МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего образования
«УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

**УТВЕРЖДАЮ**
Декан факультета информационных систем и технологий
_____ К.В. Святов
«11» 10 2021 г.

# РАБОЧАЯ ПРОГРАММА

Дисциплина (модуль): Технологии обработки и анализа больших данных в предиктивной аналитике
*наименование дисциплины (модуля)*

Уровень образования: высшее образование - магистратура
*(СПО/бакалавриат/магистратура/специалитет/подготовка кадров высшей квалификации)*

Квалификация: Магистр
*Техник/Бакалавр/Магистр/Инженер/ Исследователь. Преподаватель-исследователь*

г. Ульяновск, 2021

Рабочая программа составлена

| | |
|---|---|
| на кафедре | Информационные системы |
| факультета | информационных систем и технологий |
| в соответствии с учебным планом по направлению подготовки (специальности) | 09.04.04 Программная инженерия |
| профиль (программа / специализация) | Искусственный интеллект и предиктивная аналитика |

Составитель рабочей программы

Доцент, к.э.н. _____ Вертешев А.С.
*(должность, ученое звание, степень)* *(подпись)* *(Фамилия И. О.)*

Рабочая программа рассмотрена на заседании кафедры
Заведующий кафедрой _____ Романов А.А.
*(должность)* *(подпись)* *(Фамилия И. О.)*

СОГЛАСОВАНО:

Руководитель ОПОП
«11» октября 2021 г. _____ Романов А.А.
*(подпись)* *(Фамилия И. О.)*

Заведующий выпускающей кафедрой /научный руководитель ОПОП
«11» октября 2021г. _____ Романов А.А
*(подпись)* *(Фамилия И. О.)*

Директор библиотеки
«11» октября 2011 г. _____ Синдюкова Е.С.
*(подпись)* *(Фамилия И. О.)*

## 1 ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ) В ЗАЧЕТНЫХ ЕДИНИЦАХ С УКАЗАНИЕМ АКАДЕМИЧЕСКИХ ЧАСОВ, ВЫДЕЛЕННЫХ НА КОНТАКТНУЮ РАБОТУ ОБУЧАЮЩИХСЯ С ПРЕПОДАВАТЕЛЕМ (ПО ВИДАМ УЧЕБНЫХ ЗАНЯТИЙ) И НА САМОСТОЯТЕЛЬНУЮ РАБОТУ ОБУЧАЮЩИХСЯ

Таблица 1

Бюджет времени с учетом формы обучения, семестра и видов занятий

| **Форма обучения** | Очная | | | | Очно-заочная | | | | Заочная | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Семестр** | **3** | | | | | | | | | | | |
| **Контактная работа обучающихся с преподавателем** (по видам учебных занятий), **всего часов** | **48** | | | | | | | | | | | |
| в том числе: | | | | | | | | | | | | |
| **- занятия лекционного типа** (лекции и иные учебные занятия, предусматривающие преимущественную передачу учебной информации педагогическими работниками), часов | **16** | | | | | | | | | | | |
| **- занятия семинарского/ практического типа** (семинары, практические занятия, практикумы, коллоквиумы и иные аналогичные занятия), часов | | | | | | | | | | | | |
| **- лабораторные занятия** (включая работу обучающихся на реальных или виртуальных объектах профессиональной сферы), часов | **32** | | | | | | | | | | | |
| **Самостоятельная работа обучающихся**, часов | **204** | | | | | | | | | | | |
| в том числе: | | | | | | | | | | | | |
| - групповые и индивидуальные консультации обучающихся с преподавателями | | | | | | | | | | | | |
| - проработка теоретического курса | **90** | | | | | | | | | | | |
| - курсовая работа (проект) | | | | | | | | | | | | |
| - расчетно-графическая работа | | | | | | | | | | | | |
| - реферат | | | | | | | | | | | | |
| - эссе | | | | | | | | | | | | |
| - подготовка к занятиям семинарского/практического типа | | | | | | | | | | | | |
| - подготовка к выполнению и защите лабораторных работ | **114** | | | | | | | | | | | |
| - взаимодействие в электронной информационно-образовательной среде вуза | | | | | | | | | | | | |
| **Промежуточная аттестация** обучающихся, включая подготовку (Экзамен, Зачет, Зачет с оценкой, КП, КР) | **36** | | | | | | | | | | | |
| **Итого, часов** | **288** | | | | | | | | | | | |
| **Трудоемкость, з.е.** | **8** | | | | | | | | | | | |

## 2 ЯЗЫК ПРЕПОДАВАНИЯ

Изучение дисциплины (модуля) осуществляется на русском языке.

