

Документ подписан простой электронной подписью

Информация о владельце: **МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ**

ФИО: Гаранин Максим Александрович **ФЕДЕРАЛЬНОЕ АГЕНТСТВО ЖЕЛЕЗНОДОРОЖНОГО ТРАНСПОРТА**

Должность: Ректор

Федеральное государственное бюджетное образовательное учреждение высшего образования

Дата подписания: 06.09.2023 10:14:77

Уникальный программный ключ:

7708e3a47e66a8ee02711b298d7c78bd1e40bf88

**САМАРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ПУТЕЙ СООБЩЕНИЯ**

## **Большие данные**

### **рабочая программа дисциплины (модуля)**

Направление подготовки 09.03.02 Информационные системы и технологии

Направленность (профиль) Информационные системы и технологии на транспорте

Квалификация **бакалавр**

Форма обучения **очная**

Общая трудоемкость **5 ЗЕТ**

Виды контроля в семестрах:

экзамены 5

#### **Распределение часов дисциплины по семестрам**

Семестр (<Курс>.<Семестр на курсе>)	<b>5 (3.1)</b>		Итого	
	Неделя		Итого	
Вид занятий	уп	рп	уп	рп
Лекции	16	16	16	16
Практические	32	32	32	32
Конт. ч. на аттест. в период ЭС	2,35	2,35	2,35	2,35
Итого ауд.	48	48	48	48
Контактная работа	50,35	50,35	50,35	50,35
Сам. работа	105	105	105	105
Часы на контроль	24,65	24,65	24,65	24,65
Итого	180	180	180	180

Программу составил(и):

*к.п.н., Доцент, Додонов М.В.*

Рабочая программа дисциплины

**Большие данные**

разработана в соответствии с ФГОС ВО:

Федеральный государственный образовательный стандарт высшего образования - бакалавриат по направлению подготовки 09.03.02 Информационные системы и технологии (приказ Минобрнауки России от 19.09.2017 г. № 926)

составлена на основании учебного плана: 09.03.02-23-3-ИСТб.plm.plx

Направление подготовки 09.03.02 Информационные системы и технологии Направленность (профиль) Информационные системы и технологии на транспорте

Рабочая программа одобрена на заседании кафедры

**Цифровые технологии**

Зав. кафедрой к.т.н., доцент Авсиевич А.В.

**1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)**

1.1	Формирование у студентов необходимой теоретической базы и практических навыков анализа данных различного объема, включая предварительную обработку данных, предназначенных для решения задач кластеризации, классификации, регрессии и применение их для решения прикладных задач из различных сфер человеческой деятельности.
-----	--

**2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ**

Цикл (раздел) ОП:	Б1.О.14
-------------------	---------

**3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)**

ОПК-8	Способен применять математические модели, методы и средства проектирования информационных и автоматизированных систем.
ОПК-8.1	Применяет математические модели для проектирования информационных и автоматизированных систем
ОПК-8.3	Использует средства проектирования информационных и автоматизированных систем

**В результате освоения дисциплины (модуля) обучающийся должен**

<b>3.1</b>	<b>Знать:</b>
3.1.1	Математические модели для проектирования информационных и автоматизированных систем для работы с большими данными;
3.1.2	Основные методы применения больших данных для решения прикладных задач из различных сфер человеческой деятельности.
<b>3.2</b>	<b>Уметь:</b>
3.2.1	Применять математические модели для проектирования информационных и автоматизированных систем для работы с большими данными;
3.2.2	Использовать современные облачные сервисы для работы с большими данными. Визуализировать полученные результаты работы.
<b>3.3</b>	<b>Владеть:</b>
3.3.1	Работы с математическими моделями для проектирования информационных и автоматизированных систем для работы с большими данными;
3.3.2	Основными инструментами анализа данных на базе Google Colab или Yandex DataSphere на примере решения задач кластеризации, классификации, прогнозирования.

**4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)**

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Примечание
	<b>Раздел 1. Технологии анализа данных</b>			
1.1	Большие данные (Big Data): современные подходы к обработке и хранению. Проблема множественного сравнения данных. /Лек/	5	2	
1.2	Процесс анализа. Общая схема анализа. Извлечение и визуализация данных. Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представления наборов данных. /Лек/	5	2	
1.3	Программное обеспечение в области анализа данных. Аналитические платформы: классификация и особенности применения. Языки визуального моделирования. /Лек/	5	2	
1.4	Современные облачные сервисы для работы с большими данными. /Лек/	5	2	
1.5	Основы работы с облачными сервисами Google Colab или Yandex DataSphere /Пр/	5	2	
1.6	Понятие сценария. /Пр/	5	2	
1.7	Понятие узла обработки. /Пр/	5	4	
1.8	Консолидация данных. /Пр/	5	4	
1.9	Визуализация данных. /Пр/	5	4	
1.10	Трансформация данных. /Пр/	5	4	
1.11	Современные подходы к обработке и хранению. Проблема множественного сравнения данных. /Ср/	5	9	

