

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Цветлюк Лариса Сергеевна
Должность: Ректор
Дата подписания: 12.10.2026 16:04:19
Уникальный программный ключ:
e4e919f04dc802624637575c97796a744138b172b88dd58f9301d8c2540974f9

Автономная некоммерческая организация
высшего образования
«Институт непрерывного образования»

Рассмотрено
на заседании кафедры естественнонаучных
и общегуманитарных дисциплин
Зав. кафедрой

УТВЕРЖДАЮ:

Ректор АНО ВО «ИНО»

Трубицын А.С.
27 августа 2026 г.

Цветлюк Л.С.
27 августа 2026 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Введение в Big Data

для направления подготовки

38.03.05 Бизнес-информатика

направленность (профиль) «Информационные технологии в управлении предприятием»

Уровень бакалавриата

Квалификация выпускника

Бакалавр

Руководитель основной
профессиональной образовательной программы
доц., к.т.н. Трубицын А.С.

Москва, 2026 г.

Рабочая программа учебной дисциплины «**Введение в Big Data**» разработана Трубицыным А.С.

Рабочая программа учебной дисциплины «**Введение в Big Data**» разработана на основании федерального государственного образовательного стандарта высшего образования по направлению подготовки 38.03.05 «Бизнес-информатика» (уровень бакалавриата), утверждённого приказом Министерства образования и науки Российской Федерации от 29.07.2020 № 838, учебного плана по основной профессиональной образовательной программе высшего образования по направлению подготовки 38.03.05 Бизнес-информатика.

СОДЕРЖАНИЕ

1. Общие положения.....	4
1.1. Цель и задачи учебной дисциплины.....	4
1.2. Место учебной дисциплины в структуре основной профессиональной образовательной программы.....	4
1.3. Планируемые результаты обучения по учебной дисциплине в рамках планируемых результатов освоения основной профессиональной образовательной программы.....	4
2. Объем учебной дисциплины, включая контактную работу обучающегося с преподавателем и самостоятельную работу обучающегося.....	5
3. Содержание учебной дисциплины.....	6
3.1. Учебно-тематический план по очной форме обучения.....	6
4. Учебно-методическое обеспечение самостоятельной работы обучающихся по учебной дисциплине.....	7
5. Фонд оценочных средств для проведения текущей и промежуточной аттестации обучающихся по учебной дисциплине.....	11
5.1. Форма промежуточной аттестации обучающегося по учебной дисциплине....	11
5.2. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы.....	11
5.3. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания.....	12
5.4. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы..	13
5.5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.....	23
6. Перечень основной и дополнительной литературы для освоения учебной дисциплины.....	23
6.1. Основная литература.....	23
6.2. Дополнительная литература.....	24
7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», современные профессиональные базы данных и информационные системы, необходимые для освоения учебной дисциплины.....	24
8. Методические указания для обучающихся по освоению учебной дисциплины.....	24
9. Программное обеспечение информационно- коммуникационных технологий.....	26
9.1. Информационные технологии.....	26
9.2. Программное обеспечение.....	26
9.3. Информационно-справочные системы.....	26
10. Перечень материально-технического обеспечения, необходимого для реализации программы по учебной дисциплине	26
11. Образовательные технологии.....	26

1. Общие положения

1.1. Цель и задачи учебной дисциплины.

Цель учебной дисциплины:

- изучение математических методов и подходов, используемых в программных системах обработки и анализа больших данных.
- развитие профессиональных навыков обучающихся за счет получения практического опыта работы с IT решениями в части обработки и анализа больших данных.

Задачами учебной дисциплины являются:

- усвоение знаний о сущности, структуре и видах математических моделей принятия решений; развитие навыков создания и решения моделей, необходимых в сфере управления.

1.2. Место учебной дисциплины в структуре основной профессиональной образовательной программы.

Дисциплина является обязательным элементом обязательной части Блока 1 образовательной программы.

1.3. Планируемые результаты обучения по учебной дисциплине в рамках планируемых результатов освоения основной профессиональной образовательной программы.

Процесс освоения учебной дисциплины направлен на формирование у обучающихся следующих компетенций: ОПК-3, ОПК-4 в соответствии с основной профессиональной образовательной программой «**Бизнес-информатика**».

