

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Тольяттинский государственный университет»

Б1.В.01

(индекс дисциплины)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Методы и инструментальные средства прикладного анализа данных

(наименование дисциплины)

по направлению подготовки
09.04.03 Прикладная информатика

направленность (профиль)
Прикладной анализ данных

Форма обучения: очная

Год набора: 2022

Общая трудоемкость: 6 ЗЕ

Распределение часов дисциплины по семестрам

Семестр	2	Итого
Форма контроля	Экзамен	
Вид занятий		
Лекции	16	16
Лабораторные		
Практические	16	16
Руководство: курсовые работы (проекты) / РГР		
Промежуточная аттестация	0,35	0,35
Контактная работа	32,35	32,35
Самостоятельная работа	148	148
Контроль	35,65	35,65
Итого	216	216

Рабочую программу составил: доцент кафедры, к.п.н., Крайнова О.А.

(должность, ученое звание, степень, Фамилия И.О.)

Рецензирование рабочей программы дисциплины:



Отсутствует



Рецензент

(должность, ученое звание, степень, Фамилия И.О.)

Рабочая программа дисциплины составлена на основании ФГОС ВО и учебного плана направления подготовки

09.04.03 Прикладная информатика

Срок действия рабочей программы дисциплины до «31» августа 2024г.

УТВЕРЖДЕНО

На заседании кафедры «Прикладная математика и информатика»

(протокол заседания № 2 от «15» сентября 2021 г.).

1. Цель освоения дисциплины

Цель освоения дисциплины – формирование у студентов знаний о концепции Big Data, базовых методах и технологиях сбора, хранения, обработки и методах анализа Big Data и интерпретации результатов.

2. Место дисциплины в структуре ОПОП ВО

Дисциплины и практики, на освоении которых базируется данная дисциплина: «Цифровое общество и проблемы прикладной информатики», «Методы и технологии прикладной информатики».

Дисциплины и практики, для которых освоение данной дисциплины необходимо как предшествующее: Теоретические и прикладные основы анализа данных, Аналитический проект по исследованию больших данных 2, Методы интерпретации и визуализации анализа больших данных, Производственная практика (научно-исследовательская работа).

3. Планируемые результаты обучения

Формируемые и контролируемые компетенции (код и наименование)	Индикаторы достижения компетенций (код и наименование)	Планируемые результаты обучения
ПК-4 Способен управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации	ИД-1ПК-4 Знает способы управления этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации	Знать: положения теории принятия решений и математические методы поддержки принятия решений в условиях неопределенности и риска.
	ИД-2ПК-4 Умеет управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации	Уметь: собирать, обрабатывать и интерпретировать экспериментальные данные, применять математические и имитационные модели для принятия управленческих решений в условиях неопределенности и риска.
	ИД-3ПК-4 Владеет навыками управления этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных в организации	Владеть: навыками использования эффективных инструментальных средств математического и имитационного моделирования для принятия управленческих решений при разработке информационных ресурсов.

4. Структура и содержание дисциплины

Модуль (раздел)	Вид учебной работы	Наименование тем занятий (учебной работы)	Семестр	Объем, ч.	Баллы	Интерактив, ч.	Формы текущего контроля (наименование оценочного средства)
Модуль 1. Введение в большие данные	Ср	Тема 1. Понятие больших данных.	1	28	-	-	Промежуточный тест
	Лек	Тема 2. Обзор и средства больших данных. Используемые технологии, инструменты, программы и их особенности.	1	1	-	-	
	Пр	Подготовка данных к анализу.	1	2	-	-	Отчет по практической работе
Модуль 2. Методы многомерного статистического анализа и анализа нечисловой информации	Ср	Тема 3. Выбор метода. Факторы, влияющие на выбор метода. Настройка параметров. Оценка	1	25	-	-	Промежуточный тест
	Лек	Тема 4. Факторный анализ. Дискриминантный анализ. Кластерный анализ. Многомерное шкалирование. Методы контроля качества.	1	4	-	-	
	Пр	Обработка данных при факторном и корреляционном анализе.	1	4	-	-	Отчет по практической работе
Модуль 3. Технологии хранения и обработки больших данных	Ср	Тема 5. Основные направления развития методов обработки и хранения данных.	1	30	-	-	Промежуточный тест
	Лек	Тема 6. Проблема хранения неструктурированных данных. Проблема преобразования данных. Семантические анализаторы.	1	4	-	-	

