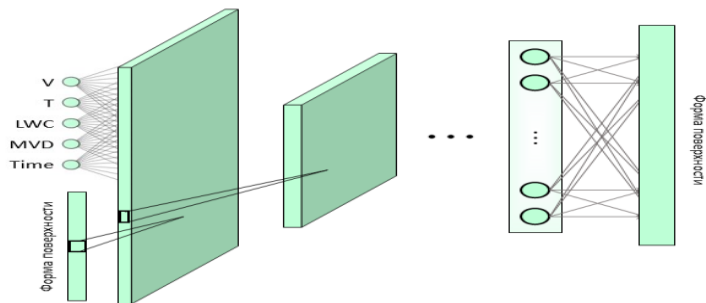
Мировые лидеры и конкуренты в данной области:

Stanford University, USA; ETH Zurich, Switzerland; Princeton University, USA; University of Toronto, Canada; McGill University, Canada; TU Delft, Netherlands; University of Michigan, USA; École Polytechnique, France; Sandia National Laboratories, USA

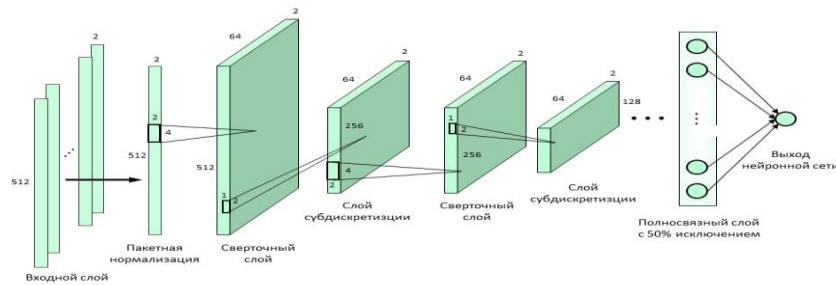
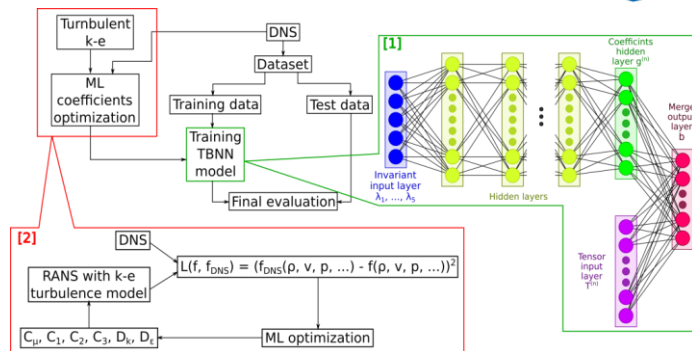
Искусственный интеллект для исследования опасных природных явлений

Алгоритм калибровки турбулентной модели потоков неньютоновской среды на склоне с использованием машинного обучения (Tensor Basis Neural Networks)

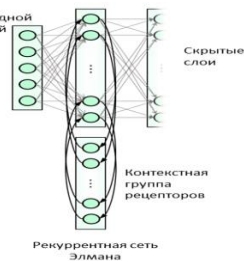
Спектр архитектур (CNN, RNN, RBNN) для задач предсказания ледяного нараста на профиле крыла и задач оценки эффективности перемешивания стратифицированной среды.



Обледенение CNN



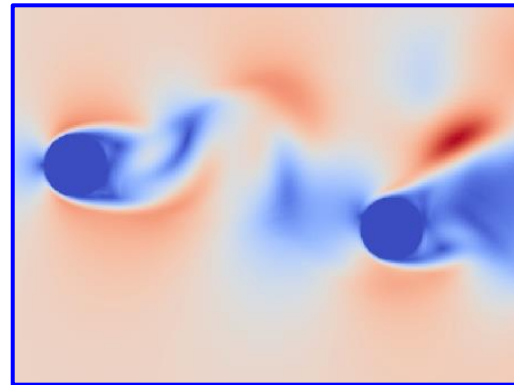
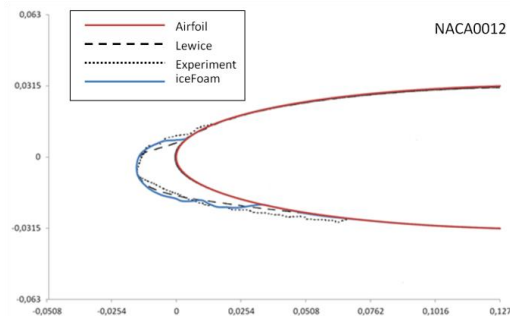
Перемешивание CNN



Элементы RNN для обеих задач

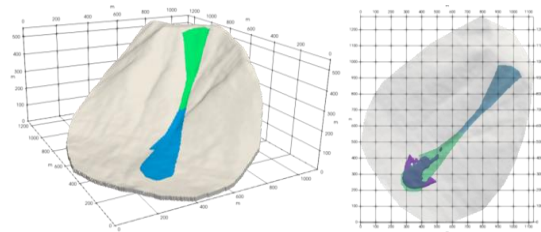
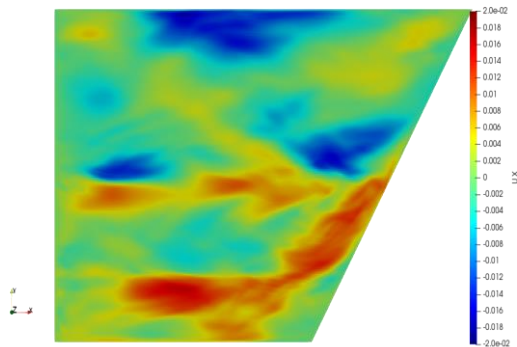
Исследование опасных природных явлений. Результаты за 2020...

