

МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)

УТВЕРЖДАЮ
Заведующий кафедрой
Вычислительной математики и
прикладных информационных технологий



Леденёва Т.М.
26.05.2023 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ
Б1.В.05 Интеллектуальный анализ данных

1. Код и наименование направления подготовки/специальности:

02.04.02 Фундаментальная информатика и информационные технологии

2. Профиль подготовки/специализация: Машинное обучение и интеллектуальные информационные технологии

3. Квалификация выпускника: магистр

4. Форма обучения: очная

5. Кафедра, отвечающая за реализацию дисциплины:

кафедра вычислительной математики и прикладных информационных технологий

6. Составители программы:

Козлова Наталья Владимировна, к.т.н., доцент кафедры ВМ и ПИТ;
Аристова Екатерина Михайловна, к.ф.-м.н., доцент кафедры ВМиПИТ

7. Рекомендована:

научно-методическим советом факультета ПММ 26.05.2023, протокол №7

8. Учебный год: 2024-2025

Семестр(ы): 4

9. Цели и задачи учебной дисциплины

Цель учебной дисциплины: сформировать у обучающихся математические и алгоритмические основы интеллектуального анализа данных как инструмента, обеспечивающего переход от данных к знаниям в интеллектуальных информационных системах различного назначения.

Задачи учебной дисциплины:

- ознакомление обучающихся с актуальными проблемами и современными технологиями обработки данных с помощью инструментов искусственного интеллекта;
- формирование навыков подготовки данных и использования современных методов интеллектуального анализа для выявления зависимостей и структуризации данных, в том числе с применением существующих программных средств;
- формирование навыков информационного поиска с использованием открытых источников информации и специализированных баз данных и анализа научно-технической информации по проблемам интеллектуального анализа данных.

10. Место учебной дисциплины в структуре ООП: (обязательная часть или часть, формируемая участниками образовательных отношений (вариативная) блока Б1, к которой относится дисциплина; требования к входным знаниям, умениям и навыкам; дисциплины, для которых данная дисциплина является предшествующей)

Дисциплина «Интеллектуальный анализ данных» входит в блок Б1 вариативной части программы магистров и изучается в 4 семестре. Данный курс непосредственно связан с дисциплиной «Основы нечеткого моделирования», изучаемой в рамках программы подготовки магистра. Изучение данного курса базируется на знании студентами материала основных математических и естественнонаучных дисциплин, изучаемых в рамках программы подготовки бакалавра по направлениям физико-математических наук.

11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями) и индикаторами их достижения:

Код	Название компетенции	Код(ы)	Индикатор(ы)	Планируемые результаты обучения
ПК-1	Способен проводить работы по обработке и анализу научно-технической информации результатов исследований	ПК-1.1	Проводит информационный поиск для решения исследовательских задач с использованием открытых источников информации и специализированных баз данных	Знать: основные методы анализа данных Уметь: применять на практике основные методы анализа данных Владеть: навыками решения практических задач
ПК-3	Способен обрабатывать, интерпретировать, оформлять и представлять профессиональному обществу результаты проведенных исследований	ПК-3.1	Использует современные методы анализа информации для обработки данных, полученных в рамках проведенных исследований	Знать: основные методы анализа и обработки информации Уметь: применять на практике основные методы анализа информации для обработки данных Владеть: навыками решения практических задач
ПК-6	Способен совершенствовать и разрабатывать различные инструменты искусственного интеллекта, модели и методы представления знаний, осуществлять генерацию и	ПК-6.3	Применяет существующие программные средства для разработки интеллектуальных информационных систем, в том числе экспертных систем.	Знать: основные методы анализа данных, методы представления знаний, основы искусственного интеллекта Уметь: применять на практике существующие программные средства для разработки интеллектуальных информационных систем

	оптимизацию баз знаний			Владеть: навыками решения практических задач
--	------------------------	--	--	----------------------------------------------

12. Объем дисциплины в зачетных единицах/часах (в соответствии с учебным планом) – 3/108.

