



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Магнитогорский государственный технический университет им. Г.И.
Носова»



УТВЕРЖДАЮ
Директор ИЭиУ
Е.С. Замбрыцкая

30.04.2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

ИНСТРУМЕНТЫ ПРОДВИНУТОЙ БИЗНЕС-АНАЛИТИКИ

Направление подготовки (специальность)
38.03.05 Бизнес-информатика

Направленность (профиль/специализация) программы
Аналитическая поддержка в управлении бизнес-процессами

Уровень высшего образования - бакалавриат

Форма обучения
очная

Институт/ факультет	Институт экономики и управления
Кафедра	Экономики
Курс	4
Семестр	7

Магнитогорск
2025 год

Рабочая программа составлена на основе ФГОС ВО - бакалавриат по направлению подготовки 38.03.05 Бизнес-информатика (приказ Минобрнауки России от 29.07.2020 г. № 838)

Рабочая программа рассмотрена и одобрена на заседании кафедры Экономики
02.04.2025, протокол № 8

Зав. кафедрой

А.Г. Васильева

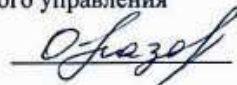
Рабочая программа одобрена методической комиссией ИЭиУ
30.04.2025 г. протокол № 4

Председатель

Е.С. Замбужицкая

Согласовано:

Зав. кафедрой Менеджмента и государственного управления



О.Л. Назарова

Рабочая программа составлена:

доцент кафедры Экономики, канд. филос. наук

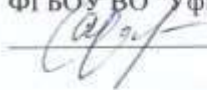


М.В. Кузнецова

Рецензент:

доцент, заместитель директора по учебной работе Сибайский институт (филиал)

ФГБОУ ВО "Уфимский университет науки и технологий", д-р экон. наук



И. А. Ситнова

Листактуализациирабочейпрограммы

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2026 - 2027 учебном году на заседании кафедры Экономики

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ А.Г. Васильева

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2027 - 2028 учебном году на заседании кафедры Экономики

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ А.Г. Васильева

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2028 - 2029 учебном году на заседании кафедры Экономики

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ А.Г. Васильева

Рабочая программа пересмотрена, обсуждена и одобрена для реализации в 2029 - 2030 учебном году на заседании кафедры Экономики

Протокол от _____ 20__ г. № ____
Зав. кафедрой _____ А.Г. Васильева

1 Цели освоения дисциплины (модуля)

Овладение современными методами и инструментами продвинутой бизнес-аналитики для эффективного решения сложных бизнес-задач, поддержки принятия обоснованных управленческих решений и повышения конкурентоспособности компании на основе анализа больших данных.

2 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина Инструменты продвинутой бизнес-аналитики входит в часть учебного плана формируемую участниками образовательных отношений образовательной программы.

Для изучения дисциплины необходимы знания (умения, владения), сформированные в результате изучения дисциплин/ практик:

Разработки приложений на платформе 1С

Рынки ИКТ и организация продаж

Информационные процессы, системы и сети

Основы алгоритмизации и программирования

Информатика

Основы экономики и управления на предприятии

Знания (умения, владения), полученные при изучении данной дисциплины будут необходимы для изучения дисциплин/практик:

Стратегический менеджмент

Управление рисками в ИТ-сфере

Управление цифровым продуктом и продуктовая аналитика

Производственная – преддипломная практика

Подготовка к защите и защита выпускной квалификационной работы

3 Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля) и планируемые результаты обучения

В результате освоения дисциплины (модуля) «Инструменты продвинутой бизнес-аналитики» обучающийся должен обладать следующими компетенциями:

Код индикатора	Индикатор достижения компетенции
УК-10	Способен принимать обоснованные экономические решения в различных областях жизнедеятельности
УК-10.1	Понимает экономические законы, категории и принципы, возможности их использования в различных областях жизнедеятельности
УК-10.2	Использует экономические знания для принятия обоснованных экономических решений в различных областях жизнедеятельности

4. Структура, объём и содержание дисциплины (модуля)

Общая трудоемкость дисциплины составляет 4 зачетных единиц 144 акад. часов, в том числе:

- контактная работа – 65,7 акад. часов;
- аудиторная – 64 акад. часов;
- внеаудиторная – 1,7 акад. часов;
- самостоятельная работа – 78,3 акад. часов;
- в форме практической подготовки – 0 акад. час;

