

Министерство науки и высшего образования
Российской Федерации

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

Учетно-финансовый факультет
Кафедра бизнес-информатики



УТВЕРЖДАЮ

проректор

П.А. Машаров

«29» марта 2024 г.

МП

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

АНАЛИТИКА БОЛЬШИХ ДАННЫХ ДЛЯ БИЗНЕСА

Укрупненная группа направлений подготовки	38.00.00 Экономика и управление
Программа высшего образования	Программа магистратуры
Направление подготовки	38.04.05 Бизнес-информатика
Магистерская программа	IT – инновации в бизнесе
Квалификация	Магистр
Форма обучения	Очная, заочная

Рабочая программа адаптирована для лиц
с ограниченными возможностями здоровья и инвалидов

Донецк 2024

Рабочая программа дисциплины «Аналитика больших данных для бизнеса» для обучающихся по направлению подготовки 38.04.05 Бизнес-информатика (Магистерская программа: IT – инновации в бизнесе) составлена на основании Федерального государственного образовательного стандарта высшего образования – магистратура по направлению подготовки 38.04.05 Бизнес-информатика, утвержденного приказом Министерства науки и высшего образования Российской Федерации от «12» августа 2020 г. № 990 (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 2024 года.

Разработчик:

доцент кафедры бизнес-информатики,
канд. экон. наук



В.В. Гридина

Рабочая программа одобрена на заседании кафедры бизнес-информатики.
Протокол от 26.03.2024 г. №8

Заведующий кафедрой



Т.О. Загорная

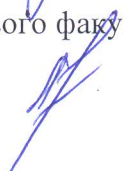
СОГЛАСОВАНО:

Декан учетно-финансового факультета
28.03.2024 г.



Н. В. Алексеенко

Учебно-методическая комиссия учетно-финансового факультета.
Протокол от 27.03.2024 г. № 7.
Председатель



А. А. Блажевич

Руководитель основной профессиональной
образовательной программы,
д-р экон. наук, проф.
26.03.2024 г.



Т.О. Загорная

1. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

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

дисциплины программы бакалавриата: Эконометрика, Информационные технологии и инструменты программирования, Программирование на Python, Анализ данных на Python.

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

Производственная практика: по получению профессиональных умений и опыта профессиональной деятельности, Производственная практика: преддипломная.

2. ОПИСАНИЕ ДИСЦИПЛИНЫ

2.1. Общая характеристика

Наименование показателя	Значение показателя
Название образовательной программы	38.04.05 Бизнес-информатика (Магистерская программа: IT – инновации в бизнесе)
Шифр и название в соответствии с учебным планом	Б1.В.ОД.3 Аналитика больших данных для бизнеса
Часть образовательной программы	Вариативная часть: выбор вуза
Количество зачетных единиц / всего часов	3 / 108

2.2. Распределение часов по формам и периодам обучения

Форма обучения	курс	семестр	Общее количество часов					Форма контроля
			лекционных	лабораторных	практических	самостоятельной работы	всего	
Очная	2	3	17	34	0	57	108	экзамен
Заочная	2	4	4	8	0	96	108	экзамен

3. ЦЕЛИ ДИСЦИПЛИНЫ

Сформировать у студентов теоретические знания и практические навыки по методологии сбора, анализа и моделирования больших данных для принятия управленческих решений.

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

Компетенции	Индикаторы	Результаты обучения
ОПК-3. Способен принимать решения, осуществлять стратегическое	ОПК-3.3. Выявляет возможности, создаваемые информационными	ОПК-3.3.1. Знает основные подходы к применению программных средств для моделирования и обработки больших данных.

планирование и прогнозирование в профессиональной деятельности с использованием современных методов и программного инструментария сбора, обработки и анализа данных, интеллектуального оборудования и систем искусственного интеллекта	и цифровыми технологиями.	ОПК-3.3.2. Умеет осуществлять автоматизированный сбор и алгоритмическую обработку больших данных в профессиональной области с использованием современных программных средств. ОПК-3.3.3. Владеет методологией сбора и анализа данных на основе методов и программного инструментария.
ПК-3. Способен осуществлять принятие решений в профессиональной деятельности на основе использования современных методов и программного инструментария сбора, обработки и анализа данных, в том числе больших данных.	ПК 3.1 Составляет график и план мероприятий по выполнению аналитических работ по сбору, обработке и анализу персональных данных и данных в организации в соответствии с действующим законодательством, в том числе больших данных	ПК-3.1.1 Знает методы планирования и организации исследований больших данных, а также методологию принятия решений на основе анализа больших данных. ПК-3.1.2 Умеет организовать исследование больших данных реального объекта, осуществить сбор и интеграцию больших данных объекта исследования. ПК-3.1.3. Владеет современными методами, подходами, алгоритмами, средствами и инструментами анализа больших данных.

