

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

УТВЕРЖДЕНО

решением Ученого совета факультета математики,
информационных и авиационных технологий
от «16» мая 2023 г. протокол 4/23

Председатель  М. А. Волков



«16» мая 2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Дисциплина	Современные методы анализа массивов и потоков данных
Факультет	Математики, информационных и авиационных технологий
Кафедра	Прикладной математики
Курс	1-2

Направление (специальность): 01.04.02 Прикладная математика и информатика

Направленность (профиль/специализация): Имитационное моделирование и анализ данных

Форма обучения: очная

Дата введения в учебный процесс УлГУ: 1 сентября 2023 г.


Программа актуализирована на заседании кафедры: протокол № ____ от ____ 20__ г.

Программа актуализирована на заседании кафедры: протокол № ____ от ____ 20__ г.

Программа актуализирована на заседании кафедры: протокол № ____ от ____ 20__ г.

Сведения о разработчиках:

ФИО	Кафедра	Должность, ученая степень, звание
Савинов Ю.Г.	ПМ	Доцент, к.ф.м.н., доцент

СОГЛАСОВАНО
Заведующий выпускающей кафедрой прикладной математики
 / Бутов А.А. / Подпись / ФИО «16» мая 2023 г.

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ:

Учебная дисциплина «Современные методы анализа массивов и потоков данных» знакомит студентов с современными методами интеллектуального анализа данных. Она является одной из обязательных дисциплин по направлению подготовки 01.04.02 Прикладная математика и информатика.

Она связана с дисциплиной "Технология хранения и обработки больших объемов информации".

Цель дисциплины: ознакомление с современными методами анализа массивов и потоков данных.

Основной **задачей** изучения дисциплины является расширение имеющихся знаний о современных методах классификации и регрессии, методах многомерного статистического анализа, методах статистической обработки нечисловых данных, машинного обучения и нейронных сетей.

Дисциплина «Современные методы анализа массивов и потоков данных» базируется на знаниях и умениях, полученных студентами при изучении курсов (по программам бакалавриата или специалитета): математический анализ, алгебра и геометрия, дифференциальные уравнения, теория вероятностей и математическая статистика, численные методы.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП:

Дисциплина «Современные методы анализа массивов и потоков данных» относится к дисциплинам Обязательной части Блока 1 «Дисциплины (модули)» Основной Профессиональной Образовательной Программы по направлению 01.04.02 Прикладная математика и информатика.

Данная дисциплина базируется на входных знаниях, умениях, навыках и компетенциях студента, полученных им при изучении предшествующих учебных дисциплин, указанных в Приложении к данной рабочей программе (в фондах оценочных средств – далее ФОС, пункт 1): Современные проблемы прикладной математики и информатики, История и методология прикладной математики и информатики, Современные компьютерные технологии, Технология хранения и обработки больших объемов информации.

Результаты освоения дисциплины будут необходимы для прохождения всех видов практик и государственной итоговой аттестации.

3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ), СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОСНОВНОЙ ПРОФЕССИОНАЛЬНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Код и наименование реализуемой компетенции	Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с индикаторами достижения компетенций
ОПК-1 Способность решать актуальные задачи фундаментальной и прикладной математики	знать основные методы анализа и статистической обработки данных различной природы

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

ОПК-2 Способность совершенствовать и реализовывать новые математические методы решения прикладных задач	владеть , иметь опыт применения современных программных пакетов статистической обработки данных различной природы.
ОПК-3 Способность разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности	уметь выбирать для анализа данных наиболее подходящие методы прикладной статистики и соответствующие программные комплексы
ОПК-4 Способность комбинировать и адаптировать существующие информационнокоммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности	уметь выбирать для анализа данных наиболее подходящие методы прикладной статистики и соответствующие программные комплексы; владеть , иметь опыт применения современных программных пакетов статистической обработки данных различной природы.

4. ОБЩАЯ ТРУДОЕМКОСТЬ ДИСЦИПЛИНЫ

4.1. Объем дисциплины в зачетных единицах (всего): 7 зачетных единиц.