Модуль (раздел)	Вид учебной работы	Наименование тем занятий (учебной работы)	Семестр	Объем, ч.	Баллы	Интерактив, ч.	Формы текущего контроля (наименование оценочного средства)
	Пр	Использование и преимущества экосистемы Hadoop.	1	4	-	-	Отчет по практической работе №3
Модуль 4. Программирование обработки и загрузки больших данных	Ср	Тема 7. Использование языков программирования для больших данных.	1	35	-	-	Промежуточный тест
	Лек	Тема 8. Фреймворки, базы данных, аналитические платформы и прочие инструменты обработки и загрузки больших данных.	1	4	-	-	
	Пр	Использование Python для анализа больших данных.	1	4	-	-	Отчет по практической работе №4
Модуль 5. Аналитика в больших данных	Ср	Тема 9. Аналитика больших данных — реалии и перспективы в России и мире.	1	30	-	-	Промежуточный тест
	Лек	Тема 10. Технологии и методы анализа, которые используются для анализа больших данных. Применение и возможности. Решения на основе больших данных.	1	3	-	-	
	Пр	Большие данные в банках, бизнесе, маркетинге.	1	2	-	-	Отчет по практической работе №5
	ПА		1	0,35	-	-	
	Контроль		1	35,65	-	-	

Модуль (раздел)	Вид учебной работы	Наименование тем занятий (учебной работы)	Семестр	Объем, ч.	Баллы	Интерактив, ч.	Формы текущего контроля (наименование оценочного средства)
Итого:				216			

5. Образовательные технологии

В рамках учебного курса предусмотрены следующие образовательные технологии:

- технология традиционного обучения: лекции и практические работы, самостоятельная работа;
- технология проектного обучения: реализация и защита отчетов по практическим работам.

6. Методические указания по освоению дисциплины

6.1. Рекомендации по подготовке к лекционным занятиям

Изучение дисциплины требует систематического и последовательного накопления знаний, следовательно, пропуски отдельных тем не позволяют глубоко освоить предмет.

В ходе лекционных следует обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации. Задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.

Студент может дополнить список использованной литературы современными источниками, не представленными в списке рекомендованной литературы, и в дальнейшем использовать собственные подготовленные учебные материалы при написании курсовых и выпускных квалификационных работ.

6.2. Рекомендации по подготовке к практическим занятиям

Студентам следует доводить каждую практическую работу до окончательного решения, демонстрировать понимание проведенных расчетов (анализов, ситуаций), в случае затруднений обращаться к преподавателю.

Для того чтобы практические занятия приносили максимальную пользу, необходимо помнить, что упражнение и решение задач проводятся по рассмотренному на лекциях материалу и связаны, как правило, с детальным разбором отдельных вопросов лекционного курса. Следует подчеркнуть, что только после усвоения лекционного материала с определенной точки зрения (а именно с той, с которой он излагается на лекциях) он будет закрепляться студентом на практических занятиях как в результате обсуждения и анализа лекционного материала, так и с помощью решения проблемных ситуаций, задач. При этих условиях студент не только хорошо усвоит материал, но и научится применять его на практике, а также получит дополнительный стимул (и это очень важно) для активной проработки лекции.

По результатам выполнения работы составляется отчет, который при необходимости нужно сопровождать комментариями, схемами, чертежами и рисунками.

Следует помнить, что выполнение каждой работы должно доводиться до окончательного логического ответа, которого требует условие, и по возможности с выводом. Полученный ответ следует проверить способами, вытекающими из существа данной задачи. Полезно также (если возможно) решать несколькими способами и сравнить полученные результаты. Решение задач данного типа нужно продолжать до приобретения твердых навыков в их решении.

6.3. Рекомендации по подготовке к экзамену

Подготовка к экзамену способствует закреплению, углублению и обобщению знаний, получаемых, в процессе обучения, а также применению их к решению практических задач. Готовясь к экзамену, студент ликвидирует имеющиеся пробелы в знаниях, углубляет, систематизирует и упорядочивает свои знания. На экзамене студент демонстрирует то, что он приобрел в процессе обучения по конкретной учебной дисциплине.

Необходимо ориентировать студентов на систематическую подготовку к занятиям в течение семестра, что позволит использовать время экзаменационной сессии для систематизации знаний.