- Улучшены классические математические модели для задач, связанных с изучением опасных природных явлений:
 - для задачи обледенения;
 - для склоновых потоков;
 - для задач перемешивания.
- Сгенерирована тестовая база данных профилей различной формы с наросшим льдом – 10000 примеров,
- Поставлены задачи для сбора обучающей выборки задачи перемешивания и схода склонового потока.



Исследование опасных природных явлений. Результаты за 2020

- Разработаны архитектуры моделей машинного обучения на базе сверточных, рекуррентных и полносвязных нейронных сетей для:
 - предсказания формы льда,
 - оценки эффективности перемешивания стратифицированной жидкости,
 - улучшения модели турбулентности в задаче склонового потока.
- Определены входные данные, гиперпараметры, функции потерь.



Исследование опасных природных явлений. Направления. Сравнение.

Направление	ИСП РАН	Конкуренты
Моделирование обледенения тел	Ансамбли CNN и RNN, обледенение вантового моста на остров Русский	École Polytechnique, France, QS rankings=61; Politecnico di Milano, Italy, QS rankings=137 Алгоритмы классификации и регрессии, FNN, CNN
Перемешивание в жидкости	Модели машинного обучения для оценки эффективности перемешивания жидкости в условиях фокусировки внутренних волн в Черном море	University of Toronto, Canada, QS rankings=25 CNN для оценки эффективности перемешивания жидкости в условиях неустойчивости Кельвина-Гельмгольца
Разработка моделей турбулентности	Модели машинного обучения для тензора напряжений с TBNN, модели турбулентности для двухфазной жидкости, снежные лавины в Хибинах и на Кавказе	ETH Zurich, QS rankings=6; TU Delft, Netherlands, QS rankings=57; Sandia National Laboratories, USA; Модели машинного обучения для тензора напряжений (TBNN), турбулентности однофазной жидкости
Обтекание тел, сильные порывы ветра, понижение размерности	Обтекание проводов потоком воздуха ЛЭП на севере и востоке РФ, POD, обработка данных, анализ структуры течений и ошибок	Princeton University, QS rankings=12; McGill University, Canada, QS rankings=31 Гидродинамика, POD, Local ROM, K-Means Clustering

Анализ климатических данных и предсказание экстремальных событий

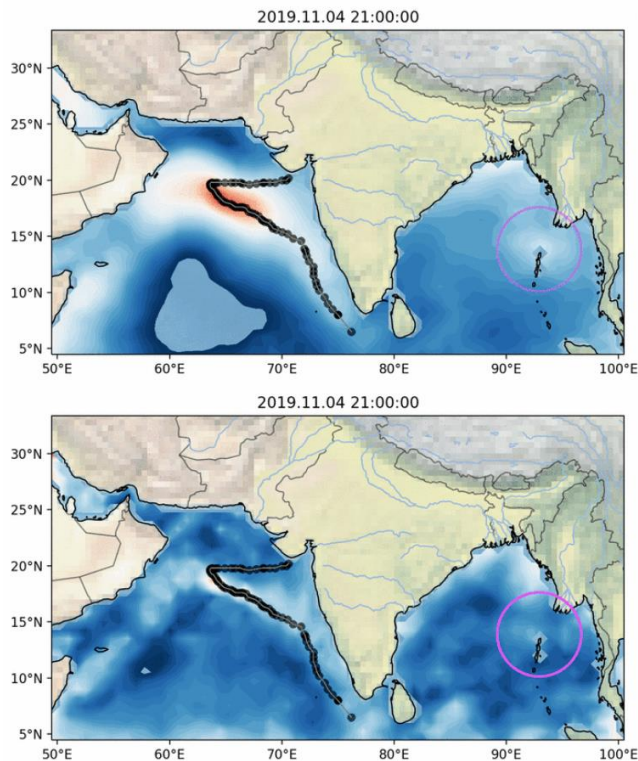
Рук. проф. Ю. Куртс (г. Потсдам, Германия)

Перспектива использования результатов: надежное предсказание экстремальных климатических событий заранее с объяснением логики предсказания.

Текущее состояние: подготовлен и предобработан большой массив климатических данных. Введён в рассмотрение новый графовый индикатор для анализа климатических данных. На основе сетевого анализа рассчитаны сетевые метрики для взвешенных корреляционных графов, снижающие размерность пространства признаков для машинного обучения.

