Минобрнауки России

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

Кафедра математических методов и моделей в экономике

РАБОЧАЯ ПРОГРАММА

ДИСЦИПЛИНЫ

«Б1.Д.В.Э.2.1 Современные информационные технологии в анализе данных и научных исследованиях»

Уровень высшего образования

БАКАЛАВРИАТ

Направление подготовки
01.03.04 Прикладная математика
(код и наименование направления подготовки)

Математическое и компьютерное моделирование
(наименование направленности (профиля) образовательной программы)

Квалификация
Бакалавр

Форма обучения
Очная

Год набора 2022

1959226

Рабочая программа дисциплины «*Б1.Д.В.Э.2.1 Современные информационные технологии в анализе данных и научных исследованиях*» рассмотрена и утверждена на заседании кафедры

Кафедра математических методов и моделей в экономике
наименование кафедры

протокол № 8 от "10" 02 2022г.

Заведующий кафедрой
Кафедра математических методов и моделей в экономике — О.Н. Яркова
наименование кафедры *подпись* *расшифровка подписи*

Исполнители:
Доцент кафедры ММиМЭ — О.С. Чудинова
должность *подпись* *расшифровка подписи*

должность *подпись* *расшифровка подписи*

СОГЛАСОВАНО:
Председатель методической комиссии по направлению подготовки
01.03.04 Прикладная математика — О.Н. Яркова
код наименование *личная подпись* *расшифровка подписи*

Заведующий отделом комплектования научной библиотеки
Н.Н. Бигалиева
личная подпись *расшифровка подписи*

Уполномоченный по качеству факультета
Н.А. Тычинина
личная подпись *расшифровка подписи*

№ регистрации _____

© Чудинова О.С., 2022
© ОГУ, 2022

1 Цели и задачи освоения дисциплины

Цель (цели) освоения дисциплины: формирование знаний, умений и навыков по выбору и применению современных информационных технологий для решения задач анализа данных и проведения научных исследований в области моделирования процессов и систем.

Задачи:

– изучение современных информационных технологий, применяемых для решения задач анализа данных и моделирования процессов и систем в научных исследованиях;

– освоение методов и алгоритмов решения задач анализа данных и моделирования процессов и систем с использованием современных информационных технологий и языков программирования;

– приобретение навыков выбора и использования современных информационных технологий для решения задач анализа данных и моделирования процессов и систем в научных исследованиях.

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к дисциплинам (модулям) по выбору вариативной части блока Д «Дисциплины (модули)»

Пререквизиты дисциплины: *Б1.Д.Б.25 Математическая статистика и анализ данных, Б1.Д.Б.28 Программирование, Б1.Д.В.7 Технологии разработки программного обеспечения, Б2.П.Б.У.1 Научно-исследовательская работа (получение первичных навыков научно-исследовательской работы)*

Постреквизиты дисциплины: *Отсутствуют*

3 Требования к результатам обучения по дисциплине

Процесс изучения дисциплины направлен на формирование следующих результатов обучения

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
ПК*-2 Способен осуществлять математическое и компьютерное моделирование для анализа рисков и выработки оптимальных решений в различных отраслях	ПК*-2-В-2 Применяет современные математические и инструментальные средства для моделирования, анализа рисков и выработки решений по оптимизации социальных, экономических, производственных процессов и систем	**Знать:** современные математические и инструментальные средства, применяемые для решения задач анализа данных при проведении научных исследований в области моделирования процессов и систем. **Уметь:** применять современные математические и инструментальные средства для решения задач анализа данных и проведения научных исследований в области моделирования процессов и систем. **Владеть:** навыками реализации методов анализа данных с помощью современных информационных технологий и интерпретации результатов для выработки решений.
ПК*-4 Способен использовать знания	ПК*-4-В-2 Применяет знания стандартных и	**Знать:** специализированные программные средства и языки

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
современных языков программирования, стандартных пакетов прикладных программ, информационно-телекоммуникационной сети "Интернет", инструментальных средств анализа данных при решении практических задач управления информацией	специализированных ППП, информационно-телекоммуникационных технологий для решения практических задач анализа данных, математического и компьютерного моделирования процессов и систем ПК*-4-В-3 Применяет знания стандартных ППП, информационно-телекоммуникационных технологий, сети "Интернет" для поиска, обработки и систематизации информации, оформления отчетов по результатам выполненных исследований, представления их в виде презентаций, докладов	программирования для анализа данных при решении исследовательских задач моделирования процессов и систем. **Уметь:** применять специализированные программные средства и языки программирования для анализа данных при решении исследовательских задач моделирования процессов и систем; оформлять результаты исследования в виде отчета. **Владеть:** навыками работы со специализированными программными средствами и языками программирования для анализа данных и научных исследований; навыками оформления результатов исследования в виде отчета.

