

Министерство науки и высшего образования РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

СОГЛАСОВАНО

Заведующий кафедрой

**Кафедра высшей и прикладной
математики (ВПМ_ИМФИ)**

наименование кафедры

подпись, инициалы, фамилия

«___» _____ 20__ г.

институт, реализующий ОП ВО

УТВЕРЖДАЮ

Заведующий кафедрой

**Кафедра высшей и прикладной
математики (ВПМ_ИМФИ)**

наименование кафедры

С.Г. Мысливец

подпись, инициалы, фамилия

«___» _____ 20__ г.

институт, реализующий дисциплину

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ПРИКЛАДНАЯ МАТЕМАТИКА
ТЕХНОЛОГИИ БОЛЬШИХ
ДАННЫХ**

Дисциплина Б1.В.ДВ.02.03.05 ПРИКЛАДНАЯ МАТЕМАТИКА
Технологии больших данных

Направление подготовки / 01.04.02 Прикладная математика и
специальность информатика Магистерская программа
01 04 02 01 Математическое моделирование

Направленность
(профиль)

Форма обучения очная

Год набора 2020

Красноярск 2021

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

составлена в соответствии с Федеральным государственным образовательным стандартом высшего образования с учетом профессиональных стандартов по укрупненной группе

010000 «МАТЕМАТИКА И МЕХАНИКА»

Направление подготовки /специальность (профиль/специализация)

Направление 01.04.02 Прикладная математика и информатика

Магистерская программа 01.04.02.01 Математическое моделирование

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

1 Цели и задачи изучения дисциплины

1.1 Цель преподавания дисциплины

Целью преподавания дисциплины является ознакомление студентов с методами и средствами хранения и работы с большими данными в научных исследованиях и решении прикладных задач, возникающих в гуманитарных и социально-экономических науках.

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

ПК-1:Способен разрабатывать и исследовать математические модели, методы и алгоритмы по тематике проводимых исследований	
Уровень 1	основные элементы процесса анализа больших данных, основные подходы к обработке больших массивов данных.
Уровень 2	возможности технологий анализа больших данных при проведении научных исследований.
Уровень 3	основные понятия, постановки задач, основные принципы и методы доказательства теорем и разработки алгоритмов анализа больших данных.
Уровень 1	выбрать подходящую технологию хранения и обработки больших данных, использовать современные системы хранения и обработки больших данных.
Уровень 2	использовать возможности отечественных и зарубежных универсальных программных средств и систем для аналитической обработки больших данных;
Уровень 3	ставить новые задачи в области прикладной математики и информатики, находить пути их решения, формулировать и доказывать теоремы, а также разрабатывать алгоритмы для написания компьютерных программ.
Уровень 1	методами интеллектуального анализа данных, в т.ч. методами оценки качества моделей, алгоритмов, методами экспериментальной

	проверки гипотез, методами обоснования гипотез.
Уровень 2	навыками выбора подходящего математического метода и программного обеспечения для решения прикладных задач.
Уровень 3	навыками использования современных информационных технологий, конкретных программных продуктов и информационных ресурсов при проведении научных исследований.

1.4 Место дисциплины (модуля) в структуре образовательной программы

Научно-исследовательская работа
 Научно-исследовательский семинар
 Современные компьютерные технологии

Научно-исследовательская работа
 Научно-исследовательский семинар
 Подготовка к сдаче и сдача государственного экзамена
 Преддипломная практика

1.5 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)	Семестр
		4
Общая трудоемкость дисциплины	4 (144)	4 (144)
Контактная работа с преподавателем:	1 (36)	1 (36)
занятия лекционного типа	0,5 (18)	0,5 (18)
занятия семинарского типа		
в том числе: семинары		
практические занятия	0,5 (18)	0,5 (18)
практикумы		
лабораторные работы		
другие виды контактной работы		
в том числе: групповые консультации		
индивидуальные консультации		
иная внеаудиторная контактная работа:		
групповые занятия		
индивидуальные занятия		
Самостоятельная работа обучающихся:	2 (72)	2 (72)
изучение теоретического курса (ТО)		
расчетно-графические задания, задачи (РГЗ)		
реферат, эссе (Р)		
курсовое проектирование (КП)	Нет	Нет
курсовая работа (КР)	Нет	Нет
Промежуточная аттестация (Экзамен)	1 (36)	1 (36)

