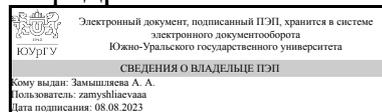


ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Заведующий выпускающей
кафедрой



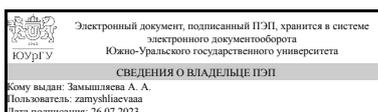
А. А. Замышляева

РАБОЧАЯ ПРОГРАММА

дисциплины 1.Ф.П0.02 Сбор, анализ и предобработка данных
для направления 01.03.02 Прикладная математика и информатика
уровень Бакалавриат
профиль подготовки Прикладная математика и искусственный интеллект
форма обучения очная
кафедра-разработчик Прикладная математика и программирование

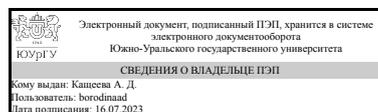
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 01.03.02 Прикладная математика и информатика, утверждённым приказом Минобрнауки от 10.01.2018 № 9

Зав.кафедрой разработчика,
д.физ.-мат.н., проф.



А. А. Замышляева

Разработчик программы,
преподаватель



А. Д. Кашеева

1. Цели и задачи дисциплины

узнать основные приемы сбора, предобработки, анализа набора данных, а также устройство систем информационного поиска. Задачами дисциплины являются изучение современных технологий поисковых систем, методов анализа и предварительной обработки данных, а также приобретение навыков программной реализации изучаемых методов с использованием различных программных сред.

Краткое содержание дисциплины

В курсе рассматриваются основные понятия информационного поиска, технологии поисковых систем, этапы подготовки данных для машинного обучения, методы анализа и предварительной обработки временных рядов.

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-3 (ПК-1 модели) Способен классифицировать и идентифицировать задачи искусственного интеллекта, выбирать адекватные методы и инструментальные средства решения задач искусственного интеллекта	Умеет: [ПК-1.3. У-1.] осуществлять сбор исходной информации с использованием платформ данных (облачных и внутрикорпоративных) Имеет практический опыт: сбора информации с использованием платформ данных (облачных и внутрикорпоративных)
ПК-10 (ПК-7 модели) Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта	Знает: [ПК-7.2. З-1.] методы редукции размерности элементов набора данных и их предварительной статистической обработки, разметки структурированных и неструктурированных данных; Умеет: [ПК-7.2. У-1.] выявлять и исключать из массива данных ошибочные данные и выбросы; [ПК-7.2. У-3.] осуществлять разметку структурированных и неструктурированных данных

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Нет	Анализ требований и проектирование систем искусственного интеллекта, Администрирование и проектирование хранилищ больших данных, Введение в искусственный интеллект и нейронные сети, Математическая статистика, Высокопроизводительные параллельные вычисления на кластерных системах, Современные архитектуры глубоких искусственных нейронных сетей, Производственная практика (практика по

получению профессиональных умений и опыта профессиональной деятельности) (8 семестр)

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Нет

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 4 з.е., 144 ч., 72,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		3	
Общая трудоёмкость дисциплины	144	144	
<i>Аудиторные занятия:</i>	64	64	
Лекции (Л)	32	32	
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	0	0	
Лабораторные работы (ЛР)	32	32	
<i>Самостоятельная работа (СРС)</i>	71,5	71,5	
Подготовка к лабораторным работам №1-8	41,5	41,5	
Подготовка к дифференцированному зачету	30	30	
Консультации и промежуточная аттестация	8,5	8,5	
Вид контроля (зачет, диф.зачет, экзамен)	-	диф.зачет	

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Сбор данных. Информационный поиск.	6	4	0	2
2	Предварительная обработка данных. Технологии анализа данных.	16	8	0	8
3	Корреляционный анализ данных	16	8	0	8
4	Регрессионный анализ данных	16	8	0	8
5	Анализ временных данных	10	4	0	6

5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Основные понятия информационного поиска. Поисковые системы. Понятие датасета для машинного обучения.	4

2	2	Python для обработки и анализа данных. Установка и настройка Jupyter Notebook, переменные и типы данных, строки, списки, цикл for, вложенные списки, условный оператор, цикл while, функции, словари, Pandas для анализа данных	4
3	2	Предварительный анализ данных. Описательная статистика. Восстановление пропусков в данных. Методология CRISP-DM для подготовки данных	4
4	3	Корреляционный анализ. Основные понятия	2
5	3	Корреляционный анализ качественных и количественных связей	4
6	3	Канонические корреляции	2
7	4	Регрессионный анализ. Основные понятия	2
8	4	Виды регрессионных моделей	6
9	5	Анализ временных данных. Основные понятия. Методы сглаживания временных данных и моделирования тенденции развития	4