Форма аттестации - зачет

Раздел/ тема дисциплины	Семестр	Аудиторная контактная работа (в акад. часах)			Самостоятельная работа студента	Вид самостоятельной работы	Форма текущего контроля успеваемости и промежуточной аттестации	Код компетенции
		Лек.	лаб. зан.	практ. зан.				
1. Раздел 1								
1.1 Продвинутое методы анализа данных: от описательной статистики к прогнозируемому моделированию.	7	7		7	20	Самостоятельное изучение учебной и научной литературы, работа с электронными библиотеками, выполнение ситуационных заданий, решение задач	Устный опрос (собеседование); семинарские занятия; участие в конференциях, олимпиадах, конкурсах, выставках; решение задач, в т.ч. ситуационных; тестирование.	УК-10.1, УК-10.2
1.2 Инструменты визуализации данных: создание информативных дашбордов и интерактивных отчетов.		7		7	20	Самостоятельное изучение учебной и научной литературы, работа с электронными библиотеками, выполнение ситуационных заданий, решение задач	Устный опрос (собеседование); семинарские занятия; участие в конференциях, олимпиадах, конкурсах, выставках; решение задач, в т.ч. ситуационных; тестирование.	УК-10.1, УК-10.2
1.3 Машинное обучение в бизнес-аналитике: алгоритмы и применение для решения бизнес-задач.		7		7	20	Самостоятельное изучение учебной и научной литературы, работа с электронными библиотеками, выполнение ситуационных	Устный опрос (собеседование); семинарские занятия; участие в конференциях, олимпиадах, конкурсах, выставках; решение задач, в т.ч.	УК-10.1, УК-10.2

						заданий, решение задач	ситуационных; тестирование.	
1.4 Ссудный процент и его роль.	7	11		11	18,3	Самостоятельное изучение учебной и научной литературы, работа с электронными библиотеками, выполнение ситуационных заданий, решение задач	Устный опрос (собеседование); семинарские занятия; участие в конференциях, олимпиадах, конкурсах, выставках; решение задач, в т.ч. ситуационных; тестирование.	УК-10.1, УК-10.2
Итого по разделу		32		32	78,3			
Итого за семестр		32		32	78,3		зачёт	
Итого по дисциплине		32		32	78,3		зачет	

5 Образовательные технологии

На сегодняшний день стали очевидны преимущества использования компьютера на лекционных и практических учебных занятиях. Объяснение нового материала с использованием презентаций, выполненных с помощью программ Microsoft PowerPoint и Microsoft FrontPage, вызывает интерес у студентов, способствует лучшему усвоению материала. Использование компьютера на учебных занятиях позволяет преподавателю экономить время, опрашивать учащихся на каждом занятии, вести статистику опроса, выявлять западающие темы. Также одним из эффективных средств информационных технологий является электронный учебник. Исходя из этого, более 20% всех занятий проводятся с применением информационных технологий.

Для обеспечения наибольшей эффективности образовательного процесса в курсе данной учебной дисциплины используются в процессе обучения передовые образовательные технологии:

- 1) традиционные образовательные технологии (информационная лекция, практические (семинарские) занятия);
- 2) технология проблемного обучения (проблемная лекция, практические занятия в форме практикума, кейс-метода);
- 3) игровые технологии (ролевые и деловые игры);
- 4) технологии проектного обучения (творческий проект);
- 5) интерактивные технологии (семинар-дискуссия);
- 6) информационно-коммуникационные образовательные технологии (лекция-визуализация, практические занятия в форме презентации)

Лекционные занятия наряду с сообщением учебной информации предполагают и решение следующих дидактических задач: заинтересовать студентов изучаемой темой, разрушить неверные стереотипы, убедить в необходимости глубокого освоения материала, побудить к самостоятельному поиску и активной мыслительной деятельности, помочь совершить переход от теоретического уровня социально-экономического планирования в муниципальных образованиях к прикладным знаниям в данной области.

Проведение групповых (семинарских и практических) занятий предполагает решение разнообразных дидактических задач: закрепление полученных знаний, формирование умения применять их на практике, совершенствование умения работать с информацией, анализировать, обобщать, принимать и обосновывать решения, аргументировано защищать собственные взгляды в дискуссии, взаимодействовать с другими членами группы в процессе разрешения конфликтных ситуаций.

6 Учебно-методическое обеспечение самостоятельной работы обучающихся

Представлено в приложении 1.

7 Оценочные средства для проведения промежуточной аттестации

Представлены в приложении 2.