Потребность в ХАИ: необходимость понимания принципов предсказания потенциальной угрозы климатического экстремального события в построенных нейросетевых моделях, а также увеличение надежности системы.

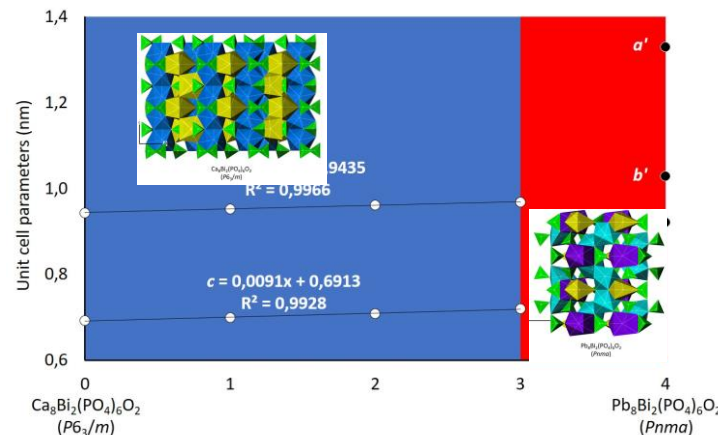
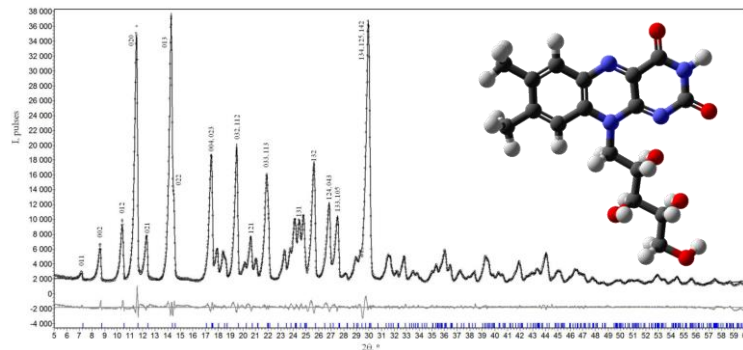
Направления работы: построение предиктивной искусственной нейронной сети на базе сетевых признаков и её упрощение до объяснимого дерева решений. Выявление причин возникновения ошибочных результатов с помощью методов ХАИ.



Методы ИИ для нахождения материалов, обладающих заранее заданными свойствами

Рук. проф. А.Р. Оганов

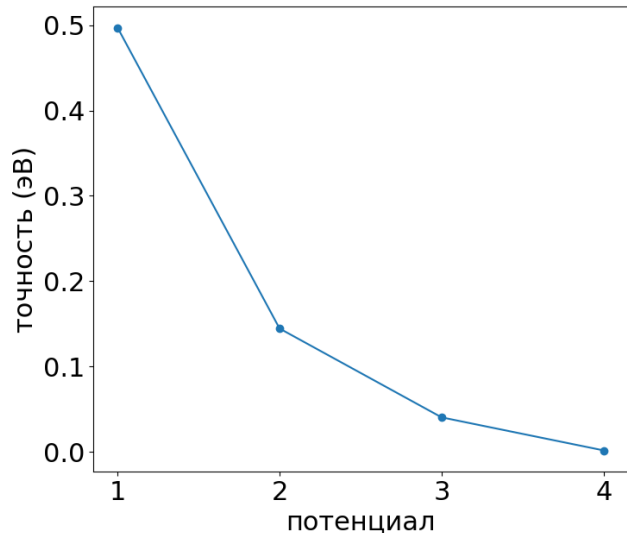
- Разработан алгоритм и программа, позволяющие определять кристаллическую структуру соединений по данным порошковой рентгенографии
- Алгоритм позволил впервые определить кристаллическую структуру витамина B₂
- Впервые синтезирован ряд предсказанных на основе машинного обучения и теории функционала плотности йод-содержащих апатитов для связывания радиоактивных изотопов йода
-



Использование машинного обучения в USPEX

Рук. проф. А. Р. Оганов

- Добавлено сопряжение между программами MLIP и USPEX, чтобы USPEX смог релаксировать структуры и рассчитывать энергию с использованием обученными потенциалами
- Разработана программа MLIP-USPEX, которая генерирует структуры с использованием топологического рандомного генератора от USPEX, обучает потенциалы с использованием MLIP и запускает USPEX для поиска
- На данный момент, программа MLIP-USPEX находится на стадии тестирования



Для релаксации структур и расчёта энергии USPEX использует каскад потенциалов с возрастающей точностью



Создание математических основ для оценки вероятностей ошибок и сертификации робастности и качества машинного обучения



Состав научной группы: 3 д.ф.-м.н., 4 к.ф.-м.н., 4 асп. Из 11 человек 6 моложе 39 лет.

Руководитель группы: д.ф.-м.н. Орлов Юрий Николаевич

ФГУ ФИЦ Институт прикладной математики им. М.В. Келдыша РАН

Содержание работ:

- построение алгоритма обнаружения предвестника заболевания и алгоритма обучения ИИ,
- разработка методики оценки вероятности ошибки,
- построение паттернов классов состояний и алгоритмов классификации

Результаты работ за 2020 г.