Форма промежуточной аттестации зачет с оценкой

13. Трудоемкость по видам учебной работы

Вид учебной работы		Трудоемкость		
		Всего	По семестрам	
			4 семестр	...
Контактная работа		36	36	
в том числе:	лекции	24	24	
	практические	0	0	
	лабораторные	12	12	
	курсовая работа	0	0	
Самостоятельная работа		72	72	
Промежуточная аттестация (для экзамена)		0	0	
Итого:		108	108	

13.1. Содержание дисциплины

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация раздела дисциплины с помощью онлайн-курса, ЭУМК *
1. Лекции			
1.1	Эконометрическое моделирование	Основные аспекты эконометрического моделирования. Линейные регрессионные модели. Линейная парная регрессия. Оценка параметров парной регрессии. Коэффициент корреляции. Оценка значимости уравнения регрессии. Коэффициент детерминации. Множественный регрессионный анализ. Классическая нормальная линейная модель множественной регрессии. Оценка параметров модели. Ковариационная матрица. Доверительные интервалы, оценка значимости множественной регрессии	Moodle (Интеллектуальный анализ данных)
1.2	Нечеткие числа и операции над ними	Основные понятия теории нечетких множеств. Понятие нечеткого числа. Нечеткие числа L-R-типа. Симметричные нечеткие числа. Нечеткая арифметика: сложение, вычитание, умножение и деление нечетких чисел	Moodle (Интеллектуальный анализ данных)
1.3	Нечеткий линейный регрессионный анализ данных	Нечеткие линейные регрессионные модели. Метод наименьших квадратов для оценки параметров нечеткой множественной линейной регрессионной модели. Оценка параметров нечеткой регрессионной модели методами линейного программирования	Moodle (Интеллектуальный анализ данных)
1.4	Поиск ассоциативных правил	Поиск ассоциативных правил. Метод Apriory. Характеристики ассоциативных правил.	Moodle (Интеллектуальный анализ данных)
2. Практические занятия			
2.1	Регрессионное моделирование	Линейные регрессионные модели. Линейная парная регрессия. Оценка параметров парной регрессии.	Moodle (Интеллектуальный анализ данных)

		Коэффициент корреляции. Оценка значимости уравнения регрессии. Коэффициент детерминации. Множественный регрессионный анализ. Классическая нормальная линейная модель множественной регрессии. Оценка параметров модели. Ковариационная матрица. Доверительные интервалы, оценка значимости множественной регрессии	ный анализ данных)
2.2	Нечеткие числа и операции над ними	Нечеткие числа L-R-типа. Симметричные нечеткие числа. Нечеткая арифметика: сложение, вычитание, умножение и деление нечетких чисел	Moodle (Интеллектуальный анализ данных)
2.3	Нечеткий линейный регрессионный анализ данных	Нечеткие линейные регрессионные модели. Метод наименьших квадратов для оценки параметров нечеткой множественной линейной регрессионной модели. Оценка параметров нечеткой регрессионной модели методами линейного программирования	Moodle (Интеллектуальный анализ данных)
2.4	Поиск ассоциативных правил	Принятие решений с помощью применения ассоциативных правил. Поиск ассоциативных правил. Метод Apriori. Нахождение характеристик при решении задач.	Moodle (Интеллектуальный анализ данных)

13.2. Темы (разделы) дисциплины и виды занятий

№ п/п	Наименование темы (раздела) дисциплины	Виды занятий (количество часов)				
		Лекции	Практические	Лабораторные	Самостоятельная работа	Всего
1	Эконометрическое моделирование	4	0	2	16	22
2	Нечеткие числа и операции над ними	6	0	3	12	21
3	Нечеткий линейный регрессионный анализ данных	6	0	4	24	34
4	Поиск ассоциативных правил	8	0	3	20	31
	Итого:	24	0	12	72	108

14. Методические указания для обучающихся по освоению дисциплины

(рекомендации обучающимся по освоению дисциплины: указание наиболее сложных разделов, работа с конспектами лекций, презентационным материалом, рекомендации по выполнению курсовой работы, по организации самостоятельной работы по дисциплине и др.)