4 Структура и содержание дисциплины

4.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 3 зачетные единицы (108 академических часов).

Вид работы	Трудоемкость, академических часов	
	6 семестр	всего
Общая трудоёмкость	**108**	**108**
Контактная работа:	**34,25**	**34,25**
Лекции (Л)	18	18
Лабораторные работы (ЛР)	16	16
Промежуточная аттестация (зачет, экзамен)	0,25	0,25
Самостоятельная работа: *- выполнение индивидуального задания;* *- самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий;* *- подготовка к лабораторным занятиям)*	**73,75**	**73,75**
Вид итогового контроля (зачет, экзамен, дифференцированный зачет)	**диф. зач.**	

Разделы дисциплины, изучаемые в 6 семестре

№ раздела	Наименование разделов	Количество часов				
		всего	аудиторная работа			внеауд. работа
			Л	ПЗ	ЛР	
1	Современные информационные технологии интеллектуального анализа данных	10	2		2	6

№ раздела	Наименование разделов	Количество часов				
		всего	аудиторная работа			внеауд. работа
			Л	ПЗ	ЛР	
2	Анализ ассоциативных правил и последовательных шаблонов	22	4		4	14
3	Деревья принятия решений: общая характеристика	14	4		-	10
4	Деревья классификации	20	2		4	14
5	Деревья регрессии	18	2		2	14
6	Ансамбли деревьев и случайный лес	24	4		4	16
	Итого:	108	18		16	74
	Всего:	108	18		16	74

4.2 Содержание разделов дисциплины

№1 Современные информационные технологии интеллектуального анализа данных

Обзор рынка информационных технологий для интеллектуального анализа данных. Обзор возможностей аналитической платформы Deductor Academic. Программная платформа обработки данных RapidMiner. Сравнение языков программирования R и Python.

№2 Анализ ассоциативных правил и последовательных шаблонов

Постановка задачи поиска ассоциативных правил, представление исходной информация для анализа. Характеристики правил: поддержка, достоверность, улучшение. Алгоритмы поиска ассоциативных правил. Определение последовательных шаблонов, понятие последовательности. Клиенториентированные базы данных.

№3 Деревья принятия решений: общая характеристика

Понятие и примеры деревьев решений. Термины, используемые для интерпретации дерева решений. Задачи, решаемые с помощью деревьев решений. Достоинства и недостатки методы деревьев принятия решений. Общий алгоритм построения деревьев решений. Критерии выбора атрибута разбиения. Прунинг дерева. Биннинг переменных. Оценка точности модели в форме дерева решений.

№4 Деревья классификации

Построение и интерпретация дерева классификации CHAID. Оценка качества модели классификации. Метод отсечения ветвей на основе стоимости-сложности с кросс-проверкой. Построение и интерпретация дерева классификации CRT. Прунинг дерева CRT. Построение ROC-кривой и оценка качества классификации. Построение и интерпретация дерева регрессии CRT.

№5 Деревья регрессии

Построение и интерпретация дерева регрессии CRT. Применение модели к новым данным. Оценка качества дерева регрессии.

№6 Ансамбли деревьев и случайный лес

Понятие ансамбля деревьев. Подходы к ансамблированию. Случайный лес: понятие, основные параметры, достоинства, оценка качества. Построение случайного леса для решения задач классификации и исследования зависимости.

