

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ УЧРЕЖДЕНИЕ
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ ИМ. М.В. КЕЛДЫША
РОССИЙСКОЙ АКАДЕМИИ НАУК»**

Утверждена
Ученым советом
ИПМ им. М.В. Келдыша РАН,
протокол № 11-23 от «26» октября 2023 г.

РАБОЧАЯ ПРОГРАММА

**УЧЕБНОЙ ДИСЦИПЛИНЫ
«Методы ИИ для анализа нестационарных
потоков данных»**

Научная специальность:
1.2.2. – «Математическое моделирование,
численные методы и комплексы
программ»

Форма обучения:
очная

Москва, 2023

Дисциплина: «Методы ИИ для анализа нестационарных потоков данных»

Научная специальность:

1.2.2 – «Математическое моделирование, численные методы и комплексы программ»

Форма обучения: очная

ИСПОЛНИТЕЛЬ (разработчик программ):

Кислицын А.А., ИПМ им. М.В. Келдыша РАН, заведующий лабораторией, с.н.с., к.ф.-м.н.

РЕЦЕНЗЕНТ: Орлов Юрий Николаевич, Институт прикладной математики им. М.В. Келдыша РАН, заведующий отделом, доктор физико-математических наук, профессор

РАБОЧАЯ ПРОГРАММА РЕКОМЕНДОВАНА

Ученым советом ФГУ ФИЦ ИПМ им. М.В. Келдыша РАН,

№ 11-23 от «26» октября 2023 г

Заведующий аспирантурой _____ / Меньшов И.С. /

Оглавление

Оглавление	2
АННОТАЦИЯ	3
1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	3
2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ	3
3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ.....	4
4. ТЕКУЩАЯ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ. ФОНДОЦЕНОЧНЫХ СРЕДСТВ	6
5. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....	8
6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ	8

АННОТАЦИЯ

Рабочая программа дисциплины «Методы ИИ для анализа нестационарных потоков данных» разработана и составлена на основании ФГТ - «Федеральные государственные требования к структуре программ подготовки научных и научно-педагогических кадров в аспирантуре (адъюнктуре), условиям их реализации, срокам освоения этих программ с учетом различных форм обучения, образовательных технологий и особенностей отдельных категорий аспирантов (адъюнктов)» (Приказ Минобрнауки № 951 от 20.10.2021г.), в соответствии с учебными планами подготовки аспирантов ФГУ «Федеральный исследовательский центр «Институт прикладной математики им. М.В. Келдыша РАН (ИПМ им. М.В. Келдыша РАН) по научным специальностям: 1.2.2. – «Математическое моделирование, численные методы и комплексы программ».

Дисциплина «Методы ИИ для анализа нестационарных потоков данных» реализуется в рамках Блока «Образовательный компонент программы подготовки научных и научно - педагогических кадров в аспирантуре ИПМ им. М.В. Келдыша РАН.

Основным источником материалов для формирования содержания программы являются: научные издания и монографические исследования и публикации, а также материалы конференций, симпозиумов, семинаров, Интернет-ресурсов.

Общая трудоемкость дисциплины по учебному плану составляет 2 зач.ед. (72 часа), из них лекций – 18 часов, семинарских занятий – 0 часов, практических занятий – 0 часов и самостоятельной работы – 54 часов. Дисциплина реализуется на 1-м курсе, в 2-м семестре, продолжительность обучения – 1 семестр.

Текущая аттестация проводится не менее двух раз в соответствии с заданиями и формами контроля, предусмотренными настоящей программой.

Промежуточная оценка знания осуществляется в период зачетно-экзаменационной сессии в форме зачета.

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цели и задачи дисциплины «Методы ИИ для анализа нестационарных потоков данных»

Цель: изучение теоретических основ методов анализа нестационарных потоков данных и практического применения статистики, называемой согласованным уровнем стационарности, к задачам, возникающим при исследованиях, связанных с анализом больших данных; углубленное изучение методов и численных алгоритмов, применяемых для сокращения описания многомерных статистических данных и аппроксимации вычисляемых статистик.

Задачи:

- освоение базовых знаний в области статистического анализа больших данных
- приобретение теоретических знаний в части представления и анализа нестационарных временных рядов в различных предметных областях
- стимулирование к самостоятельной деятельности по освоению дисциплины и формированию необходимых компетенций.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Процесс изучения дисциплины «Методы ИИ для анализа нестационарных потоков данных» направлен на формирование определенных профессиональных знаний, умений и компетенций.