Освоение дисциплины «Интеллектуальный анализ данных» включает лекционные и практические занятия, самостоятельную работу обучающихся. Рекомендуется подробно конспектировать лекционный материал, смотреть презентацию по соответствующей теме, чтобы систематизировать изучаемый материал. В течение семестра обучающимся предлагается выполнить проектное задание. К промежуточной аттестации, проводящейся на последнем занятии, предоставляется презентация по выполненному проекту.

При использовании дистанционных образовательных технологий и электронного обучения выполнять все указания преподавателей по работе на LMS-платформе, своевременно подключаться к online-занятиям, соблюдать рекомендации по организации самостоятельной работы.

15. Перечень основной и дополнительной литературы, ресурсов интернет, необходимых для освоения дисциплины (список литературы оформляется в соответствии с требованиями ГОСТ и используется общая сквозная нумерация для всех видов источников)

а) основная литература:

№ п/п	Источник
1	Ганичева А.В. Эконометрика / А.В. Ганичева, А.В. Ганичев. – Москва: Лань, 2016. – 176 с. – Режим доступа: https://e.lanbook.com/book/169792 .
2	Путко Б. А. Эконометрика / Б. А. Путко, Н. Ш. Кремер. – Москва : Юнити-Дана, 2012. –

	329 с. – Режим доступа: http://biblioclub.ru/index.php?page=book&id=118251 .
3	Тихомирова А. Н. <i>Нечеткие модели дискретной математики</i> / А. Н. Тихомирова, М. Г. Клейменова. – Москва : МИФИ, 2011. – 108 с. – Режим доступа: http://biblioclub.ru/index.php?page=book&id=231713 .

б) дополнительная литература:

№ п/п	Источник
4	Айвазян С. А. <i>Прикладная статистика и основы эконометрики : учебник для студ. экон. спец. вузов</i> / С. А. Айвазян, В. С. Мхитарян. — М. : ЮНИТИ, 1998. — 1022 с.
5	Магнус Я. Р. <i>Эконометрика. Начальный курс : учебник для студ. вузов</i> / Я. Р. Магнус, П. К. Катыхов, А. А. Пересецкий. — М. : Дело, 2005. — 503 с.
6	<i>Математико-статистические методы исследования взаимосвязей в экономике : из теории и практики статистики ГДР</i> / К. Отто [и др.]. — М. : Статистика, 1977. — 181 с.
7	Лисицин Д. В. <i>Методы построения регрессионных моделей</i> / Д. В. Лисицин. — Новосибирск : НГТУ, 2011. — 77 с. — Режим доступа: http://biblioclub.ru/index.php?page=book&id=228876 .
8	Пегат А. <i>Нечеткое моделирование и управление</i> / А. Пегат. — М. : Бином, 2009. — 798 с.

в) информационные электронно-образовательные ресурсы (официальные ресурсы интернет):

№ п/п	Ресурс
9	www.lib.vsu.ru — Зональная научная библиотека ВГУ
10	Аристова Е.М. <i>Курс «Интеллектуальный анализ данных»</i> / Образовательный портал «Электронный университет ВГУ». – Режим доступа: https://edu.vsu.ru/

16. Перечень учебно-методического обеспечения для самостоятельной работы (учебно-методические рекомендации, пособия, задачки, методические указания по выполнению практических (контрольных), курсовых работ и др.)

Самостоятельная работа обучающегося должна включать подготовку к практическим занятиям и подготовку к промежуточной аттестации. Для этого рекомендуется освоить теоретический материал, соответствующих тем, по конспектам лекций и презентационному материалу, размещенному на ЭО ресурсах, литературу из представленного ниже перечня, материалы с тематических ресурсов сети Интернет.