8 Учебно-методическое и информационное обеспечение дисциплины

а) Основная литература:

1. Боровиков, В. П. Популярное введение в современный анализ данных в системе STATISTICA: Учебное пособие для вузов / В.П. Боровиков. - Москва : Гор. линия-Телеком, 2013. - 288 с.: ил.; + CD-ROM. (обложка, cdrom) ISBN 978-5-9912-0326-5. - Текст : электронный. - URL: <https://znaniyum.ru/catalog/product/425084> (дата обращения: 18.05.2025). – Режим доступа: по подписке.

2. Гобарева, Я. Л. Бизнес-аналитика средствами Excel : учебное пособие / Я.Л. Гобарева, О.Ю. Городецкая, А.В. Золотарюк. — 3-е изд., перераб. и доп. — Москва : Вузовский учебник : ИНФРА-М, 2023. — 350 с. + Доп. материалы [Электронный ресурс]. - ISBN 978-5-9558-0560-3. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/1893969> (дата обращения: 18.05.2025). – Режим доступа: по подписке.

б) Дополнительная литература:

1) Криволапов, С. Я. Анализ данных. Методы теории вероятностей и математической статистики на языке Python : учебное пособие / С.Я. Криволапов. — Москва : ИНФРА-М, 2025. — 678 с. — (Высшее образование). — DOI 10.12737/2034420. - ISBN 978-5-16-018616-0. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2034420> (дата обращения: 18.05.2025). – Режим доступа: по подписке.

2) Кугаевских, А. В. Проектирование информационных систем. Системная и бизнес-аналитика : учебное пособие / А. В. Кугаевских. - Новосибирск : Изд-во НГТУ, 2018. - 256 с. - ISBN 978-5-7782-3608-0. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1867932> (дата обращения: 18.05.2025). – Режим доступа: по подписке.

3) Марр, Б. Ключевые инструменты бизнес-аналитики. 67 инструментов, которые должен знать каждый менеджер : учебное пособие / Б. Марр ; пер. с англ. В. Егорова. - 3-е изд. - Москва : Лаборатория знаний, 2022. - 339 с. - ISBN 978-5-00101-962-6. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/1987481> (дата обращения: 18.05.2025). – Режим доступа: по подписке.

в) Методические указания:

Назаров, Д. М. Информационные технологии в профессиональной деятельности: интеллектуальный анализ данных и бизнес-аналитика : учебно-методическое пособие / Д. М. Назаров, А. А. Копнин. — Москва : ИНФРА-М, 2025. — 326 с. - ISBN 978-5-16-019356-4. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2110964> (дата обращения: 18.05.2025). – Режим доступа: по подписке.

г) Программное обеспечение и Интернет-ресурсы:

Программное обеспечение

Наименование ПО	№ договора	Срок действия лицензии
MS Office 2007 Professional	№ 135 от 17.09.2007	бессрочно
7Zip	свободно распространяемое ПО	бессрочно
FAR Manager	свободно распространяемое ПО	бессрочно
Браузер Mozilla Firefox	свободно распространяемое ПО	бессрочно
Браузер Yandex	свободно распространяемое ПО	бессрочно
MS Office 2003 Professional	№ 135 от 17.09.2007	бессрочно
LibreOffice	свободно распространяемое ПО	бессрочно

Профессиональные базы данных и информационные справочные системы

Названиекурса	Ссылка
Электронная база периодических изданий EastView InformationServices, ООО «ИВИС»	https://dlib.eastview.com/
Национальная информационно-аналитическая система – Российский индекс научного цитирования (РИНЦ)	URL: https://elibrary.ru/project_risc.asp
Федеральное государственное бюджетное учреждение «Федеральный институт промышленной собственности»	URL: http://www1.fips.ru/
Электронные ресурсы библиотеки МГТУ им. Г.И. Носова	https://host.megaprolib.net/M/P0109/Web
Федеральный образовательный портал – Экономика. Социология. Менеджмент	http://ecsocman.hse.ru/

9 Материально-техническое обеспечение дисциплины (модуля)

Материально-техническое обеспечение дисциплины включает:

Учебные аудитории для проведения занятий лекционного типа: мультимедийные средства хранения, передачи и представления информации.

Учебные аудитории для проведения практических занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации: мультимедийные средства хранения, передачи и представления информации, комплекс тестовых заданий для проведения промежуточных и рубежных контролей

Помещения для самостоятельной работы обучающихся: персональные компьютеры с пакетом MSOffice/LibreOffice, выходом в Интернет и с доступом в электронную информационно-образовательную среду университета

Помещения для хранения и профилактического обслуживания учебного оборудования: шкафы для хранения учебно-методической документации, учебного оборудования и учебно-наглядных пособий.