7. Оценочные средства

7.1. Паспорт оценочных средств

Семестр	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
2	ПК-4	<i>Отчеты по практическим работам, Промежуточное тестирование Вопросы к экзамену</i>

7.2. Типовые задания или иные материалы, необходимые для текущего контроля

7.2.1. _____ Тестовые задания (наименование оценочного средства)

Типовые примеры тестовых заданий

1. Интервальные данные – это (подчеркните правильные ответы):

- 1) данные с интервалом;
- 2) данные об интервалах;
- 3) количество измерений в каждом интервале;
- 4) количество интервалов в каждом измерении.

2. Кластерный анализ предназначен для:

- 1) группировки объектов;
- 3) ранжирования объектов;
- 2) группировки показателей;
- 4) ранжирования показателей.

3. Дихотомическая шкала это:

- 1) состоящая из “да” и “нет”;
- 3) состоящая из двух чисел;
- 2) состоящая из “истина” и “ложь”;
- 4) состоящая из двух рангов.

4. К нечисловым шкалам относятся:

- 1) номинальная;
- 3) абсолютная;
- 2) интервалов;
- 4) ранговая.

5. То, с чем сравнивается вычисляемое значение хи-квадрат, вычисляется в EXCEL функцией:

- 1) ХИ2РАСП;

- 3) ХИ2ТЕСТ;
- 2) ХИ2ОБР;
- 4) ХИ2.

7.2.2. _____ Отчеты по практическим работам

(наименование оценочного средства)

Типовые примеры заданий

Практическая работа 1. Подготовка данных к анализу.

Форма отчета по практической работе № 1. В отчет по практической работе должны быть включены следующие пункты:

- титульный лист;
- цель работы;
- краткие теоретические сведения;
- описание хода выполнения работы;
- результаты выполненной работы.

Практическая работа 2. Обработка данных при факторном и корреляционном анализе.

Форма отчета по практической работе № 3. В отчет по практической работе должны быть включены следующие пункты:

- титульный лист;
- цель работы;
- краткие теоретические сведения;
- описание хода выполнения работы;
- результаты выполненной работы.

Практическая работа 3. Использование и преимущества экосистемы Hadoop.

Форма отчета по практической работе № 3. В отчет по практической работе должны быть включены следующие пункты:

- титульный лист;
- цель работы;
- краткие теоретические сведения;
- описание хода выполнения работы;
- результаты выполненной работы.

Практическая работа 4. Использование Python для анализа больших данных.

Форма отчета по практической работе № 4. В отчет по практической работе должны быть включены следующие пункты:

- титульный лист;
- цель работы;
- краткие теоретические сведения;
- описание хода выполнения работы;
- результаты выполненной работы.

Практическая работа 5. Большие данные в банках, бизнесе, маркетинге..

Форма отчета по практической работе № 5. В отчет по практической работе должны быть включены следующие пункты:

- титульный лист;
- цель работы;

- краткие теоретические сведения;
- описание хода выполнения работы;
- результаты выполненной работы.

Темы письменных работ

Учебным планом не предусмотрено.

7.3. Оценочные средства для промежуточной аттестации по итогам освоения дисциплины

7.3.1. Вопросы к промежуточной аттестации

Семестр 2

№ п/п	Вопросы к экзамену
1.	Технологии больших данных.
2.	Задачи и компетенции аналитиков больших данных.
3.	Основные принципы работы с большими данными.
4.	MapReduce как модель распределенной обработки данных.
5.	Инструменты извлечения данных.
6.	Интеллектуальный анализ данных.
7.	Данные и метаданные. Методы и стадии Data Mining. Задачи Data Mining.
8.	Основные характеристики и типы больших данных.
9.	Сбор, хранение, обработка, извлечение и анализ больших потоков данных.
10.	Статистическое мышление и статистический анализ при работе с большими данными.
11.	Методы классификации и прогнозирования работы с большими данными.
12.	Факторный анализ.
13.	Дискриминантный анализ при работе с большими данными.
14.	Кластерный анализ при работе с большими данными.
15.	Многомерное шкалирование при работе с большими данными.
16.	Методы контроля качества при работе с большими данными.
17.	Современные программные средства для оперативного анализа данных.
18.	Основные направления развития методов обработки и хранения больших данных.
19.	Системы управления базами данных и организация управления базами хранения больших данных.
20.	Проблема хранения неструктурированных данных.
21.	Проблема преобразования данных.
22.	Семантические анализаторы.
23.	Языки программирования, используемые для работы с большими данными.
24.	Самообучающиеся автоматы.
25.	Фреймворки, используемые для работы с большими данными.
26.	Когнитивный анализ данных при работе с большими данными.
27.	Аналитические платформы для работы с большими данными.
28.	Аналитика больших данных. Математическая статистика.
29.	Прикладные инструменты анализа данных. Корреляция.