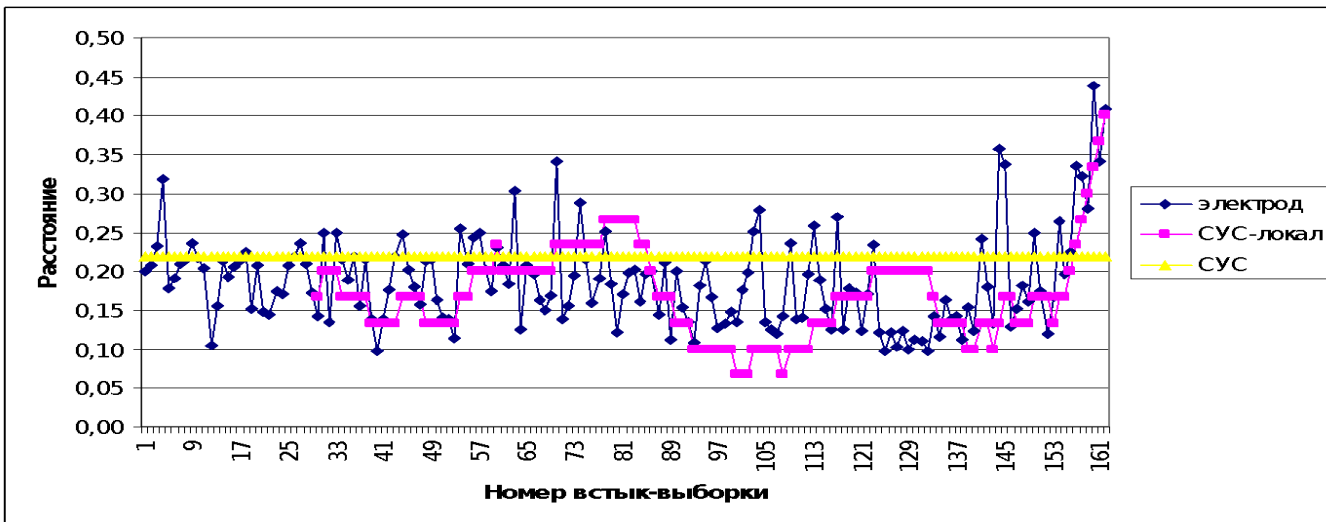
- Проведен анализ методов ИИ в задачах управления динамическими системами со случайным шумом.
- Разработана методика построения индикатора-предиктора разладки для нестационарных временных рядов.
- Построен предиктор приступа эпилепсии по анализу данных электроэнцефалограмм.
- На основе теории эквивалентных полугрупп создан численный алгоритм вычисления уровня нестационарности временного ряда.

Основная задача на 2021 г:

Анализ многомерных временных символьных рядов в текстах различной тематики, приложение методов ИИ



Пример индикации приступа эпилепсии



Каждый раз, когда локальный согласованный уровень стационарности (СУС, розовая кривая) превосходит критическую желтую линию, построенную по нормальному состоянию пациента, фиксируется разрядка.

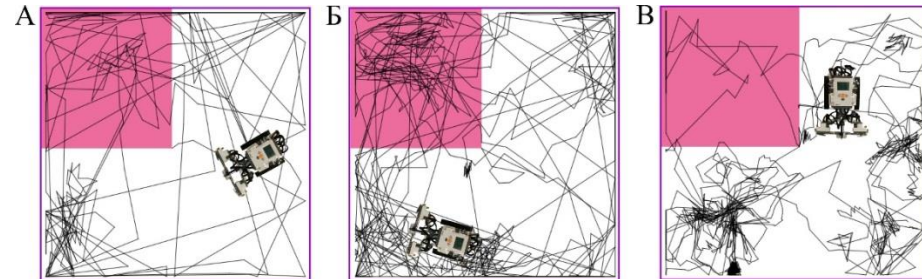
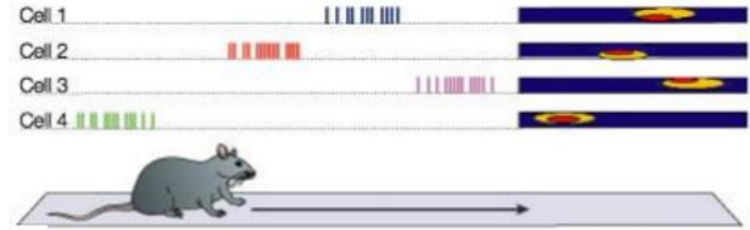
В данном примере приступ эпилепсии происходит в последний момент времени, при этом предиктор сработал на 2 минуты раньше.

Разработка моделей спайковых нейронных сетей, реализующих интеллектуальные функции

Рук. Казанцев В.Б., Макаров С.А.

