

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Хоружий Людмила Ивановна  
Должность: Директор института экономики и управления АПК  
Дата подписания: 21.03.2024 11:26:01  
Уникальный программный ключ:  
1e90b132d9b04dce67585160b015dddf2cb1e6a9



УТВЕРЖДАЮ:

Директор Института экономики и  
управления АПК

Л.И. Хоружий

2023 г.

**Лист актуализации рабочей программы дисциплины**  
**Б1.В.ДВ.03 .02 Хранилища и системы интеллектуального анализа данных**

для подготовки бакалавров

Направление: 09.03.02 Информационные системы и технологии

Направленность: Большие данные и машинное обучение (Machine Learning & Big Data)

Форма обучения очная

Год начала подготовки: 2022

Курс 3

Семестр 5

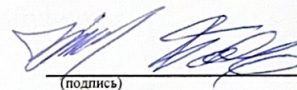
1. В оценочные материалы не вносятся изменения. Материалы актуализированы для 2023 г. начала подготовки.

2. Материалы будут распространены при организации учебного процесса на направленность (профиль): Большие данные и машинное обучение.

Разработчик (и): Демичев В.В., канд. экон. наук, доцент


Токарев В.С.

(ФИО, ученая степень, ученое звание)


  
(подпись)

«\_\_» \_\_\_\_\_ 2023 г.

Рабочая программа пересмотрена и одобрена на заседании кафедры статистики и кибернетики протокол № 11 от «28» августа 2023 г.

И.о. заведующего кафедрой статистики и кибернетики  А.В. Уколова

**Лист актуализации принят на хранение:**

И.о. заведующего кафедрой статистики и кибернетики  А.В. Уколова

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Хоружий, Владимир Иванович

Должность: директор института экономики и управления АПК

Дата подписания: 15.01.2023 19:17:06

Уникальный программный ключ:

1e90b132d9b04dce67585160b015dddf2cb1e6a9



**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 статистики и кибернетики



УТВЕРЖДАЮ:

Директор института экономики и управления АПК

*Л.И. Хоружий*  
Л.И. Хоружий  
2022 г.

## РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

**Б1.В.ДВ.03.02 Хранилища и системы интеллектуального анализа данных**

для подготовки бакалавров

ФГОС ВО

Направление: 09.03.02 «Информационные системы и технологии»

Направленность:

Большие данные и машинное обучение (Machine Learning & Big Data)

Курс 3

Семестр 5

Форма обучения очная

Год начала подготовки 2022

Москва, 2022



Разработчик (и): Харитонов А.Е., к.э.н., доцент  
(ФИО, ученая степень, ученое звание)

  
«26» 08 2022 г.

Рецензент: Коломеева Е.С., к.э.н.  
(ФИО, ученая степень, ученое звание)

(подпись)  
«26» 08 2022 г.

Программа составлена в соответствии с требованиями ФГОС ВО, профессионального стандарта и учебного плана по направлению подготовки 09.03.02 «Информационные системы и технологии».


Программа обсуждена на заседании кафедры статистики и кибернетики протокол № 11 от «26» августа 2022 г.

И.о.зав. кафедрой Уколова А.В., к.э.н., доцент  
(ФИО, ученая степень, ученое звание)

  
«26» 08 2022 г.

**Согласовано:**

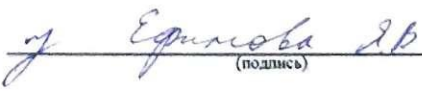
Председатель учебно-методической комиссии института экономики и управления АПК  
Корольков А.Ф., к.э.н., доцент  
(ФИО, ученая степень, ученое звание)

  
«26» 08 2022 г.

И.о.заведующего выпускающей кафедрой статистики и кибернетики  
Уколова А.В., к.э.н., доцент  
(ФИО, ученая степень, ученое звание)

  
«26» 08 2022 г.

Заведующий отделом комплектования ЦНБ

  
(подпись)

## СОДЕРЖАНИЕ

<b>АННОТАЦИЯ.....</b>	<b>4</b>
<b>1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....</b>	<b>6</b>
<b>2. МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ .....</b>	<b>6</b>
<b>3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ.....</b>	<b>6</b>
<b>4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ .....</b>	<b>7</b>
4.1 РАСПРЕДЕЛЕНИЕ ТРУДОЁМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ .....	7
ПО СЕМЕСТРАМ .....	7
4.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ.....	11
4.3 ПРАКТИЧЕСКИЕ ЗАДАНИЯ.....	13
<b>5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ .....</b>	<b>16</b>
<b>6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....</b>	<b>16</b>
6.1. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ, НЕОБХОДИМЫЕ ДЛЯ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ И НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ .....	16
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 .....	25
7.2 ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА.....	26
7.3 МЕТОДИЧЕСКИЕ УКАЗАНИЯ, РЕКОМЕНДАЦИИ И ДРУГИЕ МАТЕРИАЛЫ К ЗАНЯТИЯМ.....	27
<b>8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....</b>	<b>27</b>
<b>9. ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ.....</b>	<b>27</b>
<b>10. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ.....</b>	<b>28</b>
<b>11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ОБУЧАЮЩИМСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ.</b>	<b>29</b>
<b>12. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПРЕПОДАВАТЕЛЯМ ПО ОРГАНИЗАЦИИ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ.....</b>	<b>29</b>

## Аннотация

### рабочей программы учебной дисциплины

#### **Б1.В.ДВ.03.02 «Хранилища и системы интеллектуального анализа данных»**

**для подготовки бакалавров по направлению 09.03.02 «Информационные системы и технологии» по направленности Большие данные и машинное обучение (Machine Learning & Big Data) и**

**Цель освоения дисциплины:** Целью дисциплины «Хранилища и системы интеллектуального анализа данных» является освоение студентами теоретических и практических знаний и приобретение умений и навыков в области проведения аналитического исследования с применением технологий больших данных и выполнения работ по обеспечению функционирования баз данных и обеспечению их информационной безопасности, а также осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном языке.

**Место дисциплины в учебном плане:** включена в часть, формируемую участниками образовательных отношений (дисциплина по выбору) учебного плана по направлению подготовки 09.03.02 «Информационные системы и технологии».

**Требования к результатам освоения дисциплины:** в результате освоения дисциплины формируются следующие компетенции (индикаторы): УК-4 (УК-4.2, УК-4.3), ПКос-3 (ПКос-3.1, ПКос-3.3), ПКос-9 (ПКос-9.1, ПКос-9.2, ПКос-9.3).

