

Документ подписан простой электронной подписью

Информация о документе:

ФИО: Хоружий Лидия Ивановна

Должность: директор Института экономики и управления АПК

Дата подписания: 2022-02-23 19:31:30

Уникальный идентификатор документа:

1e90b132d9b040ce67585160b015dddf2cb1e6a9



**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ**

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ –**

**МСХА имени К.А. ТИМИРЯЗЕВА»**

**(ФГБОУ ВО РГАУ - МСХА имени К.А. Тимирязева)**

Институт Экономики и управления АПК

Кафедра Прикладной информатики



УТВЕРЖДАЮ:

Директор Института  
экономики и управления АПК

Л.И. Хоружий

« 23 » февраля 2022 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
**Б1.В.24 ТЕХНОЛОГИИ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ**

(индекс и наименование дисциплины по учебному плану)

для подготовки бакалавров

ФГОС ВО

Направление: 09.03.03 Прикладная информатика

Направленность: Прикладная информатика в экономике, Системы искусственного интеллекта

Курс 2

Семестр 3

Форма обучения: очная

Год начала подготовки: 2022

Москва, 2022

Разработчик (и): Греченева А.В. к.т.н.

(ФИО, ученая степень, ученое звание)

  
(подпись)

Быстренина И.Е., к.пед.н

(ФИО, ученая степень, ученое звание)

  
(подпись)

«29» августа 2022 г.

Рецензент: Ашмарина Т.И., к.э.н., доцент

(ФИО, ученая степень, ученое звание)

  
(подпись)

«29» августа 2022 г.

Программа составлена в соответствии с требованиями ФГОС ВО, профессионального стандарта и учебного плана по направлению подготовки 09.03.03 Прикладная информатика.

И.о. зав. кафедрой прикладной информатики Е.В. Худякова д.э.н., профессор

(ФИО, ученая степень, ученое звание)

  
(подпись)

«29» августа 2022 г.

**Согласовано:**

Председатель учебно-методической  
комиссии Института экономики и управления АПК  
Гупалова Т.Н., к.э.н., доцент

(ФИО, ученая степень, ученое звание)

  
(подпись)

«29» августа 2022 г.

И.о. заведующей выпускающей кафедрой

Е.В. Худякова, д.э.н., профессор

(ФИО, ученая степень, ученое звание)

  
(подпись)

«29» августа 2022 г.

/Заведующий отделом комплектования ЦНБ

  
(подпись)

## СОДЕРЖАНИЕ

<b>1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....</b>	<b>4</b>
<b>2. МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ .....</b>	<b>5</b>
<b>3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ Б1.В.24 ТЕХНОЛОГИИ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ.....</b>	<b>5</b>
<b>4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ .....</b>	<b>5</b>
4.1 РАСПРЕДЕЛЕНИЕ ТРУДОЁМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ .....	6
ПО СЕМЕСТРАМ .....	6
4.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ.....	10
4.3 ЛЕКЦИИ/ЛАБОРАТОРНЫЕ/ПРАКТИЧЕСКИЕ/ ЗАНЯТИЯ.....	11
<b>5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ .....</b>	<b>19</b>
<b>6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....</b>	<b>20</b>
6.1. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ, НЕОБХОДИМЫЕ ДЛЯ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ И НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ .....	20
6.2. ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ КОНТРОЛЯ УСПЕВАЕМОСТИ, ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ .....	25
<b>7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....</b>	<b>25</b>
7.1 ОСНОВНАЯ ЛИТЕРАТУРА .....	26
7.2 ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА.....	26
<b>8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ).....</b>	<b>27</b>
<b>9. ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ (ПРИ НЕОБХОДИМОСТИ) .....</b>	<b>27</b>
<b>10. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ (МОДУЛЮ).....</b>	<b>27</b>
<b>11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ОБУЧАЮЩИМСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ. 28</b>	
<b>12. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПРЕПОДАВАТЕЛЯМ ПО ОРГАНИЗАЦИИ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ.....</b>	<b>29</b>

## АННОТАЦИЯ

### рабочей программы учебной дисциплины **Б1.В.24 Технологии обработки больших данных**

для подготовки бакалавров по направлению 09.03.03 Прикладная информатика  
направленностей: Прикладная информатика в экономике, Системы искусственного интеллекта

**Цель освоения дисциплины:** формирование у студентов базовых теоретических знаний о сущности, структуре и видах математических моделей принятия решений, управления и способности к применению технологий обработки данных (в том числе big data) при проектировании технологий искусственного интеллекта (ИИ), интеллектуальных информационных технологий (ИТ) и систем в профессиональной деятельности, а также способность практического использования навыков получения и формализации знаний, а также навыков применения больших данных для моделирования сложных систем и процессов.

**Место дисциплины в учебном плане:** дисциплина включена в формируемую участниками образовательных отношений часть учебного плана по направлению подготовки 09.03.03 Прикладная информатика

**Требования к результатам освоения дисциплины:** в результате освоения дисциплины формируются следующие компетенции: ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3

**Краткое содержание дисциплины:**

Большие данные. 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. Языки программирования для Big Data (R, Python, Julia, Java, Scala, MATLAB, Go, Kafka, Hadoop). Фреймворки (Hadoop, Spark, Storm). Аналитические платформы (Rapid Miner, IBM SPSS Modeler, KNIME, Qlik Analytics Platform, STATISTICA 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).

**Общая трудоемкость дисциплины/в т.ч. практическая подготовка:**  
108/4 (часы/зач. ед.)

**Промежуточный контроль:** экзамен, курсовая работа.

### 1. Цель освоения дисциплины

Целью освоения дисциплины «Технологии обработки больших данных» является формирование у обучающихся компетенций, обеспечивающих способность к использованию больших данных при проектировании технологий искусственного интеллекта (ИИ), интеллектуальных информационных технологий (ИТ) и систем в профессиональной деятельности, а также способность практического использования навыков получения и формализации знаний, а также навыков применения больших данных для моделирования сложных систем и процессов. Учебный процесс по освоению дисциплины направлен на использование современных цифровых технологий, в том числе цифровых сервисов, моделей и программного обеспечения на основе технологий ИИ.

## **2. Место дисциплины в учебном процессе**

Дисциплина «Технологии обработки больших данных» относится к формируемой участниками образовательных части Блока 1 «Дисциплины (модули)» учебного плана. Дисциплина «Технологии обработки больших данных» реализуется в соответствии с требованиями ФГОС ВО, профессионального стандарта, ОПОП ВО и Учебного плана по направлению 09.03.03 Прикладная информатика.

(шифр, название)

Предшествующими курсами, на которых непосредственно базируется дисциплина «Технологии обработки больших данных» являются Математика, Иностранный язык, дискретная математика, теория вероятностей, математическая статистика, Алгоритмизация и программирование, Базы данных, Технологии обработки больших данных.

Дисциплина «Технологии обработки больших данных» является основополагающей для изучения следующих дисциплин: Анализ пространственно-временных данных на основе машинного обучения, Системы поддержки принятия решений, Разработка распределенных систем, Информационные системы управления производственной компанией, ВКР.