Приложение 1

Учебно-методическое обеспечение самостоятельной работы обучающихся

1. Продвинутые методы анализа данных: от описательной статистики к прогнозирующему моделированию. (Включает: регрессионный анализ, анализ временных рядов, кластерный анализ, факторный анализ, анализ выживаемости и т.д.)

Вопросы для самоконтроля:

1. Что такое регрессионный анализ и для решения каких задач он применяется?
2. Опишите разницу между простой и множественной регрессией.
3. Какие предположения необходимо проверять при проведении регрессионного анализа?
4. Что такое R-квадрат (коэффициент детерминации) и как он интерпретируется?
5. Какие методы используются для анализа временных рядов?
6. Опишите основные компоненты временного ряда (тренд, сезонность, цикл, случайная компонента).
7. В чем суть методов сглаживания временных рядов (например, скользящее среднее, экспоненциальное сглаживание)?
8. Что такое кластерный анализ и для каких целей он применяется?
9. Опишите различные методы кластеризации (например, K-средних, иерархическая кластеризация).
10. Как оценить качество кластеризации?
11. Что такое факторный анализ и для решения каких задач он используется?
12. Объясните разницу между исследовательским и подтверждающим факторным анализом.
13. Как определить количество факторов, которые следует выделить?
14. Что такое анализ выживаемости и для чего он применяется?
15. Опишите метод Каплана-Мейера для оценки функции выживаемости.

2. Инструменты визуализации данных: создание информативных дашбордов и интерактивных отчетов. (Включает: Power BI, Tableau, Qlik Sense, Python (Matplotlib, Seaborn), R (ggplot2) и принципы эффективного визуального представления информации)

Вопросы для самоконтроля:

1. Какие основные принципы эффективной визуализации данных вы знаете?
2. Опишите различные типы диаграмм и графиков (столбчатая, линейная, круговая, точечная и т.д.) и укажите, для каких данных они лучше всего подходят.
3. Что такое дашборд и какие требования предъявляются к его дизайну?
4. Какие интерактивные элементы могут быть включены в дашборд?
5. Опишите основные функции и возможности Power BI.
6. Какие типы подключений к данным поддерживает Power BI?
7. Как создать вычисляемые столбцы и меры в Power BI?
8. Опишите основные функции и возможности Tableau.
9. Как создать рабочие листы и дашборды в Tableau?
10. Как использовать фильтры и параметры в Tableau?
11. Опишите основные функции и возможности Qlik Sense.
12. Как создать приложения и истории в Qlik Sense?
13. Как использовать ассоциативный движок Qlik Sense для анализа данных?
14. Как использовать библиотеки Matplotlib и Seaborn в Python для визуализации данных?
15. Как использовать библиотеку ggplot2 в R для визуализации данных?

3. Машинное обучение в бизнес-аналитике: алгоритмы и применение для решения бизнес-задач. (Включает: классификация, регрессия, кластеризация, алгоритмы обучения с учителем и без учителя, оценка качества моделей, выбор оптимальных параметров)

Вопросы для самоконтроля:

1. Что такое машинное обучение и чем оно отличается от традиционного программирования?
2. Опишите разницу между обучением с учителем и обучением без учителя.
3. Какие алгоритмы относятся к обучению с учителем (например, линейная регрессия, логистическая регрессия, деревья решений, метод опорных векторов)?
4. Какие алгоритмы относятся к обучению без учителя (например, K-средних, иерархическая кластеризация, PCA)?
5. Для решения каких бизнес-задач может быть использован алгоритм классификации?
6. Для решения каких бизнес-задач может быть использован алгоритм регрессии?
7. Как оценить качество модели классификации (например, точность, полнота, F1-мера)?
8. Что такое переобучение (overfitting) и как с ним бороться?
9. Что такое кросс-валидация и для чего она используется?
10. Как выбрать оптимальные параметры для алгоритма машинного обучения?
11. Опишите алгоритм деревьев решений (Decision Trees).
12. Опишите алгоритм случайного леса (Random Forest).
13. Опишите метод опорных векторов (Support Vector Machines).
14. Что такое нейронные сети и как они применяются в бизнес-аналитике?
15. Как использовать Python библиотеки (Scikit-learn, TensorFlow, Keras) для машинного обучения?