3 Содержание дисциплины (модуля)

3.1 Разделы дисциплины и виды занятий (тематический план занятий)

№ п/п	Модули, темы (разделы) дисциплины	Занятия лекционного типа (акад. час)	Занятия семинарского типа		Самостоятельная работа, (акад. час)	Формируемые компетенции
			Семинары и/или Практические занятия (акад. час)	Лабораторные работы и/или Практикумы (акад. час)		
1	2	3	4	5	6	7
1	Введение	2	0	0	0	ПК-1
2	Методы и средства анализа, применимые к обработке больших данных	16	18	0	72	ПК-1
Всего		18	18	0	72	

3.2 Занятия лекционного типа

№ п/п	№ раздела дисциплины	Наименование занятий	Объем в академических часах		
			Всего	в том числе, в инновационной форме	в том числе, в электронной форме
1	1	Современная трактовка термина «большие данные» и ее развитие. Источники больших данных и особенности организации больших данных.	2	0	0

2	2	Методы и средства анализа, применимые к обработке больших данных: традиционные методы машинного обучения (ассоциативные правила, классификация, кластеризация и др.), с предварительным применением соответствующей предобработки данных.	6	0	0
3	2	Нетрадиционные методы анализа данных (поточковая обработка, онлайн-алгоритмы и т. п. на примере вычисления основных числовых моментов). Виды предобработки данных. Графические базы данных. Программно-аппаратные решения.	8	0	0
4	2	Влияние возможностей работы с большими объёмами данных на развитие математики и информационных технологий.	2	0	0
Всего			18	0	0

3.3 Занятия семинарского типа

№ п/п	№ раздела дисциплины	Наименование занятий	Объем в акад. часах		
			Всего	в том числе, в инновационной форме	в том числе, в электронной форме

1	2	Методы и средства анализа, применимые к обработке больших данных: традиционные методы машинного обучения (ассоциативные правила, классификация, кластеризация и др.), с предварительным применением соответствующей предобработки данных.	4	0	0
2	2	Нетрадиционные методы анализа данных (поточковая обработка, онлайн-алгоритмы и т. п. на примере вычисления основных числовых моментов). Виды предобработки данных. Графические базы данных. Программно-аппаратные решения.	12	0	0
3	2	Влияние возможностей работы с большими объёмами данных на развитие математики и информационных технологий.	2	0	0
Всего			18	0	0

3.4 Лабораторные занятия

№ п/п	№ раздела дисциплины	Наименование занятий	Объем в акад. часах		
			Всего	в том числе, в инновационной форме	в том числе, в электронной форме
Всего					

4 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

	Авторы, составители	Заглавие	Издательство, год
Л1.1	Быкова В. В.	Дискретная математика с использованием ЭВМ: учебное пособие	Красноярск: Красноярский университет [КрасГУ], 2006

Л1.2	Быкова В.В.	Теоретические основы анализа параметризованных алгоритмов: монография	Красноярск: СФУ, 2011
------	-------------	---	-----------------------

5 Фонд оценочных средств для проведения промежуточной аттестации

Оценочные средства находятся в приложении к рабочим программам дисциплин.