Особенностью дисциплины является получение знаний и навыков использования искусственного интеллекта в задачах оптимизации профессиональной деятельности.

Рабочая программа дисциплины «Технологии обработки больших данных» для инвалидов и лиц с ограниченными возможностями здоровья разрабатывается индивидуально с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся.

## **3. Перечень планируемых результатов обучения по дисциплине Б1.В.24 Технологии обработки больших данных, соотнесенных с планируемыми результатами освоения образовательной программы**

Образовательные результаты освоения дисциплины обучающимся, представлены в таблице 1.

## **4. Структура и содержание дисциплины**

Курс проходит в 3 семестре и рассчитан на 16 часов лекций, 48 часов практических занятий, 77 часов самостоятельной работы и 0,4 КРА.

Текущая аттестация студентов - оценка знаний и умений проводится на практических занятиях с помощью защиты практических работ и оценки самостоятельной работы студентов.

Промежуточный контроль проводится в форме курсового проекта и экзамена в 3 семестре.

#### **4.1 Распределение трудоёмкости дисциплины по видам работ по семестрам**

Общая трудоёмкость дисциплины составляет 5 зач.ед. (180 часов), их распределение по видам работ семестрам представлено в таблице 2.

Таблица 1

## Требования к результатам освоения учебной дисциплины

№ п/п	Код компетенции	Содержание компетенции (или её части)	Индикаторы компетенций	В результате изучения учебной дисциплины обучающиеся должны:		
				знать	уметь	владеть
1	ПКос-2	Способность разрабатывать и адаптировать прикладное программное обеспечение	ПКос-2.1 Знает основы написания программного кода с использованием языков программирования	принципы теоретические и прикладные основы анализа больших данных технологии хранения и обработки больших данных в организации: базы данных, хранилища данных, распределенная и параллельная обработка данных в средах программирования Google Colab, Anaconda на языке python		
			ПКос-11.2 Умеет использовать инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников, в том числе в режиме реального времени; проводить очистку, интеграцию, преобразование и анализ больших объемов данных		использовать инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников, в том числе в режиме реального времени; проводить очистку, интеграцию, преобразование и анализ больших объемов данных с помощью библиотек python NumPy, Pandas, Matplotlib, Seaborn,	

			ПКос-11.3 Владеет навыками получения и обработки больших данных с поддержкой работы в режиме реального времени		Plotly	навыками получения и обработки больших данных с поддержкой работы в режиме реального времени с помощью библиотек python NumPy, Pandas, Matplotlib, Seaborn, Plotly
2	ПКос-12	Обработка и анализ данных без предварительной разметки, в том числе машинно-генерируемых данных	ПКос-12.1 Знает принципы работы и особенности применения нейронных сетей для анализа данных, методы факторного и кластерного анализа, методы семантического анализа	принципы работы и особенности применения нейронных сетей для анализа данных, методы факторного и кластерного анализа, методы семантического анализа в средах программирования Google Colab, Anaconda на языке python		
			ПКос-12.2 Умеет производить автоматизированную обработку массивов первичных данных		производить автоматизированную обработку массивов первичных данных получать представление о наборе данных с помощью методов shape, head, describe, info библиотеки pandas, используя библиотеку sklearn, производить нормализацию данных с помощью метода RobustScaler библиотеки	



				sklearn	
			<p>ПКос-12.3  Владеет навыками выполнения автоматизированного анализа первичных данных с классификацией размеченных событий и построением прогностических моделей</p>		<p>выполнения автоматизированного анализа первичных данных с классификацией размеченных событий и построением прогностических моделей, кодировать категориальные признаки в дискретные величины с помощью метода <code>get_dummies</code></p>

Таблица 2

## Распределение трудоёмкости дисциплины по видам работ по семестрам

Вид учебной работы	Трудоёмкость (4 семестр) час. всего/*
<b>Общая трудоёмкость</b> дисциплины по учебному плану	<b>180/4</b>
<b>1. Контактная работа:</b>	<b>69,4/4</b>
<b>Аудиторная работа</b>	<b>69,4/4</b>
<i>в том числе:</i>	
<i>лекции (Л)</i>	16
<i>практические занятия (ПЗ)</i>	48/4
<i>контактная работа на промежуточном контроле (КРА)</i>	0,4
<b>2. Самостоятельная работа (СРС)</b>	<b>110,6</b>
<i>самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиумам и т.д.)</i>	77
<i>Подготовка к экзамену (контроль)</i>	39,6
Вид промежуточного контроля:	Экзамен, курсовой проект

## 4.2 Содержание дисциплины

Тематический план по очной форме обучения представлен в таблице 3.

Таблица 3

## Тематический план учебной дисциплины

Наименование разделов и тем дисциплин (укрупнённо)	Всего	Аудиторная работа			Внеаудиторная работа СР
		Л	ПЗ всего/*	ПКР	
Раздел 1. Технологии обработки больших данных	140/4	16	48/4		77
<i>контактная работа на промежуточном контроле (КРА)</i>	0,4			0,4	
<i>Подготовка к экзамену (контроль)</i>	39,6			39,6	
<b>Итого по дисциплине</b>	<b>180/4</b>	<b>16</b>	<b>48/4</b>	<b>40</b>	<b>77</b>

## Раздел 1. Технологии обработки больших данных

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

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

## 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.

**Тема 3. Языки программирования для анализа и обработки больших массивов данных.** Языки программирования для Big Data (R, Python, Julia, Java, Scala, MATLAB, Go, Kafka, Hadoop)

**Тема 4. Базы данных.**

Реляционные базы данных. Параллельные базы данных. Распределённые файловые системы. NoSQL СУБД. Технология Map-Reduce. GOOGLE BIGTABLE. MapReduce. Обычный поиск. Полнотекстовый поиск. Параллельные запросы. Технология поиска и интеграции. Базы данных (Hive, Impala, Presto, Drill).

**Тема 5. Программные средства для обработки данных и системы Data Mining**

Программа Excel. Программа Statistica. SPSS. Программные средства. ETL процесс по обработке отчётов. Понятие о технологии Data Mining. Реализация в пакетах прикладных программ. Сетевые технологии Data Mining.

**Тема 6. Платформы и фреймворки для Big Data**

Фреймворки (Hadoop, Spark, Storm). Аналитические платформы (Rapid Miner, IBM SPSS Modeler, KNIME, Qlik Analytics Platform, STATISTICA 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).

**Тема 7. Аналитика Big Data**

Технологии и методы 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).

**Тема 8. Введение в машинное обучение и разработку данных**

Машинное обучение и майнинг больших данных (Big Data). Нейронные сети как реализация алгоритмов машинного обучения

**Тема 9. Способы визуализации, интерпретации, проектная организация**

Способы визуального представления данных. Методы визуализации. Когнитивная информатика, экономика знаний, требования к специалистам в области интеллектуальной обработки данных для бизнеса.