№ п/п	Источник
11	<i>Основы нечеткого моделирования в среде MatLab : учеб. пособие / сост.: Т. М. Леденева, Д. С. Татаркин, А. С. Тарасова. — Воронеж : ЛОП ВГУ, 2006. — 51 с.</i> Режим доступа: http://www.lib.vsu.ru/elib/texts/method/vsu/sep06167.pdf

17. Образовательные технологии, используемые при реализации учебной дисциплины, включая дистанционные образовательные технологии (ДОТ), электронное обучение (ЭО), смешанное обучение): (При реализации дисциплины могут проводиться различные типы лекций (вводная, обзорная и т.д.), семинарские занятия (проблемные, дискуссионные и т.д.), применяться дистанционные образовательные технологии в части освоения лекционного материала, проведения текущей аттестации, самостоятельной работы по дисциплине или отдельным ее разделам и т.д. При применении ЭО и ДОТ необходимо в п.15 в) указать используемые ресурсы (см. пример выше)

Дисциплина реализуется с применением электронного обучения и дистанционных образовательных технологий. Для организации самостоятельной работы обучающихся используется онлайн-курс «Интеллектуальный анализ данных», размещенный на платформе Электронного университета ВГУ (LMS moodle), а также Интернет-ресурсы, приведенные в п.15в.

18. Материально-техническое обеспечение дисциплины:

Лекционная аудитория должна быть оснащена современным компьютером с подключенным к нему проектором, видеотерминалом и настенным экраном.

Лабораторные занятия должны проводиться в специализированной аудитории, оснащенной персональными компьютерами на базе ОС Windows и программным обеспечением в соответствии с тематикой изучаемого материала. Число рабочих мест в аудитории должно быть таким, чтобы обеспечивалась индивидуальная работа студента на отдельном персональном компьютере.

19. Оценочные средства для проведения текущей и промежуточной аттестаций

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

№ п/п	Наименование раздела дисциплины (модуля)	Компетенция(и)	Индикатор(ы) достижения компетенции	Оценочные средства
1	Эконометрическое моделирование	ПК-1 ПК-3 ПК-6	ПК-1.1 ПК-3.1 ПК-6.3	Практико-ориентированные задания
2	Нечеткие числа и операции над ними	ПК-1 ПК-3 ПК-6	ПК-1.1 ПК-3.1 ПК-6.3	Практико-ориентированные задания
3	Нечеткий линейный регрессионный анализ данных	ПК-1 ПК-3 ПК-6	ПК-1.1 ПК-3.1 ПК-6.3	Практико-ориентированные задания
Промежуточная аттестация форма контроля – зачет с оценкой				<i>Перечень вопросов Практическое задание</i>

20 Типовые оценочные средства и методические материалы, определяющие процедуры оценивания

20.1 Текущий контроль успеваемости

Контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств: контрольная работа

Пример заданий для контрольной работы

Тема: Множественная регрессия и корреляция

Вариант 1

1. По данным таблицы составить уравнение множественной регрессии, построить его стандартизированное уравнение и провести оценку существенных независимых переменных.

x_1	10	11	12	14	15	16	17	18	19	20	21	22	23	24
x_2	11,2	12,5	12,9	14,1	14,8	16,1	17,5	18,9	18,9	20	21,1	22,2	23,3	24,4
x_3	0,9	0,1	1	0,5	1,1	1,2	0,1	0,5	0,3	1,3	1,1	0,2	1,44	0,7
y	0,2	0,4	0,5	1,1	0,7	0,1	0,08	1,2	0,7	0,48	1,9	0,4	0,8	0,71

2. Провести корреляционный анализ для построенного уравнения множественной регрессии.