В результате освоения учебной дисциплины обучающийся должен демонстрировать следующие результаты:

Код компетенции	Содержание компетенции	Индикаторы достижения профессиональных компетенций
ОПК-3	Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации;	ОПК-3.1. Знает этапы создания продуктов, фазы их жизненного цикла, процесс использования услуг в сфере ИКТ. ОПК-3.2. Умеет разрабатывать алгоритмы и программы для практической реализации в бизнесе. ОПК-3.3. Владеет навыками управления созданием и использованием продуктов и услуг в сфере информационно коммуникационных технологий.
ОПК-4	Способен понимать принципы работы информационных технологий; использовать	ОПК-4.1. Знает методы и программные средства для сбора, обработки и анализа бизнес-информации. ОПК-4.2. Умеет обрабатывать, анализировать и моделировать информационные потоки

	информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений;	организации. ОПК-4.3. Владеет навыками обоснования принятия управленческого решения как результата информационно аналитической деятельности.
--	--	---

2. Объем учебной дисциплины, включая контактную работы обучающегося с преподавателем и самостоятельную работу обучающегося

Общая трудоемкость учебной дисциплины составляет 4 зачетных единиц.

Очная форма обучения

Вид учебной работы	Всего часов	Семестр
		6
Аудиторные учебные занятия, всего	66	66
В том числе:		
Учебные занятия лекционного типа	32	32
Практические занятия	34	34
Самостоятельная работа обучающихся, всего	42	42
В том числе:		
Самоподготовка, практические задания	42	42
Контроль: вид промежуточной аттестации (экзамен)	36	Экзамен, 36
Общая трудоемкость учебной дисциплины	4/144	4/144

3. Содержание учебной дисциплины

3.1. Учебно-тематический план по очной форме обучения

Объем аудиторных занятий составляет – 66 часов.

Объем самостоятельной работы – 42 часа.

№ п/п	Модуль, раздел (тема)	Виды учебной работы, академических часов					Формы контроля освоения обучающимися учебной дисциплины	
		Всего	Самостоятельная работа обучающегося	Контактная работа преподавателя с обучающимися			Текущий контроль освоения обучающимися учебной дисциплины	Промежуточная аттестация обучающихся
				Всего	Лекционного типа	Практические занятия		
4 семестр								
1	Тема 1. Методы многомерного статистического	26	10	16	8	8	+	+

	анализа и анализа нечисловой информации							
2	Тема 2. Технологии хранения и обработки Больших данных	26	10	16	8	8	+	+
3	Тема 3. Программирование обработки и загрузки Больших данных	26	10	16	8	8	+	+
4	Тема 4. Аналитика в больших данных	30	12	18	8	10	+	+
ВСЕГО ЧАСОВ		144	42	66	32	34		36

4. Учебно-методическое обеспечение самостоятельной работы обучающихся по учебной дисциплине

Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации

Факторный анализ. Дискриминантный анализ. Кластерный анализ. Многомерное шкалирование. Методы контроля качества.

Формы контроля самостоятельной работы обучающихся: проверка ответов на вопросы самоподготовки.

Тема 2. Технологии хранения и обработки Больших данных

Основные направления развития методов обработки и хранения данных. Volume. Закон Мура. Velocity. Variety. Фреймворк Hadoop. Проблема хранения неструктурированных данных. Проблема преобразования данных. Семантические анализаторы. Самообучающиеся автоматы.

Формы контроля самостоятельной работы обучающихся: проверка ответов на вопросы самоподготовки.

Тема 3. Программирование обработки и загрузки Больших данных

9 языков для Big Data (R, Python, Julia, Java, Scala, MATLAB, Go, Kafka, Hadoop). Фреймворки (Hadoop, Spark, Storm). Базы данных (Hive, Impala, Presto, Drill). Аналитические платформы (Rapid Miner, IBM SPSS Modeler, KNIME, Qlik Analytics Platform, STA-TISTICA Data Miner, Informatica Intelligent Data Platform, World Programming System, Deductor, SAS Enterprise Miner). Прочие инструменты (Zookeeper, Flume, IBM Watson Analytics, Dell EMC Analytic Insights Module, Windows Azure HDInsight, Microsoft Azure Machine Learning, Pentaho Data Integration, Teradata Aster Analytics, SAP BusinessObjects Predictive Analytics, Oracle Big Data Preparation).

Формы контроля самостоятельной работы обучающихся: проверка ответов на вопросы самоподготовки.