4. Работа с большими данными: технологии и инфраструктура для анализа больших объемов информации. (Включает: Hadoop, Spark, NoSQL базы данных, облачные платформы для анализа данных и инструменты для обработки и анализа больших данных)

Вопросы для самоконтроля:

1. Что такое "большие данные" (Big Data) и каковы их основные характеристики (объем, скорость, разнообразие, достоверность)?
2. Опишите архитектуру Hadoop и основные компоненты Hadoop Distributed File System (HDFS) и MapReduce.
3. Для чего используется Hadoop YARN?
4. Что такое Spark и чем он отличается от Hadoop MapReduce?
5. Какие преимущества дает использование Spark по сравнению с Hadoop MapReduce?
6. Опишите Spark Core, Spark SQL, Spark Streaming, MLlib и GraphX.
7. Что такое NoSQL базы данных и чем они отличаются от реляционных баз данных?
8. Приведите примеры NoSQL баз данных (например, MongoDB, Cassandra, Redis).
9. В каких случаях целесообразно использовать NoSQL базы данных?
10. Какие облачные платформы предоставляют услуги для анализа больших данных (например, Amazon AWS, Microsoft Azure, Google Cloud Platform)?
11. Опишите основные сервисы Amazon AWS для анализа больших данных (например, Amazon S3, Amazon EMR, Amazon Redshift).
12. Опишите основные сервисы Microsoft Azure для анализа больших данных (например, Azure Blob Storage, Azure HDInsight, Azure Synapse Analytics).

13. Опишите основные сервисы Google Cloud Platform для анализа больших данных (например, Google Cloud Storage, Google Dataproc, Google BigQuery).
14. Какие инструменты используются для обработки и анализа больших данных (например, Apache Hive, Apache Pig, Apache Kafka)?
15. Что такое Data Lake и чем он отличается от Data Warehouse?

Оценочные средства для проведения промежуточной аттестации

а) Планируемые результаты обучения и оценочные средства для проведения промежуточной аттестации:

Код индикатора	Индикатор достижения компетенции	Оценочные средства
УК-10: Способен принимать обоснованные экономические решения в различных областях жизнедеятельности		
УК-10.1	Понимает экономические законы, категории и принципы, возможности их использования в различных областях жизнедеятельности	Вопросы к зачету по дисциплине "Инструменты продвинутой бизнес-аналитики" 1. Что такое бизнес-аналитика и каковы ее основные цели? 2. Какие этапы включает в себя процесс бизнес-аналитики? 3. В чем разница между описательной, прогнозной и предписывающей аналитикой? 4. Какие основные методы анализа данных вы знаете? 5. Что такое регрессионный анализ и для каких задач он применяется? 6. Какие предположения необходимо проверять при проведении регрессионного анализа? 7. Что такое R-квадрат и как он интерпретируется? 8. Какие методы используются для анализа временных рядов? 9. Опишите основные компоненты временного ряда. 10. Что такое кластерный анализ и для каких целей он применяется? 11. Опишите различные методы кластеризации. 12. Как оценить качество кластеризации? 13. Что такое факторный анализ и для решения каких задач он используется? 14. Как определить количество факторов, которые следует выделить? 15. Какие основные принципы эффективной визуализации данных вы знаете? 16. Опишите различные типы диаграмм и графиков и укажите, для каких данных они лучше всего подходят. 17. Что такое дашборд и какие требования предъявляются к его дизайну? 18. Какие интерактивные элементы могут быть включены в дашборд? 19. Опишите основные функции и возможности Power BI (или Tableau, Qlik Sense - выберите один инструмент). 20. Что такое машинное обучение и чем оно отличается от традиционного программирования?

