

Министерство науки и высшего образования РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.ДВ.07.03 Технологии обработки и хранения больших
данных

наименование дисциплины (модуля) в соответствии с учебным планом

Направление подготовки / специальность

01.03.02 Прикладная математика и информатика

Направленность (профиль)

01.03.02.31 Математическое моделирование и вычислительная
математика

Форма обучения

очная

Год набора

2023

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Программу составили _____

к.ф.-.м.н., Доцент, Куликов Владимир Русланович

должность, инициалы, фамилия

1 Цели и задачи изучения дисциплины

1.1 Цель преподавания дисциплины

Целью преподавания дисциплины является ознакомление студентов с математическими методами и программными средствами обработки больших данных и решении прикладных задач. Курс ориентирован на применение современных программных пакетов для хранения и работы с большими данными.

1.2 Задачи изучения дисциплины

- Формирование теоретических и методологических основ в области анализа неструктурированной информации, а также практических навыков, использования алгоритмов интеллектуального анализа данных.

- Формирование теоретических основ и навыков использования парадигмы распределенных вычислений MapReduce и концепции баз данных NoSQL.

- Формирование навыков проведения сравнительного анализа основных моделей, включая когнитивный анализ данных.

1.3 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Код и наименование индикатора достижения компетенции	Запланированные результаты обучения по дисциплине
ПК-1: Способен применять базовые знания математических и естественных наук, основ программирования и информационных технологий при проведении исследования в конкретной области профессиональной деятельности	
ПК-1.1: Применяет теоретические и практические знания математических и естественных наук, основ программирования и информационных технологий для проведения в конкретной области профессиональной деятельности	основные понятия, постановки задач, основные принципы и методы разработки алгоритмов обработки больших данных; современные тенденции и направления в научных исследованиях, проводимых в мире, по теории алгоритмов обработки больших данных; классические и современные методы, применяемые при обработке больших данных, необходимые и достаточные условия их реализации. применять классические и современные методы обработки больших данных для решения задач в гуманитарных и социально-экономических науках; систематизировать научные результаты, выделять из них главное, и удалять второстепенное; самостоятельно выбирать эффективные методы решения поставленных задач и разрабатывать новые методы для получения новых научных и прикладных результатов. самостоятельно выбирать эффективные методы решения поставленных задач и разрабатывать новые методы для получения новых научных и прикладных результатов. методами математического моделирования для

	проведения научных исследований и разработок; навыками использования современных информационных технологий, конкретных программных продуктов и информационных ресурсов при проведении научных исследований;
ПК-4: Способен использовать современные методы разработки и реализации алгоритмов на базе языков программирования и пакетов прикладных программ моделирования	
ПК-4.1: Применяет знания современных методов разработки и реализации алгоритмов математических моделей на базе языков и пакетов прикладных программ моделирования при решении конкретных задач	
ПК-4.2: Разрабатывает и реализовывает алгоритмы математических моделей на базе языков и пакетов прикладных программ моделирования	

1.4 Особенности реализации дисциплины

Язык реализации дисциплины: Русский.

Дисциплина (модуль) реализуется без применения ЭО и ДОТ.

2. Объем дисциплины (модуля)

Вид учебной работы	Всего, зачетных единиц (акад.час)	е
		1
Контактная работа с преподавателем:	1 (36)	
занятия лекционного типа	0,33 (12)	
практические занятия	0,67 (24)	
Самостоятельная работа обучающихся:	1 (36)	
курсовое проектирование (КП)	Нет	
курсовая работа (КР)	Нет	

3 Содержание дисциплины (модуля)

3.1 Разделы дисциплины и виды занятий (тематический план занятий)

		Контактная работа, ак. час.							
№ п/п	Модули, темы (разделы) дисциплины	Занятия лекционного типа		Занятия семинарского типа				Самостоятельная работа, ак. час.	
				Семинары и/или Практические занятия		Лабораторные работы и/или Практикумы			
		Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС
1. Введение									
	1. Современная трактовка термина «большие данные» и ее развитие. Источники больших данных и особенности организации больших данных.	2							
2. Методы и средства анализа, применимые к обработке больших данных									
	1. Основы синтаксиса языка программирования Python 3.9.* и структуры хранения данных, применительно к большим данным.	2							
	2. Основы синтаксиса языка программирования Python 3.9.* и структуры хранения данных, применительно к большим данным.			10					
	3. Применение основных пакетов и модулей для обработки больших данных.	6							
	4. Применение основных пакетов и модулей для обработки больших данных.			12					

5. Влияние возможностей работы с большими объёмами данных на развитие математики и информационных технологий.	2							
6. Влияние возможностей работы с большими объёмами данных на развитие математики и информационных технологий.			2					
7. Работа за ПК по составлению компьютерных программ реализации алгоритмов обработки больших данных.							36	
Всего	12		24				36	

4 Учебно-методическое обеспечение дисциплины

4.1 Печатные и электронные издания:

1. Загоруйко Н. Г. Прикладные методы анализа данных и знаний: монография(Новосибирск: Институт математики СО РАН).
2. Айвазян С. А. 2011. т.2(М.: Маркет ДС).
3. Айвазян С. А. 2011. т.3(М.: Маркет ДС).
4. Быкова В. В. Дискретная математика с использованием ЭВМ: учебное пособие(Красноярск: Красноярский университет [КрасГУ]).
5. Быкова В.В. Теоретические основы анализа параметризованных алгоритмов: монография(Красноярск: СФУ).

4.2 Лицензионное и свободно распространяемое программное обеспечение, в том числе отечественного производства (программное обеспечение, на которое университет имеет лицензию, а также свободно распространяемое программное обеспечение):

1. R (The R Project for Statistical Computing)
2. Python
3. Visual Studio 2012,
4. система компьютерной вёрстки TeX,
5. MS Office,
6. Adobe Acrobat.

4.3 Интернет-ресурсы, включая профессиональные базы данных и информационные справочные системы:

1. научная библиотека СФУ <http://bik.sfu-kras.ru/>;
2. поисковые системы: Google или Яндекс.

5 Фонд оценочных средств

Оценочные средства находятся в приложении к рабочим программам дисциплин.

6 Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине (модулю)

Проектор для лекций и семинаров, классы для семинаров с компьютерами.