4.2. Объем дисциплины по видам учебной работы (в часах):

Вид учебной работы	Количество часов (форма обучения: очная)		
	Всего по плану	В т.ч. по семестрам	
		2	3
1	2	3	4
Контактная работа обучающихся с преподавателем в соответствии с УП	36/36	18	18
Аудиторные занятия	36/36	18	18
Лекции	36/36	18	18
Семинары и практические занятия	–	–	–
Лабораторные работы, практикумы	–	–	–
Самостоятельная работа	180	90	90
Форма текущего контроля знаний и контроля самостоятельной работы: тестирование, контр. работа, коллоквиум, реферат и др. (не менее 2 видов)	устный опрос, 2 реферата	устный опрос, реферат	устный опрос, реферат
Курсовая работа	–	–	–
Виды промежуточной аттестации (экзамен, зачет)	экзамен 36	зачет	экзамен 36
Всего часов по дисциплине	252	108	144

**В случае необходимости использования в учебном процессе частично/исключительно дистанционных образовательных технологий в таблице через слеш указывается количество часов работы ППС с обучающимися для проведения занятий в дистанционном формате с применением электронного обучения*

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

4.3. Содержание дисциплины (модуля). Распределение часов по темам и видам учебной работы:

Форма обучения: очная.

Название разделов и тем	Всего	Виды учебных занятий					Форма текущего контроля знаний
		Аудиторные занятия			Занятия в интерактивной форме	Самостоятельная работа	
		Лекции	Практические занятия, семинары	Лабораторные работы, практикумы			
1	2	3	4	5	6	7	8
2 семестр							
Раздел 1. ВВЕДЕНИЕ							
1.1. Предмет дисциплины. Исторические сведения о развитии методов анализа данных	6	1				5	устный опрос
1.2. Типы данных. Количество данных. Этапы анализа и обработки данных.	6	2				5	устный опрос
Раздел 2. Основные статистические методы анализа и обработки данных							
2.1. Кластерный анализ	18	3				16	устный опрос, проверка реферата
2.2. Факторный анализ	18	3				16	устный опрос, проверка реферата
2.3. Регрессионный анализ	18	3				16	устный опрос, проверка реферата
2.4. Дисперсионный анализ	18	3				16	устный опрос, проверка реферата
2.5. Корреляционный анализ	18	3				16	устный опрос, проверка реферата
3 семестр							

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

Раздел 3. Кибернетические методы анализа и обработки данных							
3.1. Деревья решений	24	4				20	устный опрос, проверка реферата
3.2. Метод опорных векторов. Метод "ближайшего соседа". Байесовская классификация	24	4				20	устный опрос, проверка реферата
3.3. Нейронные сети второго поколения	35	5				30	устный опрос, проверка реферата
3.4. Нейронные сети третьего поколения	25	5				20	устный опрос, проверка реферата
Экзамен	36						
Всего	252	36				180	

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Раздел 1. Введение

Тема 1.1 Предмет дисциплины. Исторические сведения о развитии методов анализа данных.

Тема 1.2 Типы данных. Количество данных. Этапы анализа и обработки данных.

Раздел 2. Основные статистические методы анализа и обработки данных

Тема 2.1 Кластерный анализ

Сущность, типологизация и прикладная направленность задач классификации объектов. Методы кластерного анализа. Иерархические методы кластерного анализа. Итеративные методы. Сравнительный анализ иерархических и неиерархических методов кластеризации. Новые алгоритмы и некоторые модификации алгоритмов кластерного анализа.

Тема 2.2 Факторный анализ

Основные понятия и определения. Примеры. Методы факторного анализа. Общий обзор методов факторного анализа. Метод главных компонент. Центроидный метод. Метод экстремальной группировки параметров. Критерии рационального выбора числа факторов.

Тема 2.3. Регрессионный анализ

Виды регрессионного анализа. Линейная регрессия. Оценка качества регрессионной модели. Достоинства и недостатки регрессионных моделей.

Тема 2.4. Дисперсионный анализ

Однофакторный дисперсионный анализ. Многофакторный дисперсионный анализ.

Тема 2.5. Корреляционный анализ

Изучение зависимости между случайными величинами. Оценка коэффициента

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

корреляции. Частная и множественная корреляция. Ранговая корреляция.

Раздел 3. Кибернетические методы анализа и обработки данных

Тема 3.1. Деревья решений

Преимущества деревьев решений. Процесс конструирования дерева решений. Алгоритмы. Выводы.