**Тема 10. Решения на основе Big data**

Рынок Big data. Big data в банках. Big data в бизнесе. Big data в маркетинге.

### 4.3 Лекции/лабораторные/практические/ занятия

Таблица 4

**Содержание лекций /практических занятий и контрольные мероприятия**

№ п/п	Название раздела, темы	№ и название лекций/ лабораторных/ практических/ семинарских занятий	Формируемые компетенции	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
1	<b>Раздел 1 Технологии обработки больших данных</b>		ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3		
	<b>Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации.</b>	Лекция №1 Методы многомерного статистического анализа и анализа нечисловой информации.	ПКос-12.1, ПКос-12.2, ПКос-12.3	-	2
		Практическая работа №1. Изучение основных понятий и принципов технологии обработки больших данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2/1
		Практическая работа №2. Анализ и обработка больших данных с использованием языка программирования Python	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
	<b>Тема 2. Технологии хранения и обработки Больших данных.</b>	Лекция №2 Технологии хранения и обработки Больших данных.	ПКос-12.1, ПКос-12.2, ПКос-12.3	-	2
		Практическая работа №3. Обработка данных в реальном времени с помощью Apache Kafka	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №4. Использование Apache Hadoop для обработки больших объемов данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
	<b>Тема 3. Языки программирования для анализа и обработки больших массивов данных.</b>	Лекция №3 Языки программирования для анализа и обработки больших массивов данных.	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	-	2
		Практическая работа №5. Работа с NoSQL базами данных для хранения и обработки больших данных	ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2/1

№ п/п	Название раздела, темы	№ и название лекций/ лабораторных/ практических/ семинарских занятий	Формируемые компетенции	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
		Практическая работа №6. Создание и использование кластера для обработки больших данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
	<b>Тема 4. Базы данных.</b>	Лекция №4 Базы данных. Особенности работы с базами больших данных.	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	-	2
		Практическая работа №7. Использование Apache Spark для обработки данных в распределенной среде	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №8. Обработка и анализ данных с помощью Apache Pig	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2/1
	<b>Тема 5. Программные средства для обработки данных и системы Data Mining</b>	Лекция №5 Программные средства для обработки данных и системы Data Mining	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	-	2
		Практическая работа №9. Использование Apache Hive для запросов к данным в большом масштабе	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №10. Обработка и анализ графовых данных с помощью Apache Giraph	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2/1
	<b>Тема 6. Платформы и фреймворки для Big Data</b>	Лекция №6 Платформы и фреймворки для Big Data	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	-	2

№ п/п	Название раздела, темы	№ и название лекций/ лабораторных/ практических/ семинарских занятий	Формируемые компетенции	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
		Практическая работа №11. Анализ и обработка текстовых данных с помощью Natural Language Processing (NLP)	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №12. Обработка и анализ изображений с помощью Computer Vision	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
	<b>Тема 7. Аналитика Big Data</b>	Лекция №7 Аналитика Big Data	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	-	2
		Практическая работа №13. Использование MapReduce для обработки больших объемов данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №14. Работа с инструментами Big Data Analytics для обработки и анализа больших данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
	<b>Тема 8. Введение в машинное обучение и разработку данных</b>	Лекция №8 Введение в машинное обучение и разработку данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	-	2
		Практическая работа №15. Анализ и обработка больших данных в реальном времени с помощью Apache Storm	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №16. Использование Apache Flink для обработки потоковых данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2

№ п/п	Название раздела, темы	№ и название лекций/ лабораторных/ практических/ семинарских занятий	Формируемые компетенции	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
		Практическая работа №17. Работа с инструментами машинного обучения для анализа и обработки больших данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
	Тема 9. Способы визуализации, интерпретации, проектная организация	Практическая работа №18. Использование Apache Cassandra для хранения и обработки больших объемов данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №19. Обработка и анализ временных рядов с помощью Time Series Analysis	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №20. Использование Apache Zeppelin для визуализации и анализа больших данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
	Тема 10. Решения на основе Big data	Практическая работа №21. Обработка и анализ данных с помощью R	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №22. Использование Apache NiFi для обработки и передачи данных в реальном времени	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2
		Практическая работа №23. Работа с инструментами Data Mining для анализа больших объемов данных	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2

№ п/п	Название раздела, темы	№ и название лекций/ лабораторных/ практических/ семинарских занятий	Формируемые компетенции	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
		Практическая работа №24 Анализ и обработка данных с помощью Apache Kylin.	ПКос-11.1, ПКос-11.2, ПКос-11.3, ПКос-12.1, ПКос-12.2, ПКос-12.3	устный опрос, защита практической работы	2

Перечень вопросов для самостоятельного изучения дисциплины представлен в Таблице 5.

Таблица 5

### Перечень вопросов для самостоятельного изучения дисциплины

№ п/п	Название раздела, темы	Перечень рассматриваемых вопросов для самостоятельного изучения
<b>Раздел 1 Технологии обработки больших данных</b>		
	Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации.	<ol style="list-style-type: none"> <li>1. Какие методы многомерного статистического анализа используются для анализа данных?</li> <li>2. Что такое факторный анализ и как он применяется в Big Data?</li> <li>3. Какие методы кластерного анализа используются для обработки больших объемов данных?</li> <li>4. Как проводится анализ нечисловой информации в Big Data?</li> <li>5. Какие инструменты и технологии используются для анализа нечисловой информации в Big Data?</li> <li>6. Какие преимущества имеет использование методов многомерного статистического анализа и анализа нечисловой информации в Big Data?</li> <li>7. Какие сложности могут возникнуть при проведении многомерного статистического анализа и анализа нечисловой информации в Big Data?</li> </ol>
	Тема 2. Технологии хранения и обработки Больших данных.	<ol style="list-style-type: none"> <li>1. Какие технологии используются для хранения Больших данных?</li> <li>2. Что такое Hadoop и как он используется для обработки Больших данных?</li> <li>3. Какие преимущества имеет использование NoSQL баз данных для хранения Больших данных?</li> <li>4. Какие технологии используются для обработки потоков данных в режиме реального времени?</li> <li>5. Как происходит интеграция различных источников данных в Big Data?</li> <li>6. Какие инструменты используются для визуализации и анализа Больших данных?</li> <li>7. Как обеспечивается безопасность и защита Больших данных в системах Big Data?</li> </ol>
	Тема 3. Языки программирования для анализа и обработки больших массивов данных.	<ol style="list-style-type: none"> <li>1. Какие языки программирования используются для анализа и обработки Больших данных, например, Python, R, Java?</li> <li>2. Что такое Apache Spark и как он используется для обработки Больших данных?</li> <li>3. Какие инструменты используются для машинного обучения и</li> </ol>