5. ПРОГРАММА ДИСЦИПЛИНЫ

Название темы	Краткое содержание темы (вопросы темы)
Раздел 1. Определение больших данных. Технологии хранения больших данных.	
Введение в большие данные	История и причины появления термина Big Data. Характеристики и источники Big Data. Четыре основных типа данных. Аналитика данных. Задачи, решаемые Big Data.
Классификация задач анализа данных	DataMining. Интеллектуальный анализ данных, его, отличия и задачи. Text Mining. Web Mining. Web Content Mining. Web Usage Mining. Social media mining. RapidMiner.
Жизненный цикл аналитики данных	Business Intelligence (BI). ETL (Extract, Transform, Load)-процесс. Средства BI. Online Analytical Processing (OLAP). Инструменты анализа BI. Понятие жизненного цикла аналитики данных.

Высокопроизводительные вычисления	История Hadoop и MapReduce. Hadoop Distributed File System. Технология Map Reduce. Архитектура Hadoop.
Масштабирование и многоуровневое хранение данных	NoSQL (NotOnlySQL). Масштабируемость. Репликация. CAP теорема. MongoDB.
Раздел 2. Инструменты анализа и обработки больших данных.	
Визуализация данных и результатов анализа	Типы, задачи и виды визуализации. Графики, диаграммы, инфографика. Интерактивный сторителлинг, дашборды.
Язык R и готовые инструментальные средства анализа больших данных	Язык R. Синтаксис языка R, основные типы и структуры данных. Интерфейс Rstudio.
Статистические методы анализа данных	Статистические гипотезы и критерии. Машинное обучение. Метрический и линейных классификаторы ROC-кривая. Кластерный анализ.

6. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

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

Наименования разделов и тем	Количество часов				
	Лекц.	Лабор.	Практ.	СРС	Всего
Раздел 1. Определение больших данных. Технологии хранения больших данных.	11	22	0	36	69
Введение в большие данные	2	4	0	7	13
Классификация задач анализа данных	2	4	0	7	13
Жизненный цикл аналитики данных	3	6	0	8	17
Высокопроизводительные вычисления	2	4	0	7	13
Масштабирование и многоуровневое хранение данных	2	4	0	7	13
Раздел 2. Инструменты анализа и обработки больших данных.	6	12	0	21	39
Визуализация данных и результатов анализа	2	4	0	7	13
Язык R и готовые инструментальные средства анализа больших данных	2	4	0	7	13
Статистические методы анализа данных	2	4	0	7	13
ИТОГО ЗА СЕМЕСТР	17	34	0	57	108

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

Наименования разделов и тем	Количество часов				
	Лекц.	Лабор.	Практ.	СРС	Всего
Раздел 1. Определение больших данных. Технологии хранения больших данных.	3	5	0	61	69
Введение в большие данные	1	1	0	11	13
Классификация задач анализа данных	0	1	0	12	13
Жизненный цикл аналитики данных	1	1	0	15	17
Высокопроизводительные вычисления	0	1	0	12	13
Масштабирование и многоуровневое хранение данных	1	1	0	11	13

Раздел 2. Инструменты анализа и обработки больших данных.	1	3	0	35	39
Визуализация данных и результатов анализа	0	1	0	12	13
Язык R и готовые инструментальные средства анализа больших данных	0	1	0	12	13
Статистические методы анализа данных	1	1	0	11	13
ИТОГО ЗА СЕМЕСТР	4	8	0	96	108

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

7.1. Контрольные вопросы

Раздел 1. Определение больших данных. Технологии хранения больших данных.