## 3 ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ (МОДУЛЯ)

Целью освоения дисциплины (модуля) «Технологии обработки и анализа больших данных в предиктивной аналитике» является ознакомление с основными технологиями решения задач обработки больших данных, быстро изменяющихся и плохо структурированных.

Задачами освоения дисциплины (модуля) являются формирование у обучающихся:

**Знаний,** содержащих**:**

- Базовые понятия технологии Big Data
- Базовые понятия прогнозирования
- Базовые понятия работы с хранилищами данных

**Умений:**

- применять методы анализа больших данных
- реализовывать приложения для предиктивной аналитики больших данных
-использовать инструменты анализа больших данных и интерпретировать результаты
- анализировать кластеры больших данных

В результате изучения дисциплины (модуля) «Технологии обработки и анализа больших данных в предиктивной аналитике» обучающиеся на основе приобретенных знаний, умений и навыков достигают освоения компетенций на определенном уровне.

Аннотация дисциплины (модуля) представлена в Приложении А.

## 4 ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ), СООТНЕСЕННЫХ С ИНДИКАТОРАМИ ДОСТИЖЕНИЯ КОМПЕТЕНЦИЙ

Таблица 2

Планируемые результаты обучения по дисциплине (модулю), с указанием индикатора достижения компетенций

| **Код компетенции** | **Формулировка компетенции** | **Код индикатора достижения компетенции (по данной дисциплине (модулю))** | **Индикаторы достижения компетенции (связанные с данной дисциплиной (модулем))** |
|---|---|---|---|
| **Профессиональные** | | | |
| **ПК-3** | Способен разрабатывать и применять методы и алгоритмы машинного обучения для решения задач | ИД-1 ПК-3 | Ставит задачи по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области:<br>- Знает классы методов и алгоритмов машинного обучения<br>- Умеет ставить задачи и разрабатывать новые методы и |

| | | | |
|---|---|---|---|
| | | | алгоритмы машинного обучения |
| | | ИД-2 ПК-3 | Руководит исследовательской группой по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области:<br>- Знает методы и критерии оценки качества моделей машинного обучения<br>- Умеет определять критерии и метрики оценки результатов моделирования при построении систем искусственного интеллекта в исследуемой области |
| **ПК-6** | Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях | ИД-1 ПК-6 | Осуществляет руководство проектом по построению комплексных систем на основе аналитики больших данных в различных отраслях:<br>- Знает методологию и принципы руководства проектом по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных<br>- Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных<br>- Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных |

## 5 МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина (модуль) относится к **части Б1.В, формируемая участниками образовательных отношений** образовательной программы **09.04.04 «Программная инженерия»**.

## 6 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ), СТРУКТУРИРОВАННОЕ ПО ТЕМАМ (РАЗДЕЛАМ) С УКАЗАНИЕМ ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ И ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ

### 6.1 Тематический план изучения дисциплины (модуля)

Таблица 3

Тематический план с указанием выделенных академических часов на освоение каждого из разделов и проведение промежуточной аттестации

| № | Наименование разделов (включая промежуточную аттестацию) | Очная (час) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Лекции | Практические (сем.) занятия | Лабораторные работы | Контроль | Самостоятельная работа | **Всего** |
| 1 | **Введение в большие данные.** | 4 | - | 6 | 6 | 42 | 58 |
| 2 | **Технологии хранения данных:** | 6 | - | 16 | 6 | 82 | 110 |
| 3 | **Технологии обработки и анализа больших данных в предиктивной аналитике:** | 6 | - | 10 | 6 | 62 | 84 |
| 4 | **Подготовка к промежуточной аттестации, консультации перед промежуточной аттестацией и сдача промежуточной аттестации** | - | - | - | 18 | 18 | 36 |
| | **Итого часов** | **16** | **-** | **32** | **36** | **204** | **288** |

### 6.2 Теоретический курс

Таблица 4

Основные вопросы, освещаемые на лекциях

| Раздел, тема учебной дисциплины (модуля), содержание темы |
|---|
| **Раздел 1.** Введение в большие данные:<br>Понятие Data Mining, OLTP, OLAP. Технология Apache Hadoop MapReduce .<br>Непрерывное развертывание и интеграция (github/ gitlab, jenkins). Микросервисная архитектура. |
| **Раздел 2.** Технологии хранения больших данных:<br>Распределенные хранилища, NoSql хранилища. Классификация и примеры (MongoDB, Графовая СУБД Neo4J, Redis, InfluxDB, Apache Cassandra, CouchDB, PostgreSQL, ) Elasticsearch) |
| **Раздел 3.** Технологии обработки и анализа больших данных: Жизненный цикл анализа больших данных, стандарты.<br>Когнитивный анализ данных. Визуализация больших данных (Grafana). Облачные вычисления (Apache Spark, Apache Ignite). Контейнеры (Docker, LXC, Kubernetes) Виртуализация (Hyper-V, Proxmox, Amazon) |