№ п/п	Название раздела, темы	Перечень рассматриваемых вопросов для самостоятельного изучения
		<p>анализа данных в Big Data, например, TensorFlow, Apache Mahout, Apache MXNet?</p> <p>4. Какие инструменты используются для обработки текстовых данных в Big Data, например, Apache Lucene, Apache Solr?</p> <p>5. Какие инструменты используются для обработки графовых данных в Big Data, например, Apache Giraph, Neo4j?</p> <p>6. Какие инструменты используются для обработки временных рядов в Big Data, например, Apache Kafka, Apache Flink?</p> <p>7. Какие инструменты используются для распределенного хранения и обработки Больших данных, например, Apache Cassandra, Apache HBase?</p>
	Тема 4. Базы данных.	<ol style="list-style-type: none"> <li>1. Распределенные системы хранения данных, например, Hadoop Distributed File System (HDFS).</li> <li>2. Обработка данных в реальном времени с помощью Apache Kafka и Apache Flink.</li> <li>3. Машинное обучение и анализ данных с использованием TensorFlow, Apache Mahout и Apache MXNet.</li> <li>4. Обработка текстовых данных с помощью Apache Lucene и Apache Solr.</li> <li>5. Обработка графовых данных с помощью Apache Giraph и Neo4j.</li> <li>6. Работа с временными рядами с помощью Apache Kafka и Apache Flink.</li> <li>7. Базы данных NoSQL, такие как Apache Cassandra и Apache HBase.</li> <li>8. Инструменты для визуализации данных, такие как Tableau и Power BI.</li> <li>9. Безопасность и конфиденциальность данных в Big Data.</li> <li>10. Распределенные вычисления и параллельная обработка данных.</li> </ol>
	Тема 5. Программные средства для обработки данных и системы Data Mining	<ol style="list-style-type: none"> <li>1. Основы Data Mining и машинного обучения.</li> <li>2. Алгоритмы классификации и кластеризации данных.</li> <li>3. Извлечение признаков и предобработка данных.</li> <li>4. Обработка естественного языка и анализ текстов.</li> <li>5. Методы ассоциативного анализа и поиска аномалий.</li> <li>6. Интеграция и анализ данных из различных источников.</li> <li>7. Использование графовых баз данных для Data Mining.</li> <li>8. Работа с большими объемами данных при помощи MapReduce и Hadoop.</li> <li>9. Визуализация и интерпретация результатов Data Mining.</li> <li>10. Этические и юридические аспекты использования Data Mining.</li> </ol>
	Тема 6. Платформы и фреймворки для Big Data	<ol style="list-style-type: none"> <li>1. Apache Hadoop и его экосистема: HDFS, MapReduce, YARN, Hive, Pig, Spark.</li> <li>2. NoSQL базы данных для Big Data: Cassandra, MongoDB, Couchbase, Amazon DynamoDB.</li> <li>3. Apache Kafka для обработки потоков данных.</li> <li>4. Apache Storm для обработки потоков данных в режиме реального времени.</li> <li>5. Apache Flink для обработки потоков данных и пакетных данных.</li> <li>6. Apache Beam для создания и запуска пайплайнов обработки данных на различных платформах.</li> </ol>

№ п/п	Название раздела, темы	Перечень рассматриваемых вопросов для самостоятельного изучения
		<p>7. TensorFlow и PyTorch для машинного обучения и глубокого обучения.</p> <p>8. Apache Zeppelin и Jupyter Notebook для интерактивной работы с данными и анализа результатов.</p> <p>9. Apache Druid для быстрого анализа и запросов к большим объемам данных.</p> <p>10. Kubernetes для управления кластерами и контейнерами с приложениями для Big Data.</p>
	Тема 7. Аналитика Big Data	<p>1. Основы статистики и анализа данных.</p> <p>2. Методы машинного обучения и глубокого обучения.</p> <p>3. Обработка и анализ текстовых данных.</p> <p>4. Визуализация данных и создание дашбордов.</p> <p>5. Большие данные и их влияние на бизнес-процессы.</p> <p>6. Работа с неструктурированными данными, такими как изображения и видео.</p> <p>7. Работа с данными в режиме реального времени.</p> <p>8. Анализ социальных сетей и медиа-данных.</p> <p>9. Системы рекомендаций и персонализация контента.</p> <p>10. Безопасность данных и защита от кибератак.</p>
	Тема 8. Введение в машинное обучение и разработку данных	<p>1. Основные понятия и термины в машинном обучении.</p> <p>2. Различные типы задач машинного обучения: классификация, регрессия, кластеризация и т.д.</p> <p>3. Методы предобработки данных: очистка, преобразование, масштабирование и т.д.</p> <p>4. Обучение с учителем и без учителя: примеры задач и методов решения.</p> <p>5. Модели машинного обучения: линейные модели, деревья решений, нейронные сети и т.д.</p> <p>6. Оценка качества моделей: метрики, кросс-валидация и т.д.</p> <p>7. Техники улучшения моделей: отбор признаков, настройка параметров и т.д.</p> <p>8. Применение машинного обучения в реальных проектах: от выбора модели до ее внедрения.</p> <p>9. Использование библиотек и инструментов для машинного обучения: scikit-learn, TensorFlow, Keras и т.д.</p> <p>10. Этические и правовые аспекты использования машинного обучения и обработки данных.</p>
	Тема 9. Способы визуализации, интерпретации, проектная организация	<p>11. Способы визуализации данных: графики, диаграммы, тепловые карты и т.д.</p> <p>12. Интерпретация результатов моделирования: важность признаков, влияние параметров и т.д.</p> <p>13. Организация проекта машинного обучения: сбор данных, постановка задачи, выбор модели и т.д.</p> <p>14. Работа с большими объемами данных: параллельное вычисление, распределенные вычисления и т.д.</p> <p>15. Применение машинного обучения в различных отраслях: медицина, финансы, производство и т.д.</p> <p>16. Использование библиотек и инструментов для визуализации и интерпретации данных: Matplotlib, Seaborn, Plotly и т.д.</p> <p>17. Работа в команде на проектах машинного обучения: разделение задач, контроль версий и т.д.</p>

№ п/п	Название раздела, темы	Перечень рассматриваемых вопросов для самостоятельного изучения
		<p>18. Обработка текстовых данных: предобработка, векторизация, классификация и т.д.</p> <p>19. Применение машинного обучения для анализа изображений: распознавание объектов, сегментация и т.д.</p> <p>20. Этические и правовые аспекты использования данных и моделей машинного обучения в различных сферах: конфиденциальность, дискриминация и т.д.</p>
	Тема 10. Решения на основе Big data	<p>1. Что такое Big Data и какие проблемы она решает?</p> <p>2. Какие технологии используются для обработки и анализа Big Data?</p> <p>3. Какие преимущества имеют решения на основе Big Data по сравнению с традиционными подходами к анализу данных?</p> <p>4. Какие типы данных можно обрабатывать с помощью решений на основе Big Data?</p> <p>5. Какие методы машинного обучения используются для анализа Big Data?</p> <p>6. Какие инструменты используются для хранения и обработки Big Data?</p> <p>7. Какие проблемы могут возникнуть при работе с большими объемами данных и как их можно решить?</p> <p>8. Какие примеры успешного применения решений на основе Big Data в различных отраслях?</p> <p>9. Какие требования к аппаратному обеспечению и сетевой инфраструктуре необходимы для работы с Big Data?</p> <p>10. Какие навыки и знания необходимы для работы с решениями на основе Big Data?</p>

## 5. Образовательные технологии

Таблица 6

### Применение активных и интерактивных образовательных технологий

№ п/п	Тема и форма занятия	Наименование используемых активных и интерактивных образовательных технологий	
1.	Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации. Тема 2. Технологии хранения и обработки Больших данных.	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.
2.	Тема 3. Языки программирования для анализа и обработки больших массивов данных. Тема 4. Базы данных.	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.
3.	Тема 5. Программные средства для обработки данных и системы Data Mining Тема 6. Платформы и фреймворки для Big Data	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.
4.	Тема 7. Аналитика Big Data Тема 8. Введение в машинное обучение и разработку данных	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.