Тема 3.2. Метод опорных векторов. Метод "ближайшего соседа". Байесовская классификация

Метод опорных векторов. Линейный SVM. Метод "ближайшего соседа" или системы рассуждений на основе аналогичных случаев. Решение задачи классификации новых объектов. Решение задачи прогнозирования. Байесовская классификация.

Тема 3.3. Нейронные сети второго поколения

Элементы нейронных сетей. Архитектура нейронных сетей. Обучение нейронных сетей. Модели нейронных сетей. Программное обеспечение для работы с нейронными сетями. Пример решения задачи.

Тема 3.4. Нейронные сети третьего поколения

Классификация нейронных сетей. Подготовка данных для обучения. Сверточные сети, авто-кодировщики, ограниченные машины Больцмана. Импульсная нейронная сеть. Методы обучения. Преимущества и недостатки. Применение. Выводы.

6. ТЕМЫ ПРАКТИЧЕСКИХ И СЕМИНАРСКИХ ЗАНЯТИЙ

Данный вид работы не предусмотрен учебным планом.

7. ЛАБОРАТОРНЫЕ РАБОТЫ, ПРАКТИКУМЫ

Данный вид работы не предусмотрен учебным планом.

8. ТЕМАТИКА КУРСОВЫХ, КОНТРОЛЬНЫХ РАБОТ, РЕФЕРАТОВ

Выполнение курсовых, контрольных работ не предусмотрено учебным планом.

Тематика рефератов (примерная):

1. Кластерный анализ.
2. Факторный анализ.
3. Регрессионный анализ.
4. Дисперсионный анализ.
5. Корреляционный анализ.
6. Деревья решений.
7. Метод опорных векторов.
8. Байесовская классификация.
9. Нейронные сети второго поколения.
10. Нейронные сети третьего поколения.

9. ПЕРЕЧЕНЬ ВОПРОСОВ К ЭКЗАМЕНУ (ЗАЧЕТУ)

2 семестр (зачет)

1. Типы данных. Количество данных.
2. Этапы анализа и обработки данных.
3. Методы кластерного анализа. Иерархические методы кластерного анализа.
4. Методы кластерного анализа. Итеративные методы.

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

5. Основные понятия и определения факторного анализа.
6. Метод главных компонент.
7. Виды регрессионного анализа. Оценка качества регрессионной модели. Достоинства и недостатки регрессионных моделей.
8. Однофакторный дисперсионный анализ.
9. Многофакторный дисперсионный анализ.
10. Оценка коэффициента корреляции. Частная и множественная корреляция.
11. Ранговая корреляция.


3 семестр (экзамен)

1. Деревья решений. Преимущества деревьев решений. Процесс конструирования дерева решений.
2. Метод опорных векторов.
3. Метод "ближайшего соседа".
4. Байесовская классификация.
5. Ключевые понятия и определения нейронных сетей.
6. Классификация нейронных сетей.
7. Методы обучения нейронных сетей.
8. Переобучение и регуляризация нейронных сетей.
9. Программное обеспечение для работы с нейронными сетями.
10. Свёрточные сети. Методы обучения. Преимущества и недостатки. Применение.
11. Автокодировщики. Методы обучения. Преимущества и недостатки. Применение.
12. Ограниченные машины Больцмана. Методы обучения. Преимущества и недостатки. Применение.
13. Импульсная нейронная сеть. Методы обучения. Преимущества и недостатки. Применение.

10. САМОСТОЯТЕЛЬНАЯ РАБОТА ОБУЧАЮЩИХСЯ

Форма обучения: очная.

Название разделов и тем	Вид самостоятельной работы (проработка учебного материала, решение задач, реферат, доклад, контрольная работа, подготовка к сдаче зачета, экзамена и др.)	Объем в часах	Форма контроля
2 семестр			
Раздел 1. ВВЕДЕНИЕ			
1.1. Предмет дисциплины. Исторические сведения о развитии методов анализа данных	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	5	устный опрос, проверка реферата
1.2. Типы данных. Количество данных. Этапы анализа и обработки данных.	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	5	устный опрос, проверка реферата
Раздел 2. Основные статистические методы анализа и обработки данных			
2.1. Кластерный анализ	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	16	устный опрос, проверка реферата