### 6.3 Лабораторный практикум

Таблица 6

Тематика лабораторных работ

| Номер | Наименование лабораторной работы |
|---|---|
| 1 | Знакомство с понятием Data Mining, OLTP,OLAP, технологиями Apache Hadoop MapReduce |
| 2 | Использование технологий непрерывного развертывания и интеграции (github/ gitlab, jenkins) |
| 3 | Применение микросервисной архитектуры в потоковой обработке Big Data |
| 4,5 | Использование реляционных хранилищ данных для big data (PostgreSql) |
| 6,7 | Применение NoSql хранилищ данных для big data (Neo4J,CouchDB, Redis, Apache Cassandra ) |
| 8,9 | Использование NoSql хранилищ для big data (MongoDB, InfluxDB, Elasticsearch) |
| 10 | Когнитивные методы анализа больших данных |
| 11 | Аналитика и визуализация больших данных с помощью инструмента Grafana. |
| 12 | Использование облачных вычислений при помощи Apache Spark, и Apache Ignite. |
| 13 | Знакомство с контейнерами Docker, LXC и Kubernetes. |
| 14 | Системы виртуализации Hyper-V, Proxmox, Amazon. |

### 6.4 Самостоятельная работа обучающихся

Виды самостоятельной работы распределяются в течение семестра. Подготовка к промежуточной аттестации ведется в установленные календарным учебным графиком сроки.

## 7 ПЕРЕЧЕНЬ ОЦЕНОЧНЫХ СРЕДСТВ (ОЦЕНОЧНЫХ МАТЕРИАЛОВ) ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Таблица 7

Наименование оценочных средств (оценочных материалов)

| № п/п | Код формируемой компетенции | Код индикатора достижения формируемой компетенции | Наименование оценочного средства |
|---|---|---|---|
| 1. | ПК-3 | ИД-1 ПК-3 | Лабораторные работы, экзамен |
| | | ИД-2 ПК-3 | |
| 2. | ПК-6 | ИД-1 ПК-6 | Лабораторные работы, экзамен |

## 8 ПЕРЕЧЕНЬ УЧЕБНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

1. Макшанов А. В. Большие данные. Big Data : учебник для вузов / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. СанктПетербург : Лань, 2021. 188 с.

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, А. Д. Хомоненко. — Санкт Петербург : Лань, 2020. — 236 с. : ил. — (Учебники для вузов. Специальная литература)

3. Макшанов А. В. Современные технологии интеллектуального анализа данных : учебное пособие для СПО / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. — Санкт-Петербург : Лань, 2020. — 228 с.

4. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2021. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт

5. Парфенов, Ю. П. Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. — Москва : Издательство Юрайт, 2021. — 121 с. — (Высшее образование). — ISBN 978-5-534-09837-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL:

6. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт].

7. Крутиков, В.Н. Анализ данных : учебное пособие / В.Н. Крутиков, В.В. Мешечкин ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Кемеровский государственный университет». - Кемерово : Кемеровский государственный университет, 2014. - 138 с. : ил. - Библиогр. в кн. - ISBN 978-5-8353-1770-7 ; То же [Электронный ресурс]. - URL: http://biblioclub.ru/index.php?page=book&id=278426

8. Жуковский, О.И. Информационные технологии и анализ данных : учебное пособие / О.И. Жуковский ; Министерство образования и науки Российской Федерации, Томский Государственный Университет Систем Управления и Радиоэлектроники (ТУСУР). - Томск : Эль Контент, 2014. - 130 с. : схем., ил. - Библиогр.: с. 126. - ISBN 978-5-4332-0158-3 ; То же [Электронный ресурс]. - URL: http://biblioclub.ru/index.php?page=book&id=480500

9. Волк В.К. Базы данных : учебное пособие. Ч.1. Проектирование и программирование / В.К. Волк ; Министерство науки и высшего образования Российской Федерации, Курганский государственный университет ; [науч. ред. В.А. Симахин]. - Курган : Издательство Курганского государственного университета, 2018.

10. Волк В.К. Базы данных : учебное пособие. Ч.2. Администрирование / В.К. Волк ; Министерство образования и науки Российской Федерации, Курганский государственный университет ; [науч. ред. В.А. Симахин]. - Курган : Издательство Курганского государственного университета, 2018. - 127, [1] с. - Библиогр.: с. 127. - ISBN 978-5-4217-0440-9.