№ п/п	Вопросы к экзамену
30.	Введение в Data Mining – понятие, структура, составляющие и сопутствующие науки.
31.	Способы визуализации, интерпретации, проектная организация и бизнес рекомендации.
32.	Искусственные нейронные сети при работе с большими данными.
33.	Аналитика больших данных — реалии и перспективы в России и мире.
34.	Краудсорсинг.
35.	Смещение и интеграция данных.
36.	Машинное обучение при работе с большими данными.
37.	Большие данные в банках.
38.	Большие данные в бизнесе.
39.	Большие данные в маркетинге.
40.	Прогнозная аналитика.

7.3.2. Критерии и нормы оценки

Семестр	Форма проведения промежуточной аттестации	Критерии и нормы оценки	
1	Экзамен (устно)	«отлично»	Ставится студенту на экзамене, если он исчерпывающе и грамотно дал ответы на вопросы экзаменационного билета или при ответе допустил небольшую неточность на 1 вопрос, но при этом смог грамотно ответить на дополнительные вопросы ту, проявившему полные знания в рамках требований подготовки по дисциплине, усвоившему литературу, рекомендуемую программой и показавшему систематический характер знаний. В изложении материала и ответах на дополнительные вопросы допускаются небольшие неточности
		«хорошо»	Ставится студенту на экзамене, если он исчерпывающе и грамотно дал ответ на 1 вопрос экзаменационного билета, а на другой только тезисные высказывания или допустил небольшие неточности при ответе на вопросы экзаменационного билета и дал краткие ответы на дополнительные вопросы

Семестр	Форма проведения промежуточной аттестации	Критерии и нормы оценки	
		«удовлетворительно»	Ставится студенту на экзамене, если он не смог дать ответ на один из вопросов экзаменационного билета или ответил на все вопросы, но при этом ответы содержали только тезисные высказывания
		«неудовлетворительно»	Ставится студенту на экзамене, если он не дал ответ на вопросы экзаменационного билета или в ответе содержались фундаментальные ошибки

8. Учебно-методическое и информационное обеспечение дисциплины

8.1. Обязательная литература

№ п/п	Авторы, составители	Заглавие (заголовок)	Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.)	Год издания	Количество в научной библиотеке / Наименование ЭБС
1	Бендерская, О. Б.	Бизнес-аналитика : учебное пособие / О. Б. Бендерская. — Белгород : Белгородский государственный технологический университет им. В.Г. Шухова, ЭБС АСВ, 2017. — 162 с. — ISBN 2227-8397.	учеб. пособие	2017	ЭБС "IPRbooks"
2	Ибе О.	Компьютерные сети и службы удаленного доступа	учеб. пособие	2019	ЭБС "IPRbooks"
3	Казиев, В. М.	Введение в анализ, синтез и моделирование систем : учебное пособие / В. М. Казиев. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 270 с. — ISBN 978-5-4497-0307-1.	учебное пособие	2020	ЭБС "IPRbooks"
4	Лебедев А. С.	Методы Big Data : учебно-методическое пособие / А. С. Лебедев, Ш. Г. Магомедов. — Москва : РТУ МИРЭА, 2021. — 91 с.	учебно-методическое пособие	2021	ЭБС "Лань"
5	Соловьева С. В.	Информационные технологии в профессиональной деятельности. Инструментарий бизнес-аналитики : практикум / С. В. Соловьева, Ю. П. Александровская, Ю. В. Хайрутдинова. — Казань : Казанский национальный исследовательский технологический университет, 2017. — 104 с. — ISBN 978-5-7882-2217-2.	практикум	2017	ЭБС "IPRbooks"
6	Сузи Р. А.	Язык программирования Python	учеб. пособие	2016	ЭБС "IPRbooks"