а) универсальные (УК): не предусмотрено

б) общепрофессиональных (ОПК): Владение культурой научного исследования, в том числе с использованием современных информационно-коммуникационных технологий (ОПК-2)

в) профессиональных (ПК): способность самостоятельно разрабатывать и тестировать эффективные вычислительные методы с применением компьютерных технологий (ПК-2), способность самостоятельно решать научные проблемы с применением технологии

математического моделирования вычислительного эксперимента (ПК-3).

В результате освоения дисциплины обучающийся должен:

Знать:

- основные понятия и методы статистического анализа больших данных; модели и средства представления данных.
- новые методы работы с нестационарными данными, разработанные в рамках направления «Методы ИИ для анализа нестационарных потоков данных».

Уметь:

- применять алгоритмы построения статистик, имеющих стационарное распределение для нестационарного случайного процесса, в целях принятия управляющего решения на заданном уровне значимости.
- использовать различные методов и численные алгоритмы для сокращения описания многомерных статистических данных и аппроксимации вычисляемых статистик
- внедрять методы решения задач, разработанные в рамках направления «Методы ИИ для анализа нестационарных потоков данных» в своей проблемной области;

Владеть:

- статистическими методами анализа больших данных; основными понятиями теории графов;
- навыками применения современных технологий искусственного интеллекта в прикладных задачах, связанных с анализом нестационарных потоков данных.

Приобрести опыт:

- применении изученных методик для снижения вычислительной сложности в задачах анализа нестационарных временных рядов
- практической реализации методов анализа нестационарных потоков данных на ряде тестовых задач.

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

3.1. Структура дисциплины

Распределение трудоемкости дисциплины по видам учебных работ

Вид учебной работы	Трудоемкость	
	общая	
	зач.ед.	час.
ОБЩАЯ ТРУДОЕМКОСТЬ по Учебному плану	2	72
Лекции (Л)		18
Практические занятия (ПЗ)	–	–
Семинары (С)	–	–
Самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к семинарским и практическим занятиям) и самостоятельное изучение тем дисциплины		54
<i>Вид контроля: зачет</i>		

3.2. Содержание разделов дисциплины

Общее содержание дисциплины

№ раздела	Наименование раздела	Содержание раздела	Форма текущей аттестации
1.	Согласованный уровень стационарности	Основные понятия и методы статистического анализа больших данных. Оценка вычислительной сложности задач. Эволюция выборочных функций распределения. Расстояния между выборками из стационарной генеральной совокупности. Уровень стационарности встык-выборок.	О, ДЗ

		Двухпараметрический уровень стационарности.	
2.	Эквивалентность по Чернову как метод приближенных вычислений	Теорема Чернова и области ее применения. Понятие эквивалентности по Чернову. Усреднение полугрупп и динамических потоков. Эквивалентная по Чернову стационарная точка уровня значимости. Снижение вычислительной сложности задач.	О, ДЗ
3.	Согласованный уровень стационарности как индикатор-предиктор разладки в нестационарных рядах ЭЭГ	Уровень стационарности электроэнцефалограмм. Оптимизация объема выборки данных. Согласованный уровень стационарности, как индикатор смены режима функционирования динамической системы. Эквивалентность по Чернову СУС для рядов электроэнцефалограмм. Паттерны основных состояний и практические примеры.	О, ДЗ
4.	Коррекция точности прогноза по нестационарным данным	Уровень стационарности каталога землетрясений. Выделение динамической системы из зашумленных данных. Вычисление СУС для потока событий. Аппроксимация отклонений от закона Гуттенберга-Рихтера. Нахождение однородных фрагментов данных. Фильтрация нестационарной составляющей потока. Аппроксимация параметра потока событий.	О, ДЗ
5.	Анализ влияния невесомости на локомоции человека	Особенности биометрических данных. Методики выделения паттерна шага. Индикаторы-характеристики движения. Сравнительный анализ локомоторных данных, полученных в рамках эксперимента до полета и после полета.	О, ДЗ
6.	Анализ статистик биржевых рядов.	Выбор инструмента для анализа по индексу нестационарности ряда. Методика выделения характерных длин выборок и индикация разладки. Тиковые сделки: фильтрация нестационарности. Оценка эффективности стохастического управления.	О, ДЗ
7.	Генерация ансамбля нестационарных траекторий	Эмпирическое уравнение Лиувилля. Нарушение полугруппового свойства при численном решении. Модель аппроксимации эволюции выборочной плотности. Генерация ансамбля нестационарных траекторий. Оптимизация функционала стохастического управления.	О, ДЗ
8.	Семантический анализ и структура графа толкового словаря	Основные понятия теории графов. Граф связей понятий толкового словаря. Структура сильно связной компоненты. Выделение понятийных классов. Смысл слова как циклический подграф.	О, ДЗ
9.	Графов ближайших соседей: понятия, свойства и применение.	Свойства графов ближайших соседей. Построение бенчмарка структур графов ближайших соседей, и его применение в качестве индикатора зависимости данных.	О, ДЗ