#### **Краткое содержание дисциплины:**

Отличия баз данных от Хранилищ данных. Причины появления Хранилищ данных. Хранилище - надстройка над существующими базами данных. Отличительная особенность Хранилищ. Основные требования к данным, вводимым в Хранилище. Задачи построения Хранилища. Интеграция с различными источниками данных. Предварительная обработка данных для их дальнейшего использования в алгоритмах Data Mining. Заполнение пропусков, сэмплинг, редактирование выбросов. Квантование. Конечные классы. Разбиение на множества. Редактирование выбросов. Сглаживание. Группировка. Дата и время. Дополнение данных. Замена. Кросс-таблица. Объединение. Параметры полей. Разгруппировка. Свёртка столбцов. Скользящее окно. Слияние. Соединение. Сортировка. Фильтр строк. Дубликаты и противоречия во входных данных. Автокорреляция исходных переменных. Корреляционный анализ. Коэффициент корреляции Пирсона. Линейная регрессия. Разбиение на множества. Отбор факторов и защита от переобучения. F-тест. Коэффициент детерминации. Кластеризация. Центры кластеров. Автоопределение числа кластеров. Нормализация данных. Алгоритм CLOPE. Коэффициент отталкивания. алгоритм кластеризации EM. Медианная модификация. Модель ARIMAX. Среднеквадратическая ошибка на обучающем множестве. Минимальная достоверность правила. Минимальный лифт правила. Максимальное число следствий. Использование нейросети для задачи регрессии. Разбиение на множества. Метод вали-

дации. Настройка параметров Нейросети. Параметры обучения. Критерии останова. Стратегия оптимизации. Использование нейросети для задачи классификации.

**Общая трудоемкость дисциплины составляет:** 3 зачетные единицы (108 часов).

**Промежуточный контроль:** зачет с оценкой.

## **1. Цель освоения дисциплины**

Целью дисциплины «Хранилища и системы интеллектуального анализа данных» является освоение студентами теоретических и практических знаний и приобретение умений и навыков в области проведения аналитического исследования с применением технологий больших данных и выполнения работ по обеспечению функционирования баз данных и обеспечению их информационной безопасности, а также осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном языке.

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

Дисциплина «Хранилища и системы интеллектуального анализа данных» включена в часть, формируемую участниками образовательных отношений (дисциплина по выбору) учебного плана. Дисциплина «Хранилища и системы интеллектуального анализа данных» реализуется в соответствии с требованиями ФГОС ВО, профессионального стандарта, ОПОП ВО и Учебного плана по направлению 09.03.02 «Информационные системы и технологии».

Предшествующими курсами, на которых непосредственно базируется дисциплина «Хранилища и системы интеллектуального анализа данных» являются «Введение в компьютерные науки на иностранном языке», «Математический анализ», «Математическая статистика», «Теория вероятностей», «Алгоритмизация и программирование», «Основы науки о данных (Data Science)», «Анализ экономических данных с использованием современных информационных технологий на иностранном языке», «Анализ экономических данных с использованием современных информационных технологий».

Дисциплина «Хранилища и системы интеллектуального анализа данных» является основополагающей для изучения следующих дисциплин: «Администрирование информационных систем», «Многомерные статистические методы», «Методы искусственного интеллекта», «Большие данные».

Рабочая программа дисциплины «Хранилища и системы интеллектуального анализа 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, соотнесенных с планируемыми результатами освоения образовательной программы**

Образовательные результаты освоения дисциплины обучающимся, представлены в таблице 1.

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

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

Общая трудоёмкость дисциплины составляет 3 зач.ед. (108 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.	УК-4	Способен осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном(ых) языке(ах)	УК-4.2 Уметь: применять на практике деловую коммуникацию в устной и письменной формах, методы и навыки делового общения на русском и иностранном языках		применять на практике деловую коммуникацию в сфере работы с хранилищами данных и их анализов в устной и письменной формах на русском и иностранном языках	
2.			УК-4.3 Владеть: навыками чтения и перевода текстов на иностранном языке в профессиональном общении; навыками деловых коммуникаций в устной и письменной форме на русском и иностранном языках; методикой составления суждения в межличностном деловом общении на русском и иностранном языках			навыками деловых коммуникаций в устной и письменной форме на русском и иностранном языках
3.	ПКос-3	Способность выполнять работы по обеспечению функ-	ПКос-3.1 Знать: понятие «целостность	понятие «целостность баз данных»; причины		

		ционирования баз данных и обеспечению их информационной безопасности	баз данных»; причины нарушения целостности баз данных; способы достижения безопасности баз данных	нарушения целостности баз и хранилищ данных.		
4.			ПКос-3.3 Иметь навыки: сопровождения базы данных в оперативном и обеспечивающем режимах			сопровождения базы и хранилища данных в оперативном и обеспечивающем режимах
5.	ПКос-9	Способен проводить аналитическое исследование с применением технологий больших данных	ПКос-9.1 Знать: предметную область анализа, типы больших данных, источники и методы извлечения информации, теоретические и прикладные основы анализа, технологии хранения и обработки, современные методы и инструментальные средства анализа больших данных	типы данных, источники и методы извлечения информации, теоретические и прикладные основы интеллектуального анализа, технологии хранения и обработки, современные методы и инструментальные средства анализа больших данных		
6.			ПКос-9.2 Уметь: оценивать соответствие наборов данных задачам анализа больших данных; использовать инструментальные сред-		оценивать наборы данных методами интеллектуального анализа данных; использовать инструментальные средства для извлечения, преобразования, хранения и	

			ства для извлечения, преобразования, хранения и обработки данных из разнородных источников; разрабатывать и оценивать модели больших данных; автоматизировать процесс анализа больших данных; визуализировать результаты анализа больших данных		обработки данных из разнородных источников; визуализировать результаты анализа больших данных	
7.			ПКос-9.3 Иметь навыки: выбора источников данных, оценки соответствия набора данных предметной области и задачам аналитических работ; получения и фильтрации, извлечения, проверки, очистки, агрегации и разработки представления больших объемов данных из гетерогенных источников			выбора источников данных, оценки соответствия набора данных предметной области и задачам аналитических работ; получения и фильтрации, извлечения, проверки, очистки, агрегации и разработки представления больших объемов данных из гетерогенных источников

## Распределение трудоёмкости дисциплины по видам работ по семестрам

Вид учебной работы	Трудоёмкость, 5 семестр час. всего/*
<b>Общая трудоёмкость</b> дисциплины по учебному плану	<b>108</b>
<b>1. Контактная работа:</b>	<b>50,35</b>
<b>Аудиторная работа</b>	<b>50,35</b>
<i>в том числе:</i>	
<i>лекции (Л)</i>	16
<i>практические занятия (ПЗ)</i>	34/4
<i>контактная работа на промежуточном контроле (КРА)</i>	0,35
<b>2. Самостоятельная работа (СРС)</b>	<b>57,65</b>
<i>самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к практическим занятиям и т.д.)</i>	
<i>Подготовка к зачету (контроль)</i>	9
Вид промежуточного контроля:	зачет с оценкой

\* в том числе практическая подготовка.