8.2. Дополнительная литература

№ п/п	Авторы, составители	Заглавие (заголовок)	Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.)	Год издания	Количество в научной библиотеке / Наименование ЭБС
1	Олифер В. Г.	Основы сетей передачи данных	вводный учеб. курс	2016	ЭБС "IPRbooks"
2	Букунов С. В.	Основы объектно-ориентированного программирования	учеб. пособие	2017	ЭБС "IPRbooks"
3	Безручко В. Т.	Компьютерный практикум по курсу "Информатика"	учеб. пособие	2019	ЭБС "ZNANIUM.COM"
4	Дадян Э. Г.	Современные базы данных	учеб. пособие	2017	ЭБС "ZNANIUM.COM"
5	Дороганов В. А.	Компьютерная обработка данных	учеб. пособие	2017	ЭБС "IPRbooks"

8.3. Перечень профессиональных баз данных и информационных справочных систем

- Технологии и средства хранения и обработки данных. URL: <https://www.kp.ru/guide/sistemy-khraneniya-dannykh.html>
- Оперативная обработка транзакций (OLTP). URL: <https://docs.microsoft.com/ru-ru/azure/architecture/data-guide/relational-data/online-transaction-processing>
- Базы данных [Электронный ресурс]. URL: [https://msdn.microsoft.com/ru-ru/library/ms179422\(v=sql.120\).aspx](https://msdn.microsoft.com/ru-ru/library/ms179422(v=sql.120).aspx)
- Web of Science [Электронный ресурс] : мультидисциплинарная реферативная база данных. – Philadelphia: ClarivateAnalytics, 2016– . – Режим доступа : apps.webofknowledge.com. – Загл. с экрана. – Яз. рус., англ.
- Scopus [Электронный ресурс] : реферативная база данных. – Netherlands: Elsevier, 2004– . – Режим доступа : scopus.com. – Загл. С экрана. – Яз. рус., англ.
- Elibrary [Электронный ресурс] : научная электронная библиотека. – Москва : НЭБ, 2000– . – Режим доступа : elibrary.ru. – Загл. с экрана. – Яз. рус., англ.

8.4. Перечень программного обеспечения

№ п/п	Наименование ПО	Реквизиты договора (дата, номер, срок действия)
1	СУБД MS SQL Server Express + среда MS SQL Studio (рус) 2008 R2 или выше	бесплатное ПО
2	Язык программирования Python 3.0 и среда разработки PyScripter	бесплатное ПО
3	Office Standard	Бессрочная
4	Библиотека программного обеспечения Apache Hadoop	бесплатное ПО
5	Windows	Бессрочная

8.5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

№ п/п	Наименование оборудованных учебных кабинетов, лабораторий, мастерских и др. объектов для проведения практических и лабораторных занятий, помещений для самостоятельной работы обучающихся (номер аудитории)	Перечень основного оборудования
1	Компьютерный класс. Учебная аудитория для проведения занятий лекционного типа. Учебная аудитория для проведения занятий семинарского типа. Учебная аудитория для проведения лабораторных работ. Учебная аудитория для курсового проектирования (выполнения курсовых работ). Учебная аудитория для проведения групповых и индивидуальных консультаций Учебная аудитория для проведения занятий текущего контроля и промежуточной аттестации.	Компьютер (монитор 19", системный блок Pentium (R) Dual-Core E5500 2,8 GHz / 4 Gb / 500 Gb), столы ученические, столы компьютерные, стол преподавательский, стулья. Доска аудиторная(меловая)

№ п/п	Наименование оборудованных учебных кабинетов, лабораторий, мастерских и др. объектов для проведения практических и лабораторных занятий, помещений для самостоятельной работы обучающихся (номер аудитории)	Перечень основного оборудования
2	Учебная аудитория для проведения занятий лекционного типа. Учебная аудитория для проведения занятий семинарского типа. Учебная аудитория для курсового проектирования (выполнения курсовых работ). Учебная аудитория для проведения групповых и индивидуальных консультаций. Учебная аудитория для проведения занятий текущего контроля и промежуточной аттестации.	Столы ученические двухместные (моноблок), доска аудиторная 3-х секционная (меловая), стол преподавательский , стул, проектор Acer
3	Компьютерный класс. Помещение для самостоятельной работы. Учебная аудитория для проведения занятий семинарского типа. Учебная аудитория для курсового проектирования (выполнения курсовых работ). Учебная аудитория для проведения групповых и индивидуальных консультаций. Учебная аудитория для проведения занятий текущего контроля и промежуточной аттестации.	Стол ученический, стул, ПК с выходом в сеть интернет