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

2.2. Факторный анализ	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	16	устный опрос, проверка реферата
2.3. Регрессионный анализ	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	16	устный опрос, проверка реферата
2.4. Дисперсионный анализ	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	16	устный опрос, проверка реферата
2.5. Корреляционный анализ	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	16	устный опрос, проверка реферата
Раздел 3. Кибернетические методы анализа и обработки данных			
3.1. Деревья решений	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	20	устный опрос, проверка реферата
3.2. Метод опорных векторов. Метод "ближайшего соседа". Байесовская классификация	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	20	устный опрос, проверка реферата
3.3. Нейронные сети второго поколения	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	30	устный опрос, проверка реферата
3.4. Нейронные сети третьего поколения	Проработка учебного материала, подготовка к сдаче экзамена (зачета), реферат	20	устный опрос, проверка реферата

11. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

а) Список рекомендуемой литературы

основная

1. Горяинова Е.Р., Прикладные методы анализа статистических данных : учеб. пособие / Горяинова Е.Р., Панков А.Р., Платонов Е.Н. - М. : ИД Высшей школы экономики, 2012. - 1000 с. - ISBN 978-5-7598-0866-4 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785759808664.html>
2. Вольфсон, М. Б. Анализ данных : учебное пособие / М. Б. Вольфсон. — Санкт-Петербург : СПбГУТ им. М.А. Бонч-Бруевича, 2015. — 81 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/180254>

дополнительная

1. Горелов В.И. Анализ статистических данных : практикум / В.И. Горелов, Т.Н. Ледашева. — Электрон. текстовые данные. — М. : Российская международная академия туризма, Университетская книга, 2015. — 120 с. — 978-5-98699-151-1. — Режим доступа: <http://www.iprbookshop.ru/70537.html>
2. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. —

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/450262>

3. Статистические пакеты обработки данных : учеб.- метод. пособие. Ч. 1 / И. А. Санников, Ю. Г. Савинов; УлГУ, ФМИТ. - Ульяновск : УлГУ, 2012. - Имеется печ. аналог. - Электрон. текстовые дан. (1 файл : 575 Кб). - URL: <http://lib.ulsu.ru/MegaPro/Download/MObject/234/>.

учебно-методическая

1. Савинов Ю. Г. Методические указания для самостоятельной работы студентов по дисциплине «Современные методы анализа массивов и потоков данных» для студентов магистратуры по направлению 01.04.02 «Прикладная математика и информатика» / Ю. Г. Савинов. - Ульяновск : УлГУ, 2022. - 6 с. - Неопубликованный ресурс. - URL: <http://lib.ulsu.ru/MegaPro/Download/MObject/13202>.

Согласовано:

Специалист ведущий НБ УлГУ
Должность сотрудника научной библиотеки

Боброва Н.А.
ФИО


подпись

/ _____ 2023

дата

б) Программное обеспечение: МойОфис Стандартный, Альт Рабочая станция 8.

в) Профессиональные базы данных, информационно-справочные системы

Профессиональные базы данных, информационно-справочные системы

1. Электронно-библиотечные системы:

1.1. Цифровой образовательный ресурс IPRsmart : электронно-библиотечная система : сайт / ООО Компания «Ай Пи Ар Медиа». - Саратов, [2023]. — URL: <http://www.iprbookshop.ru>. — Режим доступа: для зарегистрир. пользователей. - Текст : электронный.

1.2. Образовательная платформа ЮРАЙТ : образовательный ресурс, электронная библиотека : сайт / ООО Электронное издательство «ЮРАЙТ». — Москва, [2023]. - URL: <https://urait.ru>. — Режим доступа: для зарегистрир. пользователей. - Текст : электронный.

1.3. База данных «Электронная библиотека технического ВУЗа (ЭБС «Консультант студента») : электронно-библиотечная система : сайт / ООО «Политехресурс». — Москва, [2023]. — URL: <https://www.studentlibrary.ru/cgi-bin/mb4x>. — Режим доступа: для зарегистрир. пользователей. — Текст : электронный.

1.4. Консультант врача. Электронная медицинская библиотека : база данных : сайт / ООО «Высшая школа организации и управления здравоохранением-Комплексный медицинский консалтинг». — Москва, [2023]. — URL: <https://www.rosmedlib.ru>. — Режим доступа: для зарегистрир. пользователей. — Текст : электронный.

