

Министерство науки и высшего образования
Российской Федерации

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

Факультет математики и информационных технологий
Кафедра информационных систем управления

УТВЕРЖДАЮ
проректор

_____ П.А. Машаров
«16» апреля 2026 г.
МП

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

АНАЛИЗ БОЛЬШИХ ДАННЫХ

Укрупненная группа направлений подготовки	46.00.00 История и археология
Программа высшего образования	Программа магистратуры
Направление подготовки	46.04.02 Документоведение и архивоведение
Направленность (профиль) образовательной программы	Информационно-документационное обеспечение управления
Квалификация	Магистр
Форма обучения	Очная, очно-заочная

Рабочая программа может быть адаптирована для лиц
с ограниченными возможностями здоровья и инвалидов

Донецк 2026

Рабочая программа дисциплины «**Анализ больших данных**» для обучающихся по направлению подготовки 46.04.02 Документоведение и архивоведение (Профиль: Информационно-документационное обеспечение управления), составлена на основании Федерального государственного образовательного стандарта высшего образования – магистратура по направлению подготовки 46.04.02 Документоведение и архивоведение, утвержденный приказом Министерства науки и высшего образования Российской Федерации от 29 октября 2020 г. № 1345 (с изм. и доп.), Порядка организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры, утвержденного приказом Министерства науки и высшего образования Российской Федерации от 06 апреля 2021 г. № 245 (с изм. и доп.), в соответствии с учебным планом, утвержденным Ученым советом ФГБОУ ВО «ДонГУ» для набора 2026 года.

Разработчик:

доцент кафедры информационных систем
управления,
канд. экон. наук, доцент

А.М. Гизатулин

Рабочая программа одобрена на заседании кафедры информационных систем
управления.
Протокол от 13.04.2026 № 14.

Заведующий кафедрой

Н.Ш. Пономаренко

СОГЛАСОВАНО:

Врио декана факультета математики и
информационных технологий
16.04.2026

Н.Е. Мащенко

Учебно-методическая комиссия факультета математики и информационных технологий.
Протокол от 15.04.2026 № 3.
Председатель

Ю.В. Абраменкова

Руководитель основной образовательной
программы, д-р экон. наук, доц.
15.04.2026

Н.Ш. Пономаренко

1. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

1.1. Требования к предварительной подготовке обучающихся, предшествующие и сопутствующие дисциплины, на которых основывается изучение данной.

Моделирование управленческой деятельности, Управление знаниями в цифровой экономике, Интернет-технологии и веб-дизайн в профессиональной деятельности, Управление цифровыми инновациями.

1.2. Дисциплины, курсовые работы и практики, для которых освоение данной дисциплины необходимо как предшествующее:

Производственная практика: научно-исследовательская работа, Производственная практика: преддипломная практика, Подготовка к сдаче и сдача государственного экзамена, Подготовка к процедуре защиты и защита выпускной квалификационной работы.

2. ОПИСАНИЕ ДИСЦИПЛИНЫ

2.1. Общая характеристика

Наименование показателя	Значение показателя
Название образовательной программы (далее – ОП)	46.04.02 Документоведение и архивоведение (Профиль: Информационно-документационное обеспечение управления)
Шифр и название в соответствии с учебным планом	Б1.В.ОД.5. Анализ больших данных
Часть образовательной программы	Вариативная часть: выбор вуза
Количество зачетных единиц / всего часов	2 / 72

В случае предъявления от обучающегося или его родителя (законного представителя) заявления на обучение по адаптированной образовательной программе высшего образования, подкрепленного заключением психолого-медико-педагогической комиссии (ПМПК) или медико-социальной экспертизы (МСЭ) с рекомендациями создания индивидуальной программы реабилитации и абилитации (ИПРА), данная рабочая программа может быть адаптирована с учетом индивидуальных особенностей здоровья обучающегося.

2.2. Распределение часов по формам и периодам обучения

Форма обучения	курс	семестр	Общее количество часов					Форма контроля
			лекционных	лабораторных	практических	самостоятельной работы + контроль	всего	
Очная	2	3	17	–	17	38	72	экзамен
Очно-заочная	2	4	4	–	6	62	72	экзамен

3. ЦЕЛИ ДИСЦИПЛИНЫ

Цель дисциплины «Анализ больших данных» – формирование теоретических знаний и практических навыков, позволяющих использовать современные математико-статистические методы и программные средства для обработки и анализа данных.

4. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ КОМПОНЕНТА ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ, ИХ ИНДИКАТОРЫ И ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

4.1. Компетенции

Компетенции	Индикаторы	Результаты обучения
ПК-4. Способен осуществлять управление цифровой трансформацией документированных сфер деятельности организации (Профессиональный стандарт 07.013 «Специалист цифровой трансформации документированных сфер деятельности организации», С)	ПК-4.2. Осуществляет информационно-аналитическую поддержку анализа больших данных и интерпретирует его результаты	ПК-4.2.1. Знает основы теории анализа больших данных. ПК-4.2.2. Умеет решать задачи профессиональной деятельности на основе анализа данных. ПК-4.2.3. Владеет практическими методами анализа больших данных.

5. ПРОГРАММА ДИСЦИПЛИНЫ

Название темы	Краткое содержание темы (вопросы темы)
1. Методы классификации данных.	Постановка задачи классификации. Представление результатов классификации. Методы построения правил классификации. Деревья решений.
2. Кластерный анализ.	Постановка задачи. Меры близости, основанные на расстояниях. Иерархические методы. Итерационные методы.
3. Ассоциативные правила.	Постановка задачи. Представление результатов. Алгоритмы поиска ассоциативных правил.
4. Визуальный анализ данных (Visual Mining).	Процесс визуализации данных. Анализ средств визуализации. Методы визуализации.
5. Анализ текстовой информации (Text Mining).	Задача анализа текстов. Извлечение ключевых понятий из текста. Классификация текстовых документов. Средства анализа текстовой информации.
6. Стандарты Data Mining.	Стандарт CWM (Common Warehouse Metamodel). Стандарт CRISP-DM (CRoss-Industry Standard Process for Data Mining). Стандарт PMML (Predicted Model Markup Language). Стандарт DMX (Microsoft Data Mining eXtensions). Стандарт JDM (Java Data Mining).
7. Извлечение знаний в реальном времени.	Концепция извлечения знаний в реальном времени. Рекомендательные машины. Инструменты извлечения знаний в реальном времени.
8. Извлечение знаний из сети Интернет (Web Mining).	Проблемы анализа информации из сети Интернет. Этапы извлечения знаний из сети Интернет. Классификация задач извлечения знаний из сети Интернет. Классификация методов извлечения знаний из сети Интернет.

6. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

6.1. Форма обучения – очная, курс – 2, семестр – 3.

Наименования разделов и тем	Количество часов				
	Лекц.	Лабор.	Практ.	СРС+К	Всего
1. Методы классификации данных.	2	–	2	5	9

Наименования разделов и тем	Количество часов				
	Лекц.	Лабор.	Практ.	СРС+К	Всего
2. Кластерный анализ.	2	–	2	5	9
3. Ассоциативные правила.	2	–	2	5	9
4. Визуальный анализ данных (Visual Mining).	2	–	2	5	9
5. Анализ текстовой информации (Text Mining).	2	–	2	5	9
6. Стандарты Data Mining.	2	–	2	5	9
7. Извлечение знаний в реальном времени.	2	–	2	5	9
8. Извлечение знаний из сети Интернет (Web Mining).	3	–	3	3	9
ИТОГО ЗА СЕМЕСТР	17	–	17	38	72

6.2. Форма обучения – очно-заочная, курс – 2, семестр – 4.

Наименования разделов и тем	Количество часов				
	Лекц.	Лабор.	Практ.	СРС+К	Всего
1. Методы классификации данных.	0,5	–	1	7,5	9
2. Кластерный анализ.	0,5	–	1	7,5	9
3. Ассоциативные правила.	0,5	–	1	7,5	9
4. Визуальный анализ данных (Visual Mining).	0,5	–	1	7,5	9
5. Анализ текстовой информации (Text Mining).	0,5	–	0,5	8	9
6. Стандарты Data Mining.	0,5	–	0,5	8	9
7. Извлечение знаний в реальном времени.	0,5	–	0,5	8	9
8. Извлечение знаний из сети Интернет (Web Mining).	0,5	–	0,5	8	9
ИТОГО ЗА СЕМЕСТР	4	–	6	62	72

7. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ (СРЕДСТВА) ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

7.1 Контрольные вопросы

1. Постановка задачи классификации.
2. Представление результатов классификации.
3. Методы построения правил классификации.
4. Деревья решений.
5. Постановка задачи кластеризации.
6. Меры близости, основанные на расстояниях.
7. Иерархические методы.
8. Итерационные методы.
9. Постановка задачи ассоциации.
10. Представление результатов.
11. Алгоритмы поиска ассоциативных правил.
12. Процесс визуализации данных.
13. Анализ средств визуализации.
14. Методы визуализации.