Моделирование когнитивных нейрональных карт

- Разработана модель пространственной памяти в спайковой нейронной сети, реализующая функционал нейрональных карт животных.
- Предложен новый способ построения глобальной сетевой памяти.
- Модель использована для проектирования самообучающегося мобильного робота.



Разработка прикладных нейросетевых решений ИИ на основе спайковых нейросетей

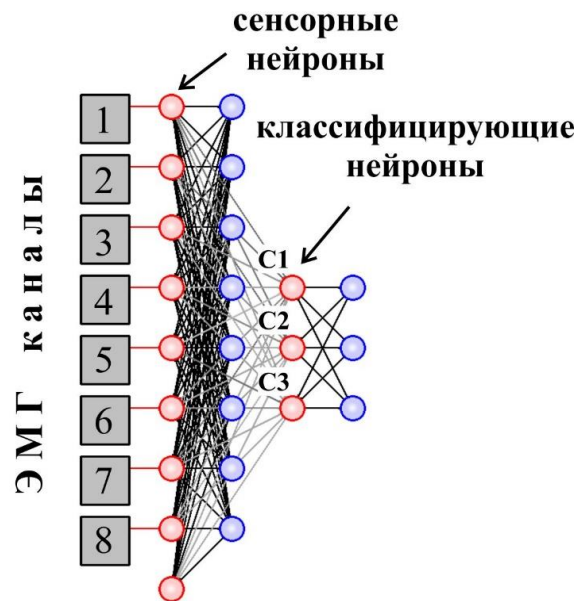
Разработка гибридного нейросетевого классификатора миографических сигналов

Рук. Казанцев В.Б.

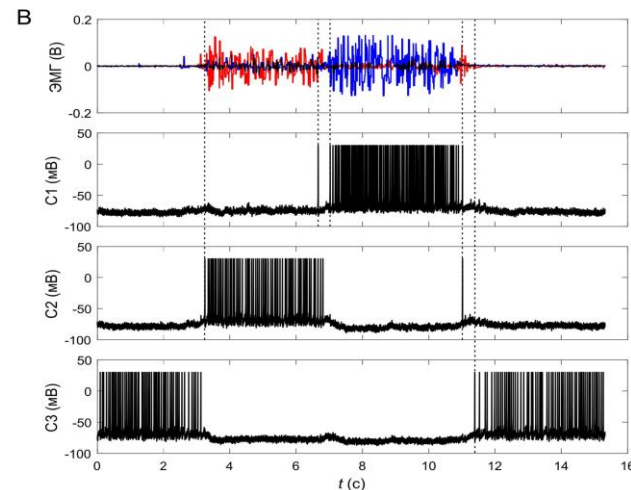
Особенности

- Классификатор реализован с помощью спайковой нейронной сети
- Применяется смешанное частотно-временное кодирование
- Динамика искусственных нейронов в сети отражает динамику реальных моторных нейронов