1.12	Применение технологий больших данных для решения прикладных задач из различных сфер человеческой деятельности /Ср/	5	10	
1.13	Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представления наборов данных. /Ср/	5	12	
1.14	Подготовка данных к анализу. Методика извлечения знаний. Data Mining. Мультидисциплинарный характер Data Mining. /Ср/	5	10	
<b>Раздел 2. Интеллектуальный анализ данных</b>				
2.1	Введение в интеллектуальный анализ данных: основные понятия, области применения современных технологий обработки и интеллектуального анализа больших данных с использованием облачных решений /Лек/	5	2	
2.2	Организация сбора подготовки и хранения больших наборов данных. Базы данных в облачных сервисах Google Colab или Yandex DataSphere. /Лек/	5	2	
2.3	Определение кластеризации. Постановка задачи кластеризации. Цели кластеризации в Data Mining. Примеры кластеризации в различных областях. Проблемы алгоритмов кластеризации /Лек/	5	2	
2.4	Применение классификации и регрессии. Обзор методов классификации и регрессии. Статистические методы. Методы, основанные на обучении, разнообразии подходов /Лек/	5	2	
2.5	Интеллектуальный анализ данных в облачных сервисах Google Colab или Yandex DataSphere /Пр/	5	4	
2.6	Кластеризация. Алгоритм кластеризации k-means. /Пр/	5	2	
2.7	Прогнозирование с помощью линейной регрессии. /Пр/	5	4	
2.8	Визуализация данных в облачных сервисах Google Colab или Yandex DataSphere. /Пр/	5	2	
2.9	Применение классификации и регрессии. Обзор методов классификации и регрессии. Статистические методы. Методы, основанные на обучении, разнообразии подходов. /Ср/	5	12	
2.10	Базы данных и СУБД. SQLиNoSQL. Интеграция баз данных и облачных сервисов. /Ср/	5	12	
<b>Раздел 3. Самостоятельная работа</b>				
3.1	Подготовка к лекциям /Ср/	5	8	
3.2	Подготовка к практическим занятиям /Ср/	5	32	
<b>Раздел 4. Аттестация</b>				
4.1	Экзамен /КЭ/	5	2,35	

### 5. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

Оценочные материалы для проведения промежуточной аттестации обучающихся приведены в приложении к рабочей программе дисциплины.

Формы и виды текущего контроля по дисциплине (модулю), виды заданий, критерии их оценивания, распределение баллов по видам текущего контроля разрабатываются преподавателем дисциплины с учетом ее специфики и доводятся до сведения обучающихся на первом учебном занятии.

Текущий контроль успеваемости осуществляется преподавателем дисциплины (модуля), как правило, с использованием ЭИОС или путем проверки письменных работ, предусмотренных рабочими программами дисциплин в рамках контактной работы и самостоятельной работы обучающихся. Для фиксирования результатов текущего контроля может использоваться ЭИОС.

### 6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

#### 6.1. Рекомендуемая литература

##### 6.1.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год	Эл. адрес
--	---------------------	----------	-------------------	-----------

	Авторы, составители	Заглавие	Издательство, год	Эл. адрес
Л1.1	Нестеров С. А.	Базы данных: учебник и практикум для вузов	Москва: Юрайт, 2021	tps://urait.ru/bcode/46951

### 6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Эл. адрес
Л2.1	Стружкин Н. П., Годин В. В.	Базы данных: проектирование: учебник для вузов	Москва: Юрайт, 2021	tps://urait.ru/bcode/46902

## 6.2 Информационные технологии, используемые при осуществлении образовательного процесса по дисциплине (модулю)

### 6.2.1 Перечень лицензионного и свободно распространяемого программного обеспечения

#### 6.2.2 Перечень профессиональных баз данных и информационных справочных систем

6.2.2.1	База книг и публикаций Электронной библиотеки "Наука и Техника" - <a href="http://www.n-t.ru">http://www.n-t.ru</a>
6.2.2.2	Крупнейший веб-сервис для хостинга IT-проектов и их совместной разработки- <a href="https://github.com/">https://github.com/</a>
6.2.2.3	Портал для разработчиков электронной техники: <a href="http://www.espec.ws/">http://www.espec.ws/</a>
6.2.2.4	База данных «Библиотека программиста» <a href="https://proglib.io/">https://proglib.io/</a>
6.2.2.5	Консультант плюс
6.2.2.6	Информационная система ГАРАНТ

## 7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1	Учебные аудитории для проведения занятий лекционного типа, укомплектованные специализированной мебелью и техническими средствами обучения: мультимедийное оборудование для предоставления учебной информации большой аудитории и/или звукоусиливающее оборудование (стационарное или переносное).
7.2	Учебные аудитории для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения: мультимедийное оборудование и/или звукоусиливающее оборудование (стационарное или переносное)
7.3	Помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета.
7.4	Помещения для хранения и профилактического обслуживания учебного оборудования