## 4.2 Содержание дисциплины

## Тематический план учебной дисциплины

Наименование разделов и тем дисциплин (укрупнёно)	Всего	Аудиторная работа			Внеаудиторная работа СР
		Л	ПЗ всего /*	ПКР	
Тема 1 Технология хранения данных при принятии решений	18	2	6/2		10
Тема 2 Предобработка и консолидация данных	11	1	4/2		6
Тема 3 Трансформация данных	7	1	2		4
Тема 4 Предварительный анализ данных	10	2	4		4
Тема 5 Регрессия	11,65	2	4		5,65
Тема 6 Кластеризация	16	2	6		8
Тема 7 ARIMAX	10	2	2		6
Тема 8 Ассоциативные правила	12	2	2		8
Тема 9 Нейросеть	12	2	4		6
Контактная работа на промежуточном контроле (КРА)	0,35			0,35	
<b>Всего за 5 семестр</b>	<b>108</b>	<b>16</b>	<b>34/4</b>	<b>0,35</b>	<b>57,65</b>
<b>Итого по дисциплине</b>	<b>108</b>	<b>16</b>	<b>34/4</b>	<b>0,35</b>	<b>57,65</b>

\* в том числе практическая подготовка

## Тема 1 Технология хранения данных при принятии решений

Отличия баз данных от Хранилищ данных. Причины появления Хранилищ данных. Хранилище - надстройка над существующими базами данных. Отличительная особенность Хранилищ. Основные требования к данным, вводимым в Хранилище. Задачи построения Хранилища. Интеграция с различными источниками данных. SOAP-запрос. REST-запрос. Выполнение программы. Извлечение XML.

## **Тема 2 Предобработка и консолидация данных**

Предварительная обработка данных для их дальнейшего использования в алгоритмах Data Mining. Заполнение пропусков, сэмплинг. Квантование. Конечные классы. Разбиение на множества. Редактирование выбросов. Сглаживание.

## **Тема 3 Трансформация данных**

Набор компонентов для первоначальной подготовки и простой обработки исходных наборов данных. Группировка. Дата и время. Дополнение данных. Замена. Кросс-таблица. Объединение. Параметры полей. Разгруппировка. Свёртка столбцов. Скользящее окно. Слияние. Соединение. Сортировка. Фильтр строк.

## **Тема 4 Предварительный анализ данных**

Дубликаты и противоречия во входных данных. Автокорреляция исходных переменных. Корреляционный анализ. Коэффициент корреляции Пирсона. Коэффициент Тау-в Кендалла. Экстремум взаимокорреляционной функции. Коэффициент корреляции Спирмена.

## **Тема 5 Линейная регрессия**

Линейная регрессия. Настройки нормализации. Разбиение на множества. Настройка линейной регрессии. Детальные настройки. Отбор факторов и защита от переобучения. F-тест. Коэффициент детерминации. Скорректированный коэффициент детерминации. Информационный критерий Акаике. Информационный критерий Байеса. Информационный критерий Ханнана-Квина. Порог значимости при добавлении фактора. Порог значимости при исключении фактора.

## **Тема 6 Кластеризация**

Кластеризация. Центры кластеров. Автоопределение числа кластеров. Нормализация данных. Алгоритм CLOPE. Коэффициент отталкивания. алгоритм кластеризации EM. Медианная модификация.

## **Тема 7 ARIMAX**

Модель ARIMAX для анализа временных рядов как интегрированная авторегрессия, скользящее среднее с возможностью учета дополнительных внешних факторов. Порядок сезонной AR части. Порядок сезонного интегрирования. Порядок сезонной MA части. Период сезонной составляющей. Среднеквадратическая ошибка на обучающем множестве. Средняя абсолютная ошибка на



обучающем множестве. Средняя относительная ошибка на обучающем множестве. Коэффициент детерминации. Скорректированный коэффициент детерминации.

### Тема 8 Ассоциативные правила

Ассоциативные правила в транзакционных данных. Минимальная поддержка. Минимальная достоверность правила. Минимальный лифт правила. Максимальное число следствий. Понятия и методы выявления закономерностей в интеллектуальном анализе данных. Использование выявленных закономерностей для предсказания неизвестных значений (прогностическое моделирование). Анализ исключений, предназначенный для выявления и толкования аномалий в найденных закономерностях.

### Тема 9 Нейросеть

Использование нейросети для задачи регрессии. Разбиение на множества. Метод валидации. Настройка параметров Нейросети. Параметры обучения. Критерии останова. Стратегия оптимизации. Использование нейросети для задачи классификации.

## 4.3 Лекции /практические занятия

Таблица 4

### Содержание лекции /практические занятия и контрольные мероприятия

№ п/п	№ раздела	№ и название практических занятий	Формируемые компетенции (индикаторы)	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
1	Тема 1 Технология хранения данных при принятии решений	Лекция 1. Технология хранения данных при принятии решений	ПКос-3.1 ПКос-9.1 УК-4.2		2
		Практическая работа № 1. Принципы работы с хранилищами данных	ПКос-3.3 ПКос-9.2 ПКос-9.3 УК-4.2 УК-4.3	Защита работы	2
		Практическая работа № 2. Интеграция с различными базами данных	ПКос-3.3 ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2/2
		Практическая работа № 3. Импорт данных в различных форматах	ПКос-3.3 ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
2	Тема 2 Предобработка и консолидация данных	Лекция 2 Предобработка и консолидация данных	ПКос-3.1 ПКос-9.1 УК-4.2		1
		Практическая работа № 4. Под-	ПКос-9.2	Защита работы	4/2

№ п/п	№ раздела	№ и название практических занятий	Формируемые компетенции (индикаторы)	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
		готовка данных для анализа	ПКос-9.3 УК-4.2 ПКос-3.3		
3	Тема 3 Трансформация данных	Лекция 3 Трансформация данных	ПКос-3.1 ПКос-9.1 УК-4.2		1
		Практическая работа № 5. Сортировка, объединение и разделение наборов данных	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	
4	Тема 4 Предварительный анализ данных	Лекция 4 Методы анализа данных	ПКос-9.1 УК-4.2		2
		Практическая работа № 6. Выявление дубликатов и противоречий во входных данных	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
		Практическая работа № 7. Предварительный анализ данных	ПКос-9.2 ПКос-9.3 УК-4.2 УК-4.3	Защита работы	2
5	Тема 5 Регрессия	Лекция 5 Регрессионный анализ	ПКос-9.1 УК-4.2		2
		Практическая работа № 8. Парная линейная регрессия	ПКос-9.2 ПКос-9.3 УК-4.2 УК-4.3	Защита работы	2
		Практическая работа № 9. Множественная линейная регрессия	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
6	Тема 6 Кластеризация	Лекция 6 Кластерный анализ	ПКос-9.1 УК-4.2		2
		Практическая работа № 10. Алгоритм CLOPE.	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
		Практическая работа № 11. Алгоритм кластеризации EM	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
7	Тема 7 ARIMAX	Лекция 7 Метод ARIMAX	ПКос-9.1 УК-4.2		2
		Практическая работа № 12. Анализ временных рядов методом ARIMAX	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
8	Тема 8 Ассоциативные правила	Лекция 8 Ассоциативные правила	ПКос-9.1 УК-4.2		2
		Практическая работа № 13 Поиск ассоциативных правил	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2