15. Задача анализа текстов.
16. Извлечение ключевых понятий из текста.
17. Классификация текстовых документов.
18. Средства анализа текстовой информации.
19. Стандарт CWM (Common Warehouse Metamodel).
20. Стандарт CRISP-DM (CRoss-Industry Standard Process for Data Mining).
21. Стандарт PMML (Predicted Model Markup Language).
22. Стандарт DMX (Microsoft Data Mining eXtensions).
23. Стандарт JDM (Java Data Mining).
24. Концепция извлечения знаний в реальном времени.
25. Рекомендательные машины.
26. Инструменты извлечения знаний в реальном времени.
27. Проблемы анализа информации из сети Интернет.
28. Этапы извлечения знаний из сети Интернет.
29. Классификация задач извлечения знаний из сети Интернет.
30. Классификация методов извлечения знаний из сети Интернет.

7.2 Темы письменных работ (типы задач)

Практические работы по темам:

1. Методы классификации данных.
2. Кластерный анализ.
3. Ассоциативные правила.
4. Визуальный анализ данных (Visual Mining).
5. Анализ текстовой информации (Text Mining).
6. Стандарты Data Mining.
7. Извлечение знаний в реальном времени.
8. Извлечение знаний из сети Интернет (Web Mining).

7.3 Образец содержания экзаменационного билета

В случае ведения учебного процесса с использованием электронного обучения и дистанционных образовательных технологий, содержание билета может отличаться от приведенного.

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 1

1. Постановка задачи кластеризации.
2. Концепция извлечения знаний в реальном времени.

8. РАСПРЕДЕЛЕНИЕ БАЛЛОВ, КОТОРЫЕ ПОЛУЧАЮТ ОБУЧАЮЩИЕСЯ

Оценка результатов обучения проводится на основе накопительной системы, которая обеспечивает учет индивидуальных достижений обучающегося и отражает уровень освоения им компетенций, предусмотренных рабочей программой дисциплины.

По дисциплине предусмотрены следующие виды контроля: текущий контроль (оценивает различные виды аудиторной и внеаудиторной учебной деятельности студента); промежуточная аттестация (оценивает уровень и качество подготовки по дисциплине в целом). Текущий контроль осуществляется в течение семестра по каждому разделу дисциплины и предполагает начисление баллов за выполнение различных видов работ, включая *организационно-учебную работу в аудитории* (выполнение лабораторных,

практических и контрольных работ, участие в обсуждениях, и др.) и *самостоятельную работу* (выполнение домашних, индивидуальных, проектных заданий и пр.).

Промежуточная аттестация по дисциплине проводится в форме экзамена, который включает ответы на теоретические вопросы и решение практических и теоретических задач.

Оценивание текущего контроля и промежуточной аттестации проводится по накопительной системе. Итоговая оценка формируется как сумма баллов, полученных студентом за текущий контроль и за промежуточную аттестацию за семестр (максимум 100).

8.1. Форма обучения – очная, курс – 2, семестр – 3.

Номера разделов	Виды работ	Максимальное количество баллов
<i>Текущий контроль</i>		60
1	Организационно-учебная работа в аудитории	20
	Самостоятельная работа (выполнение практических работ по варианту)	40
<i>Промежуточная аттестация (экзамен)</i>		40
Общий итог за семестр		100

8.2. Форма обучения – очно-заочная, курс – 2, семестр – 4.

Номера разделов	Виды работ	Максимальное количество баллов
<i>Текущий контроль</i>		60
1	Комплексная практическая работа	60
<i>Промежуточная аттестация (экзамен)</i>		40
Общий итог за семестр		100

Соответствие баллов оценке

Количество баллов из 100	ECTS	Оценка по пятибалльной шкале	
		Экзамен, дифференцированный зачет	Зачет
90-100	A	отлично	зачтено
80-89	B	хорошо	
75-79	C		
70-74	D		
60-69	E	удовлетворительно	
35-59	FX	неудовлетворительно	не зачтено
0-34	F		

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОГО ПРОЦЕССА

Учебные занятия проводятся в 8-м учебном корпусе (г. Донецк, ул. Челюскинцев, д. 198 а) университета. Для проведения лабораторных занятий требуется аудитория, оборудованная меловой или маркерной доской, мультимедийный проектор и экран, ноутбук, комплект учебной мебели для студентов, рабочее место преподавателя, выход в Интернет – проводной или с использованием Wi-Fi.

Для самостоятельной работы используются текстовые и электронные ресурсы Научной библиотеки университета и других электронных библиотечных баз данных, учебно-методическое обеспечение, представленное в учебно-методическом кабинете Главного корпуса (ауд.405).

Обучающиеся имеют возможность использовать учебные материалы по дисциплине, размещенные на платформе Moodle Центра дистанционного образования ФГБОУ ВО «ДонГУ». При изучении дисциплины могут применяться электронное обучение и дистанционные образовательные технологии.

С использованием ресурсов платформы дистанционного образования осуществляется текущий контроль знаний обучающихся на основе тестирования и проверки результатов самостоятельной работы.

10. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

10.1. Основная литература

1. Лабковская, Р. Я. Анализ больших данных : учебное пособие / Р. Я. Лабковская, П. В. Косов. — Санкт-Петербург : СПбГУТ им. М.А. Бонч-Бруевича, 2025. — 152 с. — ISBN 978-5-89160-366-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/508654> (дата обращения: 08.04.2026). — Режим доступа: для авториз. пользователей. (СЭБ ЭБС «Лань»).

2. Алетдинова, А. А. Интеллектуальный анализ больших данных : учебное пособие / А. А. Алетдинова, М. Ш. Муртазина. — Новосибирск : НГТУ, 2023. — 66 с. — ISBN 978-5-7782-4899-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/404567> (дата обращения: 08.04.2026). — Режим доступа: для авториз. пользователей. (СЭБ ЭБС «Лань»).

10.2. Дополнительная литература

1. Горожанина, Е. И. Высокопроизводительные вычисления и анализ больших данных : учебное пособие / Е. И. Горожанина. — Самара : ПГУТИ, 2022. — 132 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/411386> (дата обращения: 08.04.2026). — Режим доступа: для авториз. пользователей. (СЭБ ЭБС «Лань»).

2. Анализ и обработка больших данных на платформе QuAnt : учебное пособие / С. Г. Ермакова, К. А. Бухарова, О. С. Мандрика, С. В. Лабузов. — Санкт-Петербург : ПГУПС, 2024. — 115 с. — ISBN 978-5-7641-2002-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/439520> (дата обращения: 08.04.2026). — Режим доступа: для авториз. пользователей. (СЭБ ЭБС «Лань»).

11. ИНФОРМАЦИОННЫЕ РЕСУРСЫ

1. Научная электронная библиотека eLIBRARY.RU: сайт / НЭБ (ООО). — Москва, 2000- . — URL: <https://elibrary.ru> (дата обращения: 08.04.2026). — Режим доступа: для авторизов. пользователей. — Текст: электронный.

2. Научная электронная библиотека «КиберЛенинка»: сайт / Ассоциация «Открытая наука». — Москва, 2014- . — URL: <https://cyberleninka.ru/> (дата обращения: 08.04.2026). — Режим доступа: свободный. — Текст: электронный.

3. Национальная подписка: сайт / РЦНИ. — Москва, 2020- . — URL: <https://podpiska.rcsi.science/> (дата обращения: 08.04.2026). — Режим доступа: для авторизов.пользователей ДонГУ. — Текст: электронный.

4. Национальная электронная библиотека (НЭБ): федеральная государственная информационная система / Министерство Культуры РФ; Российская государственная

библиотека. – Москва, 2019- . – URL: <https://rusneb.ru/> (дата обращения: 08.04.2026). – Режим доступа: свободный. – Текст: электронный.

5. Образовательная платформа Юрайт: сайт. – Москва, 2013- . – URL: <https://urait.ru/library/svobodnyy-dostup/> (дата обращения: 08.04.2026). – Режим доступа: издания свободного доступа, для авторизов. пользователей. – Текст: электронный.

6. Электронная библиотека РГБ: раздел сайта / РГБ. – Текст: электронный // Российская государственная библиотека: официальный сайт. – <https://ldiss.rsl.ru/electroniclibrary>. – Режим доступа: поиск свободный, электронные документы – Виртуальный читальный зал РГБ на территории НБ ДонГУ.

7. Электронно-библиотечная система Лань: сайт / ЭБС Лань. – С.-Петербург, 2011- . – URL: <https://e.lanbook.com> (дата обращения: 08.04.2026). – Режим доступа: издания Сетевой электронной библиотеки, для авторизов. пользователей. – Текст: электронный.

8. Электронно-библиотечная система ДонГУ: сайт / ФГБОУ ВО «ДонГУ». – Донецк, 2016- . – URL: <http://library.donnu.ru/> (дата обращения: 08.04.2026). – Режим доступа: свободный. – Текст: электронный.

9. Электронный каталог Научной библиотеки ДонГУ: раздел сайта / НБ ДонГУ. – Текст: электронный // ЭБС ДонГУ: сайт. – URL: <http://library.donnu.ru/catalog/> (дата обращения: 08.04.2026). – Режим доступа: поиск свободный, электронные документы – для пользователей ДонГУ.

10. Электронный архив ДонГУ: раздел сайта / НБ ДонГУ. – Текст: электронный // ЭБС ДонГУ: сайт. – URL: <http://repo.donnu.ru/> (дата обращения: 08.04.2026). – Режим доступа: свободный.

12. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

1. РЕД ОС, Astra Linux (ОЕМ).
2. OpenOffice, LibreOffice.