Архитектура сети

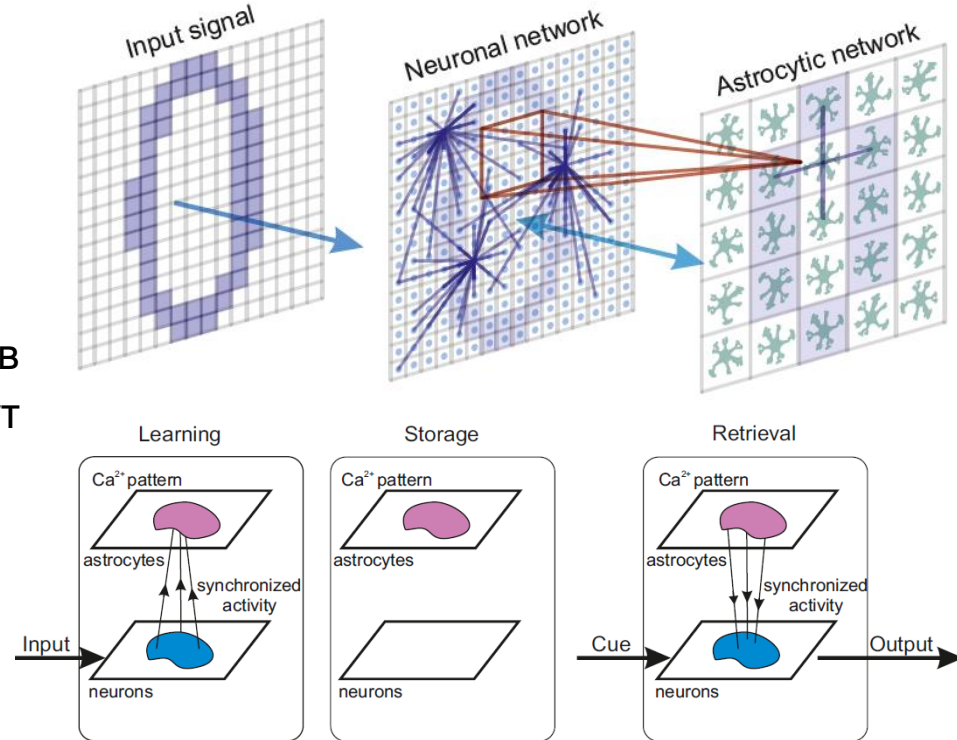


Сигналы в сети после обучения



Новые модели рабочей памяти на основе нейронно-астроцитарных взаимодействий

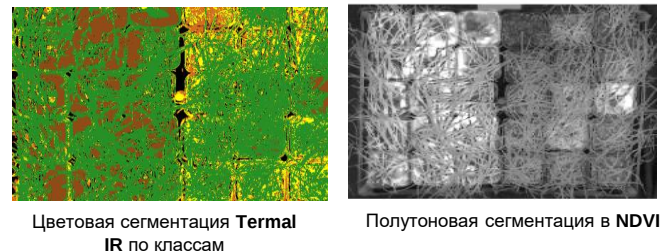
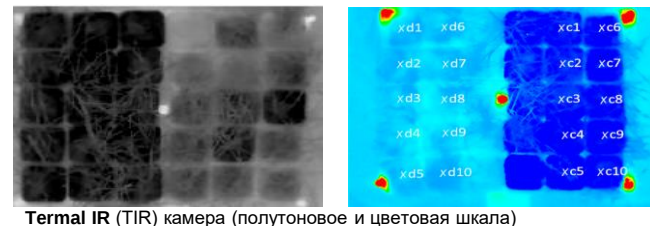
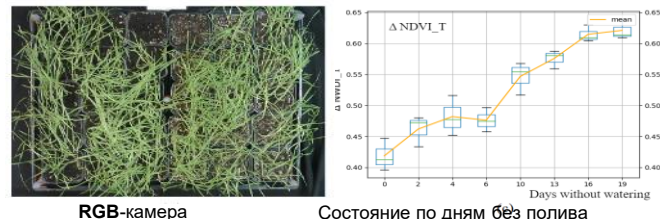
- Создана биологически правдоподобная вычислительная модель рабочей памяти, реализуемая спайкинской нейронной сетью, взаимодействующей с сетью астроцитов.
- Обнаружено, что астроциты, работающие в масштабе времени в десятки секунд, могут успешно хранить следы нейронных активаций, соответствующих информационным паттернам.



Обработка многомерных и мультисенсорных сигналов и изображений для опережающей диагностики стрессовых состояний с/х растений

Рук. Турлапов В.Е., Воденеев В.А.

- Ранняя диагностика засухи пшеницы.
- Построены три группы признаков: STAT{mean, std, min, max}; HIS – квантованные NDVI-гистограммы с бинами соответствующими TIR-бинам; GLCM – для уровней серого по числу бинов.
- Исследована прогностическая способность каждой группы и их сочетаний.
- Построены две XAI-модели: классификации засуха/норма и регрессии.



Обработка многомерных и мультисенсорных сигналов и изображений для опережающей диагностики стрессовых состояний с/х растений

Рук. Турлапов В.Е., Воденеев В.А.

Полутоновая сегментация в $NDVI_T$

Состояние по дням без полива

Эффективность групп и отдельных признаков

Эффективность групп и отдельных признаков

Полутоновая сегментация в $NDVI_G$

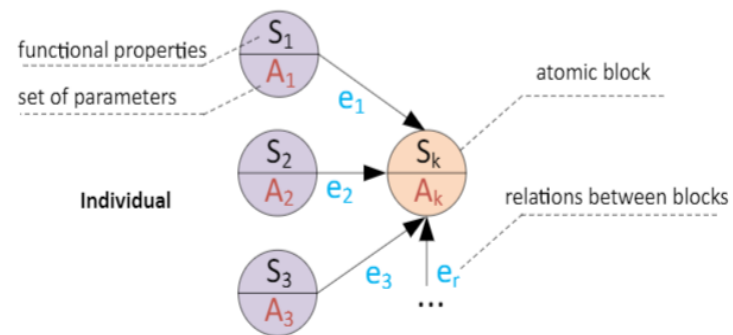
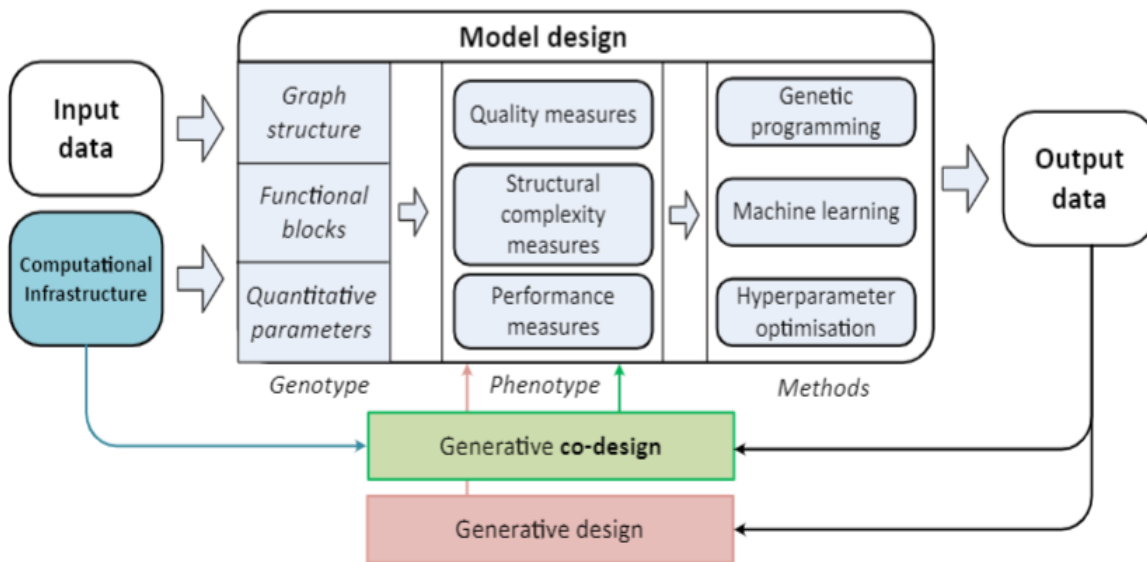
Эффективность групп от
числа нейронов