Критерии оценки

Оценка «отлично» выставляется студенту за все правильно решённые задачи.

Оценка «хорошо» выставляется студенту за 2 решённые задачи с недочётами.

Оценка «удовлетворительно» выставляется студенту за 1 правильно решённую задачу.

Оценка «неудовлетворительно» выставляется студенту, если не решено ни одной задачи.

20.2 Промежуточная аттестация

Промежуточная аттестация по дисциплине осуществляется с помощью следующих оценочных средств: вопросы к зачету

Перечень вопросов к зачету

1. Основные аспекты эконометрического моделирования.
2. Линейные регрессионные модели. Линейная парная регрессия.
3. Оценка параметров парной регрессии.
4. Коэффициент корреляции. Оценка значимости уравнения регрессии. Коэффициент детерминации.
5. Классическая нормальная линейная модель множественной регрессии.
6. Оценка параметров модели. Ковариационная матрица.
7. Доверительные интервалы, оценка значимости множественной регрессии
8. Основные понятия теории нечетких множеств.
9. Понятие нечеткого числа. Нечеткие числа L-R-типа. Симметричные нечеткие числа.
10. Нечеткая арифметика: сложение, вычитание, умножение и деление нечетких чисел.
11. Нечеткие линейные регрессионные модели.
12. Метод наименьших квадратов для оценки параметров нечеткой множественной линейной регрессионной модели.
13. Оценка параметров нечеткой регрессионной модели методами линейного программирования.

Критерии оценки ответов на вопросы зачета

Оценка «отлично» - студент демонстрирует глубокое понимание темы, умеет распространять вытекающие из них выводы для анализа данных.

Оценка «хорошо» - студент демонстрирует понимание теоретических положений темы и базовых понятий, но допускает неточности в ответах, испытывает затруднения в применении знаний к анализу информации.

Оценка «удовлетворительно» - студент отвечает не на все предложенные вопросы, но не менее чем на половину из них; не демонстрирует способности применения теоретических знаний для анализа данных.

Оценка «неудовлетворительно» - студент демонстрирует непонимание теоретических основ и базовых понятий курса.

20.3. Фонд оценочных средств сформированности компетенций студентов, рекомендуемый для проведения диагностических работ

Вопросы с вариантами ответов

1. Аналитик это ...

а) специалист в области анализа и моделирования

б) специалист в предметной области

- в) человек, решающий определенные задачи
- г) человек, который имеет опыт в программировании

Ответ: а)

2. Эксперт это ...

- а) специалист в области анализа и моделирование
- б) специалист в предметной области
- в) человек, решать определенные задачи
- г) человек, который имеет опыт в программировании

Ответ: б)

Вопросы с кратким текстовым ответом

1. Что подразумевается под классификацией объектов?

Ответ:

Классификация. Описание возможных классов, которые могут получиться.
Классификация – это установление зависимости дискретной выходной переменной от входных переменных

2. Что такое регрессия?

Ответ: Регрессия – это установление зависимости непрерывной выходной переменной от входных переменных / Зависимость одной переменной от другой по определенному закону.

3. Переменная измерена в количественной шкале. Как можно представить результаты измерений этой переменной?

Ответ: результаты измерений этой переменной можно представить в номинальной шкале измерений.

Критерии и шкалы оценивания заданий ФОС:

Для оценивания выполнения заданий используется балльная шкала:

1) закрытые задания (тестовые с вариантами ответов, средний уровень сложности):

- 1 балл – указан верный ответ;
- 0 баллов – указан неверный ответ (полностью или частично неверный).

2) открытые задания (тестовые с кратким текстовым ответом, повышенный уровень сложности):

- 2 балла – указан верный ответ;
- 0 баллов – указан неверный ответ (полностью или частично неверный).

Задания раздела 20.3 рекомендуются к использованию при проведении диагностических работ с целью оценки остаточных результатов освоения данной дисциплины (знаний, умений, навыков).