№ п/п	№ раздела	№ и название практических занятий	Формируемые компетенции (индикаторы)	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
			УК-4.3		
9	Тема 9 Нейросеть	Лекция 8 Нейронные сети	ПКос-9.1 УК-4.2		2
		Практическая работа № 14 Применение нейронной сети для регрессии	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
		Итоговый тест по дисциплине	ПКос-3.1 ПКос-3.3 ПКос-9.1 ПКос-9.2 ПКос-9.3 УК-4.2	Тестирование	2

Таблица 5

**Перечень вопросов для самостоятельного изучения дисциплины**

№ п/п	№ раздела и темы	Перечень рассматриваемых вопросов для самостоятельного изучения
1.	Тема 1 Технология хранения данных при принятии решений	Отличительная особенность Хранилищ. Основные требования к данным, вводимым в Хранилище. Задачи построения Хранилища. (УК-4.2, УК-4.3, ПКос-3.1, ПКос-3.3)
2.	Тема 2 Предобработка и консолидация данных	Квантование. Конечные классы. Сглаживание. (УК-4.2, УК-4.3, ПКос-3.1, ПКос-3.3)
3	Тема 3 Трансформация данных	Дополнение данных. Замена. Кросс-таблица. Объединение. Параметры полей. Разгруппировка. Свёртка столбцов. Скользящее окно. Слияние. Соединение. (УК-4.2, УК-4.3, ПКос-3.1, ПКос-3.3, ПКос-9.1, ПКос-9.2)
4	Тема 4 Предварительный анализ данных	Автокорреляция исходных переменных. Коэффициент Тау-b Кендалла. Экстремум взаимокорреляционной функции. Коэффициент корреляции Спирмена. (УК-4.2, УК-4.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)
5	Тема 5 Регрессия	Отбор факторов и защита от переобучения. Информационный критерий Акаике. Информационный критерий Байеса. Информационный критерий Ханнана-Квина. Порог значимости при добавлении фактора. Порог значимости при исключении фактора. (УК-4.2, УК-4.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)
6	Тема 6 Кластеризация	Коэффициент отталкивания. Медианная модификация. (УК-4.2, УК-4.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)
9	Тема 7 ARIMAX	Порядок сезонной AR части. Порядок сезонного интегрирования. Порядок сезонной MA части. Период сезонной составляющей.

№ п/п	№ раздела и темы	Перечень рассматриваемых вопросов для самостоятельного изучения
		(УК-4.2, УК-4.3, ПКос-3.1, ПКос-3.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)
10	Тема 8 Ассоциативные правила	Анализ исключений, предназначенный для выявления и толкования аномалий в найденных закономерностях. (УК-4.2, УК-4.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)
11	Тема 9 Нейросеть	Метод валидации. Настройка параметров Нейросети. Параметры обучения. Критерии останова. Стратегия оптимизации. (УК-4.2, УК-4.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)

## 5. Образовательные технологии

Таблица 6

### Применение активных и интерактивных образовательных технологий

№ п/п	Тема и форма занятия	Наименование используемых активных и интерактивных образовательных технологий	
	Лекция 6 Кластерный анализ	Л	Лекция-визуализация
1.	Практическая работа № 5. Сортировка, объединение и разделение наборов данных	ПЗ	Деловая игра

## 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

#### Пример теста

#### 1. Что такое интеллектуальный анализ данных?

- a) извлечение информации из необработанных данных
- b) хранение и обработка информации в базах данных
- c) визуализация данных
- d) статистический метод обработки данных

#### 2. Какая из концепций является базисной в технологии интеллектуального анализа данных?

- a) концепция естествознания
- b) концепция управления
- c) концепция закономерностей
- d) концепция формирования

**3. Что является результатом представления знаний в интеллектуальном анализе данных?**

- a) закономерность
- b) проект
- c) модель
- d) тренд

**4. Задача классификации – это \_\_\_\_\_ задача.**

- a) описательная
- b) предсказательная
- c) качественная
- d) количественная

**5. Задача кластеризации – это \_\_\_\_\_ задача.**

- a) описательная
- b) предсказательная
- c) качественная
- d) количественная

**6. Единицы наблюдения, которые значительно отличаются от большинства других единиц в наборе данных:**

- a) транзакция
- b) порядковое число
- c) интервалы
- d) резко выделяющиеся значения

**7. Набор конкретных примеров с известным исходом:**

- a) клиент-сервер
- b) классификатор
- c) учебный набор
- d) интеллектуальный анализ данных

**8. Преобразование данных включает в себя...**

- a) разделение данных из одного источника на несколько источников данных
- b) процесс изменения данных с обобщенного уровня на более детализированный
- c) объединение данных из одного источника с другими источниками данных
- d) процесс изменения данных с детализированного уровня на более обобщенный

**9. \_\_\_\_\_ - полезный метод обнаружения закономерностей в начале процесса интеллектуального анализа данных.**

- a) расчёт меры расстояния
- b) дерево принятия решений
- c) ассоциативные правила
- d) приёмы визуализации

**10. \_\_\_\_\_ - класс моделей, принцип которых основан на аналогии с работой человеческого мозга:**

- a) нейронные сети
- b) кластеры
- c) дерево принятия решений
- d) правило классификации



- 12. Что из перечисленного является математическим уравнением, связывающим переменные  $x$  и  $y$ ?**
- a) регрессия
  - b) интерполяция
  - c) кластеризация
  - d) экстраполяция
- 13. Уравнение вида  $y = a + bx$ :**
- a) полиномиальное уравнение
  - b) линейная регрессия
  - c) регрессия
  - d) интерполяция
- 14. Какой из этих показателей является измерением качества модели регрессии?**
- a) средняя арифметическая
  - b) дисперсия
  - c) стандартное отклонение
  - d) коэффициент детерминации
- 15. Коэффициент корреляции всегда лежит между значениями...**
- a) 0 и 1
  - b) -1 и 1
  - c) -1 и 0
  - d) 0 и 2
- 16. Изучение взаимосвязи между несколькими переменными – это задача...**
- a) парной регрессии
  - b) множественной регрессии
  - c) дерева принятия решения
  - d) моделирования
- 17. Ожидаемое значение  $y$ , когда  $X = 0$ , равно:**
- a) коэффициенту полной регрессии
  - b) условному началу
  - c) коэффициенту корреляции
  - d) коэффициенту детерминации
- 18. Собранные в разные моменты времени значения каких-либо параметров - это...**
- a) панельные данные
  - b) пространственные данные
  - c) данные временного ряда
  - d) ни один из вариантов
- 19. Изучение взаимосвязи между двумя переменными возможно благодаря...**
- a) двумерному графику
  - b) гистограммам
  - c) графикам временных рядов
  - d) ни один из вариантов
- 20. Главная цель поиска ассоциативных правил заключается в том, чтобы...**

- a) создать правила классификации
- b) проверить достоверность регрессионной модели
- c) определить главную идею той или иной текстовой информации
- d) выявить закономерности между связанными событиями в базах данных