Тема 4. Аналитика в больших данных

Аналитика Big Data — реалии и перспективы в России и мире. Технологии и методы анализа, которые используются для анализа Big Data (Data Mining; краудсорсинг; смешение и интеграция данных; машинное обучение; искусственные нейронные сети; распознавание образов; прогнозная аналитика; имитационное моделирование; пространственный анализ; статистический анализ; визуализация аналитических данных). Big Data: применение и

возможности. Решения на основе Big Data. Рынок Big Data в России. Big Data в банках. Big Data в бизнесе. Big Data в маркетинге.

Формы контроля самостоятельной работы обучающихся: проверка ответов на вопросы самоподготовки.

Содержание практических занятий по дисциплине

Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации.

1. Факторный анализ.
2. Дискриминантный анализ.
3. Кластерный анализ.
4. Многомерное шкалирование.
5. Методы контроля качества.

Тема 2. Технологии хранения и обработки Больших данных.

1. Основные направления развития методов обработки и хранения данных.
2. Проблема хранения неструктурированных данных.
3. Проблема преобразования данных.
4. Семантические анализаторы.
5. Самообучающиеся автоматы.

Тема 3. Программирование обработки и загрузки больших данных.

1. Языки для Big Data.
2. Фреймворки.
3. Базы данных.
4. Аналитические платформы.

Тема 4. Аналитика в больших данных.

1. Аналитика Big Data — реалии и перспективы в России и мире.
2. Технологии и методы анализа, которые используются для анализа Big Data.
3. Применение и возможности Big Data:
4. Решения на основе Big Data.

5. Фонд оценочных средств для проведения текущей и промежуточной аттестации обучающихся по учебной дисциплине

5.1. Форма промежуточной аттестации обучающегося по учебной дисциплине.

Контрольным мероприятием промежуточной аттестации обучающихся по учебной дисциплине является зачет с оценкой, который проводится в **устной** форме.

5.2. Перечень компетенций с указанием этапов ее формирования в процессе освоения образовательной программы.

Код компетенции	Содержание компетенции	Компоненты компетенции, степень их освоения	Индикаторы достижения профессиональных компетенций	Результаты обучения
-----------------	------------------------	---	--	---------------------

ОПК-3	Способен управлять процессами создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий, в том числе разрабатывать алгоритмы и программы для их практической реализации;	Компоненты компетенции соотносятся с содержанием дисциплины, и компетенция реализуется частично	ОПК-3.1. Знает этапы создания продуктов, фазы их жизненного цикла, процесс использования услуг в сфере ИКТ. ОПК-3.2. Умеет разрабатывать алгоритмы и программы для практической реализации в бизнесе. ОПК-3.3. Владеет навыками управления созданием и использованием продуктов и услуг в сфере информационно-коммуникационных технологий.	Знает: – этапы создания баз данных, фазы их жизненного цикла, процесс использования баз данных в сфере ИКТ Технологий. – основные этапы разработки алгоритмов и программ для практической реализации в бизнесе. – основные этапы создания и использования продуктов и услуг в сфере информационно-коммуникационных технологий. Умеет: – разрабатывать алгоритмы и программы для практической реализации в бизнесе. – формулировать этапы создания продуктов, фазы их жизненного цикла, процесс использования услуг в сфере ИКТ. – разрабатывать алгоритмы и программы для практической реализации в бизнесе. – управлять созданием и использованием продуктов и услуг в сфере информационно-коммуникационных технологий. Владеет: – навыками построения алгоритма создания продуктов, определения фаз их жизненного цикла, процесс использования услуг
-------	---	---	--	--

				в сфере ИКТ. – навыками разработки алгоритмов и программ для практической реализации в бизнесе. – навыками управления созданием и использованием баз данных в сфере информационно-коммуникационных технологий.
ОПК-4	Способен понимать принципы работы информационных технологий; использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений;	Компоненты компетенции соотносятся с содержанием дисциплины, и компетенция реализуется частично	ОПК-4.1. Знает методы и программные средства для сбора, обработки и анализа бизнес-информации. ОПК-4.2. Умеет обрабатывать, анализировать и моделировать информационные потоки организации. ОПК-4.3. Владеет навыками обоснования принятия управленческого решения как результата информационно аналитической деятельности.	Знает: – методы и программные средства для сбора, обработки и анализа бизнес-информации. – основные модели информационных потоков организации. – основные методы принятия управленческого решения как результата информационно - аналитической деятельности. Умеет: – применять программные средства для сбора, обработки и анализа бизнес-информации. – обрабатывать, анализировать и моделировать информационные потоки организации. – обосновывать принятие управленческого решения как результата информационно аналитической деятельности. Владеет: – программными средствами для сбора, обработки и анализа бизнес – информации. – методами анализа и

				моделирования информационных потоков организации. – навыками обоснования принятия управленческого решения как результата информационно аналитической деятельности.
--	--	--	--	---

5.3. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания.