1.5. Большая медицинская библиотека : электронно-библиотечная система : сайт /

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

ООО «Букап». – Томск, [2023]. – URL: <https://www.books-up.ru/ru/library/> . – Режим доступа: для зарегистрир. пользователей. – Текст : электронный.

1.6. ЭБС Лань : электронно-библиотечная система : сайт / ООО ЭБС «Лань». – Санкт-Петербург, [2023]. – URL: <https://e.lanbook.com>. – Режим доступа: для зарегистрир. пользователей. – Текст : электронный.

1.7. ЭБС **Znanium.com** : электронно-библиотечная система : сайт / ООО «Знаниум». - Москва, [2023]. - URL: <http://znanium.com> . – Режим доступа : для зарегистрир. пользователей. - Текст : электронный.

2. КонсультантПлюс [Электронный ресурс]: справочная правовая система. / ООО «Консультант Плюс» - Электрон. дан. - Москва : КонсультантПлюс, [2023].

3. Базы данных периодических изданий:

3.1. eLIBRARY.RU: научная электронная библиотека : сайт / ООО «Научная Электронная Библиотека». – Москва, [2023]. – URL: <http://elibrary.ru>. – Режим доступа : для авториз. пользователей. – Текст : электронный

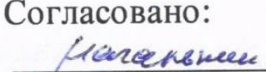
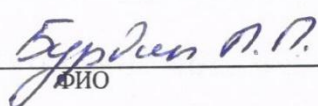
3.2. Электронная библиотека «Издательского дома «Гребенников» (Grebinnikon) : электронная библиотека / ООО ИД «Гребенников». – Москва, [2023]. – URL: <https://id2.action-media.ru/Personal/Products>. – Режим доступа : для авториз. пользователей. – Текст : электронный.

4. Федеральная государственная информационная система «Национальная электронная библиотека» : электронная библиотека : сайт / ФГБУ РГБ. – Москва, [2023]. – URL: <https://нэб.рф>. – Режим доступа : для пользователей научной библиотеки. – Текст : электронный.

5. Российское образование : федеральный портал / учредитель ФГАУ «ФИЦТО». – URL: <http://www.edu.ru>. – Текст : электронный.

6. Электронная библиотечная система УлГУ : модуль «Электронная библиотека» АБИС Мега-ПРО / ООО «Дата Экспресс». – URL: <http://lib.ulsu.ru/MegaPro/Web>. – Режим доступа : для пользователей научной библиотеки. – Текст : электронный.

Согласовано:

 УИТТ |  ФИО |  | подпись | дата

12. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ:

Аудитории для проведения лекций, для проведения текущего контроля и промежуточной аттестации, групповых и индивидуальных консультаций.

Аудитории укомплектованы специализированной мебелью, учебной доской. Аудитории для проведения лекций оборудованы мультимедийным оборудованием для предоставления информации большой аудитории. Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной инфромационно-образовательной среде, электронно-библиотечной системе.

13. СПЕЦИАЛЬНЫЕ УСЛОВИЯ ДЛЯ ОБУЧАЮЩИХСЯ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

В случае необходимости, обучающимся из числа лиц с ограниченными возможностями здоровья (по заявлению обучающегося) могут предлагаться одни из

Министерство науки и высшего образования РФ Ульяновский государственный университет	Форма	
Ф-Рабочая программа дисциплины		

следующих вариантов восприятия информации с учетом их индивидуальных психофизических особенностей:

– для лиц с нарушениями зрения: в печатной форме увеличенным шрифтом; в форме электронного документа; в форме аудиофайла (перевод учебных материалов в аудиоформат); в печатной форме на языке Брайля; индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания и консультации;

– для лиц с нарушениями слуха: в печатной форме; в форме электронного документа; видеоматериалы с субтитрами; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания и консультации;

– для лиц с нарушениями опорно-двигательного аппарата: в печатной форме; в форме электронного документа; в форме аудиофайла; индивидуальные задания и консультации.

В случае необходимости использования в учебном процессе частично/исключительно дистанционных образовательных технологий, организация работы ППС с обучающимися с ОВЗ и инвалидами предусматривается в электронной информационно-образовательной среде с учетом их индивидуальных психофизических особенностей.

Разработчик



подпись

доцент

должность

Савинов Ю.Г.

ФИО