**21. Если несколько событий связаны друг с другом, то это...**

- a) ассоциация
- b) последовательность
- c) классификация
- d) кластеризация

**22. Отношение транзакций, которые имеют набор  $F$  ( $D_F$ ) к общему количеству транзакций ( $D$ ) называется...**

- a) весомым уровнем набора  $F$
- b) моделью набора  $F$
- c) кластером набора  $F$
- d) уровнем поддержки набора  $F$

**23. Набор предметов ( $F$ ) называется частым, когда...**

- a)  $\text{Supp}(F) < \text{Supp}(\min)$
- b)  $\text{Supp}(F) > \text{Supp}(\min)$
- c)  $\text{Supp}(F) = \text{Supp}(\min)$
- d) ни один из вариантов

**24. Объекты некоторого набора предметов, которые подвергаются анализу, называются...**

- a) пропорции
- b) центроиды
- c) транзакции
- d) модели

**25. Что является последовательностью в поиске ассоциативных правил?**

- a) конечные действительные числа
- b) ранжированный ряд
- c) упорядоченное множество некоторых объектов
- d) ни один из вариантов

**26. Что из перечисленного не является оценкой полезности ассоциативных правил?**

- a) качество
- b) поддержка
- c) достоверность
- d) улучшение

**27. Поддержка любого набора объектов не может превышать минимальной поддержки любого из его подмножеств. Это основное свойство...**

- a) метода Naive Bayes
- b) алгоритма Apriori
- c) дерева принятия решений
- d) кластеризации

**28. Что из перечисленного не является методом пересчёта расстояний между кластерами?**

- a) расстояние между ближайшими соседями

- b) расстояние между дальними соседями
- c) метод медиан
- d) модальный метод

**29. Что такое кластеры?**

- a) данные временных рядов
- b) однородные группы объектов
- c) разнородные группы объектов
- d) статистическая совокупность

**30. Одной из мер близости, используемой в кластеризации, является...**

- a) расстояние Чебышева
- b) Евклидово расстояние
- c) оба варианта верны
- d) ни один из вариантов не является верным

Вопросы к зачету с оценкой

1. Какие уровни СУБД характерны в настоящее время для крупных организаций?
2. Охарактеризуйте применимость реляционных СУБД в системах ППР.
3. Охарактеризуйте понятие «хранилище данных».
4. Какие особенности имеют ненормализованные базы данных для моделирования многомерности?
5. Охарактеризуйте схему «Звезда».
6. Охарактеризуйте схему «Снежинка».
7. Что такое многомерные СУБД?
8. Приведите сравнительные характеристики многомерных и реляционных СУБД.
9. Какие операции предусматриваются для работы с измерениями в многомерных моделях БД?
10. Охарактеризуйте понятие «витрины данных».
11. Какова классификация интеллектуальных систем?
12. Что такое методы ИАД и каково их назначение?
13. Каковы особенности методов ИАД?
14. Охарактеризуйте области применения методов ИАД.
15. Каковы этапы исследований методами ИАД?
16. В чем заключается этап предварительной обработки данных при использовании ИАД?
17. Охарактеризуйте метод «Деревьев решений».
18. Охарактеризуйте особенности регрессионного анализа в методах ИАД.
19. Какие разновидности регрессионного метода ИАД Вы знаете?
20. Охарактеризуйте модели временных рядов с запаздываниями.
21. Охарактеризуйте метод «Ближайшего соседа».
22. Охарактеризуйте метод поиска правила.
23. Охарактеризуйте метод кластеризации.
24. Охарактеризуйте метод классификации.

25. Охарактеризуйте метод дискриминации.

26. Какая разница в целях и алгоритмах статистического и интеллектуального подходов.

### Пример работ

Практическая работа № 1. Принципы работы с хранилищами данных  
На основе исходных данных спроектировать хранилище данных. Установить связь между измерениями, определить факты, отражающие сущность события, соответствующие измерениям. Изобразить схему «снежинка» и загрузить данные используя аналитическую платформу Loginom. Представить данные в виде многомерных кубов, называемых также OLAP-кубами или гиперкубами. По итогам выполнения работы предоставить отчет с выводами.

*Вопросы к защите:*

1. Что Вы понимаете под хранилищем данных?
2. Что Вы понимаете под базой данных?
3. Основные отличия баз данных от хранилищ.
4. Модели организации баз данных
5. Модели организации хранилищ данных
6. Что такое OLAP-кубам?

Практическая работа № 2. Интеграция с различными базами данных

Подключить Loginom к имеющимся в исходных данных базам данным. Перед тем как использовать в сценариях узлы взаимодействия с SOAP-сервисами/REST-сервисами и импорта/экспорта из бизнес приложений, хранилищ данных, баз данных или иных источников, в пакете необходимо создать подключение к источнику данных. Подключение содержит в себе все необходимые параметры для соединения с источником данных, такие как пароль/логин, расположение источника и др. Созданное подключение в дальнейшем может использоваться в пакете многократно. По итогам выполнения работы предоставить отчет с выводами.

*Вопросы к защите:*

1. С какими базами данных можно настроить подключение Loginom
2. Какие возможности есть в меню редактирования подключения?
3. Каким образом осуществляется навигация по подключениям?

Практическая работа № 3. Импорт данных в различных форматах

Требуется импортировать исходные данные в Loginom в виде текстовых файлов, электронных таблиц, баз данных и других. Настроить параметры загрузки. Определить нужный формат импортируемых данных. По итогам выполнения работы предоставить отчет с выводами.

*Вопросы к защите:*

1. Какие типы файлов можно импортировать в Loginom?
2. Какие возможности есть в меню редактирования импорта?
3. В чем особенность импорта электронных таблиц?

Practical lesson № 3. Import data in various formats

#### Практическая работа № 4. Подготовка данных для анализа

По результатам загрузки данных прошлого занятия настроить автоматическое заполнение пропущенных значений в наборах данных. Для каждого столбца исходного набора данных пользователь может выбрать наиболее подходящий метод заполнения пропусков. Пропусками считаются Null-значения. Настроить программу для автоматической корректировки выбросов и экстремальных значений в наборах данных. Для каждого поля исходного набора данных критерии определения выбросов и экстремальных значений задаются пользователем с помощью указания допустимого стандартного отклонения или интерквартильного размаха. Осуществить разбиение данных на обучающую и тестовую выборки. По итогам выполнения работы предоставить отчет с выводами.

*Вопросы к защите:*

1. Методы заполнения пропущенных данных.
2. Что понимается под выбросами?
3. Методы определения выбросов.
4. Методы разбиения множества на выборки

#### Практическая работа № 5. Сортировка, объединение и разделение наборов данных

По данным прошлого занятия провести группировку данных, разгруппировку, замену данных, слияние и свертку столбцов. Построить кросс-таблицу. Настроить параметры полей. Осуществить сортировку и фильтрацию строк. По итогам выполнения работы предоставить отчет с выводами.

*Вопросы к защите:*

1. Что Вы понимаете под группировкой данных?
2. Для чего используется замена данных?
3. Способы слияния и свертки столбцов?
4. Что такое кросс-таблица?

