

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Тульский государственный университет»

Институт прикладной математики и компьютерных наук
Кафедра «Прикладная математика и информатика»

Утверждено на заседании кафедры
«Прикладная математика и информатика»
24 января 2022 г., протокол № 5

Заведующий кафедрой

 М.В. Грязев

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)
«Прикладной многомерный статистический анализ»

**основной профессиональной образовательной программы
высшего образования – программы магистратуры**

по направлению подготовки
01.04.02 Прикладная математика и информатика

с направленностью (профилем)

**Перспективные методы искусственного интеллекта в сетях передачи и
обработки данных**

Форма обучения: очная

Идентификационный номер образовательной программы: 010402-01-22

Тула 2022 год

1. Место дисциплины (модуля) в структуре ОПОП ВО:

Дисциплина (модуль) относится к обязательной части основной профессиональной образовательной программы.

2. Входные требования для освоения дисциплины (модуля), предварительные условия:

учащиеся должны владеть знаниями по математическому анализу, линейной алгебре и теории вероятностей в объеме, соответствующем программе обучения основных образовательных программ бакалавриата по укрупненным группам направлений и специальностей 01.00.00 «Математика и механика», 02.00.00 «Компьютерные и информационные науки».

3. Результаты обучения по дисциплине (модулю):

Планируемые результаты обучения по дисциплине (модулю)	
Формируемые компетенции (код и наименование компетенции)	Результаты обучения (знания, умения)
ОПК-2. Способен совершенствовать и реализовывать новые математические методы решения прикладных задач.	ОПК-2.1. Знает существующие математические методы и системы программирования решения прикладных задач.
	ОПК-2.2. Умеет использовать и адаптировать существующие математические методы и системы программирования для разработки и реализации алгоритмов.
	ОПК-2.3. Владеет методами и современными системами программирования для разработки и реализации алгоритмов.

4. Объем дисциплины (модуля) составляет 5 з.е., в том числе 36 академических часа контактная работа с преподавателем - 12 академических часа занятий лекционного типа, 24 академических часов занятий лабораторного типа, 144 академических часов на самостоятельную работу обучающихся.

5. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий:

5.1. Структура дисциплины (модуля) по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий (в строгом соответствии с учебным планом)

Наименование и краткое содержание разделов и тем дисциплины, форма промежуточной аттестации по дисциплине	Всего (часы)	В том числе								
		Контактная работа (работа во взаимодействии с преподавателем), часы						Самостоятельная работа учащегося, часы		
		из них						из них		
		Занятия лекционного	Занятия семинарско-	Групповые консуль-	Индивидуальные	Учебные занятия, направленные на проведение текущего контроля успеваемости: колло-	Всего	Чтение учебников и оригинальных журнальных статей	Подготовка рефератов и т.п..	Всего

						квиумы, практиче- ские кон- трольные занятия и др.				
Тема 1. Ос- новные задачи и методы при- кладного мно- гомерного ста- тистического анализа	6	2	4	-	-		6	24	-	24
Тема 2. Ре- грессионный анализ	12	4	8	-	-		12	48	-	48
Тема 3. Дис- персионный анализ.	6	2	4	-	-		6	24	-	24
Тема 4. Дис- криминантный анализ.	6	2	4	-	-		6	24	-	24
Тема 5. Кла- стерный ана- лиз.	6	2	4	-	-		6	24	24	10
Промежуточ- ная аттестация - экзамен	36						4			
Итого	36	36					144			

5.2. Содержание разделов (тем) дисциплины

№ п/п	Наименование разделов (тем) дисциплины	Содержание разделов (тем) дисциплин
1.	Тема 1. Основные задачи и ме- тоды прикладного многомерно- го статистического анализа	Обзор основных задач многомерного статистиче- ского анализа. Задача о наилучшей линейной оценке. Коэффициенты корреляции и их свойства. Типы многомерных данных. Многомерное нор- мальное распределение и его свойства.
2.	Тема 2. Регрессионный анализ	Классическая многомерная модель регрессии. Ме- тод наименьших квадратов (МНК). Свойства оце- нок параметров.

		Проверка линейных гипотез. Коэффициент детерминации и его свойства. Выбор оптимальной модели. Модели с гетероскедастичностью и автокорреляцией ошибок и их оценивание.
3.	Тема 3. Дисперсионный анализ.	Однофакторный дисперсионный анализ. Двухфакторный дисперсионный анализ.
4.	Тема 4. Дискриминантный анализ.	Анализ потока управления при оценке WCET. Анализ влияния оборудования (конвейер, кэш) на WCET. Расчет итоговой оценки WCET. Оптимизация WCET, повышение временной предсказуемости выполнения программ в ИУС РВ. Эволюционные алгоритмы поиска нижней оценки WCET.
5.	Тема 5. Кластерный анализ.	остановка задачи, описание общего метода решения задачи построения кластеров. Описание методов близости и различия кластеров. Описание алгоритма построения кластеров.