Код компетенции	Этапы формирования компетенций	Инструмент, оценивающий сформированность компетенции	Показатель оценивания компетенции
ОПК-3, ОПК-4	Этап формирования знаниевой основы компетенций (этап формирования содержательно-теоретического базиса компетенции) Лекционные и практические занятия по темам: Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации Тема 2. Технологии хранения и обработки Больших данных Тема 3. Программирование обработки и загрузки Больших данных Тема 4. Аналитика в больших данных	Практические задания* Экзамен*	А) полностью сформирована - 5 баллов Б) частично сформирована - 3-4 балла С) не сформирована- 2 и менее баллов

***Характеристики инструментов (средств), оценивающих сформированность компетенций:**

Практическое задание – это частично регламентированное задание, имеющее алгоритмическое или нестандартное решение, позволяющее диагностировать умения интегрировать знания различных научных областей, аргументировать собственную точку зрения, доказывать правильность своей позиции. Может выполняться в индивидуальном порядке или группой обучающихся.

Экзамен – контрольное мероприятие, которое проводится по учебной дисциплине в виде, предусмотренном учебным планом, по окончании изучения курса. Занятие аудиторное, проводится в устной или письменной форме с использованием фондов оценочных средств по учебной дисциплине.

5.4. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Текущая аттестация

Практические задания для самостоятельной работы студентов:

Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации.

Вопросы для самостоятельного изучения.

1. Факторный анализ.
2. Дискриминантный анализ.
3. Кластерный анализ.

Тема 2. Технологии хранения и обработки Больших данных.

Вопросы для самостоятельного изучения.

1. Основные направления развития методов обработки и хранения данных.
2. Проблема хранения неструктурированных данных.
3. Проблема преобразования данных.
4. Семантические анализаторы. Самообучающиеся автоматы.

Тема 3. Программирование обработки и загрузки Больших данных.

Вопросы для самостоятельного изучения.

1. Языки для Big Data.
2. Фреймворки.
3. Базы данных.
4. Аналитические платформы.

Тема 4. Аналитика в больших данных.

Вопросы для самостоятельного изучения.

1. Аналитика Big Data — реалии и перспективы в России и мире.
2. Big data: применение и возможности.
3. Рынок Big data в России.

Тестовые задания

Принятый способ представления данных: показатели должны быть:

- 1) по строкам;
- 3) по ячейкам;
- 2) по столбцам;
- 4) по диагонали.

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.

Среди ниже приведённых нечисловые данные следующие:

- 1) баллы;
- 2) дихотомические;
- 3) ранги;
- 4) рейтинги.

Простейшие статистические характеристики – это:

- 1) среднее;
- 3) с.к.о.;
- 2) математическое ожидание;
- 4) дисперсия.

Приведение к нормальной форме - это:

- 1) деление на с.к.о.;
- 2) округление;
- 3) деление на среднее;
- 4) деление на константу интегрирования.

Какие функции Excel имеют отношение к оцифровке:

- 1) РАНГ;
- 3) СЧЁТЕСЛИ;
- 2) КОРРЕЛ;
- 4) СУММЕСЛИ.

Многомерность в статистике - это:

- 1) переменных больше одной;
- 3) измерений больше 10;
- 2) переменных больше двух;
- 4) измерений больше 5

Следующие программы являются специализированными статистическими паке-тами:

- 1) EXCEL;
- 3) GRAPHER;
- 2) SPSS;
- 4) STATISTICA.

Проверка статистической гипотезы включает в себя:

- 1) ранжирование;
- 2) принятие уровня значимости;
- 3) вычисление эмпирического значения;
- 4) вычисление критического значения.

Кластерный анализ предназначен для:

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

Опции кластерного анализа:

- 1) расстояние между группами;
- 3) расстояние между объектами;
- 2) расстояние между показателями;

4) расстояние между телами.

Кластерный анализ реализован в программах:

- 1) EXCEL;
- 3) SPSS;
- 2) AGRAPHER;
- 4) STATISTICA.

Снижение размерности это:

- 1) уменьшение числа измерений;
- 3) уменьшение числа показателей;
- 2) уменьшение числа объектов;
- 4) уменьшение числа знаков.