		21. Опишите разницу между обучением с учителем и обучением без учителя. 22. Приведите примеры алгоритмов обучения с учителем и опишите принцип их работы. 23. Приведите примеры алгоритмов обучения без учителя и опишите принцип их работы. 24. Как оценить качество модели машинного обучения (например, классификации)? 25. Что такое переобучение (overfitting) и как с ним бороться? 26. Что такое "большие данные" и каковы их основные характеристики? 27. Опишите архитектуру Hadoop и основные компоненты HDFS и MapReduce. 28. Что такое Spark и чем он отличается от Hadoop MapReduce? 29. В каких случаях целесообразно использовать NoSQL базы данных?
УК-10.2	Использует экономические знания для принятия обоснованных экономических решений в различных областях жизнедеятельности	Тема 1: Продвинутое методы анализа данных: от описательной статистики к прогнозируемому моделированию 1. Задание: Используя датасет с информацией о продажах (например, данные о продажах автомобилей, недвижимости, товаров в интернет-магазине), проведите регрессионный анализ для определения факторов, влияющих на объем продаж. Определите наиболее значимые факторы и постройте прогноз продаж на следующий период. Интерпретируйте результаты анализа. (Использовать Excel, R или Python). 2. Задание: Используя датасет с данными о посещаемости сайта за несколько лет, проведите анализ временных рядов для выявления тренда и сезонности. Спрогнозируйте посещаемость сайта на следующий квартал. (Использовать Excel, R или Python). 3. Задание: Используя датасет с информацией о клиентах банка (например, возраст, доход, кредитная история), проведите кластерный анализ для сегментации клиентов. Опишите характеристики каждого сегмента и предложите маркетинговые стратегии для каждой группы клиентов. (Использовать R или Python). Тема 2: Инструменты визуализации данных: создание информативных дашбордов и интерактивных отчетов 1. Задание: Используя датасет с информацией о продажах (например, данные о продажах автомобилей, недвижимости, товаров в интернет-магазине), создайте интерактивный дашборд в Power BI (или Tableau, Qlik Sense). Дашборд должен включать ключевые показатели эффективности продаж, фильтры по различным параметрам (например, регион, продукт, период времени) и интерактивные элементы для анализа данных.

		2. Задание: Используя датасет с данными о посещаемости сайта, создайте интерактивный отчет в Tableau (или Power BI, Qlik Sense). Отчет должен отображать динамику посещаемости, источники трафика, поведение пользователей на сайте и другие метрики. Используйте различные типы визуализаций для представления информации. 3. Задание: Визуализируйте результаты кластерного анализа, выполненного в рамках задания по теме 1, с использованием библиотеки Matplotlib или Seaborn в Python. Представьте кластеры на графике и опишите их характеристики. Тема 3: Машинное обучение в бизнес-аналитике: алгоритмы и применение для решения бизнес-задач 1. Задание: Используя датасет с информацией о клиентах банка и их кредитной истории, постройте модель машинного обучения (например, логистическую регрессию, дерево решений) для прогнозирования вероятности невозврата кредита. Оцените качество модели и определите наиболее важные факторы, влияющие на кредитоспособность клиентов. (Использовать Python с библиотекой Scikit-learn). 2. Задание: Используя датасет с информацией о транзакциях клиентов в интернет-магазине, постройте модель машинного обучения (например, K-средних) для выявления групп товаров, которые часто покупают вместе. Предложите стратегии для повышения продаж на основе результатов анализа. (Использовать Python с библиотекой Scikit-learn). 3. Задание: Используя датасет с отзывами клиентов о продукте, постройте модель машинного обучения для автоматической классификации отзывов по тональности (положительный, отрицательный, нейтральный). Оцените качество модели и предложите способы ее улучшения. (Использовать Python с библиотеками Scikit-learn и NLTK или SpaCy). Тема 4: Работа с большими данными: технологии и инфраструктура для анализа больших объемов информации 1. Задание: Используя данные в формате CSV (например, логи веб-сервера, данные о транзакциях), загрузите их в HDFS и выполните простые запросы с использованием Hive для анализа данных. 2. Задание: Используя данные в формате JSON, загрузите их в MongoDB и выполните запросы для агрегации и анализа данных. 3. Задание: Используя облачную платформу AWS (или Azure, Google Cloud), создайте кластер
--	--	--

		Spark и выполните задачи по обработке и анализу больших данных (например, подсчет количества слов в текстовом файле, фильтрация данных по заданному критерию).
--	--	--

б) Порядок проведения промежуточной аттестации, показатели и критерии оценивания:

Промежуточная аттестация по дисциплине включает теоретические вопросы, позволяющие оценить уровень усвоения обучающимися знаний, и практические задания, выявляющие степень сформированности умений и владений, проводится в форме зачета с оценкой.

Показатели и критерии оценивания:

– на оценку «**зачтено**» – обучающийся демонстрирует высокий уровень сформированности компетенций, всестороннее, систематическое и глубокое знание учебного материала, свободно выполняет практические задания, свободно оперирует знаниями, умениями, применяет их в ситуациях повышенной сложности.

– на оценку «**незачтено**» – обучающийся демонстрирует знания не более 20% теоретического материала, допускает существенные ошибки, не может показать интеллектуальные навыки решения простых задач.