5.2. Практические занятия, семинары

Не предусмотрены

5.3. Лабораторные работы

№ занятия	№ раздела	Наименование или краткое содержание лабораторной работы	Кол-во часов
1	1	Лабораторная работа 1. Сбор данных, восстановление пропусков, замена отсутствующих данных, удаление дубликатов, удаление объектов с отсутствующими данными.	2
2,3	2	Лабораторная работа 2. Использование Excel для описательной статистики	4
4,5	2	Лабораторная работа 3. Использование Python для описательной статистики	4
6,7	3	Лабораторная работа 4. Корреляционный анализ качественных признаков	4
8	3	Лабораторная работа 5. Корреляционный анализ количественных признаков	2
9	3	Контрольная работа 1. Теоретический материал по тематикам: классификация данных, пространственные, временные данные, корреляционный анализ	2
10,11	4	Лабораторная работа 6. Линейная регрессионная модель	4
12,13	4	Лабораторная работа 7. Нелинейная регрессионная модель	4
14,15	5	Лабораторная работа 8. Анализ временных данных	4
16	5	Контрольная работа 2. Теоретический материал по тематикам: регрессионные модели, временные ряды	2

5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Подготовка к лабораторным работам №1-8	"ЭУМД, 2, гл. 1-3, 5"; "ЭУМД, 1, гл. 1-3, 5"; "ЭУМД, 5, гл. 2-8"; "ЭУМД, 3, гл. 1, 3-5, 7-14"; "ЭУМД, 4, гл. 1, 2, 4, 5"; "ЭУМД, 6, гл. 1, 2, 3, 5" "ЭУМД, 7, гл. 1-3, 5, 6";	3	41,5
Подготовка к дифференцированному зачету	"ЭУМД, 1, гл. 1-3, 5"; "ЭУМД, 5, гл. 2-8"; "ЭУМД, 3, гл. 1, 3-5, 7-14"; "ЭУМД, 4, гл. 1, 2, 4, 5"; "ЭУМД, 6, гл. 1, 2, 3, 5"	3	30

6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учитывается в ПА
1	3	Текущий контроль	Лабораторная работа №1	2	3	3 балла: Студент отвечает на все теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 2 балла: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 1 балл: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены. 0 баллов: Студент не отвечает на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены.	дифференцированный зачет
2	3	Текущий контроль	Лабораторная работа №2	2	3	3 балла: Студент отвечает на все теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы	дифференцированный зачет

					<p>выполнены. 2 балла: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 1 балл: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены. 0 баллов: Студент не отвечает на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены.</p>		
3	3	Текущий контроль	Лабораторная работа №3	4	2	<p>3 балла: Студент отвечает на все теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 2 балла: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 1 балл: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены. 0 баллов: Студент не отвечает на теоретические</p>	дифференцированный зачет

						вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены.	
4	3	Текущий контроль	Лабораторная работа №4	4	3	3 балла: Студент отвечает на все теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 2 балла: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 1 балл: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены. 0 баллов: Студент не отвечает на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены.	дифференцированный зачет
5	3	Текущий контроль	Лабораторная работа №5	6	3	3 балла: Студент отвечает на все теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 2 балла: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены.	дифференцированный зачет

						1 балл: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены. 0 баллов: Студент не отвечает на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены.	
6	3	Текущий контроль	Лабораторная работа №6	10	3	3 балла: Студент отвечает на все теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 2 балла: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 1 балл: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены. 0 баллов: Студент не отвечает на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены.	дифференцированный зачет
7	3	Текущий контроль	Лабораторная работа №7	10	3	3 балла: Студент отвечает на все теоретические вопросы по теме лабораторной работы.	дифференцированный зачет

					<p>Все задания лабораторной работы выполнены. 2 балла: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 1 балл: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены. 0 баллов: Студент не отвечает на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены.</p>		
8	3	Текущий контроль	Лабораторная работа №8	10	3	<p>3 балла: Студент отвечает на все теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 2 балла: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Все задания лабораторной работы выполнены. 1 балл: Студент отвечает с затруднениями на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены. 0 баллов: Студент не</p>	дифференцированный зачет

						отвечает на теоретические вопросы по теме лабораторной работы. Задания лабораторной работы полностью не выполнены.	
9	3	Текущий контроль	Контрольная работа 1	25	4	4 балла: Студент ответил на все теоретические вопросы контрольной работы. 3 балла: Студент ответил на 3 вопроса контрольной работы. 2 балла: Студент ответил на 2 вопроса контрольной работы. 1 балл: Студент ответил на 1 вопрос контрольной работы. 0 баллов: Студент не ответил на вопросы контрольной работы	дифференцированный зачет
10	3	Текущий контроль	Контрольная работа 2	25	4	4 балла: Студент ответил на все теоретические вопросы контрольной работы. 3 балла: Студент ответил на 3 вопроса контрольной работы. 2 балла: Студент ответил на 2 вопроса контрольной работы. 1 балл: Студент ответил на 1 вопрос контрольной работы. 0 баллов: Студент не ответил на вопросы контрольной работы	дифференцированный зачет
11	3	Промежуточная аттестация	Дифференцированный зачет	-	4	4 балла получает студент, обнаруживший всестороннее, систематическое и глубокое знание учебно-программного материала, умение свободно выполнять задания, предусмотренные билетом и свободно отвечающий на дополнительные вопросы	дифференцированный зачет