Компонентный анализ реализован в программах:

- 1) EXCEL;
- 2) SPSS;
- 3) AGRAPHER;
- 4) STATISTICA.

Методы, относящиеся к снижению размерности:

- 1) факторный анализ;
- 2) регрессия;
- 3) компонентный анализ;
- 4) корреляция.

Компонентный анализ позволяет:

- 1) сортировать;
- 3) ранжировать;
- 2) группировать;
- 4) упорядочивать.

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.

К нечисловым шкалам относятся:

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

Существует шкал для описания данных:

- 1) 4;
- 2) 6;
- 3) 5;
- 4) 7.

Количество наблюдений - это:

- 1) размерность;

- 3) ширина;
- 2) объём выборки;
- 4) поверхность выборки.

Элементы таблицы сопряжённости называются:

- 1) координаты;
- 3) скорости;
- 2) длины;
- 4) частоты.

Методы анализа таблиц сопряжённости:

- 1) Критерий Розенбаума;
- 3) хи-квадрат;
- 2) Критерий Колмогорова-Смирнова;
- 4) критерий Фишера.

В ходе анализа таблицы сопряжённости выполняется:

- 1) проверка на соответствие;
- 2) проверка на непротиворечивость;
- 3) проверка на монотонность;
- 4) проверка на значимость.

Максимальная размерность таблицы сопряжённости может быть:

- 1) 3;
- 2) 5;
- 3) 10;
- 4) какая угодно.

Вычисляемое значение критерия хи-квадрат называется:

- 1) Численное значение;
- 2) реальное значение;
- 3) экспериментальное значение;
- 4) эмпирическое значение.

Вычисляемое значение хи-квадрат сравнивается с:

- 1) критическим значением;
- 3) предельным значением;
- 2) эталонным значением;
- 4) граничным значением.

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

- 1) ХИ2РАСП;
- 3) ХИ2ТЕСТ;
- 2) ХИ2ОБР;
- 4) ХИ2.

К коэффициентам связи относятся:

- 1) коэффициент контингенции;
- 3) коэффициент ассоциации;
- 2) Коэффициент Чупрова-Крамера;
- 4) коэффициент коллигации.

К разновидности критерия хи-квадрат относятся:

- 1) критерий Вилкоксона;
- 2) информационный критерий;
- 3) критерий Джонкира;
- 4) критерий максимального правдоподобия.

Выявление вкладов, вносимых каждой клеткой таблицы, называется:

- 1) разбиение хи-квадрат;
- 2) локализация хи-квадрат;
- 3) анализ хи-квадрат;
- 4) сортировка хи-квадрат.

Лог-линейный анализ - это:

- 1) анализ синтеза таблиц;
- 2) анализ достоверности таблиц;
- 3) статистический анализ связи таблиц;
- 4) анализ разброса таблиц.

Перечень вопросов для промежуточного контроля знаний на экзамене (6 семестр):

1. Факторный анализ.
2. Дискриминантный анализ.
3. Кластерный анализ.
4. Многомерное шкалирование.
5. Методы контроля качества.
1. Основные направления развития методов обработки и хранения данных.
2. Volume.
3. Закон Мура.
4. Velocity. Variety.
5. Фреймворк Hadoop.
6. Проблема хранения неструктурированных данных.
7. Проблема преобразования данных.
8. Семантические анализаторы.
9. Самообучающиеся автоматы.
10. Языки для Big Data: R.
11. Языки для Big Data: Python.
12. Языки для Big Data: Julia.
13. Языки для Big Data: Java.
14. Языки для Big Data: Scala.
15. Языки для Big Data: MATLAB.
16. Языки для Big Data: Kafka,.
17. Языки для Big Data: Hadoop.
18. Языки для Big Data: Go.
19. Фреймворки для Big Data: Hadoop.
20. Фреймворки для Big Data: Spark.
21. Фреймворки для Big Data: Storm.
22. Базы данных для Big Data: Hive.
23. Базы данных для Big Data: Impala.
24. Базы данных для Big Data: Presto.
25. Базы данных для Big Data: Drill.
26. Аналитические платформы для Big Data: Rapid Miner.
27. Аналитические платформы для Big Data: IBM SPSS Modeler.

