

## Аннотация рабочей программы дисциплины

### К.М.01.01 «Обработка больших данных»

Направление подготовки/специальность

02.03.02 Фундаментальная информатика и информационные технологии

Курс 3 Семестр 6 Количество з.е. 4 (144 час., из них – 72,3 часов аудиторной нагрузки: лекционных 34 ч., практических 34 ч., иной контактной работы 0,3 ч 36 часов самостоятельной работы), 35,7 – контроль - экзамен.

#### Цель освоения дисциплины:

Курс «Обработка больших данных» имеет своей целью: формирование у студентов профессиональной компетенции в области разработки и использования систем обработки и анализа больших массивов данных. Данная цель соотносится с целью образовательной программой в частности с технологий разработки специализированных программных систем, отвечающих за обработку больших данных. Изучение данной дисциплины готовит выпускника к выполнению следующих профессиональных задач:

- Постановка задачи анализа данных.
- Предварительная обработка данных.
- Визуализация данных.
- Разработка, реализация и применение методов интеллектуального анализа данных к большим массивам данных.
- Представление результатов работы.
- Интерпретация.

#### Задачи дисциплины:

Основные задачи освоения дисциплины:

Студент должен **знать** методы анализа и хранения больших объемов данных, этапы жизненного цикла обработки больших данных, языки, наиболее приспособленные для обработки и аналитики больших данных, способы организации хранения и доступа к большим данным; **уметь** выполнять элементы анализа данных и интерпретировать результаты, выбирать подходящий инструмент анализа больших данных, **владеть** математическими методами анализа данных, языками и компьютерными методами обработки и визуализации.

#### Место дисциплины (модуля) в структуре образовательной программы

Курс «Обработка больших данных» относится к части блока Б1 основных дисциплин учебного плана.

Для изучения дисциплины студент должен владеть знаниями, умениями и навыками по дисциплинам:

Дискретная математика, Алгебраические структуры, Основы программирования, Алгоритмы вычислительной математики, Конструирование алгоритмов и структур данных, Теория алгоритмов и вычислительных процессов, Основы теории вероятностей и статистических методов, Алгоритмы и структуры данных, Математическая логика и теория алгоритмов, Интеллектуальный анализ данных.

Знания, получаемые при изучении дисциплины «Обработка больших данных» используются при изучении профессиональных дисциплин Распределенные задачи и алгоритмы, Программирование в компьютерных сетях, Облачные вычисления, Мультиагентные системы, а также для работ над дипломной и магистерской работой.

**Результаты обучения (знания, умения, опыт, компетенции):**

| Код и наименование индикатора*                                                                                                                                                                                                                                                                                                                                                             | Результаты обучения по дисциплине<br>( <i>знает, умеет, владеет (навыки и/или опыт деятельности)</i> )                                                                                                                                                                                         |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>ОПК-2</b> Способен применять компьютерные/суперкомпьютерные методы, современное программное обеспечение, в том числе отечественного происхождения, для решения задач профессиональной деятельности                                                                                                                                                                                      |                                                                                                                                                                                                                                                                                                |
| ОПК-2.1. Знает основные положения и концепции в области программирования, архитектуру языков программирования, теории коммуникации, знает основную терминологию, знаком с содержанием Единого Реестра Российских программ                                                                                                                                                                  | Знает стандарты обработки и анализа больших данных, и требования, связанные с созданием и использованием SQL и NoSQL систем хранения и обработки данных                                                                                                                                        |
| ОПК-2.2. Знает особенности языков программирования, теорию алгоритмов, умеет составлять программы                                                                                                                                                                                                                                                                                          | Использовать современные инструментальные и вычислительные средства – язык анализа данных R, осуществлять постановку задач анализа данных, визуализацию интерпретацию результатов                                                                                                              |
| ОПК-2.3. Имеет практический опыт решения задач анализа, интеграции различных типов программного обеспечения, анализа типов коммуникаций                                                                                                                                                                                                                                                    | Владеет способностью собирать, обрабатывать и интерпретировать данные современных научных исследований, необходимые для формирования выводов по соответствующим научным исследованиям                                                                                                          |
| <b>ОПК-3</b> Способен к разработке алгоритмических и программных решений в области системного и прикладного программирования, математических, информационных и имитационных моделей, созданию информационных ресурсов глобальных сетей, образовательного контента, прикладных баз данных, тестов и средств тестирования систем и средств на соответствие стандартам и исходным требованиям |                                                                                                                                                                                                                                                                                                |
| ОПК-3.1. Знает методы теории алгоритмов, методы системного и прикладного программирования, основные положения и концепции в области математических, информационных и имитационных моделей;                                                                                                                                                                                                 | Знает математические методы анализа данных, методы и прикладные языки для разработки программных решений в области обработки больших данных, математических, информационных и имитационных моделей.                                                                                            |
| ОПК-3.2. Умеет соотносить знания в области программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы глобальных сетей, образовательного контента, средств тестирования систем                                                                                                                                                                            | Умеет корректно построить архитектуру кроссплатформенного приложения. Реализовать программу, включающую реализацию сенсорно-моторной координации и пространственного позиционирования, алгоритмы извлечения и обработки данных, включая возможности автономного принятия решений на основе ИИ. |
| ОПК-3.3. Имеет практический опыт применения разработки программного обеспечения.                                                                                                                                                                                                                                                                                                           | Владеет языками системного и прикладного программирования для разработки математических, информационных и имитационных моделей, для обработки информационных ресурсов глобальных сетей и прикладных баз данных.                                                                                |
| <b>УК-1.</b> Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач                                                                                                                                                                                                                                                |                                                                                                                                                                                                                                                                                                |
| УК-1.1. Знает основные принципы и методы поиска данных и анализа данных                                                                                                                                                                                                                                                                                                                    | Знает методы поиска информации. Современные языки и средства обработки данных (язык R, Python), прикладные библиотеки для анализа данных                                                                                                                                                       |