## 9 ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ ДЛЯ КОНТАКТНОЙ И САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

1. Выполнение и оформление курсовых проектов (работ) : методические указания / сост. Н. В. Корунова. – Ульяновск: УлГТУ, 2014». http://venec.ulstu.ru/lib/disk/2014/129.pdf

## 10. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ РЕСУРСОВ

### 10.1 Справочные системы и современные профессиональные базы данных, к которым обучающимся обеспечен доступ (удаленный доступ), в том числе в случае применения электронного обучения, дистанционных образовательных технологий

1. Полнотекстовая база данных ScienceDirect
2. Реферативная и наукометрическая база данных Scopus
3. Национальный цифровой ресурс «Руконт»
4. Справочная система Гарант.
5. База ГОСТы и СанПиНы https://standartgost.ru/
6. База СНИПы. Нормативно-техническая документация http://snipov.net/

7. Федеральный портал Единое окно доступа к образовательным ресурсам http://window.edu.ru/library
8. Научная электронная библиотека http://elibrary.ru/defaultx.asp
9. РГБ фонд диссертаций http://diss.rsl.ru/
10. Энциклопедия http://encyclopaedia.biga.ru

**10.2 Ресурсы информационно-телекоммуникационной сети «Интернет»**

1. Электронно-библиотечная система издательства «Лань» http://e.lanbook.com/
2. Электронная библиотека по всем отраслям знаний — Режим доступа: www.iprbookshop.ru
3. Электронная библиотека по всем отраслям знаний — Режим доступа: http://biblioclub.ru
7. РГБ фонд диссертаций http://diss.rsl.ru/

## 11 ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ И ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ (МОДУЛЮ), ВКЛЮЧАЯ ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ

Таблица 8

Наименование и оснащенность помещений, используемых при осуществлении образовательного процесса по дисциплине (модулю)

| № п/п | Наименование специальных помещений и помещений для самостоятельной работы | Оснащенность специальных помещений и помещений для самостоятельной работы | Перечень лицензионного программного обеспечения (подлежит ежегодному обновлению) |
|---|---|---|---|
| 1 | Учебные аудитории для проведения лекций | Доступ к глобальной сети Интернет | Операционная система MS Windows 7 и выше; Операционная система Linux; Redis - резидентная система управления базами данных класса NoSQL; Microsoft Office; LibreOffice |
| 2 | Учебные аудитории для проведения лабораторных работ, практических работ, групповых и индивидуальных консультаций | Доступ к глобальной сети Интернет | Операционная система MS Windows 7 и выше; Операционная система Linux; Redis - резидентная система управления базами данных класса NoSQL; Microsoft Office; LibreOffice |
| 3 | Учебные аудитории для текущего контроля и промежуточной аттестации | Доступ к глобальной сети интернет | Операционная система MS Windows 7 и выше; Операционная система Linux; Redis - резидентная система управления базами данных класса NoSQL; Microsoft Office; LibreOffice |
| 4 | Помещения для самостоятельной работы (читальный зал научной библиотеки) | Доступ к глобальной сети интернет | Операционная система MS Windows 7 и выше; Операционная система Linux; Redis - резидентная система управления базами данных класса NoSQL; Microsoft Office; LibreOffice |

Приложение А

Аннотация рабочей программы

| | |
|---|---|
| Дисциплина (модуль) | «Технологии обработки и анализа больших данных в предиктивной аналитике» |
| Уровень образования | магистратура |
| Квалификация | магистр |
| Направление подготовки / специальность | 09.04.04 Программная инженерия |
| Профиль / программа / специализация | Искусственный интеллект и предиктивная аналитика |
| Дисциплина (модуль) нацелена на формирование компетенций | ПК-3; ПК-6 |
| Цель освоения дисциплины (модуля) | Целью освоения дисциплины (модуля) «Технологии обработки и анализа больших данных в предиктивной аналитике» является ознакомление с основными технологиями решения задач обработки больших данных, быстро изменяющихся и плохо структурированных. |
| Перечень разделов дисциплины | Введение в большие данные. Технологии хранения больших данных. Технологии обработки и анализа больших данных в предиктивной аналитике: |
| Общая трудоемкость дисциплины (модуля) | 8 з.е., 288 часов |
| Форма промежуточной аттестации | Лабораторные работы, экзамен |

**Лист дополнений и изменений**

к рабочей программе дисциплины (модуля)

_____

Учебный год: 20__/20__

Протокол заседания кафедры № ___ от «___» _____ 20___ г.

Принимаемые изменения:

_____
_____
_____
_____
_____
_____

Руководитель ОПОП _____ _____

личная подпись И.О. Фамилия

«____»_____20__ г.