Примечание: О – опрос, Д – дискуссия (диспут, круглый стол, мозговой штурм, ролевая игра), ДЗ – домашнее задание (эссе и пр.). Формы контроля не являются жесткими и могут быть заменены преподавателем на другую форму контроля в зависимости от контингента обучающихся. Кроме того, на занятиях семинарских может проводиться работа с нормативными документами, изданиями средств информации и прочее, что также оценивается преподавателем.

3.3. Лекционные занятия

№ занятия	№ Раздела	Краткое содержание темы занятия	Кол-во часов
1.	1	Нестационарные временные ряды: основные понятия, свойства, методы анализа. Стационарная точка уровня значимости и ее зависимость от длины выборки.	2
2.	2	Теорема Чернова и ее практическое приложение. Усреднение динамических потоков по Чернову. Применение изученных методик для снижения вычислительной сложности в задачах анализа нестационарных временных рядов.	2
3.	3	Методика проведения численных экспериментов по индикации разладки применительно к нестационарным рядам электроэнцефалограмм.	2
4.	4	Методика применения СУС для проведения фильтрации потока событий на примере каталога землетрясений и выделения нестационарных фрагментов. Аппроксимация отклонений от закона Гуттенберга-Рихтера.	2
5.	5	Статистические и нейросетевые методы машинного обучения для анализа биомедицинских данных: выделения паттернов, сегментация данных, выделение режимов.	2
6.	6	Методы анализа биржевых рядов. Составление статистики расстояний между распределениями, вычисление СУС, нахождение оптимального объема выборки для прогнозирования с заданной точностью на заданный горизонт. Составление и решение кинетического уравнения для моделирования эволюции выборочной плотности распределения.	2
7.	7	Метод генерации пучка нестационарных траекторий с заданными свойствами эволюции выборочных распределений.	2
8.	8	Введение в теорию графов. Семантический анализ и структура графа толкового словаря.	2
9.	9	Графов ближайших соседей: понятия, свойства и применение.	2
ВСЕГО			18

4. ТЕКУЩАЯ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Текущая аттестация аспирантов. Текущая аттестация аспирантов проводится в соответствии с локальным актом ФГУ ФИЦ ИПМ им. М.В. Келдыша РАН — Положением о текущей, промежуточной и итоговой аттестации аспирантов ФГУ ФИЦ ИПМ им. М.В. Келдыша РАН по программам подготовки научно-педагогических кадров в аспирантуре, — и является обязательной.

Текущая аттестация по дисциплине проводится в форме опроса, сдачи ряда учебных программ по всем разделам курса, а также оценки вопроса–ответа в рамках участия обучающихся в дискуссиях и различных контрольных мероприятиях по оцениванию фактических результатов обучения, осуществляемых преподавателем, ведущим дисциплину. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины см. ниже.

Объектами оценивания выступают:

- учебная дисциплина – активность на занятиях, своевременность выполнения различных видов заданий, посещаемость занятий;
- степень усвоения теоретических знаний и уровень овладения практическими умениями и навыками по всем видам учебной работы, проводимых в рамках семинаров, практических занятий и самостоятельной работы.

Оценивание обучающегося на занятиях осуществляется с использованием нормативных оценок по 4-х бальной системе (5 – отлично, 4 – хорошо, 3 – удовлетворительно, 2 – неудовлетворительно).

Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины.