| Код и наименование индикатора*                                                                                                                                                  | Результаты обучения по дисциплине<br>(знает, умеет, владеет (навыки и/или опыт деятельности))                                                                                                                        |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>УК-1.2.</b> Умеет применить методы анализа данных и машинного обучения для решения задач профессиональной деятельности                                                       | применить современные языки (R и Python) и прикладные библиотеки для сбора и анализа данных, применить системный подход для решения поставленных задач, сформулировать научную гипотезу и проверить ее достоверность |
| <b>УК-1.3.</b> Имеет практический опыт применения методов искусственного интеллекта для получения новых аналитических результатов в решении задач профессиональной деятельности | Владеет средствами сбора, обработки и анализа больших данных, средствами оценки эффективности решений                                                                                                                |

### Содержание и структура дисциплины (модуля)

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.  
Разделы дисциплины, изучаемые в \_6\_ семестре (очная форма)

| №  | Наименование разделов                                                                                                                   | Количество часов |                   |     |    |                             |
|----|-----------------------------------------------------------------------------------------------------------------------------------------|------------------|-------------------|-----|----|-----------------------------|
|    |                                                                                                                                         | Всего            | Аудиторная работа |     |    | Внеаудиторная работа<br>СРС |
|    |                                                                                                                                         |                  | Л                 | КСР | ЛР |                             |
| 1. | <b>Введение в большие данные.</b><br>Понятие Data Minig.<br>Прикладные инструменты для работы с Big Data. Технология MapRaduce. Hadoop. | 16               | 4                 | 2   | 4  | 6                           |
| 2. | <b>Технологии анализа данных:</b><br>Жизненный цикл анализа больших данных, стандарты. Задачи статистического обучения                  | 12               | 4                 |     | 4  | 4                           |
| 3. | Алгоритмы классификации, кластеризации.<br>Понятие корреляции и регрессионный анализ.<br>Тестирование гипотез.                          | 52               | 16                | 2   | 18 | 18                          |
| 4. | Аналитика данных на больших графах                                                                                                      | 20               | 6                 |     | 8  | 6                           |
| 5. | <b>Технологии хранения больших данных.</b><br>Распределенные хранилища, NoSql хранилища, классификация и примеры.                       | 6                | 4                 |     |    | 2                           |
|    | <i>Итого по разделам дисциплины:</i>                                                                                                    | 108              | 34                | 4   | 34 | 36                          |
|    | <b>ИКР</b>                                                                                                                              | 0,3              |                   |     |    |                             |
|    | <i>Итого:</i>                                                                                                                           | 108,3            |                   |     |    |                             |
|    | <i>Контроль</i>                                                                                                                         | 44,7             |                   |     |    |                             |
|    | <i>Итого по дисциплине:</i>                                                                                                             | 144              |                   |     |    |                             |

Примечание: Л – лекции, КСР – контрольные и самостоятельные работы, ЛР – лабораторные занятия, СРС – самостоятельная работа студента, Д-доклад, РГЗ – расчетно-графическое задание.

**Курсовые проекты или работы:** не предусмотрены **Вид аттестации:** экзамен.

Автор Приходько Т.А. – кандидат технических наук, доцент кафедры вычислительных технологий