#### Практическая работа № 6. Выявление дубликатов и противоречий во входных данных

По данным прошлого занятия провести анализ исходных данных на наличие дублирующих, противоречивых записей. По итогам выполнения работы предоставить отчет с выводами.

*Вопросы к защите:*

1. Что Вы понимаете под противоречивыми записями?
2. Для чего выявлять дубликаты в исходных данных?
3. Способы появления противоречивых и дублирующих данных?
4. Методы выявления противоречивых и дублирующих данных.

#### Практическая работа № 7. Предварительный анализ данных

По данным прошлого занятия провести анализ данных основе корреляционного анализа. Построить матрицу парных коэффициентов корреляции. Определить взаимосвязи между признаками. Проверить исходные данные на нали-



чие автокорреляции. По итогам выполнения работы предоставить отчет с выводами.

*Вопросы к защите:*

1. Что такое корреляция?
2. Что показывает коэффициент корреляции?
3. Методы отбора факторов для построения регрессионных моделей?
4. Что такое автокорреляция?
5. Способы выявления автокорреляции.
6. Способы устранения автокорреляции.

#### Практическая работа № 8. Парная линейная регрессия

На основе результатов прошлого занятия построить уравнение парной линейной регрессии. Сделать прогноз по уравнению. Определить качество построенной модели по тестовым данным. Сравнить метрики качества моделей без обучения и с обучением. Оформить отчет с выводами.

*Вопросы к защите:*

1. Назовите цель и задачи построения парной линейной модели регрессии.
2. Дайте интерпретацию коэффициента регрессии и условного начала в уравнении парной линейной регрессии.
3. Как проводить прогноз в системе LogiDom по парной линейной модели?
4. Особенности построения парной линейной регрессии в системе LogiDom.

#### Практическая работа № 9. Множественная линейная регрессия

На основе результатов занятия №7 построить множественную регрессионную модель регрессии. Построить прогноз, сделать выводы. Определить качество построенной модели по тестовым данным. Сравнить метрики качества моделей без обучения и с обучением. В результате выполнения работы необходимо оформить отчет с выводами.

*Вопросы к защите:*

1. Как проводить прогноз в системе LogiDom по множественной линейной модели.
  2. Стандартизированные коэффициенты множественной линейной регрессии. Интерпретация, особенности расчета.
  3. Способы отбора факторов в уравнение множественной линейной регрессии
3. Methods for selecting factors in the equation of multiple linear regression

#### Практическая работа № 10. Алгоритм CLOPE.

Загрузить исходные данные о посещениях веб-ресурсов в LogiDom. Применить к ним алгоритм CLOPE. Задача состоит в получении такого разбиения всего множества транзакций, чтобы похожие транзакции оказались в одном кластере, а отличающиеся друг от друга - в разных кластерах. В результате выполнения работы необходимо оформить отчет с выводами.

*Вопросы к защите:*

1. Что такое кластерный анализ?
2. В чем особенность алгоритма CLOPE?

### 3. Интерпретация результатов алгоритма CLOPE.

#### Практическая работа № 11. Алгоритм кластеризации EM

Загрузить исходные данные в Loginot. Применить к ним алгоритм кластеризации EM. Определить оптимальное число кластеров. Определить вероятность отнесения объекта к каждому кластеру. В результате выполнения работы необходимо оформить отчет с выводами.

*Вопросы к защите:*

1. Что такое кластерный анализ?
2. В чем особенность алгоритма кластеризации EM?
3. Интерпретация результатов алгоритма кластеризации EM.

#### Практическая работа № 12. Анализ временных рядов методом ARIMAX

На сайте <https://rosstat.gov.ru/> выбрать данные с 1990 года и применить к ним метод ARIMAX для прогнозирования. Оценить качество построенной модели. Оценить качества прогноза. Осуществить прогноз по ретроспективным данным и сравнить с результатами работы модели. В результате выполнения работы нужно оформить отчет с выводами.

*Вопросы к защите:*

1. Назовите методы, применяемые для выравнивания рядов динамики
2. В чем особенность метода ARIMAX?
3. Определение параметров метода ARIMAX.

#### Практическая работа № 13 Поиск ассоциативных правил

По набору данных о покупках клиентов определить наиболее часто встречающиеся товары и наибольшую вероятность совместных покупок товаров. Какие товары не встречаются в одном чеке? Оформить отчет с выводами.

*Вопросы к защите:*

1. Цель поиска ассоциативных правил
2. Примеры применения метода поиска ассоциативных правил
3. Особенности применения метода поиска ассоциативных правил
4. Качество выявленных взаимосвязей
5. Расчет и интерпретация ошибок

#### Практическая работа № 14 Применение нейронной сети для регрессии

Построить модель регрессии на основе искусственных нейронных сетей. Сравнить результаты классификации при разном уровне скрытых слоев. Сравнить качество классификации с предыдущими методами. Сравнить результаты построенных моделей. Оформить отчет с выводами.

1. Суть нейронных сетей
2. Обучение модели
3. Суть искусственного нейрона
4. Определение количества слоев
5. Определение числа нейронов

## **6.2. Описание показателей и критериев контроля успеваемости, описание шкал оценивания**

Текущий контроль знаний, умений и навыков проводится в форме тестирования и теоретическими вопросами. Оценка работ проводится по стобалльной шкале. Индивидуальные задачи, выполняемые каждым студентом на практике оцениваются по итогам защиты по аналогичной шкале. Ликвидация студентами текущих задолженностей производится также в форме выполнения индивидуальной задачи по соответствующей теме и дальнейшей ее защиты преподавателю кафедры с оценкой по стобалльной шкале.

Для получения зачета с оценкой необходимо набрать более 60%. Вид промежуточного контроля по данному направлению – зачет с оценкой.

Градации оценок:

0 – 60% - «неудовлетворительно»;

60 – 75 – «удовлетворительно»;

75 – 85 – «хорошо»;

85 – 100 – «отлично»

Формы контроля: тестовый контроль, индивидуальное собеседование, защита выполнения практического задания по индивидуальному варианту. В итоговую сумму баллов входят результаты всех контролируемых видов вашей деятельности – посещение занятий, выполнение заданий, прохождение тестов, активность на лабораторных занятиях и т.п.

В итоговый рейтинг входит: 30% - результат выполнения контрольных мероприятий (тест, самостоятельные работы и др.), 60% - баллы за сданные индивидуальные работы и 10% - посещение занятий.

При изучении каждого раздела дисциплины проводится промежуточный контроль знаний с целью проверки и коррекции хода освоения теоретического материала и практических умений и навыков.

## **7. Учебно-методическое и информационное обеспечение дисциплины**

### **7.1 Основная литература**

1. Митина, О. А. Технологии организации, обработки и хранения статистических данных : учебное пособие / О. А. Митина, И. А. Юрченков. — Москва : РТУ МИРЭА, 2019. — 163 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/171511> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.

2. Прокопенко, Н. Ю. Аналитические информационные системы поддержки принятия решений : учебное пособие / Н. Ю. Прокопенко. — Нижний Новгород : ННГАСУ, 2020. — 142 с. — ISBN 978-5-528-00395-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/164866> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.