					<p>3 балла заслуживает студент, обнаруживший полное знание учебно-программного материала, успешно выполняющий предусмотренные в билете задания, но отвечающий на дополнительные вопросы с затруднениями</p> <p>2 балла получает студент, допустивший погрешности в ответе на дифференцированном зачете и при выполнении заданий, но обладающим необходимыми знаниями для их устранения под руководством преподавателя;</p> <p>1 балл ставится студенту, обнаружившему пробелы в знаниях основного учебно-программного материала, допустившему принципиальные ошибки в выполнении предусмотренных билетом заданий.</p> <p>0 баллов ставится студенту, который не смог выполнить ни одно задание в билете.</p>	
--	--	--	--	--	--	--

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
дифференцированный зачет	<p>Рейтинг обучающегося по дисциплине формируется по результатам текущего контроля. Контрольное мероприятие дифференцированного зачета проводится в очной форме и не является обязательным, однако студент может прийти на дифференцированный зачет и повысить свой рейтинг. Студенту на дифференцированном зачете выдаётся билет. Даётся 90 минут для подготовки к ответу. Проводится собеседование по выданным вопросам.</p>	В соответствии с пп. 2.5, 2.6 Положения

6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ										
		1	2	3	4	5	6	7	8	9	10	11
ПК-3	Умеет: [ПК-1.3. У-1.] осуществлять сбор исходной информации с использованием платформ данных (облачных и внутрикорпоративных)	++									+	+
ПК-3	Имеет практический опыт: сбора информации с использованием платформ данных (облачных и внутрикорпоративных)	++						+				+
ПК-10	Знает: [ПК-7.2. З-1.] методы редукции размерности элементов набора данных и их предварительной статистической обработки, разметки структурированных и неструктурированных данных;			++			++			++		+
ПК-10	Умеет: [ПК-7.2. У-1.] выявлять и исключать из массива данных ошибочные данные и выбросы; [ПК-7.2. У-3.] осуществлять разметку структурированных и неструктурированных данных	+++			+			++				+

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

7. Учебно-методическое и информационное обеспечение дисциплины

Печатная учебно-методическая документация

а) *основная литература:*

Не предусмотрена

б) *дополнительная литература:*

Не предусмотрена

в) *отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:*

Не предусмотрены

г) *методические указания для студентов по освоению дисциплины:*

1. Методические указания к СРС

из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Методические указания к СРС

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Основная литература	Электронно-библиотечная система издательства Лань	Теофили, Т. Глубокое обучение для поисковых систем : руководство / Т. Теофили ; перевод с английского Д. А. Беликова. — Москва : ДМК Пресс, 2020. — 318 с. — ISBN 978-5-97060-776-3. — Текст : электронный // Лань : электронно-библиотечная система. https://e.lanbook.com/book/140574
2	Основная литература	Электронно-библиотечная система	Макшанов, А. В. Большие данные. Big Data / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. — 3-е изд., стер. — Санкт-Петербург : Лань, 2023. — 188 с. — ISBN

		издательства Лань	978-5-507-46866-9. — Текст : электронный // Лань : электронно-библиотечная система. https://e.lanbook.com/book/322664
3	Основная литература	Электронно-библиотечная система издательства Лань	Низаметдинов, Ш. У. Анализ данных : учебное пособие / Ш. У. Низаметдинов, В. П. Румянцев. — Москва : НИЯУ МИФИ, 2012. — 288 с. — ISBN 978-5-7262-1687-4. — Текст : электронный // Лань : электронно-библиотечная система. https://e.lanbook.com/book/75847
4	Дополнительная литература	Электронно-библиотечная система издательства Лань	Маккинни, У. Python и анализ данных / У. Маккинни ; перевод с английского А. А. Слинкина. — 2-ое изд., испр. и доп. — Москва : ДМК Пресс, 2020. — 540 с. — ISBN 978-5-97060-590-5. — Текст : электронный // Лань : электронно-библиотечная система. https://e.lanbook.com/book/131721
5	Дополнительная литература	Электронно-библиотечная система издательства Лань	Основы обработки и анализа информации : учебно-методическое пособие / Е. Г. Шмакова, А. С. Кузнецов, С. В. Разливинская [и др.]. — Москва : РТУ МИРЭА, 2022. — Часть 1 — 2022. — 118 с. — ISBN 978-5-7339-1632-3. — Текст : электронный // Лань : электронно-библиотечная система. https://e.lanbook.com/book/265562
6	Основная литература	Электронно-библиотечная система Znanium.com	Жуковский, О. И. Информационные технологии и анализ данных : учебное пособие / О. И. Жуковский. - Томск : Эль-Контент, 2014. - 130 с. - ISBN 978-5-4332-0158-3. - Текст : электронный. https://znanium.com/catalog/product/1845865
7	Основная литература	Образовательная платформа Юрайт	Анализ данных : учебник для академического бакалавриата / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2018. — 490 с. — (Бакалавр. Академический курс). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. https://urait.ru/bcode/412967

Перечень используемого программного обеспечения:

1. Microsoft-Office(бессрочно)
2. -Python(бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Лекции	336 (3б)	Компьютер для преподавателя, проектор, экран
Лабораторные занятия	327 (3б)	Персональные компьютеры для студентов