6. Фонд оценочных средств (ФОС, оценочные и методические материалы) для оценивания результатов обучения по дисциплине (модулю).

6.1. Типовые контрольные задания или иные материалы для проведения текущего контроля успеваемости, критерии и шкалы оценивания (в отсутствие утвержденных соответствующих локальных нормативных актов на факультете)

Возможные темы для домашних заданий (рефератов)

- Основные задачи многомерного статистического анализа:
 - корреляционный анализ
 - регрессионный анализ
 - снижение размерности
 - дисперсионный анализ
 - дискриминантный анализ
 - кластерный анализ
- Гильбертово пространство случайных величин.
- Что такое наилучшая линейная оценка.
- Лемма о перпендикуляре.
- Простой коэффициент корреляции и что он измеряет.
- Множественный коэффициент корреляции и что он измеряет.
- Частный коэффициент корреляции и что он измеряет.
- Множественная линейная регрессия: модель и основные ограничения.
- Описание МНК для оценки параметров.
- Явный вид оценок параметров по МНК.
- Общая схема проверки гипотезы о параметре.
- Для чего используется Т-критерий.
- Основное различие Т-критерия и F-критерия в задаче проверки значимости влияния фактора.
- Адекватность модели. Постановка задачи.
- Коэффициент детерминации и что он измеряет.

16. Основная задача в однофакторном дисперсионном анализе.
17. Основная задача в двухфакторном дисперсионном анализе.
18. Основная задача дискриминантного анализа.
19. Кластерный анализ: постановка задачи.
20. Кластерный анализ: последовательное построение факторов.

Методические материалы для проведения процедур оценивания результатов обучения

Выполнение домашних заданий для текущего контроля успеваемости и задания промежуточного контроля оценивается по пятибалльной системе. Итоговая оценка вычисляется как среднее арифметическое полученных результатов.

6.2. Типовые контрольные задания или иные материалы для проведения промежуточной аттестации по дисциплине, критерии и шкалы оценивания

Список тем вопросов для зачета.

1. Основные задачи многомерного статистического анализа.
2. Гильбертово пространство случайных величин. Задача о наилучшей линейной оценке.
3. Корреляционный и регрессионный анализ.
4. Коэффициенты корреляции.
5. Простая линейная регрессия. Метод наименьших квадратов. Свойства оценок.
6. Множественная линейная регрессия. МНК. Свойства оценок.
7. Т-критерий значимости влияния фактора.
8. Проверка линейных гипотез. F-критерий.
9. Проверка адекватности модели. Коэффициент детерминации.
10. Равенство уравнений регрессии. Тест Чоу.
11. Фиктивные переменные.
12. Модель линейной регрессии с гетероскедастичностью.
13. Модель линейной регрессии с автокорреляцией в ошибках. Критерий Дарбина-Уотсона.
14. Однофакторный дисперсионный анализ.
15. Двухфакторный дисперсионный анализ.
16. Дискриминантный анализ: постановка задачи и ее решение в случае известных параметров.
17. Решение задачи дискриминантного анализа в случае неизвестных параметров.
18. Кластерный анализ: постановка задачи и основные понятия.
19. Кластерный анализ: схема последовательного построения факторов.

ШКАЛА И КРИТЕРИИ ОЦЕНИВАНИЯ результатов обучения (РО) по дисциплине				
Оценка виды оценоч- ных средств	2 (не зачтено)	3 (зачтено)	4 (зачтено)	5 (зачтено)
Знания (виды оценоч- ных средств: опрос, тесты)	Отсутствие знаний	Фрагментарные знания	Общие, но не структурированные знания	Сформированные систематические знания

Умения (виды оценочных средств: практические задания)	Отсутствие умений	В целом успешное, но не систематическое умение	В целом успешное, но содержащее отдельные пробелы умение (допускает неточности не принципиального характера)	Успешное и систематическое умение
Навыки (владения, опыт деятельности) (виды оценочных средств: выполнение и защита курсовой работы, отчет по практике, отчет по НИР и т.п.)	Отсутствие навыков (владений, опыта)	Наличие отдельных навыков (наличие фрагментарного опыта)	В целом, сформированные навыки (владения), но используемые не в активной форме	Сформированные навыки (владения), применяемые при решении задач