№ п/п	Тема и форма занятия	Наименование используемых активных и интерактивных образовательных технологий	
			талом, электронными ресурсами.
5.	Тема 9. Способы визуализации, интерпретации, проектная организация	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.
6.	Тема 1. Методы многомерного статистического анализа и анализа нечисловой информации. Тема 2. Технологии хранения и обработки Больших данных.	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.
7.	Тема 3. Языки программирования для анализа и обработки больших массивов данных. Тема 4. Базы данных.	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.
8.	Тема 5. Программные средства для обработки данных и системы Data Mining Тема 6. Платформы и фреймворки для Big Data	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.
9	Тема 7. Аналитика Big Data Тема 8. Введение в машинное обучение и разработку данных	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.
10	Тема 9. Способы визуализации, интерпретации, проектная организация	Л	Информационные и коммуникационные технологии
		ПЗ	Работа студентов с учебно-методическим порталом, электронными ресурсами.

## 6. Текущий контроль успеваемости и промежуточная аттестация по итогам освоения дисциплины

### 6.1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности

#### 1) Вопросы для устного опроса:

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

1. Какие методы многомерного статистического анализа используются для анализа данных?
2. Что такое факторный анализ и как он применяется в Big Data?
3. Какие методы кластерного анализа используются для обработки больших объемов данных?
4. Как проводится анализ нечисловой информации в Big Data?
5. Какие инструменты и технологии используются для анализа нечисловой информации в Big Data?
6. Какие преимущества имеет использование методов многомерного статистического анализа и анализа нечисловой информации в Big Data?

7. Какие сложности могут возникнуть при проведении многомерного статистического анализа и анализа нечисловой информации в Big Data?

Тема 2. Технологии хранения и обработки Больших данных.

1. Какие технологии используются для хранения Больших данных?
2. Что такое Hadoop и как он используется для обработки Больших данных?
3. Какие преимущества имеет использование NoSQL баз данных для хранения Больших данных?
4. Какие технологии используются для обработки потоков данных в режиме реального времени?
5. Как происходит интеграция различных источников данных в Big Data?
6. Какие инструменты используются для визуализации и анализа Больших данных?
7. Как обеспечивается безопасность и защита Больших данных в системах Big Data?

Тема 3. Языки программирования для анализа и обработки больших массивов данных.

1. Какие языки программирования используются для анализа и обработки Больших данных, например, Python, R, Java?
2. Что такое Apache Spark и как он используется для обработки Больших данных?
3. Какие инструменты используются для машинного обучения и анализа данных в Big Data, например, TensorFlow, Apache Mahout, Apache MXNet?
4. Какие инструменты используются для обработки текстовых данных в Big Data, например, Apache Lucene, Apache Solr?
5. Какие инструменты используются для обработки графовых данных в Big Data, например, Apache Giraph, Neo4j?
6. Какие инструменты используются для обработки временных рядов в Big Data, например, Apache Kafka, Apache Flink?
7. Какие инструменты используются для распределенного хранения и обработки Больших данных, например, Apache Cassandra, Apache HBase?

Тема 4. Базы данных.

1. Распределенные системы хранения данных, например, Hadoop Distributed File System (HDFS).
2. Обработка данных в реальном времени с помощью Apache Kafka и Apache Flink.
3. Машинное обучение и анализ данных с использованием TensorFlow, Apache Mahout и Apache MXNet.
4. Обработка текстовых данных с помощью Apache Lucene и Apache Solr.
5. Обработка графовых данных с помощью Apache Giraph и Neo4j.
6. Работа с временными рядами с помощью Apache Kafka и Apache Flink.
7. Базы данных NoSQL, такие как Apache Cassandra и Apache HBase.
8. Инструменты для визуализации данных, такие как Tableau и Power BI.
9. Безопасность и конфиденциальность данных в Big Data.
10. Распределенные вычисления и параллельная обработка данных.

Тема 5. Программные средства для обработки данных и системы Data Mining

1. Основы Data Mining и машинного обучения.
2. Алгоритмы классификации и кластеризации данных.
3. Извлечение признаков и предобработка данных.
4. Обработка естественного языка и анализ текстов.
5. Методы ассоциативного анализа и поиска аномалий.
6. Интеграция и анализ данных из различных источников.
7. Использование графовых баз данных для Data Mining.
8. Работа с большими объемами данных при помощи MapReduce и Hadoop.
9. Визуализация и интерпретация результатов Data Mining.
10. Этические и юридические аспекты использования Data Mining.

## Тема 6. Платформы и фреймворки для Big Data

1. Apache Hadoop и его экосистема: HDFS, MapReduce, YARN, Hive, Pig, Spark.
2. NoSQL базы данных для Big Data: Cassandra, MongoDB, Couchbase, Amazon DynamoDB.
3. Apache Kafka для обработки потоков данных.
4. Apache Storm для обработки потоков данных в режиме реального времени.
5. Apache Flink для обработки потоков данных и пакетных данных.
6. Apache Beam для создания и запуска пайплайнов обработки данных на различных платформах.
7. TensorFlow и PyTorch для машинного обучения и глубокого обучения.
8. Apache Zeppelin и Jupyter Notebook для интерактивной работы с данными и анализа результатов.
9. Apache Druid для быстрого анализа и запросов к большим объемам данных.
10. Kubernetes для управления кластерами и контейнерами с приложениями для Big Data.

## Тема 7. Аналитика Big Data

1. Основы статистики и анализа данных.
2. Методы машинного обучения и глубокого обучения.
3. Обработка и анализ текстовых данных.
4. Визуализация данных и создание дашбордов.
5. Большие данные и их влияние на бизнес-процессы.
6. Работа с неструктурированными данными, такими как изображения и видео.
7. Работа с данными в режиме реального времени.
8. Анализ социальных сетей и медиа-данных.
9. Системы рекомендаций и персонализация контента.
10. Безопасность данных и защита от кибератак.