28. Аналитические платформы для Big Data: KNIME.
29. Аналитические платформы для Big Data: Qlik Analytics Platform.
30. Аналитические платформы для Big Data: STATISTICA Data Miner.
31. Аналитические платформы для Big Data: Informatica Intelligent Data Platform.
32. Аналитические платформы для Big Data: World Programming System.
33. Аналитические платформы для Big Data: Deductor.
34. Аналитические платформы для Big Data: SAS Enterprise Miner.
35. Zookeeper.
36. Flume.
37. IBM Watson Analytics.
38. Dell EMC Analytic Insights Module.
39. Windows Azure HDInsight.
40. Microsoft Azure Machine Learning.
41. Pentaho Data Integration.
42. Teradata Aster Analytics.
43. SAP BusinessObjects Predictive Analytics.
44. Oracle Big Data Preparation.
45. Аналитика Big Data — реалии и перспективы в России и мире.
46. Data Mining.
47. Краудсорсинг.
48. Смешение и интеграция данных.
49. Машинное обучение.
50. Искусственные нейронные сети.
51. Распознавание образов.
52. Прогнозная аналитика.
53. Имитационное моделирование.
54. Пространственный анализ.
55. Статистический анализ.
56. Визуализация аналитических данных.
57. Big Data: применение и возможности.
58. Решения на основе Big Data.
59. Рынок Big Data в России.
60. Big Data в банках.
61. Big Data в бизнесе.
62. Big Data в маркетинге.

Примерный перечень практических заданий на экзамене

1. Дан набор данных заданной структуры и программа SAS Data step, производящая определенную обработку и вычисления с использованием данного набора. Перепишите эту программу на SAS DS2 с использованием параллельных нитей и созданием пользовательского пакета, чтобы результат обработки сохранился тем же, но код мог выполняться в параллельной среде.
2. Дан набор заданной структуры, постройте модель прогнозирования отклика с использованием процедуры `impstat` с алгоритмом `random forest` с заданным числом деревьев. Примените полученную модель к тестовому набору данных той же структуры, визуализируйте полученный график Lift. Постройте на том же наборе модель с использованием высокопроизводительной версии метода GLM. Примените к тестовому набору. Сравните результаты GLM и Random Forest по AUC.

3. Дан текстовый корпус документов, лежащих в указанной директории. Создайте в SAS Text Miner проект, который: выберет файлы с расширением pdf; осуществит парсинг набора с определением частей речи и сохранением в признаковом пространстве только существительных и глаголов; осуществит фильтрацию документов и признаков с использованием заданной схемой определения весов лексем (например, на основе tf-idf); выделит заданное количество ключевых тематик по методу SVD. В ответе укажите топ 5 ключевых слов во второй выявленной тематике. Какой документ имеет наибольший вес в это тематике?

5.5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.

Ответы обучающегося **на экзамене** оцениваются каждым педагогическим работником по **пятибалльной системе**.

Критерии оценки ответа на вопросы экзамена:

Критерии оценки ответа на вопросы теоретического блока:

«5» – обучающийся глубоко и прочно освоил программный материал, исчерпывающе, последовательно, грамотно и логически стройно его излагает, тесно увязывает с задачами и будущей деятельностью, не затрудняется с ответом при видоизменении задания, свободно справляется с задачами и практическими заданиями, правильно обосновывает принятые решения, умеет самостоятельно обобщать и излагать материал, не допуская ошибок;

«4» - обучающийся твердо знает программный материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопрос, может правильно применять теоретические положения и владеет необходимыми умениями и навыками при выполнении практических заданий;

«3» - обучающийся освоил основной материал, но не знает отдельных деталей, допускает неточности, недостаточно правильные формулировки, нарушает последовательность в изложении программного материала и испытывает затруднения в выполнении практических заданий;

«2» - обучающийся не знает значительной части программного материала, допускает существенные ошибки, с большими затруднениями выполняет практические задания, задачи.

Критерии оценки практического задания:

«5» – задание выполнено верно, даны ясные аналитические выводы к решению задачи, подкрепленные теорией;

«4» – задание выполнено верно, отмечается хорошее развитие аргумента, однако отмечены погрешности в ответе, скорректированные при собеседовании;

«3» – задание выполнено с математическими ошибками, отсутствуют логические выводы и заключения к решению;

«2» – задание не выполнено, представленные расчеты проведены с ошибками, сделаны неверные выводы по решению задачи

6. Перечень основной и дополнительной учебной литературы для освоения учебной дисциплины.

6.1. Основная литература.