7. Ресурсное обеспечение:

7.1. Перечень основной и дополнительной литературы

Основная литература

1. Андерсон Т. Статистический анализ временных рядов / Андерсон Т. – М.: Книга по Требованию, 2012. – 752 с. ISBN 978-5-458-26140-1

Дополнительная литература

- 1) Факторный, дискриминантный и кластерный анализы. – М.: Финансы и статистика, 1989.
- 2) Шеффе Г. Дисперсионный анализ. – М.: Наука, 1980.
- 3) Болч К., Хуань К.Дж. Многомерные статистические методы для экономики. – М.: Статистика, 1979.
- 4) Hardle W., Simar L. Applied Multivariate Statistical Analysis. - <http://www.xplorestat.de/ebooks/ebooks.html>
- 5) Кулаичев А.П. Методы и средства комплексного анализа данных. – М.: Форум-Инфра-М, 2006.
- 6) Андерсон Т. Введение в многомерный статистический анализ. – М.: Наука, 1963.
- 7) Johnson R.A. and Winchern D.W. Applied Multivariate Statistical Analysis. – Pearson Prentice Hall, 2007.
- 8) Хохлов Ю.С. Эконометрика. Вводный курс: Учебное пособие. М.: Издательский отдел факультета ВМиК МГУ, 2006. – 100 с.

7.2. Перечень лицензионного программного обеспечения, в том числе отечественного производства

При реализации дисциплины может быть использовано следующее программное обеспечение:

Программное обеспечение для подготовки слайдов лекций MS PowerPoint, MS Word

Программное обеспечение для создания и просмотра pdf-документов Adobe Reader

Издательская система LaTeX

Язык программирования Python и среда разработки Jupiter Notebook (вместе с библиотеками numpy, scikit-learn, pandas)

Язык программирования R и среда разработки R Studio

7.3. Перечень профессиональных баз данных и информационных справочных систем

1. <http://www.edu.ru> – портал Министерства образования и науки РФ
2. <http://www.ict.edu.ru> – система федеральных образовательных порталов «ИКТ в образовании»
3. <http://www.openet.ru> - Российский портал открытого образования
4. <http://www.mon.gov.ru> - Министерство образования и науки Российской Федерации
5. <http://www.fasi.gov.ru> - Федеральное агентство по науке и инновациям

7.4. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

1. Math-Net.Ru [Электронный ресурс] : общероссийский математический портал / Математический институт им. В. А. Стеклова РАН ; Российская академия наук, Отделение математических наук. - М. : [б. и.], 2010. - Загл. с титул. экрана. - Б. ц.
URL: <http://www.mathnet.ru>
2. Университетская библиотека Online [Электронный ресурс] : электронная библиотечная система / ООО "Директ-Медиа" . - М. : [б. и.], 2001. - Загл. с титул. экрана. - Б. ц. URL: www.biblioclub.ru
3. Универсальные базы данных East View [Электронный ресурс] : информационный ресурс / East View Information Services. - М. : [б. и.], 2012. - Загл. с титул. экрана. - Б. ц.
URL: www.ebiblioteka.ru
4. Научная электронная библиотека eLIBRARY.RU [Электронный ресурс] : информационный портал / ООО "РУНЭБ" ; Санкт-Петербургский государственный университет. - М. : [б. и.], 2005. - Загл. с титул. экрана. - Б. ц.
URL: www.eLibrary.ru

7.5. Описание материально-технического обеспечения.

Факультет ВМК, ответственный за реализацию данной Программы, располагает соответствующей материально-технической базой, включая современную вычислительную технику, объединенную в локальную вычислительную сеть, имеющую выход в Интернет. Используются специализированные компьютерные классы, оснащенные современным оборудованием. Материальная база факультета соответствует действующим санитарно-техническим нормам и обеспечивает проведение всех видов занятий (лабораторной, практической, дисциплинарной и междисциплинарной подготовки) и научно-исследовательской работы обучающихся, предусмотренных учебным планом.

8. Соответствие результатов обучения по данному элементу ОПОП результатам освоения ОПОП указано в Общей характеристике ОПОП.

9. Рабочая программа внедрена в соответствии с Соглашением о предоставлении из федерального бюджета грантов в форме субсидий на разработку программ бакалавриата и программ магистратуры по профилю «искусственный интеллект, а также на повышение квалификации педагогических работников образовательных организаций высшего образования в сфере искусственного интеллекта, заключённым «29» сентября 2021 г. № 075-15-2021-1036 между МГУ имени М.В.Ломоносова и Минобрнауки России.