## Тема 8. Введение в машинное обучение и разработку данных

1. Основные понятия и термины в машинном обучении.
2. Различные типы задач машинного обучения: классификация, регрессия, кластеризация и т.д.
3. Методы предобработки данных: очистка, преобразование, масштабирование и т.д.
4. Обучение с учителем и без учителя: примеры задач и методов решения.
5. Модели машинного обучения: линейные модели, деревья решений, нейронные сети и т.д.
6. Оценка качества моделей: метрики, кросс-валидация и т.д.
7. Техники улучшения моделей: отбор признаков, настройка параметров и т.д.
8. Применение машинного обучения в реальных проектах: от выбора модели до ее внедрения.
9. Использование библиотек и инструментов для машинного обучения: scikit-learn, TensorFlow, Keras и т.д.
10. Этические и правовые аспекты использования машинного обучения и обработки данных.

## Тема 9. Способы визуализации, интерпретации, проектная организация

1. Способы визуализации данных: графики, диаграммы, тепловые карты и т.д.
2. Интерпретация результатов моделирования: важность признаков, влияние параметров и т.д.
3. Организация проекта машинного обучения: сбор данных, постановка задачи, выбор модели и т.д.
4. Работа с большими объемами данных: параллельное вычисление, распределенные вычисления и т.д.
5. Применение машинного обучения в различных отраслях: медицина, финансы, производство и т.д.

6. Использование библиотек и инструментов для визуализации и интерпретации данных: Matplotlib, Seaborn, Plotly и т.д.
7. Работа в команде на проектах машинного обучения: разделение задач, контроль версий и т.д.
8. Обработка текстовых данных: предобработка, векторизация, классификация и т.д.
9. Применение машинного обучения для анализа изображений: распознавание объектов, сегментация и т.д.
10. Этические и правовые аспекты использования данных и моделей машинного обучения в различных сферах: конфиденциальность, дискриминация и т.д.

#### Тема 10. Решения на основе Big data

1. Что такое Big Data и какие проблемы она решает?
2. Какие технологии используются для обработки и анализа Big Data?
3. Какие преимущества имеют решения на основе Big Data по сравнению с традиционными подходами к анализу данных?
4. Какие типы данных можно обрабатывать с помощью решений на основе Big Data?
5. Какие методы машинного обучения используются для анализа Big Data?
6. Какие инструменты используются для хранения и обработки Big Data?
7. Какие проблемы могут возникнуть при работе с большими объемами данных и как их можно решить?
8. Какие примеры успешного применения решений на основе Big Data в различных отраслях?
9. Какие требования к аппаратному обеспечению и сетевой инфраструктуре необходимы для работы с Big Data?
10. Какие навыки и знания необходимы для работы с решениями на основе Big Data?

### ***2) Примеры заданий для практических работ***

Подробный перечень заданий для практических занятий представлен в оценочных материалах дисциплины.

### ***3) Перечень тем курсовых проектов:***

1. Анализ потребностей сельского хозяйства в Big Data технологиях.
2. Разработка системы мониторинга и прогнозирования урожайности на основе Big Data.
3. Использование Big Data для оптимизации производственных процессов в сельском хозяйстве.
4. Анализ данных о погодных условиях и их влиянии на урожайность с помощью Big Data.
5. Использование машинного обучения для оптимизации использования удобрений и семян в сельском хозяйстве.
6. Разработка системы управления ресурсами и энергосбережения на основе Big Data технологий.
7. Анализ рынка сельскохозяйственной продукции и прогнозирование спроса с помощью Big Data.
8. Использование Big Data для мониторинга здоровья животных и предотвращения заболеваний.
9. Разработка системы автоматического контроля качества продукции на основе Big Data технологий.

10. Использование машинного обучения для оптимизации логистики в сельском хозяйстве.

**4) Перечень вопросов, выносимых на экзамен:**

1. Что такое Big Data и какие преимущества они могут принести в сельском хозяйстве?
2. Какие методы и инструменты используются для обработки больших данных в сельском хозяйстве?
3. Какие задачи можно решить с помощью Big Data в сельском хозяйстве?
4. Какие проблемы могут возникнуть при работе с большими объемами данных в сельском хозяйстве и как их можно решить?
5. Какие примеры успешного применения Big Data технологий в сельском хозяйстве вы знаете?
6. Какие этапы включает разработка системы мониторинга и прогнозирования урожайности на основе Big Data?
7. Какие методы машинного обучения можно использовать для оптимизации производственных процессов в сельском хозяйстве?
8. Какие данные о погодных условиях необходимы для анализа их влияния на урожайность с помощью Big Data?
9. Какие принципы лежат в основе системы управления ресурсами и энергосбережения на основе Big Data технологий?
10. Какие методы машинного обучения можно использовать для оптимизации логистики в сельском хозяйстве?
11. Какие типы датчиков используются для сбора данных в сельском хозяйстве?
12. Какие алгоритмы используются для анализа данных в сельском хозяйстве?
13. Какие принципы лежат в основе системы автоматического управления растениеводством на основе Big Data технологий?
14. Какие данные необходимы для прогнозирования спроса на продукцию сельского хозяйства с помощью Big Data?
15. Какие методы машинного обучения можно использовать для оптимизации процессов хранения и переработки продукции сельского хозяйства?
16. Какие принципы лежат в основе системы контроля качества продукции сельского хозяйства на основе Big Data технологий?
17. Какие методы машинного обучения можно использовать для прогнозирования цен на продукцию сельского хозяйства?
18. Какие принципы лежат в основе системы управления здоровьем животных на основе Big Data технологий?
19. Какие методы машинного обучения можно использовать для оптимизации процессов удобрения и обработки почвы?
20. Какие принципы лежат в основе системы управления водными ресурсами на основе Big Data технологий?
21. Какие методы машинного обучения можно использовать для оптимизации процессов выращивания растений в закрытом грунте?



22. Какие принципы лежат в основе системы управления эффективностью использования топлива на основе Big Data технологий?
23. Какие методы машинного обучения можно использовать для оптимизации процессов управления скотом на пастбищах?
24. Какие принципы лежат в основе системы управления запасами на основе Big Data технологий?
25. Какие методы машинного обучения можно использовать для оптимизации процессов сбора и обработки данных в реальном времени?
26. Какие принципы лежат в основе системы управления эффективностью использования энергии на основе Big Data технологий?
27. Какие методы машинного обучения можно использовать для оптимизации процессов управления жизненным циклом продукции сельского хозяйства?
28. Какие принципы лежат в основе системы управления здоровьем растений на основе Big Data технологий?
29. Какие методы машинного обучения можно использовать для оптимизации процессов управления рисками в сельском хозяйстве?
30. Какие принципы лежат в основе системы управления использованием пестицидов на основе Big Data технологий?