6 Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)

6.1. Основная литература			
	Авторы, составители	Заглавие	Издательство, год
Л1.1	Загоруйко Н. Г.	Прикладные методы анализа данных и знаний: монография	Новосибирск: Институт математики СО РАН, 1999
6.2. Дополнительная литература			
	Авторы, составители	Заглавие	Издательство, год
Л2.1	Айвазян С. А.	2011. т.2	М.: Маркет ДС, 2011
Л2.2	Айвазян С. А.	2011. т.3	М.: Маркет ДС, 2011
Л2.3	Маккинли У.	Python и анализ данных	Москва: ДМК Пресс, 2015
Л2.4	Коэбль Л. П.	Построение систем машинного обучения на языке Python	Москва: ДМК Пресс, 2016
6.3. Методические разработки			
	Авторы, составители	Заглавие	Издательство, год
Л3.1	Быкова В. В.	Дискретная математика с использованием ЭВМ: учебное пособие	Красноярск: Красноярский университет [КрасГУ], 2006
Л3.2	Быкова В.В.	Теоретические основы анализа параметризованных алгоритмов: монография	Красноярск: СФУ, 2011

7 Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля)

Э1	Городецкий В.И., Тушканова О.Н. Ассоциативная классификация: аналитический обзор. Ч. 1 // Труды СПИИРАН. – 2015. – № 38. – С. 183-203.	http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=trspy&paperid=782&option_lang=rus
----	--	---

Э2	Городецкий В.И., Тушканова О.Н. Ассоциативная классификация: аналитический обзор. Ч. 2 // Труды СПИИРАН. – 2015. – № 39. – С. 212-240	http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=trspy&paperid=798&option_lang=rus
Э3	Загоруйко Н.Г. Когнитивный анализ данных. □□Новосибирск: ИМ СО РАН, 2013. – 186 с.	http://www.mathnet.ru/php/person.phtml?option_lang=rus&personid=28250
Э4	Профессиональный информационно- аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных.	http://www.machinelearning.ru
Э5	Data Mining In Action	http://datamininginaction.ru/
Э6	Труды СПИИРАН	http://proceedings.spiiras.nw.ru/

8 Методические указания для обучающихся по освоению дисциплины (модуля)

В ходе лекционных занятий рассматриваются теоретические и концептуальные аспекты дисциплины, даются рекомендации для самостоятельной работы и подготовке к практическим занятиям.

В ходе практических занятий углубляются и закрепляются знания студентов по ряду рассмотренных на лекциях вопросов, развиваются навыки прикладного использования полученных знаний в рамках профессиональной деятельности.

Вопросы, не рассмотренные на аудиторных занятиях, должны быть изучены студентами в ходе самостоятельной работы. Контроль самостоятельной работы студентов над учебной программой курса осуществляется в ходе занятий методом устного опроса или посредством тестирования. В ходе самостоятельной работы каждый студент обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме, дополнить конспекты недостающим материалом, выписками из рекомендованных первоисточников. Выделить непонятные термины, найти их значение в энциклопедических словарях.

Текущий контроль предусматривает самостоятельную работу, задания на семинары.

Работа на семинаре предусматривает программную реализацию на компьютере предложенного алгоритма.

Самостоятельная работа предполагает работу за ПК по составлению компьютерных программ реализации алгоритмов обработки больших данных.

Итоговый контроль: экзамен. Проводится в устной форме или в письменной форме.

При проведении экзамена в устной форме он состоит из двух частей:

- теоретической, проводится в форме устной беседы по тематике дисциплины (30 мин.);
 - практической, связанной с обсуждением результатов самостоятельной работы (30 мин.).
- При проведении экзамена в письменной форме студенты предварительно знакомятся с примерными заданиями.

Студенты должны получить практические навыки использования современных инструментальных средств и ЭВМ при программной реализации математических моделей, рассматриваемых в рамках данного курса.

9 Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю) (при необходимости)

9.1 Перечень необходимого программного обеспечения

9.1.1	R (The R Project for Statistical Computing)
9.1.2	Python
9.1.3	Visual Studio 2012,
9.1.4	система компьютерной вёрстки TeX,
9.1.5	MS Office,
9.1.6	Adobe Acrobat.

9.2 Перечень необходимых информационных справочных систем

9.2.1	научная библиотека СФУ http://bik.sfu-kras.ru/ ;
9.2.2	поисковые системы: Google или Яндекс.

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

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