1. Параскевов, А. В. Большие данные : учебник : [12+] / А. В. Параскевов, А. Э. Сергеев. – Москва ; Вологда : Инфра-Инженерия, 2024. – 148 с. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=725632>

2. Конкина, В. В. Введение в большие данные и анализ информации : учебное пособие / В. В. Конкина, А. Б. Борисенко, И. Л. Коробова ; Тамбовский государственный технический университет. – Тамбов : Тамбовский государственный технический университет (ТГТУ), 2024. – 82 с. : ил., табл. – Режим доступа: по подписке. –

URL: <https://biblioclub.ru/index.php?page=book&id=723478>

3. Базы данных в высокопроизводительных информационных системах : учебное пособие / авт.-сост. Е. И. Николаев ; Северо-Кавказский федеральный университет. – Ставрополь : Северо-Кавказский Федеральный университет (СКФУ), 2016. – 163 с. : ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=466799>

6.2. Дополнительная литература

1. Управление данными : учебник / Ю. Ю. Громов, О. Г. Иванова, А. В. Яковлев, В. Г. Однолько ; Тамбовский государственный технический университет. – Тамбов : Тамбовский государственный технический университет (ТГТУ), 2015. – 192 с. : ил., табл., схем. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=444642>

2. Жуков, Р. А. Базы данных : учебно-методическое пособие по дисциплине «Базы данных» для направления подготовки 38.03.05 «Бизнес-информатика» (бакалавриат) : [16+] / Р. А. Жуков. – Москва ; Берлин : Директ-Медиа, 2019. – 177 с. : ил., схем., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=566814>

3. Программно-аппаратные средства сбора, хранения и обработки информации : учебное пособие : [16+] / Н. Н. Скворцова, А. К. Козак, Н. Н. Богачев [и др.]. – Москва ; Вологда : Инфра-Инженерия, 2024. – 124 с. : ил., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=725680>

7. Перечень ресурсов информационно-коммуникационной сети «Интернет», современные профессиональные базы данных и информационные системы, необходимые для освоения учебной дисциплины

1. <http://minobrnauki.gov.ru> – сайт Минобрнауки РФ
2. <http://biblioclub.ru> – ЭБС Университетская библиотека
3. <https://ro-edu.ru/> – Медиапортал «Российское образование»
4. <http://www.prlib.ru> – Президентская библиотека
5. <http://www.rusneb.ru> – Национальная электронная библиотека
6. <http://lib.icone.ru> - Электронно-библиотечная система АНО ВО «Институт непрерывного образования»
7. <http://elibrary.rsl.ru/> – сайт Российской государственной библиотеки (раздел «Электронная библиотека»)
8. <http://elibrary.ru> – научная электронная библиотека «Elibrary»
9. <https://zbmath.org> – Реферативный математический журнал и также реферативная электронная база данных Zentralblatt MATH
10. <http://www.problems.ru> – Интернет-проект «Задачи» по Математике
11. <http://eqworld.ipmnet.ru/> Научно-образовательный сайт EqWorld — Мир математических уравнений

8. Методические указания для обучающихся по освоению учебной дисциплины

Освоение обучающимся учебной дисциплины «Введение в Big Data» предполагает изучение материалов дисциплины на аудиторных занятиях и в ходе самостоятельной работы. Аудиторные занятия проходят в форме лекций и практических занятий. Самостоятельная работа включает разнообразный комплекс видов и форм работы обучающихся.

Для успешного освоения учебной дисциплины и достижения поставленных целей необходимо внимательно ознакомиться настоящей рабочей программы учебной дисциплины. Ее может представить преподаватель на вводной лекции или самостоятельно обучающийся использует информацию на официальном Интернет-сайте Института.

Следует обратить внимание на список основной и дополнительной литературы, которая имеется в электронной библиотечной системе Института, на предлагаемые преподавателем ресурсы информационно-телекоммуникационной сети Интернет. Эта информация необходима для самостоятельной работы обучающегося.

При подготовке к аудиторным занятиям необходимо помнить особенности каждой формы его проведения.

Подготовка к учебному занятию лекционного типа заключается в следующем.

С целью обеспечения успешного обучения обучающийся должен готовиться к лекции, поскольку она является важнейшей формой организации учебного процесса, поскольку:

- знакомит с новым учебным материалом;
- разъясняет учебные элементы, трудные для понимания;
- систематизирует учебный материал;
- ориентирует в учебном процессе.