4.3 Лабораторные работы

№ ЛР	№ раздела	Наименование лабораторных работ	Кол-во часов
1	1	Обзор современных информационных технологий интеллектуального анализа данных	2
2	2	Построение ассоциативных правил	2
3	2	Построение последовательных шаблонов	2
4	4	Построение и интерпретация дерева классификации CHAID	2
5	4	Построение и интерпретация дерева классификации CRT	2

№ ЛР	№ раздела	Наименование лабораторных работ	Кол-во часов
6	5	Построение и интерпретация дерева регрессии	2
7	6	Классификация на основе случайного леса	2
8	6	Исследование зависимости на основе случайного леса	2
		Итого:	16

5 Учебно-методическое обеспечение дисциплины

5.1 Основная литература

1. Седова, Е.Н. Ассоциативные правила в социально-экономических и экологических исследованиях: учебное пособие [Электронный ресурс]/ Е.Н. Седова, А.В. Раменская, Р.М. Безбородникова; Оренбургский гос. ун-т. – Оренбург: ОГУ, 2015. – 170 с. - Режим доступа: http://artlib.osu.ru/web/books/metod_all/7861_20150427.pdf
2. Федотова, Е.Л. Информационные технологии в науке и образовании : учеб. пособие / Е.Л.Федотова, А.А.Федотов. — Москва : ИД «ФОРУМ»; ИН-ФРА-М, 2015. — 336 с. — (Высшее образование). - ISBN 978-5-8199-0434-3 (ИД «ФОРУМ») ; ISBN 978-5-16-004266-4 (ИНФРА-М, print) ; ISBN 978-5-16-103184-I (ИНФРА-М, online). - Текст : электронный. - URL: https://znanium.com/catalog/product/487293
3. Золотарюк, А. В. Язык и среда программирования R : учебное пособие / А. В. Золотарюк. — Москва : ИНФРА-М, 2020. — 162 с. — (Высшее образование: Бакалавриат). - ISBN 978-5-16-016021-4. - Текст : электронный. - URL: https://znanium.com/catalog/product/1077985

5.2 Дополнительная литература

1. Методы и модели эконометрики [Электронный ресурс] : учебное пособие для студентов, обучающихся по программам высшего образования по направлениям подготовки 01.03.04 Прикладная математика, 38.04.01 Экономика, 38.03.05 Бизнес-информатика / О.И. Бантикова, В.И. Васянина, Ю.А. Жемчужникова, А.Г. Реннер, Е.Н. Седова, О.И. Стебунова, Л.М. Туктамышева, О.С. Чудинова /под ред. А. Г. Реннера. - Ч. 1. Анализ данных. - Оренбург : ОГУ, 2015. - - Режим доступа: http://artlib.osu.ru/web/books/metod_all/8283_20150703.pdf
2. Советов, Б. Я. Информационные технологии [Текст] : учебник для прикладного бакалавриата / Б. Я. Советов, В. В. Цехановский; С.-Петерб. гос. электротехн. ун-т "ЛЭТИ" им. В. И. Ульянова (Ленина).- 6-е изд., перераб. и доп. - Москва : Юрайт, 2012, 2016. - 263 с.
3. Зададаев, С.А. Математика на языке R : учебник / С.А. Зададаев ; Финансовый университет при Правительстве РФ, Департамент анализа данных, принятия решений и финансовых технологий. - Москва : Прометей, 2018. - 324 с. : схем., ил. - (Учебники для вузов. Специальная литература). - ISBN 978-5-907003-59-0 ; То же [Электронный ресурс]. - Режим доступа: http://biblioclub.ru/index.php?page=book&id=494941

5.3 Периодические издания

1. Информационные технологии : журнал. - М. : Агентство "Роспечать", 2022.
2. Вестник компьютерных и информационных технологий : журнал. - М. : Агентство "Роспечать", 2022.

5.4 Интернет-ресурсы

1. http://www.r-project.org/ - The R Project for Statistical Computing - сайт проекта R;
2. http://journal.r-project.org/ - «The R Journal» - журнал по R;
3. http://herba.msu.ru/shipunov/software/r/rplus1.htm Шипунов А.Б. n+1 вопрос про R [Электронный ресурс]. 2002.