3. Скляр, А. Я. Технология хранения и интерактивных обработки данных : учебное пособие / А. Я. Скляр. — Москва : РТУ МИРЭА, 2020. — 69 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/163914> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.

4. Орешков, В. И. Хранилища данных и OLAP-технологии : учебное пособие / В. И. Орешков. — Рязань : РГРТУ, 2017. — 64 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/167981> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.

5. Data Science / Francesco Palumbo, Angela Montanari, Maurizio Vichi. Springer International Publishing AG, 2017 – Текст : электронный // Springer: электронно-библиотечная система. URL: <https://link.springer.com/book/10.1007/978-3-319-55723-6#editorsandaffiliations> (дата обращения: 27.11.2022).

## 7.2 Дополнительная литература

1. Бессмертный, И. А. Интеллектуальные системы : учебник и практикум для вузов / И. А. Бессмертный, А. Б. Нугуманова, А. В. Платонов. — Москва : Издательство Юрайт, 2022. — 243 с. — (Высшее образование). — ISBN 978-5-534-01042-8. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/490020> (дата обращения: 27.11.2022).

2. Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практикум для среднего профессионального образования / Л. А. Станкевич. — Москва : Издательство Юрайт, 2022. — 397 с. — (Профессиональное образование). — ISBN 978-5-534-11659-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/495988> (дата обращения: 27.11.2022).

3. New Advances in Statistics and Data Science / Ding-Geng, Chen Zhezhen, Jin Gang, Li Yi, Li Aiyi, Liu Yichuan, Zhao. Springer International Publishing AG, 2017 – Текст : электронный // Springer: электронно-библиотечная система. URL: <https://link.springer.com/book/10.1007/978-3-319-69416-0#editorsandaffiliations> (дата обращения: 27.11.2022).

4. Точилкина, Т. Е. Хранилища данных и средства бизнес-аналитики : учебное пособие / Т. Е. Точилкина, А. А. Громова. — Москва : Финансовый университет, 2017. — 161 с. — ISBN 978-5-7942-1387-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/208367> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.

### 7.3 Методические указания, рекомендации и другие материалы к занятиям

1. Харитонова, А.Е. Хранилища и системы интеллектуального анализа данных: методические указания / А.Е. Харитонова. – М.: РГАУ-МСХА им. К.А. Тимирязева, 2016. – 25 с.

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

1. Loginom. URL: <https://loginom.ru/>(открытый доступ)
2. Справочник Loginom. URL: <https://help.loginom.ru/userguide/>(открытый доступ)
3. Официальный сайт Росстата. URL: <https://rosstat.gov.ru/> (открытый доступ)
4. Официальный сайт Центрального Банка России. URL: <http://www.cbr.ru> (открытый доступ)
5. Bureau of Economic Analysis. URL: <http://www.bea.gov> (открытый доступ)
6. Московская международная валютная биржа. <http://www.micex.ru> (открытый доступ)
7. Официальный сайт Всемирного банка . URL: [http:// www.worldbank.org](http://www.worldbank.org) (открытый доступ)
8. Официальный сайт Министерства финансов РФ. URL: <http://www.minfin.gov.ru> (открытый доступ)
9. Официальный сайт Национального бюро экономических исследований США. URL: [http:// www.nber.org](http://www.nber.org) (открытый доступ)

#### 9. Перечень программного обеспечения и информационных справочных систем

Таблица 9

##### Перечень программного обеспечения

№ п/п	Наименование раздела учебной дисциплины	Наименование программы	Тип программы	Автор	Год разработки
1	Тема 1 Технология хранения данных при принятии решений Тема 2 Предобработка и консолидация данных Тема 3 Трансформация данных Тема 4 Предварительный анализ данных Тема 5 Регрессия Тема 6 Кластеризация Тема 7 ARIMAX Тема 8 Ассоциативные правила Тема 9 Нейросеть	Loginom	Аналитическая платформа	Loginom Company	Текущая версия

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

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
<i>учебная аудитория для проведения занятий лекционного типа, учебная аудитория для проведения занятий семинарского типа, учебная аудитория для групповых и индивидуальных консультаций, учебная аудитория для текущего контроля и промежуточной аттестации</i>	<ol style="list-style-type: none"> <li>1. Экран с электроприводом 1 шт. (Инв. №558771/2)</li> <li>2. Проектор 1 шт. (без инв. №) – приобретался не за счет средств вуза</li> <li>3. Вандалоустойчивый шкаф 1 шт. (Инв.№558850/7)</li> <li>4. Системный блок с монитором 1 шт. (Инв. №558777/9)</li> <li>5. Стенд «Сергеев Сергей Степанович 1910-1999» 1 шт. (Инв.№591013/25)</li> <li>6. Огнетушитель порошковый 1 шт. (Инв. №559527)</li> <li>7. Подвесное крепление к огнетушителю 1 шт. (Инв. № 559528)</li> <li>8. Жалюзи 2шт. (Инв. №1107-221225, Инв. №1107-221225)</li> <li>9. Лавка 20 шт.</li> <li>10. Стол аудиторный 20 шт.</li> <li>11. Стол для преподавателя 1 шт.</li> <li>12. Стул 2 шт.</li> <li>13. Доска маркерная 1 шт.</li> <li>14. Трибуна напольная 1 шт. (без инв. №)</li> </ol>
<i>учебная аудитория для проведения занятий семинарского типа, учебная аудитория для групповых и индивидуальных консультаций, учебная аудитория для текущего контроля и промежуточной аттестации, помещение для самостоятельной работы</i>	<ol style="list-style-type: none"> <li>1. Системный блок Intel Core Intel Core i3-2100/4096Мб/500Gb/DVD-RW 10 шт. (Инв.№601997, Инв.№601998, Инв.№601999, Инв.№602000, Инв.№602001, Инв.№602002, Инв.№602003, Инв.№602004, Инв.№602005, Инв.№602006)</li> <li>2. Монитор 10 шт. (без инв. №) - приобретались не за счет средств вуза</li> <li>3. Шкаф 2 шт. (Инв.№594166, Инв.№594167)</li> <li>4. Тумба 1 шт. (Инв.№594168)</li> <li>5. Подвесное крепление к огнетушителю 1 шт. (Инв. № 559528)</li> <li>6. Огнетушитель порошковый 1 шт. (Инв. №559527)</li> <li>7. Жалюзи 1 шт. (Инв.№551557)</li> <li>8. Доска магнитно-маркерная 1 шт.</li> <li>9. Стол 5 шт.</li> <li>10. Стол компьютерный 12 шт.</li> <li>11. Стул офисный 21 шт.</li> <li>12. Сейф 1 шт. (без Инв.№).</li> </ol>
<i>учебная аудитория для проведения занятий лекционного типа, учебная аудитория для проведения занятий семинарского типа, учебная аудитория для групповых и индивидуальных консультаций, учебная аудитория для текущего контроля и промежуточной аттестации, помещение для самостоятельной работы</i>	<ol style="list-style-type: none"> <li>1. Трибуна напольная 1 шт. (Инв.№ 599205)</li> <li>2. Шкаф для документов 3 шт. (Инв.№593633, Инв.№593634, Инв.№559548/18)</li> <li>3. Вешалка напольная 2 шт. (Инв.№1107-333144, Инв.№1107-333144)</li> <li>4. Жалюзи 1 шт. (Инв.№591110)</li> <li>5. Доска магнитно-маркерная 1 шт.</li> <li>6. Стол 15 шт.</li> <li>7. Скамейка 14 шт.</li> <li>8. Стол эрго 1 шт.</li> <li>9. Стул 2 шт.</li> </ol>
<i>Центральная научная библиотека имени</i>	Читальные залы библиотеки