Точность регрессии от числа
нейронов и уровней квантования

Точность полученного решения по дням

Обобщающая информационная технология: генеративный дизайн математических моделей с элементами ИИ

Рук. А.В. Бухановский А.В.



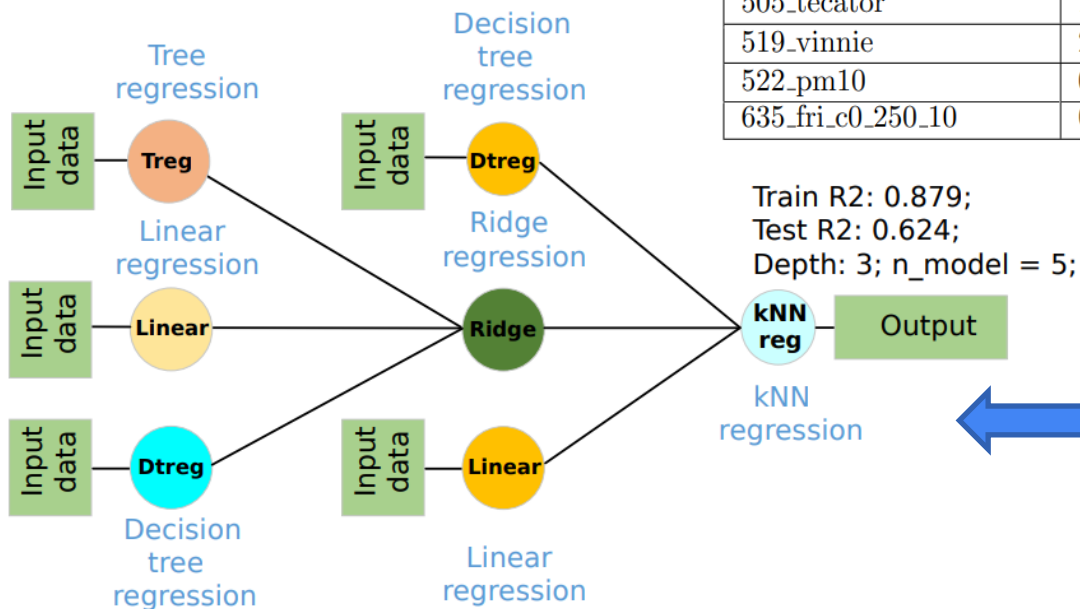
functions					parameters							edges		
S_1	S_2	S_3	$\dots$	$S_{ S_p }$	a_{11}	$\dots$	$a_{1 A_1 ,1}$	$\dots$	$a_{1, S_p }$	$\dots$	$a_{ A_p , S_p }$	e_1	$\dots$	e_r

Структурная формализация идеи генеративного дизайна композитных моделей

Метрики качества по сравнению с решением TPOT:

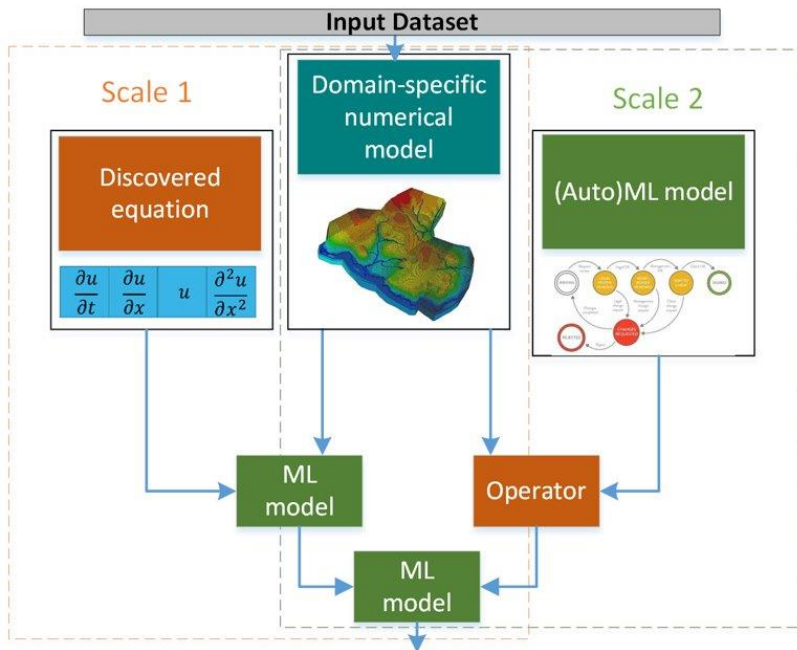


Dataset_name	TPOT.mse	TPOT.r2	Fedot.mse	Fedot.r2
1029_LEV	0.469	0.450	0.492	0.424
201_pol	38.885	0.977	20.503	0.988
210_cloud	0.103	0.933	0.157	0.904
228_elusage	135.500	0.491	100.202	0.624
485_analcatdata_vehicle	52316.870	0.463	28183.72	0.711
505_tecator	1.128	0.995	0.824	0.996
519_vinnie	2.701	0.705	2.486	0.730
522_pm10	0.572	0.214	0.551	0.244
635_fri_c0_250_10	0.171	0.818	0.193	0.800

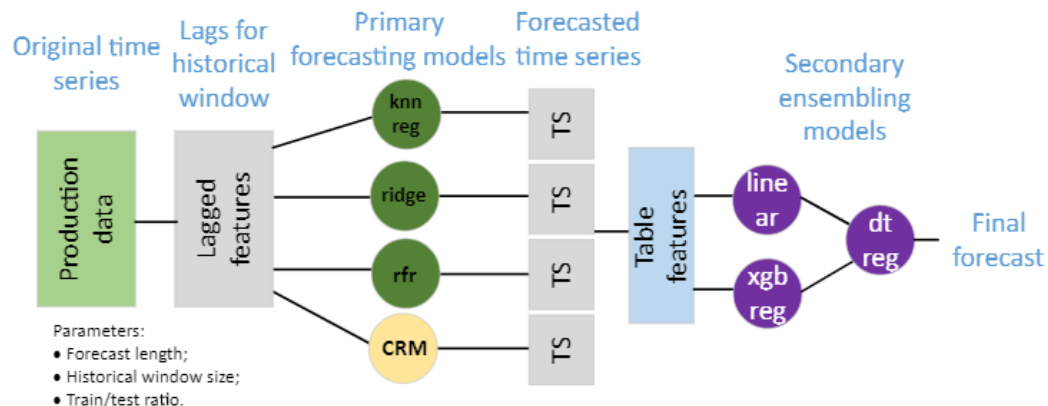


Структура лучшей композитной модели для одного из тестов, найденной генетическим алгоритмом

Развитие: дизайн гибридных моделей в форме ОДУ/ДУЧП



Пример: структура гибридной модели для балансовых процессов



«Выращивание» структуры уравнений на основе комбинации дифференциальных операторов и различных нелинейных членов непосредственно по данным

1. Формализована проблема генеративного дизайна математических моделей в форме направленных ациклических графов с узлами сложной структуры
2. Разработаны программные библиотеки алгоритмов генеративного дизайна композитных моделей
3. Разработанные алгоритмы апробированы на задачах математической физики
4. Обоснован и подготовлен набор данных по эпидемиологии, включая данные по ВИЧ, гриппу и коронавирусу из российских и международных источников – для выращивания моделей эпидемиологической динамики

Показатели

Наименование показателя	Промежуточный результат		
	Год	Плановое значение	Фактическое значение
Количество статей по тематике проекта в изданиях Q1, Q2	2020	6	14
	2021	18	4 опубликовано +8 направлено
Количество диссертаций (канд. и докт.), планируемых к защите	2020	0	0
Доля аспирантов и молодых ученых (до 39 лет)	2020	40%	Всего 63% (В ННГУ 72%)

Расходование средств

<i>Вид расходования средств</i>	<i>Сумма, млн. руб.</i>
Заработная плата с налогами	56,1
Заключение договоров с членами консорциума	35,0
Развитие инфраструктуры, приобретение оборудования, материалов, лицензионного ПО	5,5
Прочие расходы, связанные с содержанием инфраструктуры	2,5
Оплата публикаций Open Access	0,9
<i>Итого</i>	100,0

Что надо сделать

- Апрель-май – приготовить план на год в соответствии с календарным планом (до 15 мая). В частности, указать основные достижения поквартально (примерно) – milestones.
- Подготовить не менее 3 Q1-Q2 публикаций к выходу в 2021. Есть два спецвыпуска: Entropy (Q2) (для авторов, которых в этом спецвыпуске еще не было) и Sensors (Q1). Для них срок представления – 10 октября (чтобы вышло в 2021). Желательно не ограничиваться этими публикациями.
- Встретиться для обсуждения хода работ на Volga Neuroscience.
- Будет работать общий семинар (дважды в месяц). Просим доклады уже по ходу работ – семинар из представительского должен преобразоваться в общий.