## 6.2. Описание показателей и критериев контроля успеваемости,

Таблица 7

Шкала оценивания	Экзамен/ Зачет с оценкой
85-100	Отлично
70-84	Хорошо
60-69	Удовлетворительно
0-59	Неудовлетворительно

При использовании традиционной системы контроля и оценки успеваемости студентов должны быть представлены критерии выставления оценок по четырехбалльной системе «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

### Критерии оценивания результатов обучения

Таблица 8

Оценка	Критерии оценивания
Высокий уровень «5» (отлично)	оценку <b>«отлично»</b> заслуживает студент, освоивший знания, умения, компетенции и теоретический материал без пробелов; выполнивший все задания, предусмотренные учебным планом на высоком качественном уровне; практические навыки профессионального применения освоенных знаний сформированы. <b>Компетенции, закреплённые за дисциплиной, сформированы на уровне – высокий.</b>
Средний уровень «4» (хорошо)	оценку <b>«хорошо»</b> заслуживает студент, практически полностью освоивший знания, умения, компетенции и теоретический материал, учебные задания не оценены максимальным числом баллов, в основном сформировал практические навыки.

	<b>Компетенции</b> , закреплённые за дисциплиной, <b>сформированы на уровне – хороший (средний)</b> .
Пороговый уровень «3» (удовлетворительно)	оценку « <b>удовлетворительно</b> » заслуживает студент, частично с пробелами освоивший знания, умения, компетенции и теоретический материал, многие учебные задания либо не выполнил, либо они оценены числом баллов близким к минимальному, некоторые практические навыки не сформированы. <b>Компетенции</b> , закреплённые за дисциплиной, <b>сформированы на уровне – достаточный</b> .
Минимальный уровень «2» (неудовлетворительно)	оценку «неудовлетворительно» заслуживает студент, не освоивший знания, умения, компетенции и теоретический материал, учебные задания не выполнил, практические навыки не сформированы. <b>Компетенции</b> , закреплённые за дисциплиной, <b>не сформированы</b> .

## 7. Учебно-методическое и информационное обеспечение дисциплины

### 7.1 Основная литература

1. Москвитин, А.А. Данные, информация, знания: методология, теория, технологии : монография / А.А. Москвитин. — Санкт-Петербург : Лань, 2019. — 236 с. — ISBN 978-5-8114-3232-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/113937> — Режим доступа: для авториз. пользователей.
2. Макшанов, А.В. Технологии интеллектуального анализа данных : учебное пособие / А.В. Макшанов, А.Е. Журавлев. — 2-е изд., стер. — Санкт-Петербург : Лань, 2019. — 212 с. — ISBN 978-5-8114-4493-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/120063> — Режим доступа: для авториз. пользователей.
3. Рындина, С. В. Цифровая трансформация бизнеса: использование аналитики на основе больших данных : учебное пособие / С. В. Рындина. — Пенза : ПГУ, 2019. — 182 с. — ISBN 978-5-907262-04-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/162301> (дата обращения: 15.05.2022). — Режим доступа: для авториз. пользователей.

### 7.2 Дополнительная литература

1. Железнов, М. М. Методы и технологии обработки больших данных : учебно-методическое пособие / М. М. Железнов. — Москва : МИСИ – МГСУ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/145102> (дата обращения: 17.05.2022). — Режим доступа: для авториз. пользователей.
2. Методы и модели исследования сложных систем и обработки больших данных : монография / И. Ю. Парамонов, В. А. Смагин, Н. Е. Косых, А. Д.

Хомоненко ; под редакцией В. А. Смагина и А. Д. Хомоненко. — Санкт-Петербург : Лань, 2020. — 236 с. — ISBN 978-5-8114-4006-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/126938> (дата обращения: 17.05.2022). — Режим доступа: для авториз. пользователей.

3. Методы искусственного интеллекта в обработке данных и изображений : монография / А. Ю. Дёмин, А. К. Стоянов, В. Б. Немировский, В. А. Дорофеев. — Томск : ТПУ, 2016. — 130 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/106257> (дата обращения: 17.05.2022). — Режим доступа: для авториз. пользователей.

## 8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля)

1. [ANACONDA Documentation \(mathworks.com\)](https://docs.continuum.io/anaconda/) (открытый доступ)
2. Аналитическая платформа Polymatica | GETCRM +74957254376 ([polymatica-service.ru](http://polymatica-service.ru)) (открытый доступ)
3. Kaggle

## 9. Перечень программного обеспечения и информационных справочных систем (при необходимости)

1. [Google Colab](https://colab.research.google.com/)

Таблица 9

### Перечень программного обеспечения

№ п/п	Наименование раздела учебной дисциплины (модуля)	Наименование программы	Тип программы	Автор	Год разработки
1	Раздел 1 Технологии обработки больших данных	Anaconda	программирование	<a href="https://anaconda.com">anaconda.com</a>	2020

## 10. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Таблица 10

### Сведения об обеспеченности специализированными аудиториями, кабинетами, лабораториями

Наименование специальных* помещений и помещений для самостоятельной работы (№ учебного корпуса, № аудитории)	Оснащенность специальных помещений и помещений для самостоятельной работы**
1	2
Учебная аудитория для проведения занятий лекционного типа, групповых и индивидуальных консультаций (1 корпус, 110 аудитория)	проектор, экран настенный, компьютер
Компьютерный класс (1 корпус,	Персональные компьютеры 20 шт. с доступом к

201аудитория)	интернету, Парты 20 шт. Стулья 20 шт. Доска маркерная 1 шт.
ЦНБ им. Н.И. Железнова	Читальный зал (25 компьютеров)
Общежитие	Комната для самоподготовки

## 11. Методические рекомендации обучающимся по освоению дисциплины

Основными видами обучения студентов по дисциплине являются лекции, практические занятия в компьютерном классе и самостоятельная работа студентов.

Самостоятельная работа студентов по дисциплине «Технологии обработки больших данных» направлена на углубление и закрепление знаний, полученных на лекциях и практических занятиях, на развитие практических умений и включает такие виды работ, как:

- работа с лекционным материалом;
- работа с рекомендованной литературой при подготовке к практическим занятиям;
- подготовка к зачету.

При изучении дисциплины «Технологии обработки больших данных» используется рейтинговая система оценивания знаний студентов, которая позволяет реализовать непрерывную и комплексную систему оценивания учебных достижений студентов. Непрерывность означает, что текущие оценки не усредняются (как в традиционной технологии), а непрерывно складываются на протяжении семестра при изучении дисциплины. Комплексность означает учет всех форм учебной и самостоятельной работы студента в течение семестра.

Принципы рейтинга: непрерывный контроль (на каждом из аудиторных занятий) и получение более высокой оценки за работу, выполненную в срок. При проведении практических занятий предусмотрено широкое использование активных и интерактивных форм (разбор конкретных ситуаций, устный опрос, защита практических работ).

Бально–рейтинговая система повышает мотивацию студентов.

Промежуточным контролем по дисциплине является зачет.

В результате изучения дисциплины формируются знания и умения в области инструментальных средств, студенты получают опыт в проектирование информационных систем. Каждому студенту во время практических занятий предоставляется полная возможность быть индивидуальным пользователем компьютера, самостоятельно отрабатывать учебные вопросы и выполнять индивидуальные учебные задания преподавателя.