Форма контроля знаний	Вид аттестации	Примечание
проверочные работы в течение всего курса, прием домашних заданий в форме практической реализации изученных методов	текущая	Ниже приведены перечни рекомендуемых задач и контрольных вопросов
зачет	итоговая	

Примерный перечень рекомендуемых контрольных вопросов для оценки **текущего** уровня успеваемости аспиранта:

1. Согласованный уровень стационарности.
2. Непараметрический критерий Колмогорова-Смирнова
3. Эквивалентность по Чернову
4. Методы выделения динамической системы из зашумленных данных
5. Оптимизация длин выборок в задачах стохастического управления
6. Статистические и нейросетевые методы машинного обучения для анализа биомедицинских данных
7. Основные понятия теории графов.
8. Граф ближайших соседей: понятия и основные свойства

Итоговая аттестация аспирантов. Итоговая аттестация аспирантов по дисциплине проводится в соответствии с локальным актом ФГУ ФИЦ ИПМ им. М.В. Келдыша РАН – Положением о текущей, промежуточной и итоговой аттестации аспирантов ИПМ им. М.В. Келдыша РАН по программам подготовки научно-педагогических кадров в аспирантуре и является обязательной.

Итоговая аттестация по дисциплине осуществляется в форме зачета в период зачетно-экзаменационной сессии в соответствии с Графиком учебного процесса по приказу (распоряжению заместителю директора по научной работе). Обучающийся допускается к зачету в случае выполнения аспирантом всех учебных заданий и мероприятий, предусмотренных настоящей программой. В случае наличия учебной задолженности (пропущенных занятий и (или) невыполненных заданий) аспирант отрабатывает пропущенные занятия и выполняет задания.

Оценивание обучающегося на промежуточной аттестации осуществляется с использованием нормативных оценок на зачете – зачет, незачет.

Список вопросов к зачету:

1. Согласованный уровень стационарности.
2. Непараметрический критерий Колмогорова-Смирнова
3. Эквивалентность по Чернову
4. Методы выделения динамической системы из зашумленных данных
5. Оптимизация длин выборок в задачах стохастического управления
6. Статистические и нейросетевые методы машинного обучения для анализа биомедицинских данных
7. Основные понятия теории графов.
8. Граф ближайших соседей: понятия и основные свойства

Оценивание аспиранта на промежуточной аттестации в форме зачета

Оценка	Требования к знаниям и критерии выставления оценок
--------	----------------------------------------------------

Незачет	основное содержание учебного материала не раскрыто; допущены грубые ошибки в определении понятий и при использовании терминологии; не даны ответы на дополнительные вопросы.
Зачет	раскрыто содержание материала, даны корректные определения понятий; допускаются незначительные нарушения последовательности изложения; допускаются небольшие неточности при использовании терминов или в логических выводах; при неточностях задаются дополнительные вопросы.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература

1. Кслицын А.А., Орлов Ю.Н. Моделирование индикаторов разладки в нестационарных временных рядах: Монография. – М.: РУДН, 2021. – 241 с
2. Ширяев А.Н. Стохастические задачи о разладке. – М.: МЦНМО, 2016. – 392 с.
3. Орлов Ю.Н. Кинетические методы исследования нестационарных временных рядов. – М.: МФТИ, 2014. – 277 с.

Дополнительная литература и Интернет-ресурсы

4. Колчин В.Ф. Случайные графы. – М.: URSS, 2021. – 255 с.
5. Уилкс С. Математическая статистика (пер. с англ.). – М.: Наука, 1967. – 632 с.
6. Гутенберг Б., Рихтер К.Ф. Сейсмичность Земли и связанные с ней явления (пер. с англ.). – М.: ГИИЛ, 1948. – 160 с.
7. Biau G., Devroye L., Lectures on the Nearest Neighbor Method. – М.: Springer Series in the Data Sciences, 2015. – 284 p.
8. Srivastava A., Sahami M. Text mining: classification, clustering, and applications. – М.: Chapman & Hall/CRC, 2009. – 308 p.
9. Кислицын А.А. Анализ каталога землетрясений // Математическое моделирование, 2023.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Для обеспечения интерактивных методов обучения для чтения лекций требуется аудитория с мультимедиа (возможен вариант с интерактивной доской) или белая доска под фломастеры.

Для проведения дискуссий и круглых столов, возможно, использование аудиторий со специальным расположением столов и стульев.

ИСПОЛНИТЕЛИ (разработчики программы):

Кислицын А.А., ИПМ им. М.В. Келдыша РАН, заведующий лабораторией, с.н.с., к.ф.-м.н.