1. История и причины появления термина Big Data.
2. Характеристики и источники Big Data.
3. Четыре основных типа данных.
4. Аналитика данных.
5. Задачи, решаемые Big Data.
6. DataMining.
7. Интеллектуальный анализ данных, его, отличия и задачи.
8. Text Mining.
9. Web Mining.
10. Web Content Mining.
11. Web Usage Mining.
12. Social media mining.
13. RapidMiner.
14. Business Intelligence (BI).
15. ETL (Extract, Transform, Load)-процесс.
16. Средства BI.
17. Online Analytical Processing (OLAP).
18. Инструменты анализа BI.
19. Понятие жизненного цикла аналитики данных.
20. История Hadoop и MapReduce.
21. Hadoop Distributed File System.
22. Технология Map Reduce.
23. Архитектура Hadoop.
24. NoSQL (NotOnlySQL).
25. Масштабируемость.
26. Репликация.
27. CAP теорема.
28. MongoDB.

Раздел 2. Инструменты анализа и обработки больших данных.

29. Типы, задачи и виды визуализации.
30. Графики, диаграммы, инфографика.
31. Интерактивный сторителлинг, дашборды.
32. Язык R.
33. Синтаксис языка R, основные типы и структуры данных.
34. Интерфейс Rstudio.
35. Статистические гипотезы и критерии.
36. Машинное обучение.
37. Метрический и линейных классификаторы ROC-кривая.

38. Кластерный анализ.

7.2. Образец содержания экзаменационного билета.

Экзаменационный билет № 1

1. Четыре основных типа данных.
2. Графики, диаграммы, инфографика.

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

8. РАСПРЕДЕЛЕНИЕ БАЛЛОВ, КОТОРЫЕ ПОЛУЧАЮТ ОБУЧАЮЩИЕСЯ

Общая оценка знаний обучающихся по дисциплине проводится по 100-балльной шкале исходя из максимума, приведенного в таблице ниже. Организационно-учебная работа в аудитории оценивается на основе таких критериев как посещаемость занятий, своевременное и качественное выполнение домашних заданий, активность во время проведения лекционных и практических занятий (участие в обсуждении текущего и пройденного материала, решение задач и т.п.).

8.1. Форма обучения – очная.

Номера разделов	Виды работ	Максимальное количество баллов
1-2	Организационно-учебная работа в аудитории	15
	Самостоятельная работа	15
ИТОГО		60
Экзамен		40
Общий итог за семестр		100

8.2. Форма обучения – заочная.

Номера разделов	Виды работ	Максимальное количество баллов
1-2	Организационно-учебная работа в аудитории	15
	Самостоятельная работа	15
ИТОГО		60
Экзамен		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	зачтено
75-79	C		зачтено

70-74	D	удовлетворительно	зачтено
60-69	E		зачтено
35-59	FX	неудовлетворительно	не зачтено
0-34	F		не зачтено

9. ОБЕСПЕЧЕНИЕ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ДЛЯ ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ И ИНВАЛИДОВ

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- 1) для слепых и слабовидящих:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств;
 - письменные задания оформляются увеличенным шрифтом.
- 2) для глухих и слабослышащих:
 - лекции оформляются в виде электронного документа;
 - письменные задания выполняются на компьютере в письменной форме;
 - экзамен проводится в письменной форме на компьютере; возможно проведение в форме тестирования.
- 3) для лиц с нарушениями опорно-двигательного аппарата:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере;
 - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- 1) для слепых и слабовидящих:
 - в печатной форме увеличенным шрифтом;
 - в форме электронного документа;
- 2) для глухих и слабослышащих:
 - в печатной форме;
 - в форме электронного документа.
- 3) для обучающихся с нарушениями опорно-двигательного аппарата:
 - в печатной форме;
 - в форме электронного документа.

10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОГО ПРОЦЕССА

Учебные занятия проводятся в 8-м учебном корпусе ДонГУ (г. Донецк, ул. Челюскинцев, д. 198 а). 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 8-го учебного корпуса (ауд. 105), материально-техническую базу учебных лабораторий кафедры бизнес-информатики (ауд. 101-103).

Обучающиеся имеют возможность использовать учебные материалы по дисциплине, размещенные на платформе Moodle Центра дистанционного образования ФГБОУ ВО «ДонГУ». При изучении дисциплины применяются 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.

11. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

11.1. Основная литература

1. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining [Электронный ресурс]: учебное пособие/ Федин Ф.О., Федин Ф.Ф.— Электрон. текстовые данные. — Москва: Московский городской педагогический университет, 2017.— 308 с.— URL: <https://1lib.domains/?redirectUrl=/book/3590439/181f68>

2. Воронов В.И. Data Mining - технологии обработки больших данных [Электронный ресурс]: учебное пособие/ Воронов В.И., Воронова Л.И., Усачев В.А.— Электрон. текстовые данные. — Москва: Московский технический университет связи и информатики, 2018.— 47 с. — URL: <https://knigogid.ru/books/1888855-big-data-metody-i-sredstva-analiza-uchebnoe-posobie/download/pdf>

11.2. Дополнительная литература

3. Крутиков, В. Н. Анализ данных [Электронный ресурс] : учебное пособие по направлению подготовки 010400.62 «Прикладная математика и информатика» / В. Н. Крутиков, В. В. Мешечкин ; Кемеровский государственный университет. - Кемерово : Кемеровский государственный университет, 2014. - 137 с.

4. Радченко, И.А. Технологии и инфраструктура Big Data [Электронный ресурс] / И.А. Радченко, И.Н. Николаев. – СПб: Университет ИТМО, 2018. – 52 с. – URL: <https://books.ifmo.ru/file/pdf/2326.pdf>

5. Мастицкий С.Э. Статистический анализ и визуализация данных с помощью R / С.Э. Мастицкий, В.К. Шитиков. – URL: https://r-analytics.blogspot.com/p/blog-page_46.html

6. Мастицкий С.Э. Анализ временных рядов с помощью R / С.Э. Мастицкий. – URL: https://r-analytics.blogspot.com/p/blog-page_46.html

12. ИНФОРМАЦИОННЫЕ РЕСУРСЫ

1. **Национальная электронная библиотека (НЭБ):** федеральная государственная информационная система / Министерство Культуры РФ; Российская государственная библиотека. – Москва, 2019- . – URL: <https://rusneb.ru/> (дата обращения: 01.09.2023). – Режим доступа: свободный, подписка. Необходима установка программного обеспечения. – Текст: электронный.

2. **eLIBRARY.RU:** научная электронная библиотека: сайт. – Москва, 2000- . – URL: <https://elibrary.ru> (дата обращения: 01.09.2023). – Режим доступа: для авторизов. пользователей. –Текст: электронный.

3. Научная электронная библиотека «КиберЛенинка»: сайт / Ассоциация «Открытая наука». – Москва, 2014- . – URL: <https://cyberleninka.ru/>. – Режим доступа: свободный. – Текст: электронный.
4. Электронно-библиотечная система «Лань»: [сайт]. – URL: <https://e.lanbook.com> (дата обращения: 01.09.2023). – Режим доступа: для авторизов. пользователей. – Текст: электронный.
5. ЭБС Юрайт: электронная библиотечная система: сайт. – Москва, 2013. – URL: <https://biblio-online.ru> (дата обращения: 01.09.2023). – Режим доступа: для авторизов. пользователей. – Текст: электронный.
6. Электронно-библиотечная система ДонГУ: сайт / ФГБОУ ВО «ДонГУ». – Донецк, 2016- . – URL: <http://library.donnu.ru/> (дата обращения: 01.09.2023). – Режим доступа: свободный. – Текст: электронный.
7. Электронный каталог Научной библиотеки ДонГУ: раздел сайта / НБ ДонГУ. – Текст: электронный // ЭБС ДонГУ: сайт. – URL: <http://library.donnu.ru/catalog/> (дата обращения: 01.09.2023). – Режим доступа: поиск свободный, электронные документы – для пользователей ДонГУ.
8. Электронный архив ДонГУ: раздел сайта / НБ ДонГУ. – Текст: электронный // ЭБС ДонГУ: сайт. – URL: <http://repo.donnu.ru/> (дата обращения: 01.09.2023). – Режим доступа: свободный.

13. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

1. Windows 7 PRO (корпоративная лицензия ДонГУ № 46484614)
2. Microsoft Office (корпоративная лицензия ДонГУ № 46472919)
3. Microsoft Visual Studio (лицензия программы Dream Spark для высших учебных заведений)
4. Антивирус Касперского, Adobe Acrobat Reader, xPDF (лицензии GPL, Apache, BSD для свободного программного обеспечения).