Основная рекомендация сводится к обеспечению равномерной активной работы студентов над дисциплиной в течение всего семестра: студенты должны прорабатывать курс прослушанных лекций, готовиться к выполнению и защите практических работ, а также выполнять задания, вынесенные на самостоятельную работу. Рекомендуется перед каждой лекцией просматривать содержание предстоящей лекции по учебнику и конспекту с тем, чтобы лучше воспринять материал лекции. Важно помнить, что ни одна дисциплина не может быть изучена в необходимом объеме только по конспектам. Для хорошего усвоения

курса нужна систематическая работа с учебной и научной литературой, а конспект может лишь облегчить понимание и усвоение материала.

В подготовке к занятиям по дисциплине студенты должны активно использовать дополнительную литературу, поскольку именно с ее помощью можно получить наиболее полное и верное представление о происходящих в стране и в мире процессах.

### **Виды и формы отработки пропущенных занятий**

Студент, пропустивший занятия обязан его отработать:

- лекцию отрабатывают путем устного ответа по пропущенной теме;
- практическое занятие путем выполнения практической работы, которая выполнялась на данном практическом занятии.

### **12. Методические рекомендации преподавателям по организации обучения по дисциплине**

В процессе обучения по дисциплине «Технологии обработки больших данных» используются лекционно-практические занятия, разбор конкретных ситуаций, организуется работа с методическими и справочными материалами, целесообразно применение современных технических средств обучения и информационных технологий. Освоение учебной дисциплины предполагает осмысление её разделов и тем на практических занятиях, в процессе которых студент должен закрепить и углубить теоретические знания.

Дисциплина «Технологии обработки больших данных» имеет прикладной характер, её теоретические положения и практические навыки могут быть использованы в будущей практической деятельности.

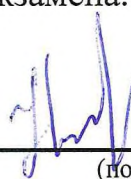
Промежуточный контроль – экзамен.

Рекомендуется определять сроки проведения контрольных мероприятий, максимальная оценка за каждое из них и правила перевода общего количества баллов, полученных при изучении дисциплины, в итоговый результат (экзамен).

Выполнение практических заданий является обязательным для всех обучающихся. Студенты, не выполнившие в полном объеме работы, предусмотренные учебным планом, не допускаются к сдаче экзамена.

**Программу разработал:**

Греченева А.В., к.т.н.



(подпись)

Быстренина И.Е., к.пед.н.



(подпись)

## РЕЦЕНЗИЯ

на рабочую программу дисциплины Б1.В.24 «Технологии обработки больших данных» ОПОП ВО по направлению 09.03.03 Прикладная информатика, направленности: Прикладная информатика в экономике, Системы искусственного интеллекта (квалификация выпускника – бакалавр)

Ашмарина Татьяна Игоревна, к.э.н., ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева, доцент (далее по тексту рецензент), проведена рецензия рабочей программы дисциплины Б1.В.25 «Технологии обработки больших данных» ОПОП ВО по направлению 09.03.03 - Прикладная информатика, направленности «Прикладная информатика в экономике», «Системы искусственного интеллекта» (бакалавриат) разработанной в ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», на кафедре прикладной информатики, разработчики – Греченева Анастасия Владимировна, к.т.н., Быстренина И.Е., к.пед.н..

Рассмотрев представленные на рецензирование материалы, рецензент пришел к следующим выводам:

1. Предъявленная рабочая программа дисциплины «Технологии обработки больших данных» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению 09.03.03 - Прикладная информатика. Программа содержит все основные разделы, соответствует требованиям к нормативно-методическим документам.

2. Представленная в Программе актуальность учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится к формируемой участниками образовательных отношений части учебного цикла – Б1.

3. Представленные в Программе цели дисциплины соответствуют требованиям ФГОС ВО направления 09.03.03 - Прикладная информатика .

4. В соответствии с Программой за дисциплиной «Технологии обработки больших данных» закреплено 2 компетенции. Дисциплина «Технологии обработки больших данных» и представленная Программа способна реализовать их в объявленных требованиях. Результаты обучения, представленные в Программе в категориях знать, уметь, владеть соответствуют специфике и содержанию дисциплины и демонстрируют возможность получения заявленных результатов.

5. Общая трудоёмкость дисциплины «Технологии обработки больших данных» составляет 6 зачётных единиц (108 часов/из них практическая подготовка 4).

6. Информация о взаимосвязи изучаемых дисциплин и вопросам исключения дублирования в содержании дисциплин соответствует действительности. Дисциплина «Технологии обработки больших данных» взаимосвязана с другими дисциплинами ОПОП ВО и Учебного плана по направлению 09.03.03 - Прикладная информатика и возможность дублирования в содержании отсутствует.

7. Представленная Программа предполагает использование современных образовательных технологий, используемые при реализации различных видов учебной работы. Формы образовательных технологий соответствуют специфике дисциплины.

8. Программа дисциплины «Технологии обработки больших данных» предполагает 8\_ занятий в интерактивной форме.

9. Виды, содержание и трудоёмкость самостоятельной работы студентов, представленные в Программе, соответствуют требованиям к подготовке выпускников, содержащимся во ФГОС ВО направления 09.03.03 - Прикладная информатика .

10. Представленные и описанные в Программе формы текущей оценки знаний (опрос, как в форме обсуждения отдельных вопросов и выступлений, а также контроль выполнения и проверка отчетности по практическим работам), соответствуют специфике дисциплины и требованиям к выпускникам.

Форма промежуточного контроля знаний студентов, предусмотренная Программой, осуществляется в форме зачета, что соответствует статусу дисциплины, как дисциплины

вариативной части учебного цикла – Б1 ФГОС ВО направления **09.03.03 - Прикладная информатика**.

11. Формы оценки знаний, представленные в Программе, соответствуют специфике дисциплины и требованиям к выпускникам.

Учебно-методическое обеспечение дисциплины представлено: основной литературой – 2 источника (базовый учебник), дополнительной литературой – 3 наименования, периодическими изданиями – 2 источника со ссылкой на электронные ресурсы, Интернет-ресурсы – 2 источника и соответствует требованиям ФГОС ВО направления **09.03.03 - Прикладная информатика**.

12. Материально-техническое обеспечение дисциплины соответствует специфике дисциплины **«Технологии обработки больших данных»** и обеспечивает использование современных образовательных, в том числе интерактивных методов обучения.

13. Методические рекомендации студентам и методические рекомендации преподавателям по организации обучения по дисциплине дают представление о специфике обучения по дисциплине **«Технологии обработки больших данных»**.

#### **ОБЩИЕ ВЫВОДЫ**

На основании проведенного рецензирования можно сделать заключение, что характер, структура и содержание рабочей программы дисциплины **«Технологии обработки больших данных»** ОПОП ВО по направлению **09.03.03 - Прикладная информатика**, направленность **«Системы искусственного интеллекта»** (квалификация выпускника – бакалавр), разработанная Греченовой А.В., к.т.н., Быстрениной И.Е, к.пед.н. соответствует требованиям ФГОС ВО, современным требованиям экономики, рынка труда и позволит при её реализации успешно обеспечить формирование заявленных компетенций.

Рецензент: Ашмарина Т.И., к.э.н., доцент  « 29 » августа 2022 г.  
(подпись)