С этой целью:

- внимательно прочитайте материал предыдущей лекции;
- ознакомьтесь с учебным материалом по учебнику и учебным пособиям с темой прочитанной лекции;
- внесите дополнения к полученным ранее знаниям по теме лекции на полях лекционной тетради;
- запишите возможные вопросы, которые вы зададите лектору на лекции по материалу изученной лекции;
- постарайтесь уяснить место изучаемой темы в своей подготовке;
- узнайте тему предстоящей лекции (по тематическому плану, по информации лектора) и запишите информацию, которой вы владеете по данному вопросу

Подготовка к практическому занятию

При подготовке к практическому занятию следует обратить внимание на следующие моменты: на процесс предварительной подготовки, на работу во время занятия, обработку полученных результатов, исправление полученных замечаний.

Предварительная подготовка к практическому занятию заключается в изучении теоретического материала в отведенное для самостоятельной работы время, ознакомление с инструктивными материалами с целью осознания задач практического занятия.

Работа во время проведения практического занятия включает несколько моментов:

- консультирование студентов преподавателями с целью предоставления исчерпывающей информации, необходимой для самостоятельного выполнения предложенных преподавателем задач, ознакомление с правилами техники безопасности при работе в аудитории;
- самостоятельное выполнение заданий согласно обозначенной учебной программой тематики.

Самостоятельная работа.

Для более углубленного изучения темы задания для самостоятельной работы рекомендуется выполнять параллельно с изучением данной темы. При выполнении заданий по возможности используйте наглядное представление материала. Более подробная информация о самостоятельной работе представлена в разделах «Учебно-методическое обеспечение самостоятельной работы по дисциплине (модулю)», «Методические указания к самостоятельной работе по дисциплине (модулю)».

Подготовка к экзамену.

К экзамену необходимо готовиться целенаправленно, регулярно, систематически и с первых дней обучения по данной дисциплине. При подготовке к экзамену обратите внимание на защиту практических заданий на основе теоретического материала.

При подготовке к экзамену по теоретической части выделите в вопросе главное, существенное (понятия, признаки, классификации и пр.), приведите примеры, иллюстрирующие теоретические положения.

После предложенных указаний у обучающихся должно сформироваться четкое представление об объеме и характере знаний и умений, которыми надо будет овладеть по дисциплине.

9. Программное обеспечение информационно-коммуникационных технологий

9.1. Информационные технологии

1. Персональные компьютеры.
2. Доступ в Интернет.
3. Проектор.
4. Система VOTUM

9.2. Программное обеспечение.

1. Windows 7
2. LibreOffice Writer,
3. LibreOffice Calc,
4. LibreOffice Impress

9.3. Информационно-справочные системы.

1. Справочно-правовая система по законодательству Российской Федерации Гарант - <http://www.garant.ru/>
2. Справочная правовая система КонсультантПлюс -<http://www.consultant.ru/>

10. Перечень материально-технического обеспечения, необходимого для реализации программы по учебной дисциплине.

Наименование помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом, в том числе помещения для самостоятельной работы, с указанием перечня основного оборудования, учебно-наглядных пособий и используемого программного обеспечения	Адрес (местоположение) помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом (в случае реализации образовательной программы в сетевой форме дополнительно указывается наименование организации, с которой заключен договор)
Учебная аудитория №5 Оснащение: Стул с пюпитром – 48 шт. Стол преподавателя – 1 шт. Стул преподавателя – 1 шт. Экран большого размера– 1 шт. Меловая доска – 1 шт. Ноутбук - 1 шт. Проектор INFocus -1 шт.	115114, г. Москва, вн.тер.г. муниципальный округ Даниловский, наб. Дербеневская, д. 11, этаж № 1, нежилое помещение, площадью 479.2 м2 (часть нежилого помещения)
Помещение для самостоятельной работы Оснащение: Стул с пюпитром – 25 шт. Стол преподавателя – 1 шт., Стул преподавателя – 1 шт. Персональный компьютер с периферией, с возможностью подключения к сети «Интернет» и	115114, г. Москва, вн.тер.г. муниципальный округ Даниловский, наб. Дербеневская, д. 11, этаж № 1, нежилое помещение, площадью 479.2 м2 (часть нежилого помещения)

обеспечением доступа в электронную информационно-образовательную среду Организации. – 1 шт. МФУ - 1 шт.	
--	--

11. Образовательные технологии

При реализации учебной дисциплины «Введение в Big Data» применяются различные образовательные технологии.

Освоение учебной дисциплины «Введение в Big Data» предусматривает использование в учебном процессе активных и интерактивных форм проведения учебных занятий в форме разбора конкретных ситуаций и практических задач в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков обучающихся.