4. http://www.statmethods.net/index.html - справочный ресурс по R;
5. http://pj.freefaculty.org/R/Rtips.html - Под Э. Джонсон Rtips. Revival! - Советы по использованию R;
6. https://r-analytics.blogspot.com – R: Анализ и визуализация данных;
7. http://www.basegroup.ru/ сайт разработчика аналитической платформы Deductor
8. http://www.gks.ru - Федеральная служба государственной статистики РФ;
9. http://sophist.hse.ru - Единый архив экономических и социальных данных Высшей Школы Экономики;
10. http://www.economicsnetwork.ac.uk/links/data_free - Перечень открытых источников статистических данных на сайте, поддерживаемым университетом Бристоля;
11. http://quantile.ru/06/06-AT.pdf - Анатольев Станислав и Цыплаков Александр (2009) «Где найти данные в сети?», Квантиль, №6, с.59-71. Эссе, содержащее кратко аннотированные списки веб-сайтов, на которых прикладной экономист может найти данные для исследований.
12. https://loginom.ru/blog/decision-tree-p1 - Блог Low-code платформа Loginom (Деревья решений: общие принципы)
13. https://rapidminer.com/ - Официальный сайт компании разработчика программной платформыобработки данных RapidMiner
14. https://cran.r-project.org/ - Официальный сайт языка R (Описание пакетов и библиотек, документация для R)
15. https://www.rstudio.com/products/rstudio/ - Свободная среда разработки программного обеспечения с открытым исходным кодом для языка программирования R RStudio [Электронный ресурс].

5.5 Программное обеспечение, профессиональные базы данных и информационные справочные системы

Лицензионное программное обеспечение

1. Операционная система Microsoft Windows
2. Пакет настольных приложений Microsoft Office (Word, Excel, PowerPoint, OneNote, Outlook, Publisher, Access)

Свободно распространяемое программное обеспечение

1. Бесплатное средство просмотра файлов PDF Adobe Reader. Доступна бесплатно. Условия распространения, тип лицензия: Бесплатно после принятия лицензионного соглашения на ПО Adobe. Разработчики: Adobe Systems. Режим доступа: https://get.adobe.com/ru/reader/
2. Свободное статистическое программное обеспечение для анализа данных (с открытым исходным кодом в рамках проекта GNU) R. Доступна бесплатно. Условия распространения, тип лицензия: GNU GPL 2. Разработчики: сотрудники статистического факультета Оклендского университета Росс Айхэка (англ. Ross Ihaka) и Роберт Джентлмен (англ. Robert Gentleman). Режим доступа: https://www.r-project.org/
3. Свободная среда разработки программного обеспечения с открытым исходным кодом для языка программирования R Rstudio. Доступна бесплатно. Условия распространения, тип лицензия: GNU Affero General Public License v3. Режим доступа: https://www.rstudio.com/

Профессиональные базы данных

1. SCOPUS [Электронный ресурс]: реферативная база данных / компания Elsevier. – Режим доступа: https://www.scopus.com/, в локальной сети ОГУ
2. Web of Science [Электронный ресурс]: реферативная база данных / компания Clarivate Analytics. – Режим доступа : http://apps.webofknowledge.com/, в локальной сети ОГУ

Информационные справочные системы

1. Законодательство России [Электронный ресурс]: информационно-правовая система. – Режим доступа: http://pravo.fso.gov.ru/ips/, в локальной сети ОГУ

2. Консультант Плюс [Электронный ресурс] : справочно-правовая система / Компания Консультант Плюс. – Электрон. дан. – Москва, [1992–2022]. – Режим доступа: \\fileserver1\!CONSULT\cons.exe, в локальной сети ОГУ

3. Гарант [Электронный ресурс] : справочно-правовая система / НПП Гарант-Сервис. – Электрон. дан. – Москва, [1990–2022]. – Режим доступа: \\fileserver1\GarantClient\garant.exe, в локальной сети ОГУ

6 Материально-техническое обеспечение дисциплины

Учебные аудитории для проведения занятий лекционного типа, для проведения групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.

Аудитории оснащены комплектами ученической мебели, техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Для проведения лабораторных занятий используется компьютерный класс, оснащенный комплектами ученической мебели, компьютерами с подключением к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду ОГУ.

Помещение для самостоятельной работы обучающихся оснащены компьютерной техникой подключенной к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду ОГУ.