Н.И. Железнова	
Студенческое общежитие	Комната для самоподготовки

## 11. Методические рекомендации обучающимся по освоению дисциплины

Все виды учебных работ должны быть выполнены точно в сроки, предусмотренные программой обучения. Если студент не выполнил какое-либо из учебных заданий по неуважительной причине (пропустил тестовый контроль, не выполнили домашнего задания, выполнил работу не по своему варианту и т.п.), то за данный вид учебной работы баллы рейтинга не начисляются, а подготовленные позже положенного срока работы оцениваются с понижающим коэффициентом. Если же невыполнение учебных работ произошло по уважительной причине, то следует представить преподавателю подтверждающий документ, и защитить пропущенные занятия в часы, отведенные для еженедельных консультаций.

### Виды и формы отработки пропущенных занятий

Студент, пропустивший занятия обязан выполнить самостоятельно индивидуальную работу, выполняемую на занятиях по своему варианту.

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

Курс должен давать не абстрактно-формальные, а прикладные знания. Данная цель может быть реализована только при условии соблюдения в учебных планах преемственности учебных дисциплин. Базовые знания для изучения дисциплины дают такие предметы, как экономическая теория, информатика.

Преподаватель должен указывать, в какой последовательности следует изучать материал дисциплины, обращать внимание на особенности изучения отдельных тем и разделов, помогать отбирать наиболее важные и необходимые сведения из учебных пособий, а также давать объяснения вопросам программы курса, которые обычно вызывают затруднения. При этом преподавателю необходимо учитывать следующие моменты:

1. Не следует перегружать студентов творческими заданиями.
2. Чередовать творческую работу на занятиях с заданиями во внеаудиторное время.
3. Давать студентам четкий инструктаж по выполнению самостоятельных заданий: цель задания; условия выполнения; объем; сроки; требования к оформлению.
4. Осуществлять текущий учет и контроль за самостоятельной работой.
5. Давать оценку обобщать уровень усвоения навыков самостоятельной, творческой работы.

Программу разработал (и):

Харитонов А.Е., к.э.н., доцент

  
 \_\_\_\_\_  
 (подпись)



## РЕЦЕНЗИЯ

на рабочую программу дисциплины Б1.В.ДВ.03.02 «Хранилища и системы интеллектуального анализа данных»

ОПОП ВО по направлению 09.03.02 «Информационные системы и технологии», направленность «Большие данные и машинное обучение (Machine Learning & Big Data)»

(квалификация выпускника – бакалавр)

Колосова Елена Сергеевна, доцент кафедры финансов ФГБОУ ВО г. Москвы «РГАУ-МСХА имени К.А. Тимирязева», кандидатом экономических наук (далее по тексту рецензент), проведена рецензия рабочей программы дисциплины «Хранилища и системы интеллектуального анализа данных» ОПОП ВО по направлению 09.03.02 «Информационные системы и технологии», направленность «Большие данные и машинное обучение (Machine Learning & Big Data)» (бакалавриат) разработанной в ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», на кафедре статистики и кибернетики (разработчик – Харитоновна Анна Евгеньевна, кандидат экономических наук, доцент кафедры статистики и кибернетики).

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

1. Предъявленная рабочая программа дисциплины «Хранилища и системы интеллектуального анализа данных» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению 09.03.02 «Информационные системы и технологии». 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 дисциплинам по выбору части, формируемой участниками образовательных отношений учебного цикла – Б1.В.ДВ.

3. Представленные в Программе цели дисциплины соответствуют требованиям ФГОС ВО направления 09.03.02 «Информационные системы и технологии».

4. В соответствии с Программой за дисциплиной «Хранилища и системы интеллектуального анализа данных» закреплено 3 компетенции (7 индикаторов). Дисциплина «Хранилища и системы интеллектуального анализа данных» и представленная Программа способна реализовать их в объявленных требованиях.

5. Общая трудоёмкость дисциплины «Хранилища и системы интеллектуального анализа данных» составляет 3 зачётных единицы (108 часов/из них практическая подготовка 4 ч.).

6. Информация о взаимосвязи изучаемых дисциплин и вопросам исключения дублирования в содержании дисциплин соответствует действительности. Дисциплина «Хранилища и системы интеллектуального анализа данных» взаимосвязана с другими дисциплинами ОПОП ВО и Учебного плана по направлению 09.03.02 «Информационные системы и технологии» 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 «Хранилища и системы интеллектуального анализа данных» предполагает 2 часа 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 09.03.02 «Информационные системы и технологии».

10. Представленные и описанные в Программе формы текущей оценки знаний (опрос, как в форме обсуждения отдельных вопросов, так и участие в деловых играх), соответствуют специфике дисциплины и требованиям к выпускникам.



Форма промежуточного контроля знаний студентов, предусмотренная Программой, осуществляется в форме зачета с оценкой, что соответствует статусу дисциплины, как дисциплины по выбору части, формируемой участниками образовательных отношений учебного цикла – Б1.В.ДВ ФГОС ВО направления *09.03.02 «Информационные системы и технологии»*.

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

11. Учебно-методическое обеспечение дисциплины представлено: основной литературой – 5 источников (базовый учебник), дополнительной литературой – 4 наименования, Интернет-ресурсы – 9 источников и соответствует требованиям ФГОС ВО направления *09.03.02 «Информационные системы и технологии»*.

12. Материально-техническое обеспечение дисциплины соответствует специфике дисциплины «Хранилища и системы интеллектуального анализа 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 «Хранилища и системы интеллектуального анализа данных».

### ОБЩИЕ ВЫВОДЫ

На основании проведенного рецензирования можно сделать заключение, что характер, структура и содержание рабочей программы дисциплины «Хранилища и системы интеллектуального анализа данных» ОПОП ВО по направлению *09.03.02 «Информационные системы и технологии»*, направленность **«Большие данные и машинное обучение (Machine Learning & Big Data)»** (квалификация выпускника – бакалавр), разработанная Харитоновой А. Е., к.э.н., доцентом кафедры статистики и кибернетики, соответствует требованиям ФГОС ВО, современным требованиям экономики, рынка труда и позволит при её реализации успешно обеспечить формирование заявленных компетенций.

Рецензент: Коломеева Е.С., доцент кафедры финансов ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», кандидат экономических наук \_\_\_\_\_ « 26 » 08 2022г.

(подпись)